



OPEN ACCESS

EDITED BY

Fabien Sénéchal,
Université de Picardie Jules Verne, France

REVIEWED BY

Diana Moreira,
Centre for Biotechnology and Fine Chemistry
(CBQF), Portugal
Kaifeng Zheng,
Beijing Normal University, China

*CORRESPONDENCE

Els J. M. Van Damme
✉ elsjm.vandamme@ugent.be

RECEIVED 06 March 2025

ACCEPTED 23 April 2025

PUBLISHED 22 May 2025

CITATION

De Coninck T, Verbeke I, Rougé P, Desmet T
and Van Damme EJM (2025) OsAPSE
modulates non-covalent interactions
between arabinogalactan protein O-glycans
and pectin in rice cell walls.
Front. Plant Sci. 16:1588802.
doi: 10.3389/fpls.2025.1588802

COPYRIGHT

© 2025 De Coninck, Verbeke, Rougé, Desmet
and Van Damme. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

OsAPSE modulates non-covalent interactions between arabinogalactan protein O-glycans and pectin in rice cell walls

Tibo De Coninck¹, Isabel Verbeke¹, Pierre Rougé²,
Tom Desmet³ and Els J. M. Van Damme^{1*}

¹Department of Biotechnology, Laboratory for Biochemistry & Glycobiology, Ghent University, Ghent, Belgium, ²Unité Mixte de Recherche (UMR) 152 PharmaDev, Université Toulouse III Paul Sabatier, Institut de Recherche et Développement, Toulouse, France, ³Department of Biotechnology, Centre for Synthetic Biology, Ghent University, Ghent, Belgium

Flexibility of cell walls is crucial to accommodate cell elongation and growth, typically associated with the reorganization of cell wall polysaccharides. Seed germination is a fast-paced developmental process in which cell wall adaptability is highly required. The plant cell utilizes multiple strategies to obtain a flexible cell wall and in part relies on cell wall-active enzymes to loosen both covalent and non-covalent interactions between cell wall polysaccharides. OsAPSE is an example of a cell wall-active enzyme originating from Japanese rice (*Oryza sativa* subsp. Japonica) belonging to the glycoside hydrolase family 27 (GH27), potentially active on the pectin–arabinogalactan protein O-glycan junction. We provide insights into the biochemical and enzymatic properties of this protein, characterized by the presence of a GH27 domain linked to a ricin-B-like domain. Using small-scale production experiments in a cell-free protein synthesis system, we demonstrated the catalytic activity of the recombinant OsAPSE towards synthetic and natural substrates. Furthermore, subcellular localization analysis and *in silico* data suggest that OsAPSE may undergo unconventional secretion to the cell surface. We hypothesize that OsAPSE plays a role during rice seed germination by removing terminal α -D-Galp and β -L-Arap moieties along the pectin–arabinogalactan protein O-glycan network. This activity may abolish non-covalent interactions between pectic rhamnogalacturonan I and O-glycans of arabinogalactan proteins, contributing to cell wall relaxation for growth during germination.

KEYWORDS

cell wall, glycoside hydrolase, α -D-galactopyranosidase, β -L-arabinopyranosidase, *Oryza sativa*, rice, arabinogalactan protein, germination

1 Introduction

Glycoside hydrolases (GHs) are carbohydrate-active enzymes (CAZymes), catalyzing the hydrolytic cleavage of glycosidic bonds (Henrissat and Davies, 1997). Today, 1.9 million modules are classified in almost 190 GH families in the CAZy database (Drula et al., 2022). Members from the same GH family are evolutionarily related, show a conserved protein structure and act mechanistically similar on substrates.

The GH family of interest in this study is the GH27 family, which is present in every kingdom of life (Naumoff, 2004). Members of the GH27 family can display several activities, including α -D-galactopyranosidase (AGAL)/melibiase, N-acetylgalactosaminidase (NAGA) or β -L-arabinopyranosidase (ARAP) activity. In general, *bona fide* GH27 enzymes catalyze the hydrolysis of glycosidic bonds between α -D-galactopyranosyl (α -D-Galp), α -1,3-N-acetyl-D-galactosaminyl (α -D-GalNAc) and/or β -L-arabinopyranosyl (β -L-Arap) residues and other carbohydrates in a wide range of substrates (Supplementary File S1). Several GH27 enzymes are bifunctional proteins and display both AGAL and ARAP activity (Sakamoto et al., 2010; Kotake et al., 2016; Imaizumi et al., 2017; Kikuchi et al., 2017). This property is attributed to the structural similarities between α -D-Galp and β -L-Arap (Kotake et al., 2016), but also to the presence of conserved residues in the catalytic pocket of GH27 enzymes (Imaizumi et al., 2017) that make use of the Koshland double displacement mechanism and retain the anomeric configuration of the substrate upon hydrolysis (McCarter and Stephen Withers, 1994). The catalytic residues are aspartic acid residues and are strongly conserved within the GH27 family (Zhu et al., 1995; Hart et al., 2000; Ly et al., 2000; Garman et al., 2002; Fujimoto et al., 2003; Guce et al., 2010; Okazawa et al., 2015; Kytidou et al., 2018). Non-canonical activities have been reported sporadically, including glucan- α -1,6-isomaltosidase and galactan:galactosyltransferase activity (Supplementary File S1). These aforementioned activities, whether or not canonical, have also been observed in other GH families. Families GH27, GH31 and GH36 constitute the GH-D clan, a GH superfamily with mechanistic and structural resemblances (Comfort et al., 2007). In eukaryotes, the canonical activities are confined to the GH-D clan, while in prokaryotes these activities are also found in families outside the GH-D clan, *i.e.* in GH4, GH31, GH57, GH97, GH109, GH110 and GH129.

GH27 enzymes are of interest for various applications (Katrolia et al., 2014). In human medicine, several debilitating disorders, including Fabry, Schindler and Kanzaki disease, are associated with mutations in AGAL and NAGA genes, causing accumulation of glycosphingolipids and glycoproteins (Garman et al., 2002; Guce et al., 2010). Enzyme replacement therapy and gene therapy are employed to treat the aforementioned diseases (Kytidou et al., 2018; Umer and Kalra, 2023). Furthermore, NAGA and AGAL can be used to convert blood type A and B antigens respectively, to the universal donor type O blood (Rahfeld and Withers, 2020). In animal feed industry, AGALs are used to degrade raffinose family oligosaccharides (RFOs) in legumes, since non-ruminants are unable to digest RFOs (Di Stefano et al., 2007; Elango et al.,

2022). RFOs are fermented by gut bacteria, causing abdominal discomfort, flatulence and diarrhea (Mutuyemungu et al., 2023). In plants, GH27 enzymes have been implicated in several developmental processes including seed germination (Guimaraes et al., 2001; Blöchl et al., 2008; Jia et al., 2015; Lien et al., 2018; Arunraj et al., 2020; Zhang et al., 2021a; Okazawa et al., 2022; Gojto, 2023), fruit development (Soh et al., 2006; Tsaniklidis et al., 2016; Hua et al., 2021; Liu et al., 2022), and senescence (Chrost et al., 2006; Lee et al., 2009; Zhang et al., 2021b), but also in the response towards biotic (Evers et al., 2006) and abiotic stresses (Pennycooke et al., 2003; Tapernoux-Lüthi et al., 2004; Zhao et al., 2006; Gu et al., 2018; Chen et al., 2023). The physiological roles for GH27 enzymes are multifarious and mostly associated with AGAL/ARAP-mediated degradation of storage oligosaccharides/polysaccharides (*i.e.* RFOs, galactomannan) or modification of structural glycoconjugates (*i.e.* galactolipids, O-glycans of arabinogalactan proteins (AGP)).

GH domains occur often in combination with a carbohydrate-recognition domain (CRD), which supports their function as a catalyst by enhancing substrate binding (Boraston et al., 2004). In plants, GH27 sequences often encode multidomain proteins in which the catalytic domain is coupled to a carbohydrate binding module (CBM) of family 13 or a ricin-B(-like) domain (Van Holle et al., 2017; Van Holle and Van Damme, 2019; De Coninck et al., 2024b).

The subject of this study is OsAPSE, a GH27 enzyme from Japanese rice (*O. sativa* subsp. Japonica), which was named after its characterized homolog AtAPSE from *Arabidopsis thaliana* (Imaizumi et al., 2017). The goal of this study is to provide clues about the enzymatic properties of the GH27 domain towards synthetic and natural substrates, the biological function of this bifunctional enzyme in relation to seed germination and cell wall metabolism, and its occurrence and phylogeny in the plant kingdom.

2 Materials and methods

2.1 Cloning, protein production and analysis

2.1.1 Cloning of the GH27 domain of OsAPSE

The native coding sequence of the GH27 domain, flanked by 5' *NcoI* and 3' *KpnI* restriction sites, an N-terminal His₆-tag and double stop codon, was synthetically produced and cloned into a shuttle vector using the GeneArt Gene Synthesis service (Thermo Fisher Scientific, Waltham (MA), USA). The GH27 domain of OsAPSE was cloned into the *pALiCE02* expression vector for cell-free protein production (LenioBio GmbH, Düsseldorf, Germany) by means of a double restriction digest using 5 μ g shuttle vector or 5 μ g expression vector, 2.5 U *NcoI* and 2.5 U *KpnI* in 10X rCutSmart buffer (New England Biolabs, Ipswich (MA), USA) for 1 hour at 37°C, and 20 min heat inactivation of the restriction enzymes at 80°C. The double digests were purified using the QIAquick[®] PCR & Gel Cleanup Kit (Qiagen, Hilden, Germany). The GH27 insert was

ligated into the expression vector in a 3/1 insert-to-plasmid ratio using 5 U T4 DNA ligase (Thermo Fischer Scientific), 0.5 mM dithiothreitol (Thermo Fisher Scientific) and 10X ligase buffer (Thermo Fisher Scientific). The resulting expression plasmid was transformed into heat-shock competent *Escherichia coli* TOP10 cells (Thermo Fisher Scientific). Putatively transformed colonies were selected on lysogeny broth agar plates containing 80 µg/mL carbenicillin (Duchefa Biochemie, Haarlem, The Netherlands) and analyzed by colony PCR using Taq DNA polymerase (VWR, Radnor (PA), USA) and gene-specific primers (Supplementary File S2), with 5 min initial denaturation at 95°C, 35 cycles (30s at 95°C, 30s at 53°C, 1 min at 72°C) and 5 min final elongation at 72°C. Transformed TOP10 cells were propagated in lysogeny broth with 80 µg/mL carbenicillin and plasmids were purified at ultra-high purity using the NucleoBond Xtra Midi kit (Macherey-Nagel, Düren, Germany). Finally, the recombinant expression vector was analyzed by Sanger sequencing (Biosearch/LGC Genomics GmbH, Berlin, Germany) with plasmid-specific primers (Supplementary File S2).

2.1.2 Cell-free production of the GH27 domain of OsAPSE

Protein synthesis was executed using the ‘Almost Living Cell-free Expression’ (ALICE) cell-free production system (CFPS). Multiple reactions were initiated, by adding 500 ng of purified *pALiCE02::GH27_OsAPSE* per reaction at a final volume of 50 µL. Reactions with the *pALiCE02* empty vector were used as a control. After 48 hours of incubation at 25°C and 700 rpm on a thermomixer, the produced proteins were collected as described by the manufacturer’s protocol (Buntru et al., 2022). Due to the small scale of the CFPS reactions, no further purification was undertaken.

2.1.3 Protein analysis

Protein concentrations were determined using the Bradford assay (Bio-Rad, Hercules (CA), USA) (Bradford, 1976) with bovine serum albumin (BSA) (MP Biomedicals, Irvine (CA), USA) as reference protein (0–1 mg/mL) in 96-well plates using a TECAN Infinite 200 PRO (TECAN, Männedorf, Switzerland) plate reader.

Discontinuous acrylamide gels containing 0.01% SDS (MP Biomedicals) and different concentrations of acrylamide/bisacrylamide ROTIPHORESE® Gel 30 (37.5:1) (Carl Roth GmbH, Karlsruhe, Germany) in the stacking gel (pH 6.8, 4% acrylamide) and separating gel (pH 8.8, 15% acrylamide) respectively, were prepared. Polymerization was initiated with TEMED (Carl Roth GmbH) and 10 V% ammonium persulfate (Thermo Fisher Scientific). Protein samples were heat-treated (98°C) for 10 minutes with 4X sample buffer containing 1 M Tris-HCl pH 6.8 (MP Biomedicals), 8% SDS (MP Biomedicals), 40% glycerol (Chem-Lab), 0.4% bromophenol blue (Sigma-Aldrich, Saint Louis (MO), USA) and 1.125 M 2-mercaptoethanol (Sigma-Aldrich). Proteins were analyzed in a continuous electric field (180 V) for 1 hour in the presence of running buffer containing 25 mM Tris, 200 mM glycine (MP Biomedicals) and 0.1% SDS using a Mini-PROTEAN Tetra cell (Bio-Rad). Afterwards, acrylamide gels were

stained with acidic Coomassie solution containing 0.1% Coomassie Brilliant Blue R250 (Merck, Darmstadt, Germany), 2.9 M glacial acetic acid (Chem-Lab) and 10.2 M HPLC-grade methanol (Chem-Lab), and destained with acidic destaining solution, containing 2.5 M technical ethanol (Chem-Lab) and 1.3 M glacial acetic acid for 2–3 hours.

Western blotting on methanol-activated Amersham Hybond™-P PVDF membranes (GE Healthcare, Chicago (IL), USA) was performed by semi-dry electroblotting (Bio-Rad) in Towbin buffer containing 25 mM Tris, 2.45 M HPLC-grade methanol and 192 mM glycine. After blotting, the membranes were incubated in 5% non-fat milk powder solution (AppliChem GmbH, Darmstadt, Germany). Immunodetection was executed with subsequent incubation steps (1 hour at room temperature) in consecutively 1/5000 THE™ His-tag monoclonal antibody (GenScript, Piscataway (NJ), USA), 1/1000 polyclonal rabbit anti-mouse antibody conjugated with horseradish peroxidase (Agilent/DAKO, Santa Clara (CA), USA), 1/300 peroxidase anti-peroxidase antibody (Sigma-Aldrich) and final detection in 100 mM Tris-HCl pH 7.6 buffer containing 1 mM 3,3'-diaminobenzidine (DAB) (Thermo Fisher Scientific) and 320 µM H₂O₂ (Acros Organics, Geel, Belgium). Trissaline containing 10 mM Tris, 150 mM NaCl (Chem-Lab) and 0.1 V% Triton-X100 (Sigma Aldrich) was used as diluent for all antibodies and for membrane washes (3x5 min) in between antibody incubations.

2.2 Enzymatic assays

2.2.1 Experimental set-up

Different experimental set-ups were applied for the enzymatic assays including the initial screening for AGAL and ARAP activity, determination of the pH/temperature optima, determination of K_M and V_{max} and the activity on natural substrates (Table 1). The pH/temperature optima and enzymatic characteristics were determined using synthetic substrates, *i.e.* pNP- α -D-Galp and pNP- β -L-Arap (Sigma-Aldrich) and detection through absorbance measurements at 405 ± 10 nm using a TECAN Infinite 200 PRO plate reader. Activity assays on natural substrates made use of 50 mM melibiose monohydrate (Merck), 50 mM raffinose pentahydrate (Sigma-Aldrich), 50 mM verbascose (Megazyme, Wicklow, Ireland), 2.5 mg/mL arabinogalactan from larch wood (Sigma-Aldrich), 2.5 mg/mL carob bean galactomannan (Megazyme) and 2.5 mg/mL AGPs from *A. thaliana* PSB-D plant cell suspension cultures (Van Leene et al., 2011; Tryfona et al., 2012). Released α -D-Galp and β -L-Arap moieties were detected using the K-ARGA kit (Megazyme), which makes use of a galactose mutarotase and β -galactose dehydrogenase to convert L-Ara and D-Gal to their β -anomeric form and to oxidize the β -sugars to L-arabinonic acid, D-galactonic acid and NADH + H⁺. The amount of NADH formed is measured spectrophotometrically at a wavelength of 340 nm (Fukimura, 1988; Sturgeon, 1988) using a GENESYS150 UV/Vis spectrophotometer (Thermo Fisher Scientific) with 1.5 mL disposable 1 cm cuvettes (BRAND GmbH, Wertheim, Germany).

TABLE 1 Experimental setup for the different enzymatic assays.

Experiment	Initial activity screening	pH and temperature stability	Determination of kinetic parameters	Activity on di/oligosaccharides	Activity on polysaccharides
Substrate ^a	50 mM pNP- α -D-Galp or 50 mM pNP- β -L-Arap (100 μ L)	50 mM pNP- α -D-Galp (80 μ L)	1–100 mM pNP- α -D-Galp (120 μ L)	50 mM melibiose monohydrate 50 mM raffinose pentahydrate 50 mM verbascose (240 μ L)	2.5 mg/mL arabinogalactan from larch wood 2.5 mg/mL galactomannan 2.5 mg/mL arabinogalactan protein O-glycans from <i>A. thaliana</i> PSB-D (240 μ L).
Assay type	Continuous	Discontinuous	Discontinuous	Discontinuous	
Control reaction	<i>pALICE02</i> + pNP- α -D-Galp	<i>pALICE02</i> + pNP- α -D-Galp	<i>pALICE02</i> + pNP- α -D-Galp	<i>pALICE02</i> + pNP- α -D-Galp	
Total reaction volume	200 μ L	160 μ L	240 μ L	480 μ L	
Buffer	50 μ L 200 mM Tris-HCl pH 7.5	40 μ L 200 mM Tris-HCl (pH 6, 7, 7.5, 8, 9).	60 μ L 200 mM Tris-HCl (pH 8)	120 μ L 200 mM Tris-HCl (pH 8)	
Enzyme	50 μ L	40 μ L	60 μ L	120 μ L	
Temperature	22°C (room temperature)	25°C, 30°C, 35°C, 40°C	25°C	25°C	
Sampling volume	Not applicable	50 μ L		100 μ L	
Sampling points	Every 5 min for 2 h	5 min, 1 h, 2 h	5 min, 30 min, 1 h, 2 h	5 min, 30 min, 1 h, 2 h	
Inactivation	Not applicable	50 μ L 500 mM Na ₂ CO ₃ (pH 11) (Sigma-Aldrich)	50 μ L 500 mM Na ₂ CO ₃ (pH 11) (Sigma-Aldrich)	5 min at 95°C	
Assay	Absorbance at 405 \pm 10 nm			K-ARGA enzymatic/colorimetric kit (Megazyme), absorbance measurement at 340 nm.	

^aconcentrations correspond to initial concentrations. The final concentration equals half of the initial concentration.

2.2.2 Calculation of kinetic parameters

Initial reaction velocities (v_0) in mol·L⁻¹·s⁻¹ are calculated by measuring the release of pNP or NADH using the Lambert-Beer law (Equation 1).

$$v_0 = \frac{d[P]}{dt} = \frac{dA}{\epsilon \cdot L \cdot dt} \quad (1)$$

With: d[P]: increase of product concentration in mol·L⁻¹; dt: time coordinate in s; dA: increase of absorbance at 405 \pm 10 nm (pNP) or 340 nm (NADH), unitless; ϵ : molecular extinction coefficient (pNP: 18000 mol·L⁻¹·cm⁻¹; NADH: 6300 mol·L⁻¹·cm⁻¹); L: pathlength in cm.

The catalytic activity expressed as katal (1 kat = 1 mol·s⁻¹) is calculated by multiplication of the initial velocity with the reaction volume as indicated in Table 1. We define 1 unit (U) as the release of 1 μ mol product (*i.e.* pNP or NADH) in 1 minute. For the determination of K_M and V_{max} , the Hanes-Woolf linearization method was used (Hanes, 1932). The calculated the K_M and V_{max} were used to construct a theoretical Michaelis-Menten plot according to Equation 2 (Michaelis and Menten, 1913). The resulting hyperbole

was compared with the obtained experimental values for v_0 and the quality of the fit was evaluated by R² values.

$$v_0 = \frac{V_{max} \cdot [S]}{K_M + [S]} \quad (2)$$

With: [S]: substrate concentration in mol·L⁻¹; V_{max} : maximum reaction velocity in mol·L⁻¹·s⁻¹; K_M : Michaelis constant in mol·L⁻¹.

2.3 Protein modelling, molecular dynamics and phylogeny

2.3.1 Determination of substrate-binding affinities

Modeling of OsAPSE was performed with AlphaFold (Jumper et al., 2021) (RRID: SCR_025454), while the thermodynamic quality of the model was assessed in Swiss-Model (Waterhouse et al., 2018) (RRID: SCR_018123). Molecular structures of monosaccharides (L-Arap, D-Xylp, D-Glcp, D-Galp, D-GalNAc) and oligosaccharides

1 <https://pubchem.ncbi.nlm.nih.gov/>.

(melibiose, raffinose) used in the docking experiments, were retrieved from PubChem¹ (RRID: SCR_004284). Carbohydrates were docked to the GH27 domain of OsAPSE with SwissDock (Grosdidier et al., 2011; Bugnon et al., 2024) (RRID: SCR_022564), using the Attractive cavities method (Zoete et al., 2016; Röhrig et al., 2023). Binding affinities between the GH27 domain of OsAPSE and the carbohydrates were calculated. Molecular cartoons were drawn with the Chimera software (Pettersen et al., 2004) (RRID: SCR_004097).

2.3.2 Molecular dynamics simulations

The number of hydrogen bonds arising between the GH27 domain of OsAPSE and different carbohydrate structures was determined for simulations of 100 ns duration (Osterne et al., 2024). For these analyses, the AlphaFold structure of OsAPSE was used, combined with a set of additional carbohydrates compared to previous simulations, downloaded from PubChem or ChemSpider² (RRID: SCR_006360) (D-Galp, L-Arap, D-GalNAc, D-GlcNAc, pNP-D-Galp, pNP-L-Arap, melibiose, raffinose, stachyose, verbascose and ajugose). Carbohydrates were docked into the catalytic site of the GH27 domain using the GOLD software v2023 (RRID: SCR_000188) within the Hermes suite using standard settings (Jones et al., 1997; Verdonk et al., 2003). Docking grids of 7 Å and 12 Å, were established around the catalytic site for docking of monosaccharides and oligosaccharides respectively. CHEMPLP scores were used to evaluate the most favorable protein-carbohydrate interactions. CHARMM-GUI (RRID: SCR_025037) was used to prepare a TIP3P solution system, with neutralizing Na⁺ and Cl⁻ ions (Jorgensen et al., 1983; Jo et al., 2008; Brooks et al., 2009; Lee et al., 2020). The simulations were executed using the pmemd.cuda module of AMBER23 with the ff19SB and GLYCAM_06j force fields for the protein and carbohydrates, respectively (Kirschner et al., 2008; Tian et al., 2020; Case et al., 2023). Simulations were performed under isobaric (Monte Carlo barostat, 1 bar) and isothermal (Langevin thermostat, 300 K) conditions (Berendsen et al., 1984; Loncharich et al., 1992). Pressure and temperature were equilibrated by 500 ps in the NPT and NVT ensembles, respectively. Every simulation was run for 50 ns, collecting 5000 frames. Intermolecular hydrogen bonds were analyzed using of the Cpptraj module and visualized in Xmgrace (Roe and Cheatham, 2013).

2.3.3 Phylogenetic analyses

Phylogenetic analyses of GH27 and APSE sequences from plant species (*Viridiplantae*), grasses and cereals (*Poales*) and across kingdoms were executed. GH27 sequences were obtained through the Conserved Unique Peptide Patterns (CUPP) database (Supplementary File S3) (Barrett et al., 2020) (RRID: SCR_026501). The CANDY tool for carbohydrate active enzyme domain analysis was employed to analyze the protein domain modularity (Windels et al., 2024). Additional phylogenetic analyses based on the GH27 domain sequences were executed using phylogeny.fr (Dereeper et al., 2008) (RRID: SCR_010266).

Sequences were aligned using MUSCLE v3.8.31 (RRID: SCR_011812). Phylogenetic analysis was performed using the Maximum Likelihood method implemented in PhyML v3.1/3.0 aLRT (RRID: SCR_014629) within the phylogeny.fr pipeline. Tree reconstruction employed the WAG substitution model with empirical amino acid frequencies. Rate heterogeneity among sites was modeled using a gamma distribution and included a proportion of invariant sites. Amino acid frequencies were estimated from the alignment and used in the Maximum Likelihood calculations. Support for clades was evaluated using both bootstrap analysis and approximate likelihood-ratio tests. Trees were rendered using TreeDyn v198.3 (RRID: SCR_015946) and the resulting phylograms were visualized using the interactive Tree of Life v6.0 (Letunic and Bork, 2024) (RRID: SCR_018174) and formatted using inkscape v1.3.2 (RRID: SCR_014479). Multiple sequence alignments were performed in Clustal Omega (Madeira et al., 2024) (RRID: SCR_001591) and used as input to generate a WebLogo using WebLogo3 (Crooks et al., 2004) (RRID: SCR_010236). PyMOL v2.5.4 (RRID: SCR_000305) was used for structural comparisons of 3D models, either from crystallization data (PDB) or AlphaFold models obtained via UniProt (RRID: SCR_002380) (Supplementary File S4). Domain coordinates were extracted from InterPro (RRID: SCR_006695). Root-mean square deviation (RMSD) values were used to assess the structural alignment quality (Shindyalov and Bourne, 1998; Kufareva and Abagyan, 2011).

2.3.4 In silico prediction of biochemical protein properties

Biochemical protein properties including molecular weight, isoelectrical point, amino acid distribution, stability index and hydrophobicity index were calculated using ExPASy ProtParam (Gasteiger et al., 2003) (RRID: SCR_018087). Prediction of post-translational modifications (PTMs) was done through NetNGlyc v1.0 for N-glycosylation (Gupta and Brunak, 2002) (RRID: SCR_001570), NetOGlyc v4.0 for O-glycosylation (Steenftoft et al., 2013) (RRID: SCR_009026), diANNA v1.1 for disulfide bridges (Ferre and Clote, 2005) (RRID: SCR_018529), SignalP v6.0 for signal peptides (SP) (Teufel et al., 2022) (RRID: SCR_015644), TargetP v2.0 for other transfer peptides (Almagro Armenteros et al., 2019) (RRID: SCR_019022), NucPred for the presence of Nuclear Localization Signals (NLS) (Brameier et al., 2007) (RRID: SCR_026502) and DeepLoc v2.1 for membrane association (Ødum et al., 2024) (RRID: SCR_026503).

2.4 Expression of OsAPSE during rice seed germination

2.4.1 Cultivation of transgenic, mutant and wild type rice

Transgenic *O. sativa* subsp. Japonica cv. Kitaake lines were created by means of *Agrobacterium*-mediated transformation, including 3 overexpression *pUBI::OsAPSE* and 3 knock-out *osapse* lines. The overexpression *pUBI::OsAPSE* lines were generated using a binary vector harboring a hygromycin resistance gene, in which

² <https://www.chemspider.com/>.

the *OsAPSE* coding sequence is under the control of the constitutive maize ubiquitin promoter and a nopaline synthase terminator. Mutant *osapse* lines were generated using the CRISPR-Cas9 system, with 2 guide RNAs (gRNA) directed against the coding sequence of *OsAPSE*: 5'-CTTGCTGAGTTTCCACCAAGAGG-3' and 5'-CATCATCCAGAATTGATAAAGGG-3' (*i.e.* single gene, dual target) (Jiang and Doudna, 2017).

Rice seeds were de-husked with coarse sandpaper, sterilized by incubation on a rotary wheel (10 rpm) in 70% ethanol (Chem-Lab) for 5 min followed by 45 min in 5% commercial bleach (Carrefour supermarket), washed 7–10 times with sterile water and incubated overnight in sterile water on a rotary wheel. Afterwards, rice seeds were sown on Murashige and Skoog (MS) medium (pH 5.7–5.8) with modified vitamins (Duchefa Biochemie), 3% sucrose (Chem-Lab) and 1.5% micro agar (Duchefa Biochemie). Seeds were germinated for 10 days inside a controlled Adaptis growth cabinet (Conviron, Winnipeg (MB), Canada) at 28°C using a 16/8 photoperiod with photon flux density of 310 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$. Afterwards, rice seeds were brought to greenhouses of the Institute for Agriculture and Fishery Research (Instituut voor Landbouw en Visserijonderzoek) in Melle, Belgium (50° 59'35.667" N, 3°47'4.902" O) for seed multiplication. Rice plantlets were transferred from MS medium to general potting soil in 30 cm diameter pots. The rice plants were cultivated aerobically at 25–30°C and were watered daily using a tidal irrigation system. Additional iron and ammonium were supplemented with 0.18% FeSO_4 (Carl Roth GmbH) and 0.09% $(\text{NH}_4)_2\text{SO}_4$ (Chem-Lab) during the first weeks of growth. After 6–8 months, rice seeds were harvested and dried at 28°C for 2 weeks prior to further usage.

2.4.2 Characterization of transgenic and mutant rice plants

Wild type (WT), transgenic overexpression *pUBI::OsAPSE* and mutant *osapse* plants were grown as described above. At the age of 1 month, 3–4 cm samples of young rice leaves were collected in duplicate in sterile round-bottom safe-lock Eppendorf tubes and stored on dry ice during transportation and at -80°C until further usage. Rice material was ground using a Tissue Lyser II (Qiagen) with magnetic beads (\varnothing 3 mm) and prior cooling on liquid nitrogen. Afterwards, 0.1 g crushed leaf material was mixed with 1 mL DNA extraction buffer containing 2% hexadecyl-trimethyl ammonium bromide (CTAB) (Sigma-Aldrich), 0.1 M Tris-HCl, pH 7.5, 1.4 M NaCl and 2 mM Na_2EDTA (Sigma-Aldrich) followed by extraction using a mixture of chloroform (Chem-Lab) and isoamyl-alcohol (Carl Roth GmbH) in 24:1 ratio. Total genomic DNA (gDNA) was precipitated with 100% isopropanol (Chem-Lab) and washed with mixtures of 76% ethanol + 0.2 M NaOAc pH 8 (Merck) and 76% ethanol + 10 mM NH_4OAc pH 6 (Chem-Lab). The DNA pellet was dissolved in 50 μL sterile water and stored at -20°C.

PCR analyses using gDNA extracted from *pUBI::OsAPSE* plants, allowed to amplify a fragment of the hygromycin resistance gene using Taq DNA polymerase (VWR), 2 μL DNA with initial denaturation at 95°C for 5 min, 35 cycles (95°C for 30s, 52°C for 30s, 72°C for 30s) and final elongation at 72°C for 5 min

(Supplementary File S2). Similarly PCR using gDNA from mutant *osapse* plants, aimed to amplify the target region for CRISPR knock-out using ALLin™ Mega HiFi Red Mastermix (highQu GmbH, Kraichtal, Germany), 2 μL DNA with initial denaturation at 95°C for 5 min, 35 cycles (95°C for 30s, 60°C for 45s, 72°C for 30s) and final elongation at 72°C for 5 min (Supplementary File S2). The resulting PCR amplicons were purified using the QIAquick® PCR & Gel Cleanup Kit (Qiagen) and sequenced (Biosearch/LGC Genomics GmbH). Amplified sequences of WT and *osapse* plants were aligned to screen for mutations, caused by non-homologous end-joining after Cas9 endonuclease-mediated double-stranded breaks. The effect of mutations on the resulting polypeptides was assessed using AlphaFold (Jumper et al., 2021).

2.4.3 *OsAPSE* expression during rice seed germination

WT seeds, transgenic overexpression *pUBI::OsAPSE* and mutant *osapse* seeds from the F₃ generation were de-husked, sterilized and sown on non-selective MS medium as described above. The number of germinating and non-germinating (dead) seeds was counted at 1, 4, 7 and 11 days post imbibition (dpi), with 20 seeds per time point. Germination rates were calculated. Total seedling material, including roots, shoots and seeds from 8–10 plantlets per biological replicate were collected at 3, 7 and 10 dpi for the *pUBI::OsAPSE* and *osapse* lines, and at 1, 4, 7 and 11 dpi for WT, with minimum 3 biological replicates per sampling point. Different samples were used for RNA extraction and for germination assays. The *OsAPSE* transcript levels for WT at 4–7–11 dpi, *pUBI::OsAPSE* overexpression lines and mutant *osapse* lines at 3–7–10 dpi were correlated to the germination rates of WT, *pUBI::OsAPSE* overexpression lines and mutant *osapse* lines at 4–7–11 dpi. The germination rates at 1 dpi were excluded since *OsAPSE* transcript levels in *pUBI::OsAPSE* overexpression lines and mutant *osapse* lines were not determined at 1 dpi.

Plant material was crushed to a fine powder using a mortar, pestle and liquid nitrogen. All materials were decontaminated and rinsed between samples, using 70% ethanol and RNase AWAY (Thermo Fisher Scientific). Crushed materials were stored at -80°C until further usage. Total RNA was extracted using the Spectrum™ kit (Sigma-Aldrich), treated with RNase-free DNase I (Thermo Fisher Scientific) and RiboLock RNase inhibitor (Thermo Fisher Scientific) to remove co-extracted gDNA. Complementary DNA (cDNA) was synthesized from 500 ng RNA using Maxima Reverse Transcriptase (Thermo Fisher Scientific) according to the manufacturers' protocol. The obtained cDNAs were diluted 5 times in ultrapure water prior to further usage, and stored at -20°C. In between operations, RNA quality and quantity were analyzed using a NanoDrop2000 spectrophotometer (Thermo Fisher Scientific). RT-PCR for quality control of the cDNA samples was executed, amplifying reference genes as controls (Supplementary File S2), using 2 μL cDNA and Taq DNA polymerase with initial denaturation at 95°C for 5 min, 40 cycles (95°C for 30s, 58°C for 30s, 72°C for 30s) and final elongation at 72°C for 5 min. Finally, RT-qPCR was performed with 8 μL 4x diluted cDNA, 1 μL of each primer and 10 μL iQ™ SYBR® Green Supermix (Bio-Rad) in a CFX

Duet Real-Time PCR System (Bio-Rad) using the following amplification protocol: 95°C for 3 min followed by 42 cycles (95°C for 15s, 60°C for 25s, 72°C for 20s). Melting curves from 65°C to 95°C were generated with 0.5°C increments and fluorescence measurements every 5 s and analyzed using the CFX Maestro v2.3 software. The generated output was analyzed in qBase+ (Hellemans et al., 2007). The list of reference genes is included in [Supplementary File S2](#). Primers were designed using Primer3Plus (Rozen and Skaletsky, 1999) (RRID: SCR_003081). Primer amplification efficiency and stability (Bustin et al., 2009) were analyzed using the GeNorm algorithm in qBase+ (Vandesompele et al., 2002) (RRID: SCR_003370).

2.4.4 Analysis of agronomical traits

Agronomical traits such as number of (im)mature seeds per panicle, seed setting rate, panicle mass and seed mass were determined by counting and weighing F₃ seeds of WT, *osapse* and *pUBI::OsAPSE* plants. Panicles from 4–5 individual plants per line were used. Images of wild type, *pUBI::OsAPSE* and *osapse* panicles were taken using a Canon EOS 70D digital camera (Canon Inc., Shimomaruko (Tokyo), Japan) on a statue with fixed height.

2.5 Subcellular localization of OsAPSE

2.5.1 Transient expression of OsAPSE-EGFP in *Nicotiana benthamiana* leaves

The *OsAPSE* coding sequence was codon-optimized for expression in *N. benthamiana* and synthetically produced through the GeneArt Gene Synthesis service (Thermo Fisher Scientific). The *OsAPSE* sequence was cloned into the Gateway™-compatible (Invitrogen, Carlsbad (CA), USA) pK7FWG2 vector for C-terminal fusion with the Enhanced Green Fluorescent Protein (EGFP) (Karimi et al., 2002) under control of the constitutive 35S Cauliflower Mosaic Virus promoter, following the cloning procedure as described earlier (Van Hove et al., 2011). The expression plasmid was confirmed through sequencing (Biosearch/LGC Genomics GmbH) and transformed (300–500 ng of plasmid DNA) in electrocompetent *A. tumefaciens* EHA105 cells through electroporation (2.5 kV, 25 µF, 400 Ω, time constant 5–6 ms). *Agrobacterium* cells were selected on yeast extract medium containing 5 g/L beef extract (Lab M Ltd., Heywood, United-Kingdom), 5 g/L peptone (Merck), 1 g/L yeast extract (Merck), 5 g/L sucrose, 15 g/L bacterial agar (Thermo Fisher Scientific), 200 µg/mL rifampicin (Duchefa Biochemie) and 50 µg/mL spectinomycin (Duchefa Biochemie), for 2 days at 28°C. Putatively transformed *Agrobacterium* colonies were analyzed through colony PCR using Taq DNA polymerase and gene-specific primers ([Supplementary File S2](#)), with 5 min initial denaturation at 95°C, 35 cycles (30s at 95°C, 30s at 50°C, 2 min at 72°C) and 5 min final elongation at 72°C.

Recombinant *Agrobacterium* cells were cultured overnight at 28°C (180 rpm) in selective yeast extract broth. Similarly, *A. tumefaciens* GV3105 cells harboring the empty vector *pK7FWG2* plasmid as free-EGFP positive control was cultured. All *Agrobacterium* cells were cultured until OD₆₀₀ = 0.75 – 0.85.

Thereafter, cells were washed using infiltration medium (pH 5.6) containing 10 mM 2-(N-morpholino)-ethanesulfonic acid (Carl Roth GmbH), 2 mM Na₂HPO₄ (VWR), 0.5% glucose (Carl Roth GmbH), till OD₆₀₀ = 0.4 and prepared for infiltration by adding acetosyringone (Sigma Aldrich) to a final concentration of 100 µM.

The abaxial side of 3–5 weeks old *N. benthamiana* leaves (Bally et al., 2018) was transiently transformed with a suspension of *Agrobacterium* cells harboring either the *pK7FWG2::OsAPSE* plasmid or the empty vector *pK7FWG2* control (Sparkes et al., 2006). The *Agrobacterium* suspensions were administered using 2 mL syringes without a needle. Thereafter, the infiltrated area was highlighted with thin marker and the plants were incubated at 28°C for 2–3 days prior to microscopic analysis. Nuclear colocalization was visualized with 4',6-diamidino-2-phenylindole (DAPI) (Thermo Fisher Scientific), whereby a working solution at final concentration of 10 µg/mL was infiltrated 30 minutes prior to microscopy analysis.

2.5.2 Confocal fluorescence microscopy and image acquisition

A Nikon A1R confocal laser scanning microscope mounted on a Nikon Ti-E inverted epifluorescence body (Nikon instruments, Shinjuku, Japan) was used to capture confocal images. EGFP was excited at a wavelength of 488 nm using an argon laser and detected using an emission filter (515–530 nm). Microscopy analysis and image acquisition with the Fiji ImageJ software (Schindelin et al., 2012) were executed as described earlier (Dubiel et al., 2020).

2.5.3 In silico prediction of subcellular localization of OsAPSE

Subcellular localization was also predicted *in silico* using online web servers including MultiLoc2 (Blum et al., 2009) (RRID: SCR_003151), Plant-mPLoc (Chou and Shen, 2010) (RRID: SCR_023014), CELLO v2.5 (Yu et al., 2014) (RRID: SCR_011968), DeepLoc v1/v2.1 (Almagro Armenteros et al., 2019; Ødum et al., 2024) and MuLocDeep (Jiang et al., 2023) (RRID: SCR_026504).

2.6 Statistical analyses and visualizations

Statistical analyses were performed using SPSS v29.0 (RRID: SCR_002865). Throughout this study, significance levels at $p < 0.05$ were enforced. Comparison of means with v degrees of freedom (df) was executed using the Student's T-test for comparison between 2 samples or one-way ANOVA for comparison between >2 samples. Prior analysis of normality using the Shapiro-Wilk test and homogeneity of variance using the Levene's test was executed when applicable. In case the normality criterium was violated, the non-parametric Mann-Whitney U test or pairwise Kruskal-Wallis test were performed. The Welch test was executed when the homoskedasticity criterium was violated. Effect sizes for comparison of means was assessed based on the reported η^2 value, explaining a proportion of the observed variance per dataset. For the germination assays, a generalized linear model

(GLM) was fitted with a binomial distribution and logit link function to model the probability of seed germination as a function of transgenic line, time point and their interaction. For each time point, differences in germination rates between transgenic lines, knock-out lines and WT were assessed using separate binomial logistic regression models, with WT as the reference. Bar charts were generated by means of Microsoft Excel (RRID: SCR_016137) and reformatted in Inkscape v1.3.2. Diagrams were generated in BioRender (RRID: SCR_018361).

3 Results and discussion

3.1 Phylogeny of GH27 sequences from grasses and cereals

Most of the GH27 sequences found in the CUPP database belong to the taxonomical division of *Bacteria* (60.0%), followed by *Fungi* (24.1%), *Metazoa* (7.4%) and *Viridiplantae* (7.3%) (Supplementary File S3.1). A total of 140 GH27 sequences from grasses and cereals (*Poales*) were retrieved. After removal of 38 duplicate sequences or gene fragments (Supplementary File S3.2), the CANDY tool was employed for modularity analysis (Windels et al., 2024). Several recurring InterPro domains were found across the GH27 sequences (Supplementary File S3.3). All sequences were attributed with the GH27-related IPR002241 and GH superfamily-related IPR017853 identifiers. The IPR041233 and IPR013780 identifiers both refer to the same domain, although the IPR041233 identifier is used specifically for MELs and the IPR013780 identifier applies to GH-all-beta domains in general. Almost every retrieved GH27 sequence contained the C-terminal GH-all-beta domain, which is a common terminal domain for a number of GH families, including GH5, GH13 and GH42. Structurally, this C-terminal domain resembles a Greek key β -sandwich (Hutchinson and Thornton, 1993). Furthermore, several sequences were provided with the IPR035992 identifier describing ricin-B-like lectin domains (Van Holle et al., 2017; De Coninck et al., 2024b).

Phylogenetic analyses based on the full-length sequences of GH27 enzymes were executed and enabled the identification of distinct clades (Figure 1A). The gamma shape parameter was estimated at $\alpha = 4.44$ indicating moderate rate variation across sites. Approximately 9.4% of the sites were inferred to be invariant. Intriguingly, ricin-B-like domains have only been identified in a subpopulation of the GH27 sequences from *Poales*. As shown in the phylogram (Figure 1A), ricin-B-like domains are present only in the so-called “APSE clade”, which is distant from other GH27 sequences without a ricin-B-like domain. This is also observed when GH27 sequences across kingdoms are studied (Supplementary File S3.1). Members within the APSE clade in *Poales* species show high sequence similarity (median >70%; Supplementary File S3.4) towards the characterized APSE from *Arabidopsis thaliana* (Imaizumi et al., 2017) and have highly conserved amino acid sequences, judging from weblogs

(Supplementary File S3.5) and multiple sequence alignments (Supplementary File S3.6).

The presence of a ricin-B-like domain is a distinctive trait to categorize the GH27 family. Almost all members from the APSE clade have been designated as CUPP group GH27:14, while all other GH27 sequences from *Poales* are found in CUPP groups GH27:6 and GH27:23 (Figure 1A). Belonging to different CUPP groups indicates that unique peptide patterns exist amongst APSE proteins and other GH27 proteins. These peptide patterns are highly conserved across species and are indicative for a unique protein structure and biological function (Barrett et al., 2020). Next to the APSE clade, three other clades have been identified, numbered with Greek letters I, II and III, and coincide mostly with CUPP groups GH27:6 and GH27:23. The apparent separation of the APSE clade from the other clades is not only attributed to the ricin-B(-like) domain, but also to the GH27 domain itself. When phylogenetic analysis is executed only on the GH27 domains of *Poales*, the APSE clade remains phylogenetically isolated from the other GH27 sequences, but the formation of clades I, II and III is now lost (Supplementary File S3.7), emphasizing the highly dissimilar nature of the APSE GH27 domain compared to other GH27 sequences.

Striking differences are apparent between GH27 domains from the APSE clade compared to GH27 domains from clades I, II and III. Although GH27 domains from clades I, II and III are mutually remarkably similar, they contain several large gaps compared to GH27 domains from the APSE clade (Figure 1B, Supplementary File S3.8). Despite the important differences between APSE and clade I, II, III GH27 sequences, the catalytic residues are conserved, which is a known characteristic of the GH27 family (Figure 1C, Supplementary File S3.8) (Hart et al., 2000; Ly et al., 2000; Garman et al., 2002; Fujimoto et al., 2003; Guce et al., 2010).

The dissimilarities between APSE members and other GH27 sequences are also present at the structural level. Structural comparison between OsAPSE and a selection of GH27 AGAL AlphaFold or PDB models from *Z. mays*, *T. aestivum*, *O. sativa* and *S. viridis* (Supplementary File S4.1) yields bad structural alignments with mean RMSD = 5.97 ± 0.06 Å. Surprisingly, also the alignment between OsAPSE and the other OsAGALs from rice yields bad structural alignments with mean RMSD = 4.57 ± 1.54 Å (Figure 1D) (Supplementary File S4.2). Likewise, poor RMSD values (5.37 ± 1.15 Å) are obtained when the OsAPSE GH27 domain was compared to the GH27 domains of characterized GH27 proteins from other kingdoms (Supplementary File S4.1). In contrast, structural comparison between OsAPSE and APSE sequences from the characterized *A. thaliana* AtAPSE (Figure 1E) and *H. vulgare*, *A. comosus*, *Z. mays*, *Triticum monococcum* (Figure 1F) yields close structural alignments (0.74 ± 0.51 Å) (Supplementary File S4.2).

3.2 Modelling and molecular dynamics of OsAPSE

3.2.1 Modelling of OsAPSE

The OsAPSE protein is composed of 3 domains: an N-terminal GH27 catalytic domain, linked by a long loop (L1) to a β -trefoil

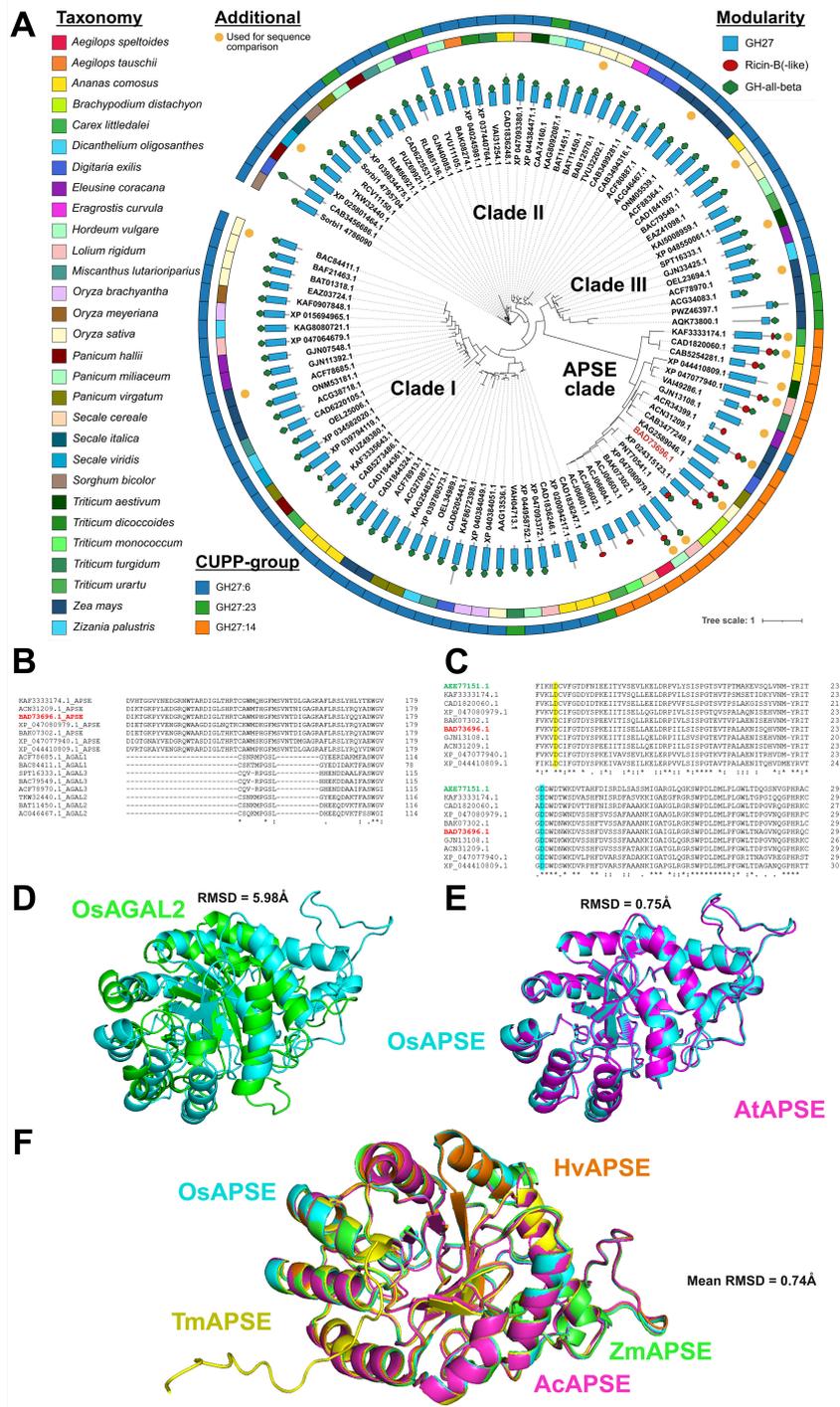


FIGURE 1

Phylogenetic and structural analysis of GH27 domain sequences from *Poales*. Modularity of the full-length GH27 sequences, represented by their corresponding GenBank IDs, is displayed by blue rectangles (GH27 domain), red circles (Ricin-B-(like) domain) or green hexagons (GH-all-beta domain). The outer strips represent taxonomy and CUPP-group and are colored as indicated in the legend. OsAPSE (GenBank ID: BAD73696.1) is highlighted in red. GH27 domain sequences used in the sequence alignment are highlighted with an orange circle (A). Partial alignment of a selection of GH27 domains from the APSE clade and defined clades I, II and III denoted as AGAL1, AGAL2 and AGAL3 (B). Partial alignment of GH27 domains from the APSE clade with the characterized APSE domain from *Arabidopsis thaliana* (highlighted in green) and OsAPSE (highlighted in red), illustrating the conservation of the catalytic residues. The catalytic nucleophile is highlighted in yellow and the catalytic acid/base in cyan (C). Explanation of symbols: gap (-), conserved residue (.), highly conserved residue (:), identical residue (*). The OsAPSE sequence is highlighted in red. Structural alignment of OsAPSE (cyan) with OsAGAL2 (green; GenBank ID: BAC79549.1) (D). Structural alignment of OsAPSE (cyan) with AtAPSE (magenta) (E). Multiple structural alignments of OsAPSE (cyan) with other APSE clade members of *Hordeum vulgare* (orange; GenBank ID: BAK07302.1), *Zea mays* (green; GenBank ID: ACN31209.1), *Ananas comosus* (magenta; GenBank ID: CAD182006.1) and *Triticum monococcum* (yellow; GenBank ID: ACJ06602.1) (F). RMSD values were calculated using the cealign algorithm.

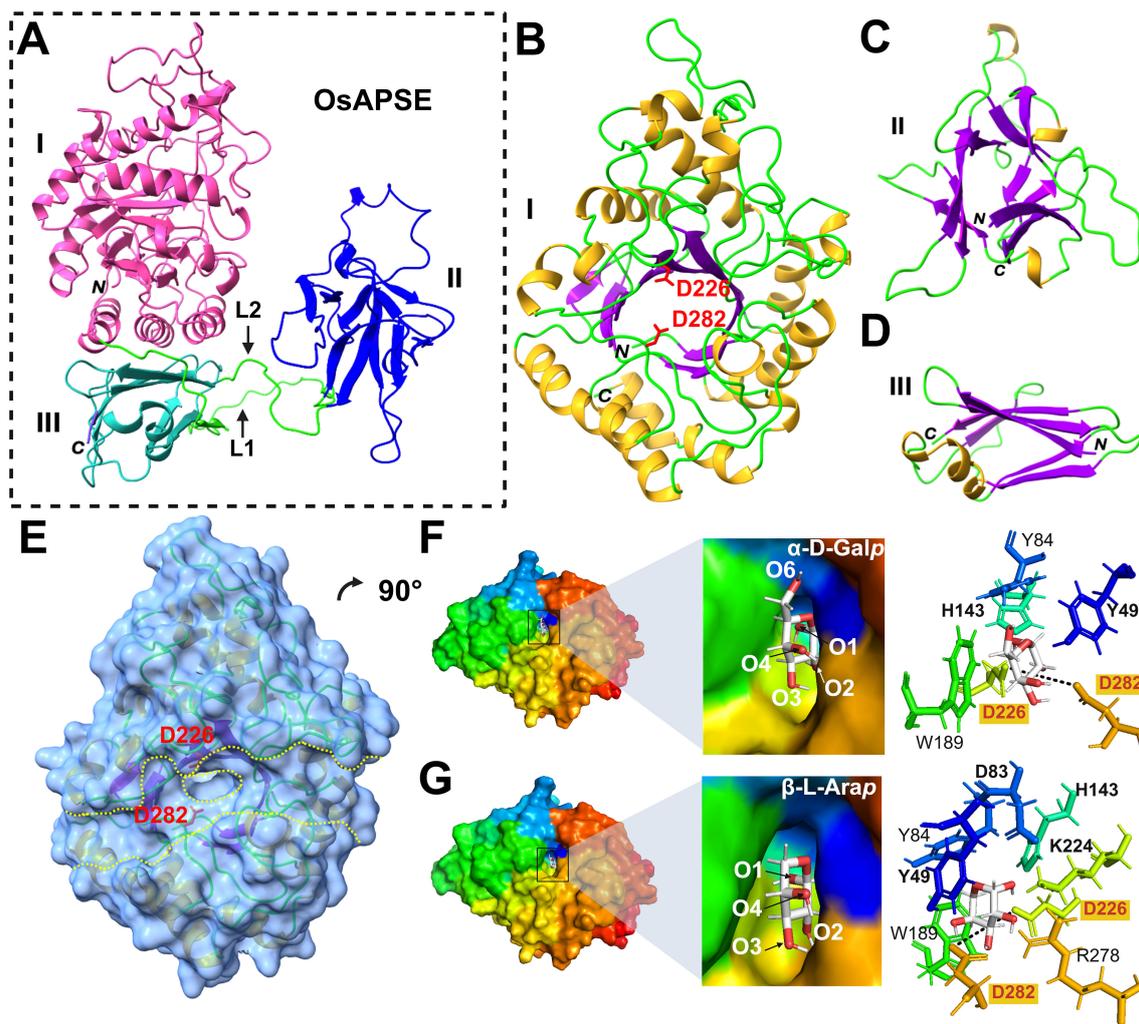


FIGURE 2

Structure of OsAPSE and interaction with monosaccharide substrates. OsAPSE is a 3-domain protein (A) comprising of a GH27 domain (B) with a catalytic cleft (E). The other subdomains are the ricin-B-like domain (C) and the GH-all- β domain (D). The Roman numbers I, II and III enumerate the subdomains of OsAPSE. L1 and L2 are two linkers connecting the GH27 domain with the ricin-B-like domain and the ricin-B-like domain and the GH-all- β domain respectively. The capital italic letters *N* and *C* indicate the *N* and *C*-terminal ends of OsAPSE and its subdomains. OsAPSE is shown without its native signal peptide. The catalytic residues D226 and D282 are indicated in red in sub-figures (B, E). In sub-figure (E), the protein surface is represented with a transparency of 40% to show localization of α -helices, β -strands and catalytic residues D226 and D282 in the vicinity of the entry of the active site (encircled yellow dashed line) within the catalytic groove of the domain (parallel yellow dashed lines). The monosaccharide substrates α -D-Galp (F) and β -L-Arap (G) were docked in the catalytic site of the GH27 domain. The black dashed line is indicative for the distance between the two catalytic residues D226 and D282 and measures 6.8 Å.

ricin-B-like domain with a putative CRD, and a C-terminal GH-all- β domain of unknown function linked by a shorter loop (L2) to the ricin B(-like) domain (Figure 2A). Modeling of the structure of OsAPSE resulted in a low Qualitative Model Energy Analysis (QMEAN) value of 0.54. The low QMEAN value mostly depends on the occurrence of extended loops L1 and L2 connecting the ricin-B-like domain with the *N*-terminal GH27 and *C*-terminal GH-all-beta domain, respectively (Benkert et al., 2011; Studer et al., 2020). Modelling of extended loops typically gives conformations of poorly reliable geometric and thermodynamic quality (Fiser et al., 2000), in our case accounting for 1.85% and 1.78% of the Ramachandran and rotamer outliers, respectively. In addition, the QMEAN value of the individual GH27 domain was determined, yielding a QMEAN value

of 0.69. The reported QMEAN value is considerably higher compared to the value of the complete OsAPSE protein, due to the presence of rigid α -helices and β -sheets in the TIM barrel of the GH27 domain. However, the value was somewhat lowered due to the presence of flexible loops connecting the α -helices and β -sheets.

The catalytic domain exhibits a canonical TIM ($\alpha_8\beta_8$) barrel structure, made of a central crown of 8 β -strands, linked by short loops to a peripheral crown of 8 α -helices (Figure 2B), which is a common structure in enzymes (Vega et al., 2003). The catalytic cleft occupies the center of the TIM barrel and two aspartic acid residues (D226 and D282) located at the center of the catalytic cleft, form the active site of the enzyme (Figure 2B). These catalytic residues are highly conserved (Figure 2E) within the GH27 family (Zhu et al.,

1995; Hart et al., 2000; Ly et al., 2000; Garman et al., 2002; Fujimoto et al., 2003; Guce et al., 2010; Okazawa et al., 2015; Kytidou et al., 2018). The ricin-B(-like) domain comprises 3 bundles of β -sheets organized in a typical β -trefoil lectin structure with putative carbohydrate-binding activity (Figure 2C) (Hazes, 1996; Steeves et al., 1999). The short C-terminal all-beta domain is made of 2 anti-parallel β -sheets forming a β -sandwich (Figure 2D). The function of all-beta domains is multifarious: they provide structural stability and assist in protein folding by acting as a nucleation site (Boissinot et al., 1997; Kemplen et al., 2015).

3.2.2 Docking of substrates and molecular dynamics

Because of its close structural similarity to AtAPSE, it was expected that OsAPSE will display AGAL and ARAP activity against carbohydrate structures from the cell wall. Therefore, it was hypothesized that the main substrates for OsAPSE would be molecules with α -D-Galp and/or β -L-Arap side chains. Docking experiments performed with the GH27 domain (Supplementary File S5.1) showed that all the assayed mono- and oligosaccharides bind to the catalytic pocket via a network of hydrogen bonds with the catalytic residues D226 and D282 and surrounding hydrophilic residues (N47, D83, H143, K224, S255, S257, R278, D318, D320, M321), although these residues varied in number and type depending on the ligand (Supplementary File S5.1). The substrates α -D-Galp and β -L-Arap were docked in stable chair conformation. For α -D-Galp (Figure 2F), O1/2, and O3/4 are predicted to make contact with the catalytic residues D226 and D282, respectively. In addition, stacking interactions between aromatic residues Y49 (O4/6), Y84 and W189 located around the catalytic pocket, and the pyranose ring of the saccharides complete and reinforce the interaction with the GH27 domain (Asensio et al., 2013; Spiwok, 2017). Similar stacking residues were observed in other GH27 enzymes from fungi (Brumer et al., 1999), chicken (Garman et al., 2002), human (Garman and Garboczi, 2004; Guce et al., 2010), rice (Fujimoto et al., 2003) and tobacco (Kytidou et al., 2018). It was expected that more or less the same residues would be involved in substrate binding to α -D-Galp and β -L-Arap (Ichinose et al., 2009). However, despite β -L-Arap being a smaller molecule, the *in silico* docking yielded two additional residues, D83 (O1) and K224 (O2), to be involved in substrate binding next to Y49 (O4), H143 (O1), D226 (O1) and D282 (O2/3) (Figure 2G). These results showcase the ability of the OsAPSE GH27 domain to accommodate substrates with terminal α -D-Galp and β -L-Arap residues.

GH27 enzymes, including OsAPSE, adhere to the classical Koshland double-displacement retaining mechanism, characterized by two consecutive displacement steps resulting in the retention of the anomeric configuration of the released sugar (Sinnott, 1990). In OsAPSE, D282 functions as the general acid, while D226 serves as the catalytic nucleophile. Upon substrate binding, D282 protonates the aglycone, facilitating glycosidic bond cleavage and generating an oxocarbenium ion-like transition state. Subsequently, the carboxylate group of D226 attacks the C1 atom of the sugar moiety, forming a covalent galactosyl/arabinosyl-OsAPSE intermediate (Vocadlo et al., 2001). In the second step, D282 deprotonates a water molecule,

activating it for nucleophilic attack on C1, leading to a second oxocarbenium ion-like intermediate. This results in the cleavage of the catalytic bond between D226 and the galactosyl/arabinosyl group, releasing free galactose or arabinose, and restoring the enzyme to its initial state.

All saccharides displayed energetically favorable interactions ($\Delta G < 0$), within the same range ($K_D = 8\text{--}59\ \mu\text{M}$), (Supplementary File S5.1), even with carbohydrates for which no particular interaction was expected, such as D-Glcp and D-Xylp. The binding affinity for D-Glcp is lower compared to other saccharides. The limited differences in binding affinity between the assayed carbohydrates is probably due to the strong structural resemblance between D-Galp and D-Glcp, which are only differing in their configuration at C4 (Homolak et al., 2024), and similarly, L-Arap and D-Xylp are only different at the C4 anomeric configuration (Guo et al., 2025). However, it should be emphasized that the possibility to dock alternative ligands into the catalytic site does not necessarily mean that OsAPSE will cleave off these moieties. The structural flexibility of carbohydrates during molecular dynamics simulations is often exaggerated and may distort genuine protein-carbohydrate interactions (Boonstra et al., 2016), hence experimental validation is always required (Lerbret et al., 2007). Docking of carbohydrates to the ricin-B-like domain was outside the scope of this study.

Additional saccharides with varying degree of polymerization (DP) (*i.e.* stachyose, verbascose, ajugose, GalNAc, pNP- α -D-Galp and pNP- β -L-Arap) were docked in the catalytic pocket of the GH27 domain of OsAPSE, within a docking grid of 7Å and the number of hydrogen bonds during molecular dynamics simulations was determined (Supplementary File S5.2). Most stable interactions were observed for the smaller substrates L-Arap, D-GalNAc, pNP- α -D-Galp and D-Galp, while the interactions were least stable for verbascose (DP = 5) and ajugose (DP = 6). It should be emphasized that the results should be interpreted with reservation, as ligand docking and molecular dynamics simulations were performed with modelled protein structures.

3.3 Production and activity of the recombinant GH27 domain of OsAPSE

Several attempts have been undertaken to produce OsAPSE or its GH27 domain recombinantly in multiple prokaryotic and eukaryotic hosts and strains, under a wide range of experimental conditions and construct designs, but these assays were mostly unsuccessful due to protein insolubility. The lack of soluble recombinant proteins is widely recognized as a main bottleneck in proteomics research (Bhatwa et al., 2021; Beygmoradi et al., 2023). CFPS platforms derived from cell lysates can be considered when cell-based strategies are inadequate (White et al., 2013). Although, prokaryotic CFPS systems are often preferred due to the low production costs, high productivity and scalability, they may not be the platform of choice for eukaryotic proteins due to the absence of appropriate PTMs and chaperones, which impact protein folding, structure and activity (Harbers, 2014; Zemella

et al., 2015). About 10 years ago a CFPS based on tobacco BY-2 lysates was developed, facilitating oxidative folding, PTMs and assembly of multidomain enzymes/antibodies (Buntru et al., 2014, 2015). The performance of this platform was already demonstrated by successfully producing an array of eukaryotic glycoproteins with disulfide bridges and proper *N*-glycosylation (Buntru et al., 2022).

3.3.1 Cell-free production of the GH27 domain of OsAPSE and screening for GH27 activity

The GH27 domain of OsAPSE consists of 352 amino acids and has a predicted molecular weight of 39.9 kDa and pI = 5.99.

Furthermore, the protein is predicted to be stable (instability index = 34.79) and moderately hydrophilic (Grand Average of Hydropathicity index = -0.365). Protein concentrations were estimated (Figure 3A) using BSA as reference protein (Supplementary File S6). GH27_OsAPSE was detected after Western blot analysis (Figure 3B). A distinct and unique protein band can be observed for the reaction with *pALiCE02::GH27_OsAPSE*. The protein polypeptide appears somewhat smaller (< 5 kDa) compared to the predicted size, but is not attributed to protease activity (Buntru et al., 2014). Deviating protein sizes are sometimes observed and are attributed to their charge distribution and more compact protein folding (Rath et al.,

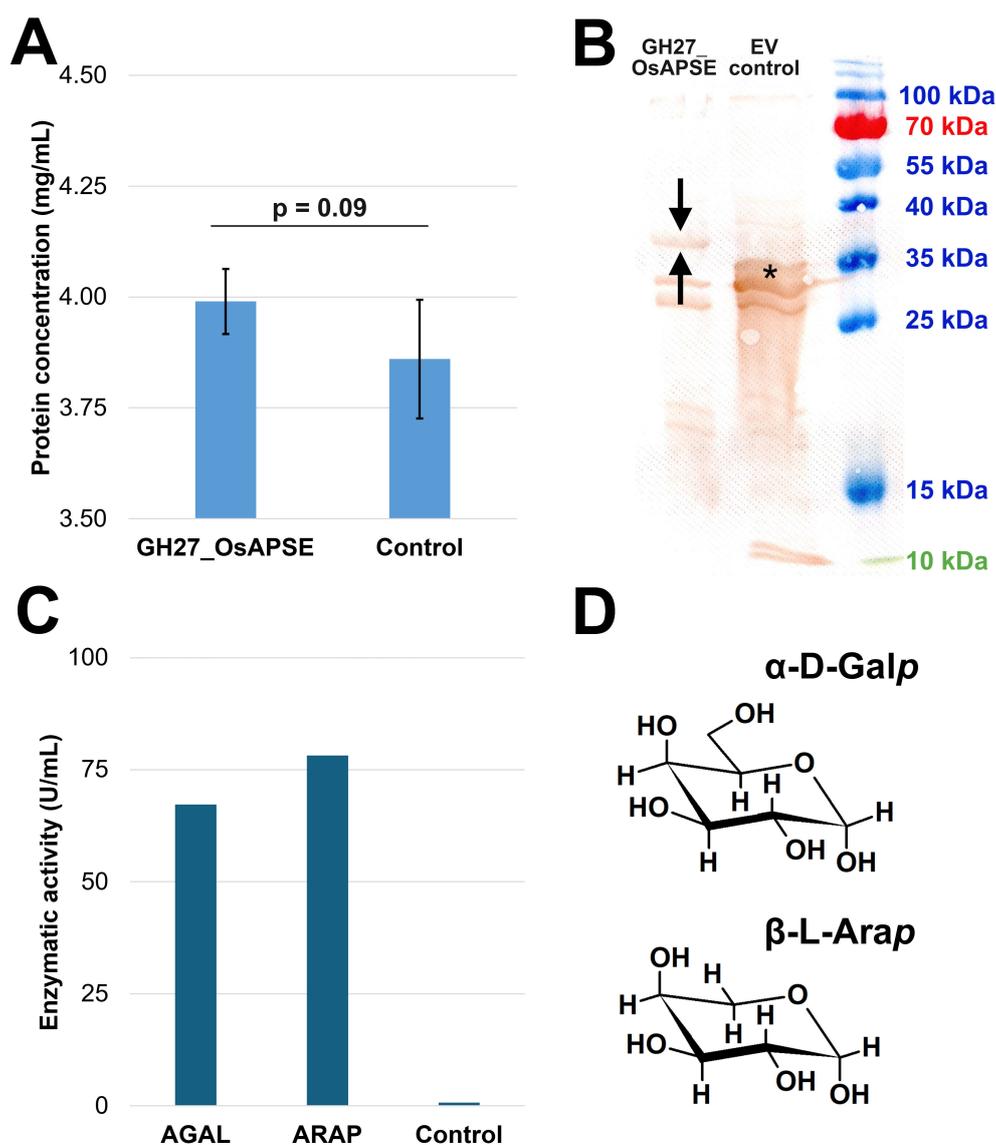


FIGURE 3

Quantification and screening for GH27 activities. Proteins were quantified using the Bradford assay. Error bars represent standard deviations based on four independent biological replicates (A). Western blot with DAB detection using anti-His₆ antibodies. The arrows indicate the protein of interest (i.e. GH27_OsAPSE) while the asterisk indicates the yellow fluorescent protein, present in the control reaction as reporter (B). AGAL and ARAP activity expressed in U/mL, measured in a continuous assay. The control is the CFPS reaction with empty vector *pALiCE02* incubated at 25°C at pH 7.5 (C). Structural comparison between α -D-Galp and β -L-Arap. Structural images were drawn using ChemSketch (D). Abbreviations: EV (empty vector).

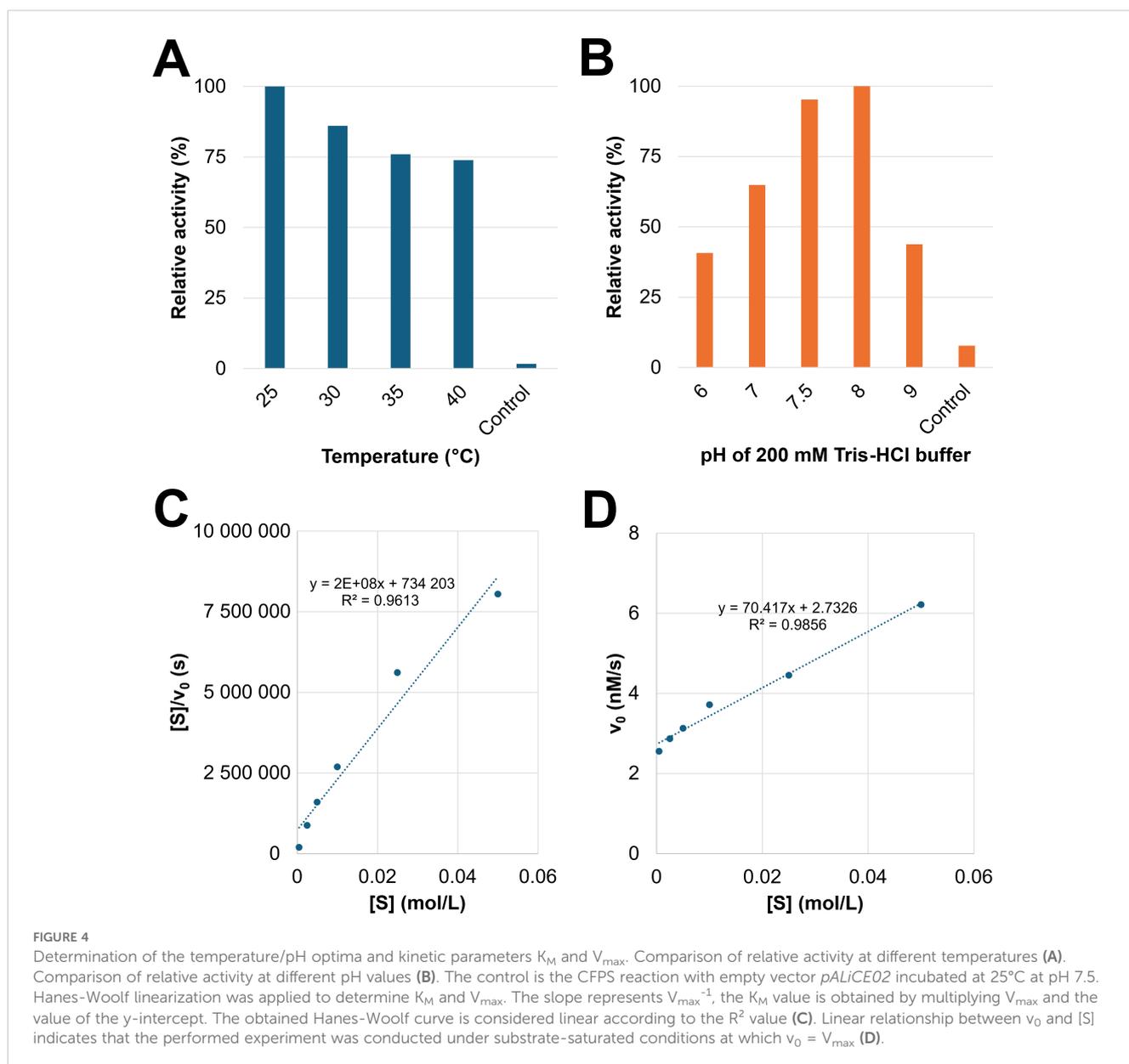
2009; Shi et al., 2012). The *pALiCE02* plasmid used in the control reactions contains a His₆-tagged yellow fluorescent protein (YFP) reporter sequence, yielding a polypeptide with estimated size of 33 kDa (Figure 3B). The YFP sequence is removed from the *pALiCE02::GH27_OsAPSE* construct during the cloning process (Supplementary File S7).

A continuous enzymatic assay detected both AGAL and ARAP activity (Figure 3C) in the CFPS protein fractions (Supplementary File S8.1). This is not surprising since α -D-Galp and β -L-Arap are structurally very similar (Figure 3D) (Kotake et al., 2016), and both fit in the active site of the GH27_OsAPSE domain (Figure 2). OsAPSE shows high sequence identity (65%) (Supplementary File S3.4) and structural similarity (RMSD = 0.75Å) towards the characterized AtAPSE (Figure 1E), although the latter mainly demonstrated ARAP activity. However, it has been reported that AGALs may display both ARAP and AGAL activity (Imaizumi et al., 2017).

Absorbance measurements were used to calculate reaction velocities and activities (Equation 1). The reactions with *pALiCE02::GH27_OsAPSE* yielded an enzymatic AGAL and ARAP activity of 67.2 U/mL and 78.2 U/mL, respectively, whereas the control reaction with *pALiCE02* only released negligible pNP moieties (Figure 3C), most probably due to spontaneous degradation over the course of the enzymatic assay. Background activity originating from deglycosylases should not be present, as it was already demonstrated before that the ALiCE CFPS system gives rise to intact N-glycans (Buntru et al., 2022).

3.3.2 Determination of temperature/pH optima and enzyme kinetics

Discontinuous enzymatic assays with pNP- α -D-Galp revealed the temperature optimum of GH27_OsAPSE at 25°C (Figure 4A) and pH optimum at pH = 8 (Figure 4B) (Supplementary File S8.2).



The relative activity decreases with increasing temperature. At temperatures around 50°C, the reaction mixtures turned opaque, due to protein denaturation and precipitation. However, the relative activity at various pH conditions shows a typical bell-shaped curve, with >95% of the activity retained between pH 7.5–8 and sharp decline when deviating from the optimum.

GH27 enzymes from plant origin are often categorized based on their pH optimum (Supplementary File S8.5), as there are acidic and alkaline AGALs present in plants. Plant AGALs usually display pH optima around 4.5–8.5 and temperature optima around 30–40°C although there are also AGALs with somewhat extreme optima, for instance in fava bean ($\text{pH}_{\text{opt}} = 2.5$) (Dey and Pridham, 1972) and maize and melon ($\text{pH}_{\text{opt}} = 8.5$) (Gao and Schaffer, 1999; Zhao et al., 2006). However, the pH_{opt} of OsAPSE (7.5–8) is not in accordance with its supposed biological environment, being the apoplast with typical apoplastic pH values (pH_{apo}) between 5.5–6, although pH_{apo} values as low as 3.5 and as high as 8.5 have been reported before (Yu et al., 2000). The pH_{apo} in rice tissues has not been reported yet but is likely to be within the same range as closely related organisms such as barley leaves ($\text{pH}_{\text{apo}} = 5.6$ – 6.6) and maize coleoptiles ($\text{pH}_{\text{apo}} = 5.7$ – 6.0). Several cell wall-active enzymes display pH_{opt} values that differ from their surrounding physiological pH. For instance, expansins typically have a $\text{pH}_{\text{opt}} = 4$, which is far below the acidic pH_{apo} of 5.5–6 (Sampedro and Cosgrove, 2005). In addition, pectin methylsterases from *Arabidopsis* and citrus often have an $\text{pH}_{\text{opt}} = 7$ – 8 despite residing in generally acidic cell wall environments (Xu et al., 2022; Hocq et al., 2023). It is suggested that such pH optimum discrepancies imply enzyme dormancy until pH shifts occur. In this way, the enzyme activity is inhibited and becomes active only upon cell wall acidification or alkalinization. This built-in pH discrepancy is thought to enable rapid regulation of wall-loosening by the cell but also prevents excessive enzyme action until local conditions are adequate. Alternatively, pH discrepancies may also arise from artefacts of recombinant expression (altered folding, missing processing). An upward shift of the optimal pH due to the changed electrostatic environment of the catalytic site has been reported before (Montor-Antonio et al., 2017; Hofer et al., 2020). It is, however, difficult to predict whether or not OsAPSE was produced with an aberrant catalytical site.

AGALs with a $T_{\text{opt}} = 60$ – 65°C were reported in papaya (Soh et al., 2006) and sugarcane (Chinen et al., 1981). Stability towards pH is usually ± 2 – 3 pH values around the optimum. Likewise, the Q_{10} temperature coefficient for most plant enzymes is typically 2–4, meaning that the reaction rate decreases 2–4 fold with a 10°C temperature increase (Bernacchi et al., 2002; Elias et al., 2014).

Kinetic parameters including the K_M value and V_{max} were determined (Supplementary File S8.3). The absorbance did not further increase after 1 hour of incubation. The Hanes-Woolf linearization calculations were performed with the data point between substrate concentrations of 0.5–50 mM and yielded a linear relationship ($R^2 = 0.9895$) (Figure 4C). It was calculated that $K_M = 0.67$ mM and $V_{\text{max}} = 3.9$ nM·s⁻¹. The initial fit to the hyperbolic Michaelis-Menten plot, according to Equation 2, yielded a poor R^2 0.237 ($n = 6$ data points) or $R^2 = 0.399$ ($n = 4$ data points).

$$v_0 = \frac{V_{\text{max}} \cdot [S]}{K_M + [S]} = \frac{3.9 \cdot 10^{-9} \cdot [S]}{0.67 \cdot 10^{-3} + [S]}$$

It is obvious that the relationship between v_0 and $[S]$ is linear ($R^2 = 0.9858$) (Figure 4D). This is not surprising, since the experiment made use of a range of substrate concentrations between 0.5–50 mM, which are far above the calculated K_M , under substrate-saturated conditions ($v_0 = V_{\text{max}}$). The value for V_{max} mainly depends on the enzyme concentration. The K_M value, on the other hand, is an intrinsic enzyme parameter, independent of the enzyme concentration. The observed K_M is in line with values reported in scientific literature (Supplementary File S8.5) for experiments with comparable substrates and enzymes. For instance, an AGAL from rice was produced recombinantly in *P. pastoris* and yielded a $K_M = 0.47$ mM for pNP- α -D-Galp as substrate, which is only slightly lower compared to the K_M value in this study (Chien et al., 2008). Furthermore, K_M values are correlated with the complexity of the used substrate (Supplementary File S8.5). Larger substrates (*fi.* RFOs), typically yield higher K_M values, *i.e.* lower enzyme-substrate affinity. We did not determine the K_M and V_{max} for pNP- β -L-Arap as substrate due to limited availability of the used CFPS system and the very high probability of achieving similar values. It would not be unreasonable to assume that the K_M value for pNP- β -L-Arap would be in the same order of magnitude, since we demonstrated that GH27_OsAPSE displays similar AGAL and ARAP activity at identical concentrations of substrate and protein (Figure 3C).

3.3.3 Activity on natural substrates

A wide range of putative substrates decorated with α -D-Galp and/or β -L-Arap side chains were submitted to enzymatic hydrolysis by the GH27 domain of OsAPSE (Figure 5A). Highest activities were observed for the disaccharide melibiose, the trisaccharide raffinose and the polysaccharide galactomannan (Supplementary File S8.4). An activity slightly higher compared to the empty vector *pALiCE02* control was observed for verbascose (+36.3%), arabinogalactan from larch wood (+24.2%) and AGPs from *A. thaliana* cell suspension cultures (+30.6%) (Figure 5B). The modest activity may be due to the inherently low abundance of α -D-Galp and β -L-Arap moieties in AGP O-glycans (Tryfona et al., 2012). Furthermore, AGAL activity is typically lower for RFOs with a higher DP (Supplementary File S8.5).

Similar to AtAPSE, GH27_OsAPSE shows high activity for substrates which are not colocalizing with OsAPSE at the cell surface (Imaizumi et al., 2017). RFOs do not occur at the cell surface of rice cells (De Coninck et al., 2024a), but are stored in vacuoles (Van den Ende, 2013; Elango et al., 2022). Galactomannan is speculated to be present in low quantities in the cell walls of rice endosperm and the aleurone layer (Ren et al., 2007). However, GH27 activity was also detected for AGPs with O-glycans (Figure 5B).

GH27 enzymes from plants typically accommodate hydrolytic cleavage of monosaccharides from storage polysaccharides or cell wall structures (Gao and Schaffer, 1999; Imaizumi et al., 2017; Chuankhayan et al., 2023). Most of the arabinose and galactose at

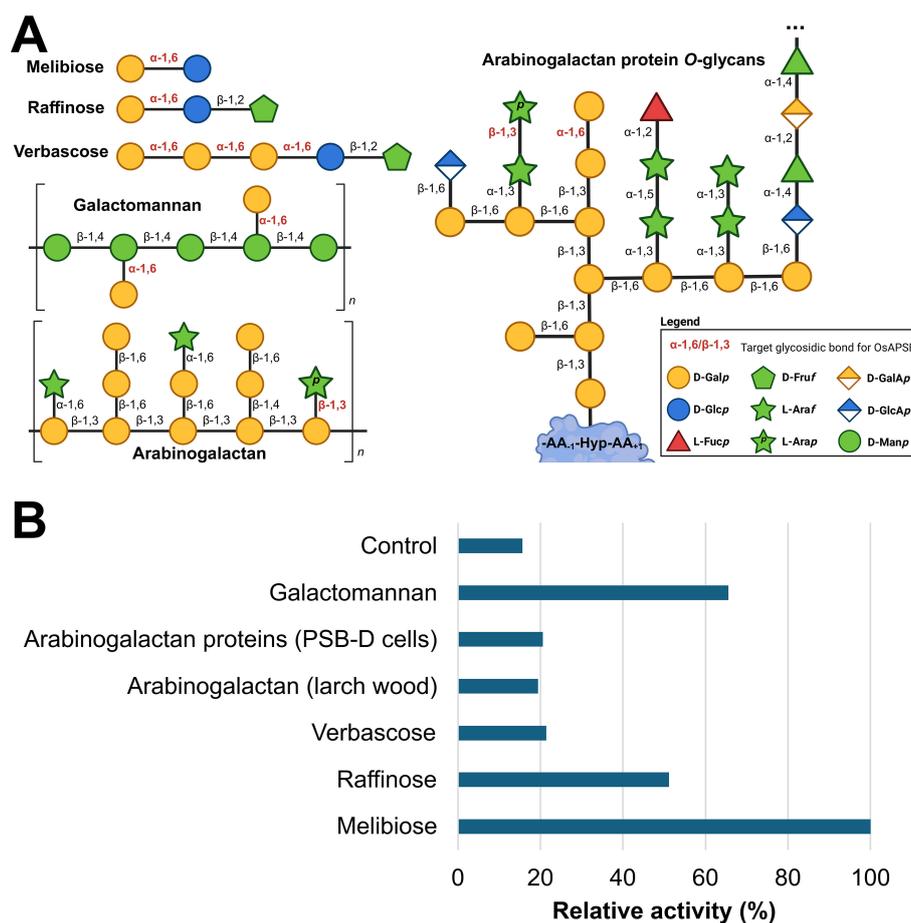


FIGURE 5

Enzymatic activity of GH27_OsAPSE on natural substrates. Structure of the considered natural substrates. The structures for galactomannan, arabinogalactan and arabinogalactan protein O-glycans are average structures based on theoretical models (Knoch et al., 2014; Seifert, 2020; Strasser et al., 2021; Voiniciuc, 2022; Tan et al., 2023) (A). Bar chart representing relative activities of GH27_OsAPSE on natural substrates. The control is the CFPS reaction with empty vector *pALICE02* incubated at 25°C at pH 7.5 (B).

the cell surface occurs as α / β -L-Araf and β -D-Galp (Ghosh et al., 2023) and is present in arabinogalactan, rhamnogalacturonan I and xylan side chains. Theoretical models depicting cell wall structure do not always include β -L-Arap and α -D-Galp (Seifert, 2020; Delmer et al., 2024) despite the fact that these residues have been detected by NMR at the extremities of AGP O-glycans (Odonmazig et al., 1994; Ponder and Richards, 1997; Strasser et al., 2021) or in side chains of pectin rhamnogalacturonan-I (Perez, 2003; Caffall and Mohnen, 2009; Goetz et al., 2016). The structure (and function) of AGP O-glycans and pectic polysaccharides depends on the activity of CAZymes, such as glycosyltransferases for synthesis and GHs for trimming and degradation (Leszczuk et al., 2023). Both AGP O-glycans and pectic polysaccharides are highly complex and heterogeneous structures, and the relationship between glycan/polysaccharide structure and biological function is not fully understood (Strasser et al., 2021). In plants, most of the transferases for glycan/pectin synthesis are known (Caffall and Mohnen, 2009; Silva et al., 2020) while far less information is available concerning O-glycan degradation in plants (Ellis et al.,

2010), since only 5 AGP O-glycan degrading GHs are currently known (Knoch et al., 2014).

The composition and modification of AGP O-glycans vary significantly across cell types, tissues and species, suggesting that plants modify these structures to their environment and physiological needs (Leszczuk et al., 2023). Although not fully understood, evidence indicates that AGP O-glycan structure is linked to biological function. AGP O-glycans help plants respond to environmental stresses. For example, the seagrass *Zostera marina* produces O-glycans rich in 4-O-methylglucuronic acid, which provides a polyanionic interface and contributes to osmotic adjustment to salinity stress (Pfeifer et al., 2020). Such O-glycosylation patterns are not observed in land plants, which mostly obey to the typical ‘type-II AGP’ structure (Strasser et al., 2021). In land plants, fungal degradation of AGP O-glycans can impair cellulose production and growth, highlighting the importance of glycan length and branching for cell wall integrity (Kikuchi et al., 2022). AGP O-glycans also stabilize the cell wall by binding ions like Ca^{2+} through negatively charged residues (*f.i.*

glucuronic acid). Plants with reduced glucuronidation show severe developmental issues, which can be alleviated by Ca^{2+} supplementation, indicating that ion binding plays a structural role (Lopez-Hernandez et al., 2020). Additionally, AGP O-glycans form covalent links with other cell wall components like pectic rhamnogalacturonan-I and arabinoxylan, contributing to structural stability of the cell wall matrix (Tan et al., 2013, 2023). Finally, AGP O-glycans are also crucial in development, influencing cell division, elongation and differentiation. For instance, trimming of AGP O-glycans affects apple fruit ripening (Leszczuk et al., 2020), and proper glycosylation by hydroxyproline-O-galactosyltransferases is essential for pollen development, as shown by sterility in *Arabidopsis* mutants lacking these enzymes (Kaur et al., 2022).

3.4 OsAPSE may be unconventionally secreted to the cell surface

3.4.1 Arguments for unconventional protein secretion

OsAPSE is predicted to be synthesized with an N-terminal SP. No other localization signals are detected (Supplementary File S9.1). Because of the presence of a SP, it can be expected that OsAPSE will follow the secretory pathway involving protein synthesis on the endoplasmic reticulum (ER). Furthermore, PTMs such as the addition of N-glycans and the formation of disulfide bridges are likely to occur. Asparagine residues N269, N372 and N380 occur in a sequon and are predicted as N-glycosylation sites with high confidence (Supplementary File S9.1). However, N380 is part of a NPT sequon and will therefore not be recognized by the oligosaccharyltransferase complex (Matsumoto et al., 2017). The sequons at N269 and N372 are likely to accommodate N-glycans as these asparagine residues are correctly oriented at the protein surface (Supplementary File S9.2). Several disulfide bridges are predicted in the structure of OsAPSE, although only 1 disulfide bridge is likely to occur in the GH27 domain of OsAPSE. Cysteine residues C187 and C227 are positioned under a favorable dihedral angle and inter-atomic distance of 2.03 Å, which is within the average disulfide bridge length of 1.8–3.0 Å.

It remains speculative whether or not secretion of OsAPSE occurs conventionally or unconventionally. By default, it is assumed that proteins with a SP are secreted at the cell surface through conventional secretion (Rose and Lee, 2010). However, it was shown that several secretory proteins do not necessarily possess a SP (Wang et al., 2018). Unconventional protein secretion in multivesicular bodies has been demonstrated for several cell wall-active enzymes, including xyloglucan endotransglucosylase/hydrolases from *Arabidopsis* (De Caroli et al., 2021) and HaAPSE (HanXRQChr08g0208381), the *Helianthus annuus* (sunflower) homologue of OsAPSE (50.2% sequence identity) (Regente et al., 2017). Retrieving cell wall-active enzymes from multivesicular bodies is significant and remarkable, as it was supposed for a long time that these enzymes are trafficked by a SP and secreted conventionally (Rose and Lee, 2010). However, it has been hypothesized that some cell wall-active enzymes with SP may be

delivered at the cell surface using a vesicular “type IV-UPS” bypass of the Golgi apparatus (Rabouille, 2017; Maricchiolo et al., 2022). Alternatively, unconventionally secreted cell wall-active enzymes with SP may bypass both the Golgi apparatus and vesicular transport through ER-plasma contact sites (Bellucci et al., 2018). Unfortunately, in contrast to mammalian proteins, tools for predicting unconventional protein secretion, such as SecretomeP (RRID: SCR_026505), are not available for plants (Bendtsen et al., 2004).

3.4.2 Localization of OsAPSE according to *in silico* prediction tools

OsAPSE was predicted to be located at the cell surface or cell wall (Supplementary File S9.3). The results were mainly uniform, predicting the localization of OsAPSE. This is in agreement with the hypothesized function of OsAPSE being a cell wall-active enzyme. Localization data from the WallProtDB-2 database (San Clemente et al., 2022) (RRID: SCR_026506) reveals that OsAPSE has already been detected in the cell wall proteome of rice callus cultures (Chen et al., 2009).

3.4.3 Localization of OsAPSE-EGFP in transiently transformed *N. benthamiana* leaves

Subcellular localization studies of OsAPSE in *N. benthamiana* leaves, transiently transformed with the *pK7FWG2::OsAPSE-EGFP* fusion construct, reveal that the OsAPSE-EGFP fusion protein is localized at the cell surface (Figures 6A–E). No fluorescent signals were detected in the cytosol. Transient expression of cytosolic EGFP-fusion proteins typically shows cytoplasmic strands, attributed to the voluminous (*i.e.* 80–90% of total volume) vacuoles of the tobacco epidermal cells (Cui et al., 2020). Cytoplasmic strands were abundantly observed in the free-EGFP control treatments (Figures 6F–J), but only sporadically and faintly in *pK7FWG2::OsAPSE-EGFP* treated plants (Figure 6A). In several images, we also observed colocalization between OsAPSE-EGFP and the DAPI signal, suggesting a nuclear localization (Figures 6B, C) although no NLS was identified in the coding sequence of OsAPSE (Supplementary File S9.2).

Heterologous expression of non-native SPs is often less efficient (Jarvis et al., 1993; Wilbers et al., 2016), due to the occurrence of certain amino acids in the SP (*f.i.* double arginine and multiple repeated proline residues) that hinder helix formation, lower the affinity towards the SP recognition particle and inhibit SP peptidases (Nilsson and Von Heijne, 1992; Snapp et al., 2017). The SP of OsAPSE contains such disturbing elements (-PPPWRRLRLRCALLPP-). RR motifs in SPs are ER retention signals, due to their ability to interfere with vesicle formation that would otherwise transport the protein along the secretory pathway (Schutze et al., 1994). Disturbed recognition or incomplete SP cleavage results in subpopulations of the protein of interest being either processed adequately or accumulating in the ER (Wilbers et al., 2016). Protein accumulation in the ER in turn leads to ER stress and unfolded protein responses (Hicks, 2013; Srivastava et al., 2014), which involve ER-associated degradation and retro-translocation of the misfolded protein to the cytosol for

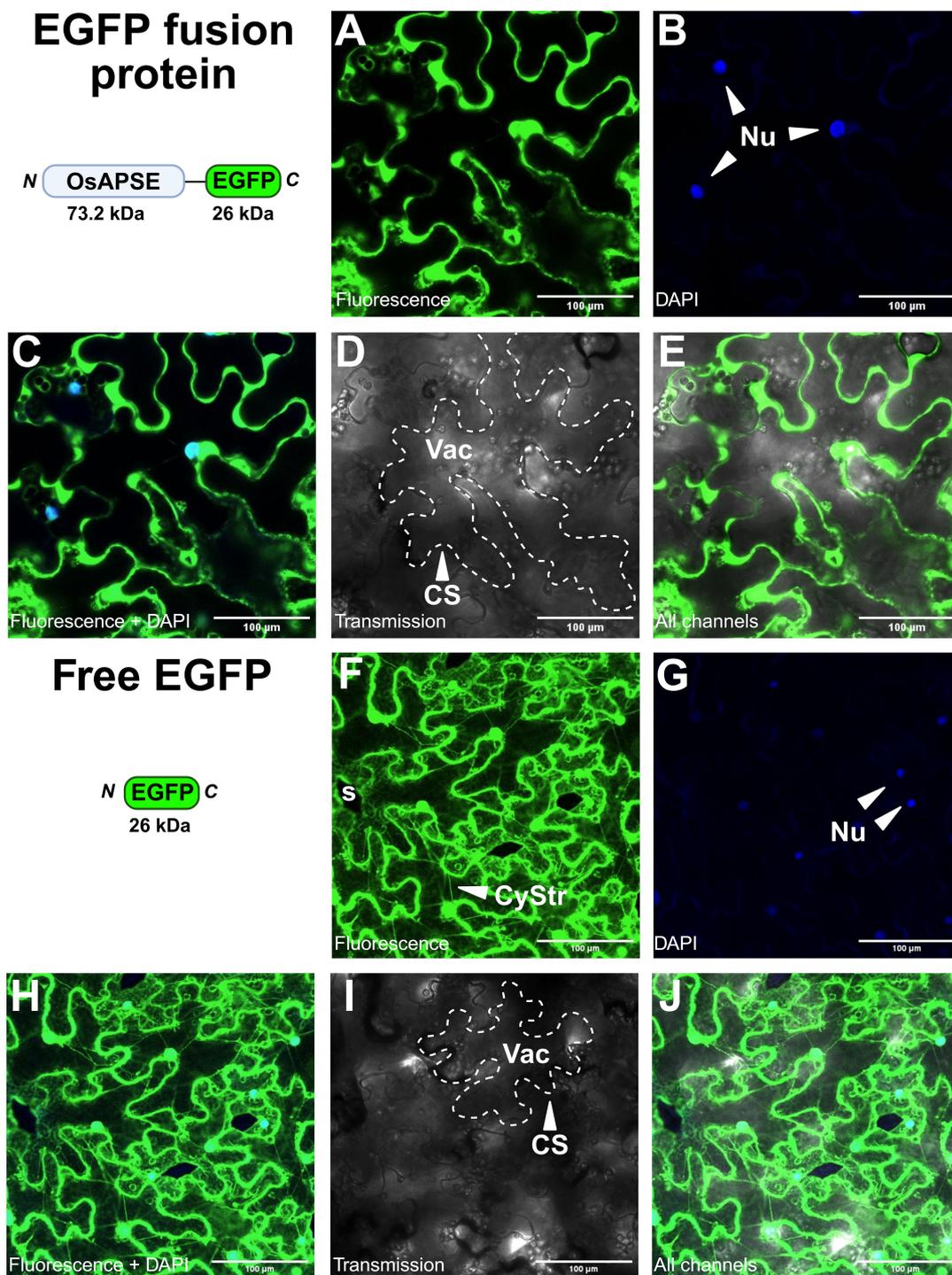


FIGURE 6
Subcellular localization of OsAPSE-EGFP in epidermal cells of *Nicotiana benthamiana* leaves. (A-E) show *pk7FWG2::OsAPSE-EGFP* infiltrated tissues, while (F-J) show tissues infiltrated with the *pk7FWG2* empty vector control. The domain architecture, orientation and size of the transiently produced protein is shown. (A, F) show the green fluorescent signal. (B, G) show the DAPI signal. (C, H) show the combined green fluorescent and DAPI signal. (D, I) show the transmission channel. (E, J) show the combination of all channels. Abbreviations: CS (cell surface), CyStr (cytoplasmic strand), Nu (nucleus), s (stomata), Vac (vacuole).

degradation (Johnson and Haigh, 2000). Upon proteolytic cleavage of the misfolded proteins, hidden NLS sequences are often exposed, causing nuclear import (Srivastava et al., 2014). Likewise, cytosolic degradation of (partially) misfolded OsAPSE-EGFP could result in

the release of EGFP, which localizes by default to the nucleus (Seibel et al., 2007). Such ‘protein reflux’ has been described in the context of ER stress (El Meskini et al., 2001). Noteworthy, it has been established that EGFP(-fusion) proteins with size up to 110 kDa

may diffuse spontaneously into the nucleus (Wang and Brattain, 2007).

Alternatively, it is possible that the native SP of OsAPSE is recognized both as a SP and a NLS. NLS sequences are typically hydrophobic (Kiefer et al., 1994). The activity of background proteases in *N. benthamiana* (Jutras et al., 2020) may trim the native SP, explaining NLS recognition. Similar cases in which *N*-terminal sequences were recognized simultaneously as SP and NLS have been reported, albeit in animals (Kiefer et al., 1994), although NLS sequences are functionally conserved amongst higher eukaryotes (Wagner and Hall, 1993; Hicks, 2013). In these cases, the protein of interest had a dual fate, in both the secretory pathway and nucleus (Kiefer et al., 1994; Iwata et al., 2008). It should be emphasized that not all NLS sequences in plants have been discovered, as they are often non-canonical and not defined by a consensus sequence (Hicks, 2013; Lu et al., 2021).

Further investigation of the subcellular localization at the cell surface using propidium iodide as organelle marker for the plasma membrane did not deliver reliable results (data not shown). The combination of *in silico* predictions for OsAPSE localization, occurrence of OsAPSE in cell wall proteome databases and the microscopy images suggest that OsAPSE is most likely localized at the cell surface.

3.5 OsAPSE is involved in germination of rice seeds by acting on cell walls

3.5.1 Screening of transgenic and mutant rice plants

Transgenic rice lines with overexpression (*pUBI::OsAPSE*) or gene knock-out (*osapse*) were genotyped (Supplementary File S10) and analyzed for agronomical qualities. *pUBI::OsAPSE* panicles contained less mature seeds and showed lower seed setting rates compared to WT and *osapse* seeds, while panicle mass and seed mass were highest for WT seeds compared to *osapse* and *pUBI::OsAPSE* panicles and seeds (Figure 7). The agronomical parameters are in accordance with the model of Smith & Fretwell (Smith and Fretwell, 1974) and indicate the high plasticity of seed number and limited variation in seed mass (Supplementary File S11).

3.5.2 OsAPSE is involved in rice seed germination and seedling development

The involvement of OsAPSE in seedling development and seed formation was inferred from both *in vitro* rice seed germination assays as well as RT-qPCR experiments on germinating and developing seeds and seedlings. Results from RNA extraction,

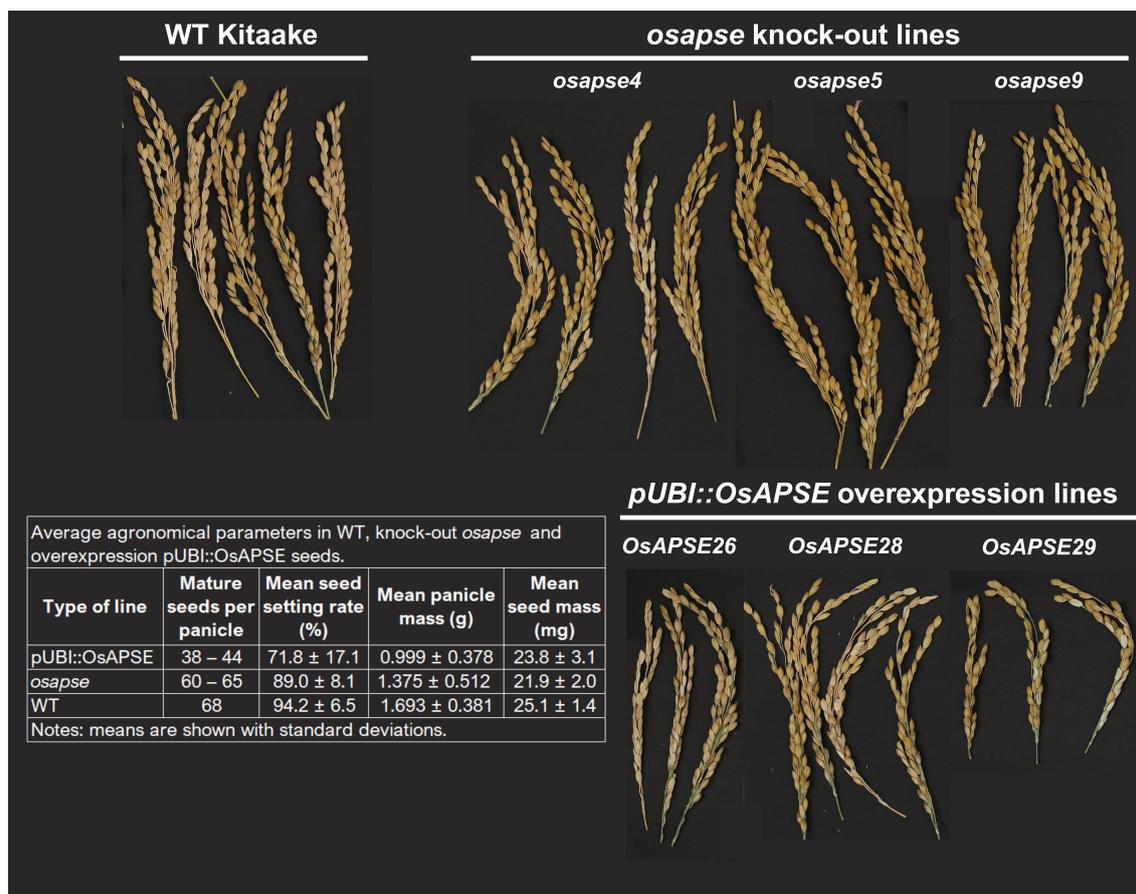


FIGURE 7

Representative panicles and average agronomical parameters of wild type, overexpression *pUBI::OsAPSE* and knock-out *osapse* rice seeds.

cDNA synthesis, RT-qPCR data and normalization of gene expression are included in [Supplementary File S12](#).

Figure 8 shows the *OsAPSE* transcript levels in developing WT (**Figure 8A**), mutant *osapse* and overexpressing *pUBI::OsAPSE* seedlings (**Figure 8B**). In WT seedlings, the transcript levels for *OsAPSE* are low but show an increase over time (**Figure 8A**). The *OsAPSE* transcript levels are drastically lowered in *osapse* mutants, reaching an overall average transcript level of 0.22 ± 0.12 CNRQ (calibrated normalized relative quantities) during rice seed germination. The low *OsAPSE* transcript levels in knock-out mutants illustrate that *osapse* plants are true knock-out mutants (one-way ANOVA, $F = 60.280$; $\nu = 75$ df, $p < 0.001$, $\eta^2 = 0.840$) compared to WT and *pUBI::OsAPSE*. Similar to WT seedlings, the *OsAPSE* transcript levels in *osapse* mutants increase slightly over time (**Figure 8B**) (one-way ANOVA, $F = 20.037$; $\nu = 14$ df; $p < 0.001$, $\eta^2 = 0.845$). Elevated and generally stable *OsAPSE* transcript levels (on average 18.23 ± 5.94 CNRQ) are observed for *pUBI::OsAPSE* overexpression lines, although the transcript levels in *pUBI::OsAPSE* 28 show a decrease over time. An apparent increase in *OsAPSE* transcription, though not statistically different, was observed in WT and is mainly attributed to the increased metabolic activity in developing seedlings. Indeed, it was already reported that the transcript levels for several CAZymes including AGALs and β -D-mannosidases increase during rice seed germination (Ren et al., 2007).

Knock-out mutants of O-glycan-active enzymes and cell wall-active enzymes often yield aberrant phenotypes (*f.i.* prolonged roots, reduced hypocotyls) (Eudes et al., 2008; Nibbering et al., 2020; Dash et al., 2023), as was also observed for *atapse* knock-out plants (Imaizumi et al., 2017). Despite the strong structural similarity between *OsAPSE* and *AtAPSE*, we did not observe differences in coleoptile length in *osapse* or *pUBI::OsAPSE* seedlings (data not shown).

Differences in germination rates were apparent between *osapse*, *pUBI::OsAPSE* and WT seeds across different time points ([Supplementary File S13](#)) (**Figure 9**). The Omnibus Test of Model Coefficients revealed that the constructed GLM, comprising the relationship between sampling time, line and observed germination rate was significant ($\chi^2 = 343.460$, $\nu = 27$ df, $p < 0.001$). The GLM displayed main effects from the sampling time ($\chi^2 = 84.672$, $\nu = 3$ df, $p < 0.001$) and type of line ($\chi^2 = 251.142$, $\nu = 6$ df, $p < 0.001$), but not the interaction ($\chi^2 = 18.488$, $\nu = 18$ df, $p < 0.424$). Interestingly, *pUBI::OsAPSE* and knock-out *osapse* lines displayed a lower germination rate compared to WT, which consistently achieved highest germination rates (80-97%) at the different time points (**Figure 9A**). The transgenic *pUBI::OsAPSE* lines and knock-out *osapse* lines demonstrated aberrant germination phenotypes. Especially at time points 1 dpi and 4 dpi, large fractions of the *pUBI::OsAPSE* seeds (40-70%) and *osapse* seeds (40-75%) displayed seed lethality, no radicle emergence and disturbed root development, arrested coleoptile elongation and seedling etiolation, while WT seeds showed normal development, with emerging leaves and root formation. Knock-out line *osapse5* displayed germination rates that were most similar to WT, while lines *pUBI::OsAPSE28*, *pUBI::OsAPSE29*, *osapse9* and *osapse4* demonstrated the most deviating germination rates (**Figure 9A**). In general, the seed germination rates of *osapse* knock-out lines were higher compared to *pUBI::OsAPSE* overexpression lines (**Figure 9B**), but still lower compared to WT seeds. The germination behavior of *osapse4* and *osapse5* seeds differed considerably, despite bearing identical mutations ([Supplementary File S10](#)). The differences between the overexpression lines are attributed to the independent nature of these lines, as each overexpression line is created by a single transformation event, leading to specific *OsAPSE* transcript levels. Interestingly, data from the germination assay and RT-qPCR experiments suggest that

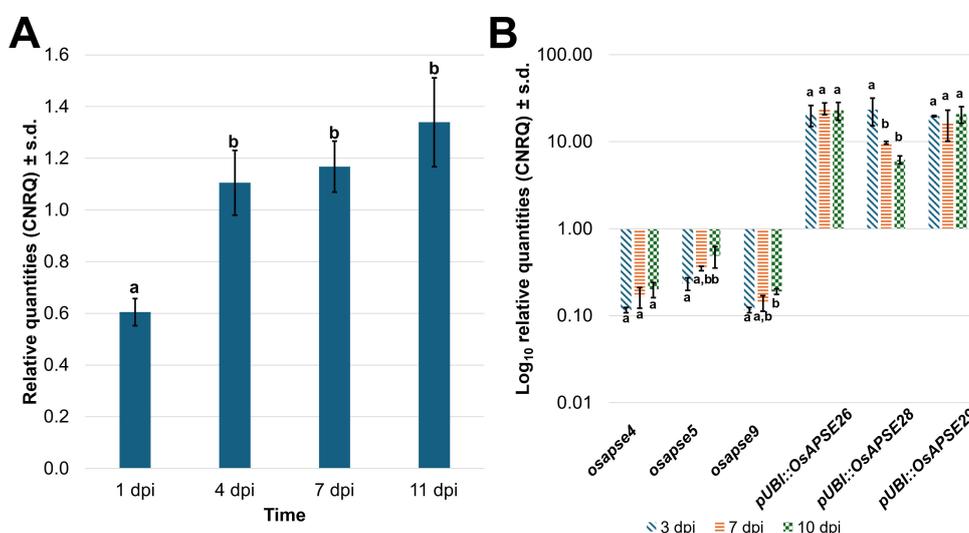


FIGURE 8

OsAPSE transcript levels in developing rice seedlings. Transcript levels for *OsAPSE* in germinating WT seeds and seedlings (A). Transcript levels for *OsAPSE* in germinating seeds and seedlings of knock-out and overexpression plants (B). Transcript levels are shown as calibrated normalized relative quantities. Homogenous subsets, based on Tukey *post-hoc* ANOVA analysis are indicated with letters a and b.

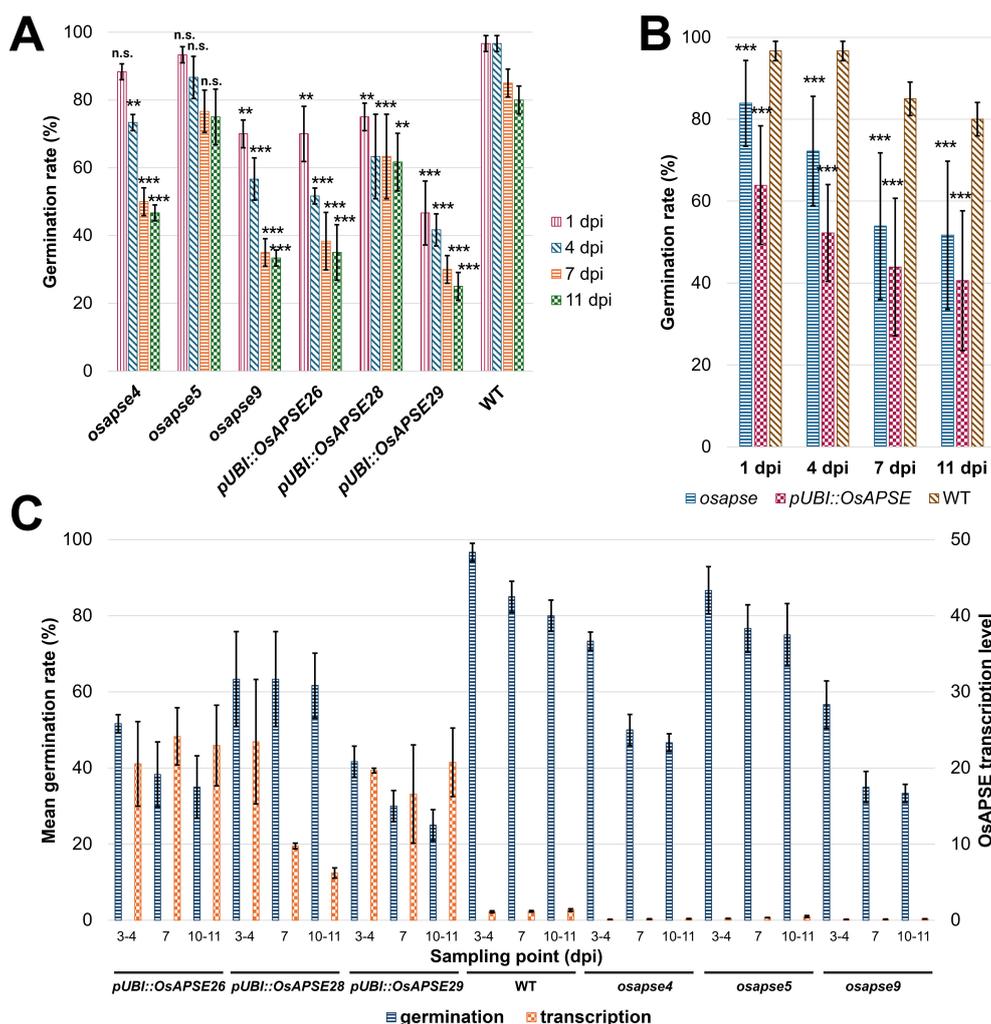


FIGURE 9

Seed germination rates of mutant *osapse* and overexpressing *pUBI::OsAPSE* rice compared to wild type plants. Germination rates across mutant *osapse*, overexpressing *pUBI::OsAPSE* and wild type rice at 1 dpi, 4 dpi, 7 dpi and 11 dpi (A). Average germination rates for knock-out lines and overexpression lines compared to WT (B). Correlation between the transcript level of *OsAPSE* and the observed rice seed germination rate at time points 3–4 dpi, 7 dpi and 10–11 dpi (C). Asterisks indicate the significance level: * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$). Error bars are standard deviations originating from 3 biological replicates of 20 rice seeds per replicate. Abbreviations: n.s. (non-significant difference: $p \geq 0.05$).

optimal *OsAPSE* transcription levels are required for rice seed germination (Figure 9C). There are several examples of genes with similar expression regulation patterns. *OsMADS1* from rice involved in flower development and pollen morphology (Liu et al., 2023, 2025), *AtWUS* from *Arabidopsis* involved in shoot apical meristem maintenance and floral development (Stephenson et al., 2019; Li et al., 2023) and *ZmCCT10* from maize involved in photoperiod sensitivity and regulation of flowering time (Ikeda et al., 2009; Yadav et al., 2011), exemplify genes whose expression must be finetuned to ensure normal development. In each case, both knock-out and overexpression lead to aberrant phenotypes, ranging from sterility and disrupted organ formation to severe architectural defects, highlighting their dosage-sensitive nature, similar to *OsAPSE*.

3.5.3 Involvement of *OsAPSE* in cell wall metabolism during rice seed germination

The involvement of GH27 proteins in seed germination or seedling development has been reported in pea (Blöchl et al., 2008), chickpea (Arunraj et al., 2020), cluster bean (Hughes et al., 1988), soybean (Guimaraes et al., 2001; Lien et al., 2018) and vetch (Gojlo, 2023). Often, AGALs are involved in the germination process through their ability to hydrolyze storage polysaccharides (Elango et al., 2022) and galactomannan (Sharma et al., 2022), after which the released D-Galp moieties are epimerized to D-Glcp and used as energy source (Zhang et al., 2006) or act as extracellular signaling molecules (Showalter, 2001).

We hypothesize that adequate *OsAPSE* transcription levels are required for normal rice seed germination. Deviating transcript

levels, as observed in *osapse* and *pUBI::OsAPSE* plants, likely cause metabolic imbalances, with repercussions for seed germination capacity. The importance of OsAPSE for the germination process is explained by dual ARAP/AGAL activity on cell wall structures. We assume that OsAPSE displays AGAL/ARAP activity on α -D-Galp and β -L-Arap residues occurring along the continuous pectin (rhamnogalacturonan-I)-AGP O-glycan network (Figure 10) (Tan et al., 2013, 2023). The ricin-B-like domain can enhance substrate binding since ricin-B(-like) domains and related CBM13 modules are known to recognize specifically D-Galp and D-GalNAc residues (Hazes, 1996; Steeves et al., 1999), but can also recognize L-Arap residues due to the structural similarity between D-Galp and L-Arap as demonstrated in a GH27 ARAP from *Streptomyces avermitilis* (Ichinose et al., 2009; Fujimoto, 2013).

Removal of terminal α -D-Galp and β -L-Arap moieties from AGP O-glycans and rhamnogalacturonan-I likely abolishes non-covalent interactions between the AGP O-glycan and the pectin fraction of the primary cell wall (Figure 10). Removal of these residues contributes to structural reorganization of the cell wall, as pectic polysaccharides and AGPs are in turn directly tethered to cellulose microfibrils and hemicellulose polysaccharides (Ellis et al., 2010; Peaucelle et al., 2012; Jamet and Dunand, 2020). Weakened interactions between AGP O-glycans and pectin reduce cell wall tension and allow cells to elongate and expand under influence of turgor pressure (Figure 10). The required degree of mobility between compounds of the primary cell wall and AGPs could be regulated (*i.e.* tightening and loosening of non-covalent interactions) through well-dosed activity of CAZymes

like OsAPSE. Flexibility and extensibility of cell wall constituents is of paramount importance to attain normal cell elongation, cell expansion and growth in general (Cosgrove, 2024), especially in fast-paced processes like seed germination (Wolny et al., 2018; Yang et al., 2020).

4 Conclusions

This study focused on the biological function of OsAPSE, a member of the GH27 family from Japanese rice. Using a multi-perspective approach, we aimed to decipher the biological relevance of this protein.

We have demonstrated the intriguing phylogeny of OsAPSE, which is based on the presence of a ricin-B-like domain (De Coninck et al., 2024b). The presence of this lectin domain is only observed in a subset of GH27 sequences from *Viridiplantae* and distinguishes APSE homologues from other GH27 members. A subdivision as suggested by the CUPP database would therefore be appropriate (Barrett et al., 2020). Likewise, the GH27 domain of OsAPSE and other APSE homologues is structurally different compared to regular GH27 domains, with RMSD values $> 5\text{Å}$.

We also provided insights into the biochemical characteristics and putative biological function of OsAPSE. Although we did not succeed in obtaining soluble recombinant OsAPSE or its subdomains in bacterial or yeast cells, we were successful in the production of small quantities of the OsAPSE GH27 domain using a cell-free system. The produced proteins showed clear AGAL and

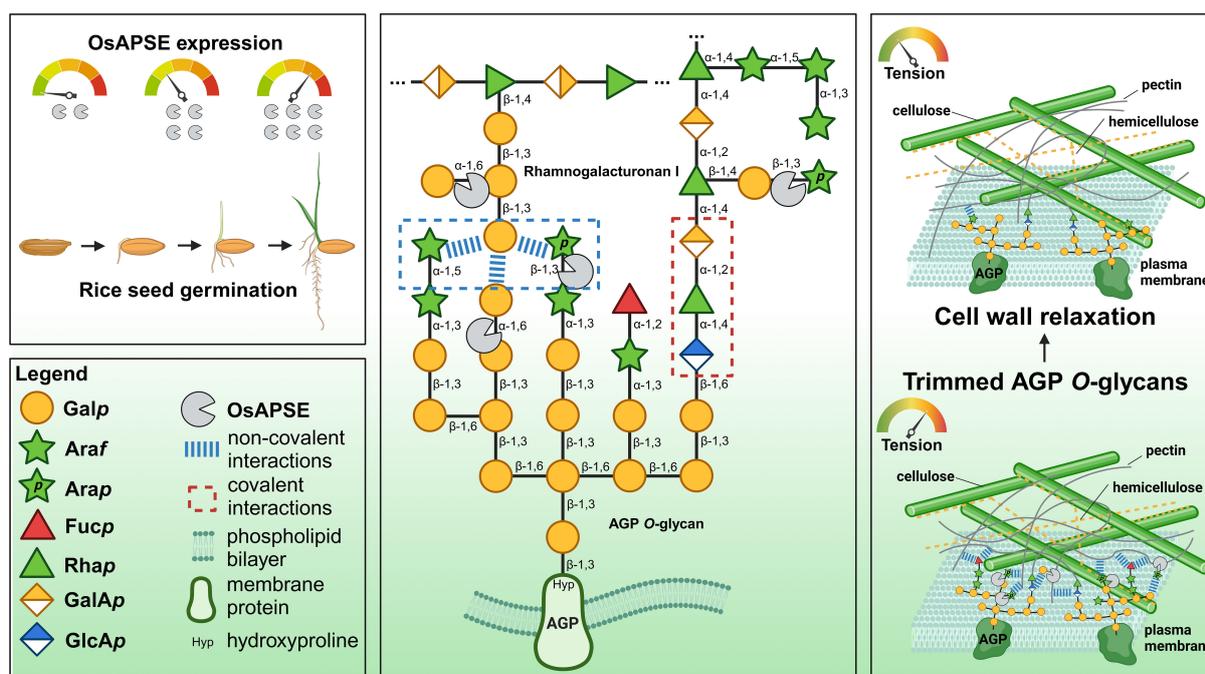


FIGURE 10

Working hypothesis for OsAPSE involvement in cell wall metabolism during rice seed germination. OsAPSE displays AGAL and ARAP activity, acting on the pectin-AGP O-glycan network, thereby affecting non-covalent interactions and impacting cell wall flexibility required for cell growth and expansion. AGP O-glycan structure and representation of the primary cell wall are based on existing theoretical models (Carpita and Gibeau, 1993; Odonmazig et al., 1994; Ponder and Richards, 1997; Perez, 2003; Caffall and Mohnen, 2009; Goetz et al., 2016; Seifert, 2020; Strasser et al., 2021; Tan et al., 2023). This diagram was created using BioRender.com.

ARAP activity, with optimal activity at pH = 8 and 25°C. Although the optimum pH differed from the typical apoplastic pH (5.5-6), we assume this is attributable to enzyme dormancy and activation upon cell wall alkalization, typically associated with regulation of cell wall loosening and expansion (Geilfus et al., 2017). We calculated the Michaelis constant for pNP- α -D-Galp as substrate, *i.e.* $K_M = 0.67$ mM, in line with observations from other plant GH27 enzymes. To our knowledge this study is the first to report enzymatic parameters for plant GH27 enzymes produced by a CFPS platform. Highest GH27 activity was present for melibiose, galactomannan and raffinose, while lower activities were obtained for verbascose, glycosylated arabinogalactan proteins and arabinogalactan. The activities observed for natural substrates were somewhat unexpected, as higher activities were obtained for substrates that will never make contact with OsAPSE at the cell surface. Melibiose does not occur in rice and galactomannan and raffinose are only present in low quantities in specific cells (*f.i.* endosperm cells, vascular tissue) (Ren et al., 2007; Van den Ende, 2013; Li et al., 2018; Elango et al., 2022). A similar observation for AtAPSE was made in the past (Imaizumi et al., 2017).

Subcellular localization analyses *in planta* suggest the localization of OsAPSE-EGFP at the cell surface but needs confirmation in future experiments. Co-localization studies with organelle reporters for the ER, Golgi apparatus and plasma membrane could be considered, in combination with plasmolysis assays, to pinpoint the exact subcellular localization of OsAPSE (Serna, 2005; Stellmach et al., 2022).

Transcriptomics analyses in WT rice revealed that OsAPSE transcript levels double during the germination process, indicating the need for GH27 activity during seed germination. The germination rate of rice is negatively affected when OsAPSE transcript levels are decreased or increased. Interestingly, the agronomical traits including mature seeds per panicle, the setting rate, the panicle mass and average seed mass were usually lower compared to WT plants. The importance of OsAPSE for the rice seed germination process is mainly explained by its proposed dual enzymatic activity along the continuous pectin-AGP O-glycan network (Tan et al., 2023). Based on the observed AGAL and ARAP activity and the ability of OsAPSE to cleave off β -L-Arap and α -D-Galp moieties from cell wall structures, we hypothesize that OsAPSE displays dual activities on β -L-Arap and α -D-Galp moieties along the continuous pectin-AGP O-glycan network at the cell surface, while the ricin-B-like domain might enhance substrate binding. The activity of OsAPSE likely affects non-covalent interactions between the pectin fraction and AGP O-glycans, either *in muro* or anchored in the plasma membrane (Ellis et al., 2010; Nibbering et al., 2020). As a result, altered non-covalent interactions would allow increased or decreased attraction between AGP O-glycans and pectin, thereby enhancing cell wall loosening and relaxation, which is a key driver in cellular expansion and growth (Cosgrove, 2024). Seed germination is *par excellence* a fast-paced developmental process, characterized by rapid cellular growth, and therefore a strong need for rapid cell wall metabolism (Wolny et al., 2018). Rapid and developmentally controlled AGP turnover, mediated by CAZymes, has been described in germinating rice seeds (Lu et al., 2001; Yuan et al., 2008). We have shown that the absence or excess of

OsAPSE transcripts leads to aberrant germination phenotypes. Shortage of OsAPSE transcripts may lead to persisting α -D-Galp and β -L-Arap residues and therefore also more persisting non-covalent interactions between AGP O-glycans and pectin, resulting in reduced cell wall flexibility, while increased OsAPSE transcripts could result in excessive detachment. We therefore believe that OsAPSE may be a key enzyme in enabling cell wall relaxation through regulating non-covalent interactions between pectin and AGP O-glycans during the rice seed germination process. Future experiments will be required to further investigate the involvement of OsAPSE in seed germination. For instance, effects of defective arabinosyl- and galactosyltransferase activity during seed development and germination should be explored. The role of several GTs in AGP processing have been elucidated in *Arabidopsis*, but remain elusive in rice (Silva et al., 2020; Strasser et al., 2021). Furthermore, we recommend future research to focus on the enzymatic properties of AGP O-glycan active enzymes. For instance, the characterization of AGP O-glycans and the activity of OsAPSE on AGP O-glycans can be investigated, combined with HPAEC-PAD for quantification and NMR and MS/MS for determining the saccharide composition and glycosidic linkages (Tan et al., 2024). It should be emphasized that the function of OsAPSE may not be exclusively connected to seed germination, as AGP O-glycans are involved in a plethora of physiological processes (Ellis et al., 2010; Mareri et al., 2019; Strasser et al., 2021) including growth and development (Hromadová et al., 2021; Lampion et al., 2021), cell differentiation (Borassi et al., 2020), cellular communication (Lopez-Hernandez et al., 2020; Teh et al., 2022), reproduction (Kaur et al., 2022) and signaling of biotic (Kikuchi et al., 2022) and abiotic (Pfeifer et al., 2020) stresses.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author/s.

Author contributions

TD: Conceptualization, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft. IV: Formal analysis, Investigation, Writing – review & editing. PR: Formal analysis, Investigation, Writing – review & editing. TD: Methodology, Supervision, Writing – review & editing. EV: Writing – review & editing, Conceptualization, Funding acquisition, Methodology, Project administration, Supervision.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by funding from Fonds voor Wetenschappelijk Onderzoek (FWO) Vlaanderen, grant number G008619N.

Acknowledgments

The authors wish to thank Koen Gistelincx from the Laboratory for Biochemistry & Glycobiology (Ghent University) for preparing the arabinogalactan protein extracts from *A. thaliana* PSB-D plant cell suspension cultures. The authors also acknowledge Alex Windels from the Centre for Synthetic Biology (Ghent University) for using the CANDy tool for domain modularity analysis. Finally, the authors also want to acknowledge the contributions of Ellen Guldemont and Zoé Suffys (in framework of bachelor training, Katholieke Universiteit Leuven and Odisee Hogeschool) for their contributions in the enzymatic assays and RT-qPCR experiments, and Anne-Sophie Chys (student from Hogeschool West-Vlaanderen) and Dr. Vinicius da Silva-Osterne for their help with the molecular dynamics simulations.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z
- Arunraj, R., Skori, L., Kumar, A., Hickerson, N. M. N., Shoma, N., M., V., et al. (2020). Spatial regulation of alpha-galactosidase activity and its influence on raffinose family oligosaccharides during seed maturation and germination in *Cicer arietinum*. *Plant Signal. Behav.* 15, 1709707. doi: 10.1080/15592324.2019.1709707
- Asensio, J. L., Ardá, A., Cañada, F. J., and Jiménez-Barbero, J. (2013). Carbohydrate-aromatic interactions. *Acc. Chem. Res.* 46, 946–954. doi: 10.1021/ar300024d
- Bally, J., Jung, H., Mortimer, C., Naim, F., Philips, J. G., Hellens, R., et al. (2018). The rise and rise of *Nicotiana benthamiana*: A plant for all reasons. *Annu. Rev. Phytopathol.* 56, 405–426. doi: 10.1146/annurev-phyto-080417-050141
- Barrett, K., Hunt, C. J., Lange, L., and Meyer, A. S. (2020). Conserved unique peptide patterns (CUPP) online platform: peptide-based functional annotation of carbohydrate active enzymes. *Nucleic Acids Res.* 48, W110–W115. doi: 10.1093/nar/gkaa375
- Bellucci, M., De Marchis, F., and Pompa, A. (2018). The endoplasmic reticulum is a hub to sort proteins toward unconventional traffic pathways and endosymbiotic organelles. *J. Exp. Bot.* 69, 7–20. doi: 10.1093/jxb/erx286
- Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., and Brunak, S. (2004). Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356. doi: 10.1093/protein/gzh037
- Benkert, P., Biasini, M., and Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27, 343–350. doi: 10.1093/bioinformatics/btq662
- Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81, 3684–3690. doi: 10.1063/1.448118
- Bernacchi, C. J., Portis, A. R., Nakano, H., Von Caemmerer, S., and Long, S. P. (2002). Temperature response of mesophyll conductance. Implications for the determination of rubisco enzyme kinetics and for limitations to photosynthesis *in vivo*. *Plant Physiol.* 130, 1992–1998. doi: 10.1104/pp.008250
- Beymoradi, A., Homaei, A., Hemmati, R., and Fernandes, P. (2023). Recombinant protein expression: Challenges in production and folding related matters. *Int. J. Biol. Macromol.* 233, 123407. doi: 10.1016/j.ijbiomac.2023.123407
- Bhatwa, A., Wang, W., Hassan, Y. I., Abraham, N., Li, X.-Z., and Zhou, T. (2021). Challenges associated with the formation of recombinant protein inclusion bodies in *Escherichia coli* and strategies to address them for industrial applications. *Front. Bioeng. Biotechnol.* 9. doi: 10.3389/fbioe.2021.630551
- Blöchl, A., Peterbauer, T., Hofmann, J., and Richter, A. (2008). Enzymatic breakdown of raffinose oligosaccharides in pea seeds. *Planta* 228, 99–110. doi: 10.1007/s00425-008-0722-4
- Blum, T., Briesemeister, S., and Kohlbacher, O. (2009). MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. *BMC Bioinf.* 10, 274. doi: 10.1186/1471-2105-10-274
- Boissinot, M., Karnas, S., Lepock, J. R., Cabelli, D. E., Tainer, J. A., Getzoff, E. D., et al. (1997). Function of the Greek key connection analysed using circular permutants of superoxide dismutase. *EMBO J.* 16, 2171–2178. doi: 10.1093/emboj/16.9.2171
- Boonstra, S., Onck, P. R., and van der Giessen, E. (2016). CHARMM TIP3P water model suppresses peptide folding by solvating the unfolded state. *J. Phys. Chem. B* 120, 3692–3698. doi: 10.1021/acs.jpcc.6b01316
- Borassi, C., Gloazzo Dorosz, J., Ricardi, M. M., Carignani Sardoy, M., Pol Fachin, L., Marzol, E., et al. (2020). A cell surface arabinogalactan-peptide influences root hair cell fate. *New Phytol.* 227, 732–743. doi: 10.1111/nph.16487
- Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* 382, 769–781. doi: 10.1042/BJ20040892
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3
- Brameier, M., Krings, A., and MacCallum, R. M. (2007). NucPred—Predicting nuclear localization of proteins. *Bioinformatics* 23, 1159–1160. doi: 10.1093/bioinformatics/btm066
- Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., et al. (2009). CHARMM: The biomolecular simulation program. *J. Comput. Chem.* 30, 1545–1614. doi: 10.1002/jcc.21287
- Brumer, H., Sims, P. F. G., and Sinnott, M. L. (1999). Lignocellulose degradation by *Phanerochaete chrysosporium*: purification and characterization of the main α -galactosidase. *Biochem. J.* 339, 43–53. doi: 10.1042/bj3390043
- Bugnon, M., Röhrig, U. F., Goullieux, M., Perez, M. A. S., Daina, A., Michielin, O., et al. (2024). SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina. *Nucleic Acids Res.* 52, W324–W332. doi: 10.1093/nar/gkae300
- Buntru, M., Vogel, S., Finnern, R., and Schillberg, S. (2022). “Plant-based cell-free transcription and translation of recombinant proteins,” in *Recombinant Proteins in Plants*. Eds. S. Schillberg and H. Spiegel (Springer US, New York, NY), 113–124. doi: 10.1007/978-1-0716-2241-4_8

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript.

AI was used for English writing.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1588802/full#supplementary-material>

- Buntru, M., Vogel, S., Spiegel, H., and Schillberg, S. (2014). Tobacco BY-2 cell-free lysate: an alternative and highly-productive plant-based *in vitro* translation system. *BMC Biotechnol.* 14, 37. doi: 10.1186/1472-6750-14-37
- Buntru, M., Vogel, S., Stoff, K., Spiegel, H., and Schillberg, S. (2015). A versatile coupled cell-free transcription-translation system based on tobacco BY-2 cell lysates. *Biotechnol. Bioeng.* 112, 867–878. doi: 10.1002/bit.25502
- Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55, 611–622. doi: 10.1373/clinchem.2008.112797
- Caffall, K. H., and Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* 344, 1879–1900. doi: 10.1016/j.carres.2009.05.021
- Carpita, N. C., and Gibeau, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30. doi: 10.1111/j.1365-313X.1993.tb00007.x
- Case, D. A., Aktulga, H. M., Belfon, K., Cerutti, D. S., Cisneros, G. A., Cruzeiro, V. W. D., et al. (2023). AmberTools. *J. Chem. Inf. Model.* 63, 6183–6191. doi: 10.1021/acs.jcim.3c01153
- Chen, W., Cui, Y., He, Y., Zhao, L., Cui, R., Liu, X., et al. (2023). Raffinose degradation-related gene GhAGAL3 was screened out responding to salinity stress through expression patterns of GhAGALS family genes. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1246677
- Chen, X.-Y., Kim, S. T., Cho, W. K., Rim, Y., Kim, S., Kim, S.-W., et al. (2009). Proteomics of weakly bound cell wall proteins in rice calli. *J. Plant Physiol.* 166, 675–685. doi: 10.1016/j.jplph.2008.09.010
- Chien, S.-F., Chen, S.-H., and Chien, M.-Y. (2008). Cloning, expression, and characterization of rice α -galactosidase. *Plant Mol. Biol. Rep.* 26, 213–224. doi: 10.1007/s11105-008-0035-6
- Chinen, I., Nakamura, T., and Fukuda, N. (1981). Purification and Properties of α -Galactosidase from Immature Fruits of *Saccharum officinarum* (Sugar Cane). *J. Biochem.* 90, 1453–1461. doi: 10.1093/oxfordjournals.jbchem.a133612
- Chou, K.-C., and Shen, H.-B. (2010). Plant-mPLoc: A top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS One* 5, e11335. doi: 10.1371/journal.pone.0011335
- Chrost, B., Kolukisaoglu, U., Schulz, B., and Krupinska, K. (2006). An α -galactosidase with an essential function during leaf development. *Planta* 225, 311–320. doi: 10.1007/s00425-006-0350-9
- Chuanhayan, P., Lee, R.-H., Guan, H.-H., Lin, C.-C., Chen, N.-C., Huang, Y.-C., et al. (2023). Structural insight into the hydrolase and synthase activities of an alkaline α -galactosidase from *Arabidopsis* from complexes with substrate/product. *Acta Crystallogr. Sect. Struct. Biol.* 79, 154–167. doi: 10.1107/S2059798323000037
- Comfort, D. A., Bobrov, K. S., Ivanen, D. R., Shabalin, K. A., Harris, J. M., Kulminskaya, A. A., et al. (2007). Biochemical analysis of *Thermotoga maritima* GH36 α -galactosidase (*Tm* GalA) confirms the mechanistic commonality of clan GH-D glycoside hydrolases. *Biochemistry* 46, 3319–3330. doi: 10.1021/bi061521n
- Cosgrove, D. J. (2024). Structure and growth of plant cell walls. *Nat. Rev. Mol. Cell Biol.* 25, 340–358. doi: 10.1038/s41580-023-00691-y
- Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Cui, Y., Zhao, Q., Hu, S., and Jiang, L. (2020). Vacuole biogenesis in plants: how many vacuoles, how many models? *Trends Plant Sci.* 25, 538–548. doi: 10.1016/j.jplants.2020.01.008
- Das Gupta, M., Flaskamp, Y., Roentgen, R., Juergens, H., Gimenez, J. A., Albrecht, F., et al. (2022). ALiCE[®]: A versatile, high yielding and scalable eukaryotic cell-free protein synthesis (CFPS) system. doi: 10.1101/2022.11.10.515920
- Dash, L., Swaminathan, S., Shimura, J., Gonzales, C. L. P., Montes, C., Solanki, N., et al. (2023). Changes in cell wall composition due to a pectin biosynthesis enzyme GAUT10 impact root growth. *Plant Physiol.* 193, 2480–2497. doi: 10.1093/plphys/kiad465
- De Caroli, M., Manno, E., Piro, G., and Lenucci, M. S. (2021). Ride to cell wall: *Arabidopsis* XTH11, XTH29 and XTH33 exhibit different secretion pathways and responses to heat and drought stress. *Plant J.* 107, 448–466. doi: 10.1111/tpj.15301
- De Coninck, T., Desmet, T., and Van Damme, E. J. M. (2024a). Carbohydrate-active enzymes involved in rice cell wall metabolism. *J. Exp. Bot.* 75, 6206–6227. doi: 10.1093/jxb/erae295
- De Coninck, T., Gippert, G. P., Henrissat, B., Desmet, T., and Van Damme, E. J. M. (2024b). Investigating diversity and similarity between CBM13 modules and ricin-B lectin domains using sequence similarity networks. *BMC Genomics* 25, 643. doi: 10.1186/s12864-024-10554-1
- Delmer, D., Dixon, R. A., Keegstra, K., and Mohnen, D. (2024). The plant cell wall—dynamic, strong, and adaptable—is a natural shapeshifter. *Plant Cell* 36, 1257–1311. doi: 10.1093/plcell/koad325
- Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469. doi: 10.1093/nar/gkn180
- Dey, P. M., and Pridham, J. B. (1972). “Biochemistry of α -galactosidases,” in *Advances in Enzymology - and Related Areas of Molecular Biology*. Ed. A. Meister (New Jersey, Hoboken: Wiley), 91–130. doi: 10.1002/9780470122815.ch3
- Di Stefano, M., Miceli, E., Gotti, S., Missanelli, A., Mazzocchi, S., and Corazza, G. R. (2007). The effect of oral α -galactosidase on intestinal gas production and gas-related symptoms. *Dig. Dis. Sci.* 52, 78–83. doi: 10.1007/s10620-006-9296-9
- Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 50, D571–D577. doi: 10.1093/nar/gkab1045
- Dubiel, M., De Coninck, T., Osterne, V. J. S., Verbeke, I., Van Damme, D., Smaghe, G., et al. (2020). The arathEULS3 lectin ends up in stress granules and can follow an unconventional route for secretion. *Int. J. Mol. Sci.* 21, 1659. doi: 10.3390/ijms21051659
- Elango, D., Rajendran, K., van der Laan, L., Sebastiar, S., Raigne, J., Thaiparambil, N. A., et al. (2022). Raffinose family oligosaccharides: friend or foe for human and plant health? *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.829118
- Elias, M., Wiczorek, G., Rosenne, S., and Tawfik, D. S. (2014). The universality of enzymatic rate-temperature dependency. *Trends Biochem. Sci.* 39, 1–7. doi: 10.1016/j.tibs.2013.11.001
- Ellis, M., Egelund, J., Schultz, C. J., and Bacic, A. (2010). Arabinogalactan-proteins: key regulators at the cell surface? *Plant Physiol.* 153, 403–419. doi: 10.1104/pp.110.156000
- El Meskini, R., Jin, L., Marx, R., Bruzzaniti, A., Lee, J., Emeson, R. B., et al. (2001). A signal sequence is sufficient for green fluorescent protein to be routed to regulated secretory granules. *Endocrinology* 142, 864–873. doi: 10.1210/endo.142.2.7929
- Eudes, A., Mouille, G., Thevenin, J., Goyallon, A., Minic, Z., and Jouanin, L. (2008). Purification, Cloning and Functional Characterization of an Endogenous beta-Glucuronidase in *Arabidopsis thaliana*. *Plant Cell Physiol.* 49, 1331–1341. doi: 10.1093/pcp/pcn108
- Evers, D., Ghislain, M., Hoffmann, L., Hausman, J. F., and Dommes, J. (2006). A late light resistant potato plant overexpresses a gene coding for α -galactosidase upon infection by *Phytophthora infestans*. *Biol. Plant* 50, 265–271. doi: 10.1007/s10535-006-0017-1
- Ferre, F., and Clote, P. (2005). DiANNA: a web server for disulfide connectivity prediction. *Nucleic Acids Res.* 33, W230–W232. doi: 10.1093/nar/gki412
- Fiser, A., Do, R. K. G., and Šali, A. (2000). Modeling of loops in protein structures. *Protein Sci.* 9, 1753–1773. doi: 10.1110/ps.9.9.1753
- Fujimoto, Z. (2013). Structure and function of carbohydrate-binding module families 13 and 42 of glycoside hydrolases, comprising a β -trefoil fold. *Biosci. Biotechnol. Biochem.* 77, 1363–1371. doi: 10.1271/bbb.130183
- Fujimoto, Z., Kaneko, S., Momma, M., Kobayashi, H., and Mizuno, H. (2003). Crystal structure of rice α -galactosidase complexed with D-galactose. *J. Biol. Chem.* 278, 20313–20318. doi: 10.1074/jbc.M302292200
- Fukimura, Y. (1988). “D-galactose,” in *Methods of Enzymatic Analysis* (VCH Publishers Ltd, Cambridge, UK), 288–296.
- Gao, Z., and Schaffer, A. A. (1999). A novel alkaline α -galactosidase from melon fruit with a substrate preference for raffinose. *Plant Physiol.* 119, 979–988. doi: 10.1104/pp.119.3.979
- Garman, S. C., and Garboczi, D. N. (2004). The molecular defect leading to fabry disease: structure of human α -galactosidase. *J. Mol. Biol.* 337, 319–335. doi: 10.1016/j.jmb.2004.01.035
- Garman, S. C., Hannick, L., Zhu, A., and Garboczi, D. N. (2002). The 1.9 Å structure of α -N-acetylgalactosaminidase: molecular basis of glycosidase deficiency diseases. *Structure* 10, 425–434. doi: 10.1016/s0969-2126(02)00726-8
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563
- Geilfus, C.-M., Tenhaken, R., and Carpentier, S. C. (2017). Transient alkalization of the leaf apoplast stiffens the cell wall during onset of chloride salinity in corn leaves. *J. Biol. Chem.* 292, 18800–18813. doi: 10.1074/jbc.M117.799866
- Ghosh, K., Takahashi, D., and Kotake, T. (2023). Plant type II arabinogalactan: Structural features and modification to increase functionality. *Carbohydr. Res.* 529, 108828. doi: 10.1016/j.carres.2023.108828
- Goetz, S., Rejzek, M., Nepogodiev, S. A., and Field, R. A. (2016). The impact of aminopyrene trisulfonate (APTS) label in acceptor glycan substrates for profiling plant pectin β -galactosyltransferase activities. *Carbohydr. Res.* 433, 97–105. doi: 10.1016/j.carres.2016.07.017
- Gojlo, E. (2023). Activity of α -D-galactosidase in long-stored seeds of *Vicia hirsuta*. *Agriculture* 13, 1306. doi: 10.3390/agriculture13071306
- Grosdidier, A., Zoete, V., and Michielin, O. (2011). Fast docking using the CHARMM force field with EADock DSS. *J. Comput. Chem.* 32, 2149–2159. doi: 10.1002/jcc.21797
- Gu, H., Lu, M., Zhang, Z., Xu, J., Cao, W., and Miao, M. (2018). Metabolic process of raffinose family oligosaccharides during cold stress and recovery in cucumber leaves. *J. Plant Physiol.* 224–225, 112–120. doi: 10.1016/j.jplph.2018.03.012
- Guce, A. I., Clark, N. E., Salgado, E. N., Ivanen, D. R., Kulminskaya, A. A., Brumer, H., et al. (2010). Catalytic mechanism of human α -galactosidase. *J. Biol. Chem.* 285, 3625–3632. doi: 10.1074/jbc.M109.060145
- Guimaraes, V. M., Tavares de Rezende, S., Alves Moreira, M., Gonçalves de Barros, E., and Felix, C. R. (2001). Characterization of α -galactosidases from germinating soybean seed and their use for hydrolysis of oligosaccharides. *Phytochemistry* 58, 67–73. doi: 10.1016/s0031-9422(01)00165-0

- Guo, S., Du, J., Li, D., Xiong, J., and Chen, Y. (2025). Versatile xylose and arabinose genetic switches development for yeasts. *Metab. Eng.* 87, 21–36. doi: 10.1016/j.jmb.2024.11.004
- Gupta, R., and Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* 7, 310–322.
- Hanes, C. S. (1932). Studies on plant amylases: The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. *Biochem. J.* 26, 1406–1421. doi: 10.1042/bj0261406
- Harbers, M. (2014). Wheat germ systems for cell-free protein expression. *FEBS Lett.* 588, 2762–2773. doi: 10.1016/j.febslet.2014.05.061
- Hart, D. O., He, S., Chany, C. J., Withers, S. G., Sims, P. F. G., Sinnott, M. L., et al. (2000). Identification of asp-130 as the catalytic nucleophile in the main α -galactosidase from *Phanerochaete chrysosporium*, a family 27 glycosyl hydrolase. *Biochemistry* 39, 9826–9836. doi: 10.1021/bi0008074
- Hazes, B. (1996). The (QxW)₃ domain: A flexible lectin scaffold. *Protein Sci.* 5, 1490–1501. doi: 10.1002/pro.5560050805
- Hellems, J., Mortier, G., De Paepe, A., Speleman, F., and Vandensompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* 8, R19. doi: 10.1186/gb-2007-8-2-r19
- Henrissat, B., and Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* 7, 637–644. doi: 10.1016/s0959-440x(97)80072-3
- Hicks, G. R. (2013). “Nuclear import of plant proteins,” in *Madame Curie Bioscience Database* (Landes Bioscience, Austin (TX)). Available at: <https://www.ncbi.nlm.nih.gov/books/NBK6124/>.
- Hocq, L., Habrylo, O., S n chal, F., Voxel, A., Pau-Roblot, C., Safran, J., et al. (2023). Mutation of AtPME2, a pH-dependent pectin methylesterase, affects cell wall structure and hypocotyl elongation. *Plant Cell Physiol.* 65, 301–318. doi: 10.1093/pcp/pcad154
- Hofer, F., Kraml, J., Kahler, U., Kamenik, A. S., and Liedl, K. R. (2020). Catalytic Site p K_a Values of Aspartic, Cysteine, and Serine Proteases: Constant pH MD Simulations. *J. Chem. Inf. Model.* 60, 3030–3042. doi: 10.1021/acs.jcim.0c00190
- Homolak, J., Babic Perhoc, A., Virag, D., Knezovic, A., Osmanovic Barilar, J., and Salkovic-Petrisic, M. (2024). D-galactose might mediate some of the skeletal muscle hypertrophy-promoting effects of milk—A nutrient to consider for sarcopenia? *BioEssays* 46, 2300061. doi: 10.1002/bies.202300061
- Hromadova, D., Soukup, A., and Tylova, E. (2021). Arabinogalactan proteins in plant roots – an update on possible functions. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.674010
- Hua, B., Zhang, M., Zhang, J., Dai, H., Zhang, Z., and Miao, M. (2021). CsAGA1 and csAGA2 mediate RFO hydrolysis in partially distinct manner in cucumber fruits. *Int. J. Mol. Sci.* 22, 13285. doi: 10.3390/ijms222413285
- Hughes, S. G., Overbeeke, N., Robinson, S., Pollock, K., and Smeets, F. L. M. (1988). Messenger RNA from isolated aleurone cells directs the synthesis of an alpha-galactosidase found in the endosperm during germination of guar (*Cyamopsis tetragonoloba*) seed. *Plant Mol. Biol.* 11, 783–789. doi: 10.1007/BF00019518
- Hutchinson, E. G., and Thornton, J. M. (1993). The Greek key motif: extraction, classification and analysis. *Protein Eng. Des. Sel.* 6, 233–245. doi: 10.1093/protein/6.3.233
- Ichinose, H., Fujimoto, Z., Honda, M., Harazono, K., Nishimoto, Y., Uzura, A., et al. (2009). A β -L-arabinopyranosidase from streptomyces avermitilis is a novel member of glycoside hydrolase family 27. *J. Biol. Chem.* 284, 25097–25106. doi: 10.1074/jbc.M109.022723
- Ikeda, M., Mitsuda, N., and Ohme-Takagi, M. (2009). *Arabidopsis* WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *Plant Cell* 21, 3493–3505. doi: 10.1105/tpc.109.069997
- Imazumi, C., Tomatsu, H., Kitazawa, K., Yoshimi, Y., Shibano, S., Kikuchi, K., et al. (2017). Heterologous expression and characterization of an *Arabidopsis* β -L-arabinopyranosidase and α -D-galactosidases acting on β -L-arabinopyranosyl residues. *J. Exp. Bot.* 68, 4651–4661. doi: 10.1093/jxb/erx279
- Iwata, Y., Fedoroff, N. V., and Koizumi, N. (2008). *Arabidopsis* bZIP60 is a proteolysis-activated transcription factor involved in the endoplasmic reticulum stress response. *Plant Cell* 20, 3107–3121. doi: 10.1105/tpc.108.061002
- Jamet, E., and Dunand, C. (2020). Plant cell wall proteins and development. *Int. J. Mol. Sci.* 21, 2731. doi: 10.3390/ijms21082731
- Jarvis, D. L., Summers, M. D., Garcia, A., and Bohlmeyer, D. A. (1993). Influence of different signal peptides and prosequences on expression and secretion of human tissue plasminogen activator in the baculovirus system. *J. Biol. Chem.* 268, 16754–16762. doi: 10.1016/S0021-9258(19)85481-9
- Jia, C.-F., Hu, W.-H., Chang, Z., and Gao, H.-L. (2015). Acid α -galactosidase is involved in D-chiro-inositol accumulation during tartary buckwheat germination. *Acta Soc Bot. Pol.* 84, 53–58. doi: 10.5586/asbp.2015.002
- Jiang, F., and Doudna, J. A. (2017). CRISPR–Cas9 structures and mechanisms. *Annu. Rev. Biophys.* 46, 505–529. doi: 10.1146/annurev-biophys-062215-010822
- Jiang, Y., Jiang, L., Akhil, C. S., Wang, D., Zhang, Z., Zhang, W., et al. (2023). MULocDeep web service for protein localization prediction and visualization at subcellular and suborganellar levels. *Nucleic Acids Res.* 51, W343–W349. doi: 10.1093/nar/gkad374
- Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29, 1859–1865. doi: 10.1002/jcc.20945
- Johnson, A. E., and Haigh, N. G. (2000). The ER translocon and retrotranslocation: is the shift into reverse manual or automatic? *Cell* 102, 709–712. doi: 10.1016/s0092-8674(00)00059-3
- Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 267, 727–748. doi: 10.1006/jmbi.1996.0897
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi: 10.1063/1.445869
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi: 10.1038/s41586-021-03819-2
- Jutras, P. V., Dodds, I., and van der Hoorn, R. A. (2020). Proteases of *Nicotiana benthamiana*: an emerging battle for molecular farming. *Curr. Opin. Biotechnol.* 61, 60–65. doi: 10.1016/j.copbio.2019.10.006
- Karimi, M., Inz , D., and Depicker, A. (2002). GATEWAYTM vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.* 7, 193–195. doi: 10.1016/S1360-1385(02)02251-3
- Katrolia, P., Rajashekara, E., Yan, Q., and Jiang, Z. (2014). Biotechnological potential of microbial α -galactosidases. *Crit. Rev. Biotechnol.* 34, 307–317. doi: 10.3109/07388551.2013.794124
- Kaur, D., Moreira, D., Coimbra, S., and Showalter, A. M. (2022). Hydroxyproline-O-galactosyltransferases synthesizing type II arabinogalactans are essential for male gametophytic development in *Arabidopsis*. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.935413
- Kempen, K. R., De Sancho, D., and Clarke, J. (2015). The response of Greek key proteins to changes in connectivity depends on the nature of their secondary structure. *J. Mol. Biol.* 427, 2159–2165. doi: 10.1016/j.jmb.2015.03.020
- Kiefer, P., Acland, P., Pappin, D., Peters, G., and Dickson, C. (1994). Competition between nuclear localization and secretory signals determines the subcellular fate of a single CUG-initiated form of FGF3. *EMBO J.* 13, 4126–4136. doi: 10.1002/j.1460-2075.1994.tb06730.x
- Kikuchi, A., Hara, K., Yoshimi, Y., Soga, K., Takahashi, D., and Kotake, T. (2022). *In vivo* structural modification of type II arabinogalactans with fungal endo- β -1,6-galactanase in *Arabidopsis*. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1010492
- Kikuchi, A., Okuyama, M., Kato, K., Osaki, S., Ma, M., Kumagai, Y., et al. (2017). A novel glycoside hydrolase family 97 enzyme: Bifunctional β -L-arabinopyranosidase/ α -galactosidase from *Bacteroides thetaiotaomicron*. *Biochimie* 142, 41–50. doi: 10.1016/j.biochi.2017.08.003
- Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., et al. (2008). GLYCAM06: A generalizable biomolecular force field. *Carbohydrates. J. Comput. Chem.* 29, 622–655. doi: 10.1002/jcc.20820
- Knoch, E., Dilokpimol, A., and Geshi, N. (2014). Arabinogalactan proteins: focus on carbohydrate active enzymes. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00198
- Kotake, T., Yamanashi, Y., Imaizumi, C., and Tsumuraya, Y. (2016). Metabolism of L-arabinose in plants. *J. Plant Res.* 129, 781–792. doi: 10.1007/s10265-016-0834-z
- Kufareva, I., and Abagyan, R. (2011). “Methods of protein structure comparison,” in *Homology Modeling*. Eds. A. J. W. Orry and R. Abagyan (Humana Press, Totowa, NJ), 231–257. doi: 10.1007/978-1-61779-588-6_10
- Kytidou, K., Beekwilder, J., Artola, M., Van Meel, E., Wilbers, R. H. P., Moolenaar, G. F., et al. (2018). *Nicotiana benthamiana* α -galactosidase A1.1 can functionally complement human α -galactosidase A deficiency associated with Fabry disease. *J. Biol. Chem.* 293, 10042–10058. doi: 10.1074/jbc.RA118.001774
- Lampert, D. T. A., Tan, L., and Kieliszewski, M. J. (2021). A molecular pinball machine of the plasma membrane regulates plant growth—A new paradigm. *Cells* 10, 1935. doi: 10.3390/cells10081935
- Lee, J., Hitzberger, M., Rieger, M., Kern, N. R., Zacharias, M., and Im, W. (2020). CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* 153, 035103. doi: 10.1063/5.0012280
- Lee, R., Hsu, J., Huang, H., Lo, S., and Grace Chen, S. (2009). Alkaline α -galactosidase degrades thylakoid membranes in the chloroplast during leaf senescence in rice. *New Phytol.* 184, 596–606. doi: 10.1111/j.1469-8137.2009.02999.x
- Lerbret, A., Bordat, P., Affouard, F., H doux, A., Guinet, Y., and Descamps, M. (2007). How do trehalose, maltose, and sucrose influence some structural and dynamical properties of lysozyme? Insight from molecular dynamics simulations. *J. Phys. Chem. B* 111, 9410–9420. doi: 10.1021/jp071946z
- Leszczuk, A., Cybulska, J., Skrzypczak, T., and Zdunek, A. (2020). Properties of arabinogalactan proteins (AGPs) in apple (*Malus × Domestica*) fruit at different stages of ripening. *Biology* 9, 225. doi: 10.3390/biology9080225
- Leszczuk, A., Kalaitzis, P., Kulik, J., and Zdunek, A. (2023). Review: structure and modifications of arabinogalactan proteins (AGPs). *BMC Plant Biol.* 23, 45. doi: 10.1186/s12870-023-04066-5

- Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 52, W78–W82. doi: 10.1093/nar/gkae268
- Li, T., Dong, M., Xie, W., Zhang, Y., Tao, D., and Li, S. (2018). Kinetic properties of raffinose synthase from rice (*Oryza sativa* L.). *Food Biosci.* 25, 39–43. doi: 10.1016/j.food.2018.07.006
- Li, B., Wang, Z., Jiang, H., Luo, J., Guo, T., Tian, F., et al. (2023). *ZmCCT10*-related photoperiod sensitivity regulates natural variation in the arithmetical formation of male germinal cells in maize. *New Phytol.* 237, 585–600. doi: 10.1111/nph.18559
- Lien, D. T. P., Tram, P. T. B., Trung, T. S., and Toan, H. T. (2018). Changes in α -galactosidase activity and oligosaccharides during germination of soybean seeds. *Can. Tho Univ. J. Sci.* 54, 8. doi: 10.22144/ctu.jsi.2018.089
- Liu, Z., Li, P., Yu, L., Hu, Y., Du, A., Fu, X., et al. (2023). OsMADS1 regulates grain quality, gene expressions, and regulatory networks of starch and storage protein metabolisms in rice. *Int. J. Mol. Sci.* 24, 8017. doi: 10.3390/ijms24098017
- Liu, H., Liu, X., Zhao, Y., Nie, J., Yao, X., Lv, L., et al. (2022). Alkaline α -galactosidase 2 (CsAGA2) plays a pivotal role in mediating source-sink communication in cucumber. *Plant Physiol.* 189, 1501–1518. doi: 10.1093/plphys/kiac152
- Liu, R., Zhao, D., Li, P., Xia, D., Feng, Q., Wang, L., et al. (2025). Natural variation in OsMADS1 transcript splicing affects rice grain thickness and quality by influencing monosaccharide loading to the endosperm. *Plant Commun.* 6, 101178. doi: 10.1016/j.xplc.2024.101178
- Loncharich, R. J., Brooks, B. R., and Pastor, R. W. (1992). Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanine-N'-methylamide. *Biopolymers* 32, 523–535. doi: 10.1002/bip.360320508
- Lopez-Hernandez, F., Tryfona, T., Rizza, A., Yu, X. L., Harris, M. O. B., Webb, A. A. R., et al. (2020). Calcium binding by arabinogalactan polysaccharides is important for normal plant development. *Plant Cell* 32, 3346–3369. doi: 10.1105/tpc.20.00027
- Lu, H., Chen, M., and Showalter, A. M. (2001). Developmental expression and perturbation of arabinogalactan-proteins during seed germination and seedling growth in tomato. *Physiol. Plant* 112, 442–450. doi: 10.1034/j.1399-3054.2001.1120319.x
- Lu, J., Wu, T., Zhang, B., Liu, S., Song, W., Qiao, J., et al. (2021). Types of nuclear localization signals and mechanisms of protein import into the nucleus. *Cell Commun. Signal.* 19, 60. doi: 10.1186/s12964-021-00741-y
- Ly, H. D., Howard, S., Shum, K., He, S., Zhu, A., and Withers, S. G. (2000). The synthesis, testing and use of 5-fluoro-a-D-galactosyl fluoride to trap an intermediate on green coffee bean α -galactosidase and identify the catalytic nucleophile. *Carbohydr. Res.* 329, 539–547. doi: 10.1016/S0008-6215(00)00214-7
- Madeira, F., Madhusodanan, N., Lee, J., Euseibi, A., Niewielska, A., Tivey, A. R. N., et al. (2024). The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res.* 52, W521–W525. doi: 10.1093/nar/gkae241
- Mareri, L., Romi, M., and Cai, G. (2019). Arabinogalactan proteins: actors or spectators during abiotic and biotic stress in plants? *Plant Biosyst. - Int. J. Deal. Asp. Plant Biol.* 153, 173–185. doi: 10.1080/11263504.2018.1473525
- Maricchiolo, E., Panfilii, E., Pompa, A., De Marchis, F., Bellucci, M., and Pallotta, M. T. (2022). Unconventional pathways of protein secretion: mammals vs. Plants. *Front. Cell Dev. Biol.* 10. doi: 10.3389/fcell.2022.895853
- Matsumoto, S., Taguchi, Y., Shimada, A., Igura, M., and Kohda, D. (2017). Tethering an N-glycosylation sequon-containing peptide creates a catalytically competent oligosaccharyltransferase complex. *Biochemistry* 56, 602–611. doi: 10.1021/acs.biochem.6b01089
- McCarter, J. D., and Stephen Withers, G. (1994). Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* 4, 885–892. doi: 10.1016/0959-440X(94)90271-2
- Michaelis, L., and Menten, M. L. (1913). The kinetics of invertase action. *Biochem. Z.* 49, 333–369. doi: 10.1016/j.febslet.2013.07.015
- Montor-Antonio, J. J., Hernández-Heredia, S., Ávila-Fernández, Á., Olvera, C., Sachman-Ruiz, B., and Del Moral, S. (2017). Effect of differential processing of the native and recombinant α -amylase from *Bacillus amyloliquefaciens* JJC33M on specificity and enzyme properties. *3 Biotech.* 7, 336. doi: 10.1007/s13205-017-0954-8
- Mutyemungu, E., Singh, M., Liu, S., and Rose, D. J. (2023). Intestinal gas production by the gut microbiota: A review. *J. Funct. Foods* 100, 105367. doi: 10.1016/j.jff.2022.105367
- Naumoff, D. G. (2004). Phylogenetic analysis of α -Galactosidases of the GH27 family. *Mol. Biol.* 38, 388–400. doi: 10.1023/B:MBIL.0000032210.97006.de
- Nibbering, P., Petersen, B. L., Motawia, M. S., Jørgensen, B., Ulvskov, P., and Niittylä, T. (2020). Golgi-localized exo- β 1,3-galactosidases involved in cell expansion and root growth in *Arabidopsis*. *J. Biol. Chem.* 295, 10581–10592. doi: 10.1074/jbc.RA120.013878
- Nilsson, I., and Von Heijne, G. (1992). A signal peptide with a proline next to the cleavage site inhibits leader peptidase when present in a *sec*-independent protein. *FEBS Lett.* 299, 243–246. doi: 10.1016/0014-5793(92)80124-Y
- Odonmazig, P., Erbringerová, A., Machova, E., and Alföldi, J. (1994). Structural and molecular properties of the arabinogalactan isolated from Mongolian larchwood (*Larix dahurica* L.). *Carbohydr. Res.* 252, 317–324. doi: 10.1016/0008-6215(94)90028-0
- Ødum, M. T., Teufel, F., Thumulari, V., Almagro Armenteros, J. J., Johansen, A. R., Winther, O., et al. (2024). DeepLoc 2.1: multi-label membrane protein type prediction using protein language models. *Nucleic Acids Res.* 52, W215–W220. doi: 10.1093/nar/gkae237
- Okazawa, A., Baba, A., Okano, H., Tokunaga, T., Nakae, T., Ogawa, T., et al. (2022). Involvement of α -galactosidase OmAGAL2 in plantose hydrolysis during seed germination of *Orobanche minor*. *J. Exp. Bot.* 73, 1992–2004. doi: 10.1093/jxb/erab527
- Okazawa, Y., Miyazaki, T., Yokoi, G., Ishizaki, Y., Nishikawa, A., and Tonozuka, T. (2015). Crystal structure and mutational analysis of isomalto-dextranase, a member of glycoside hydrolase family 27. *J. Biol. Chem.* 290, 26339–26349. doi: 10.1074/jbc.M115.680942
- Osterne, V. J. S., Pinto-Junior, V. R., Oliveira, M. V., Nascimento, K. S., Van Damme, E. J. M., and Cavada, B. S. (2024). Computational insights into the circular permutation roles on ConA binding and structural stability. *Curr. Res. Struct. Biol.* 7, 100140. doi: 10.1016/j.crsbti.2024.100140
- Peaucelle, A., Braybrook, S., and Höfte, H. (2012). Cell wall mechanics and growth control in plants: the role of pectins revisited. *Front. Plant Sci.* 3. doi: 10.3389/fpls.2012.00121
- Pennycooke, J. C., Jones, M. L., and Stushnoff, C. (2003). Down-regulating α -galactosidase enhances freezing tolerance in transgenic petunia. *Plant Physiol.* 133, 901–909. doi: 10.1104/pp.103.024554
- Perez, S. (2003). A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. *Biochimie* 85, 109–121. doi: 10.1016/S0300-9084(03)00053-1
- Petersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084
- Pfeifer, L., Shafee, T., Johnson, K. L., Bacic, A., and Classen, B. (2020). Arabinogalactan-proteins of *Zostera marina* L. contain unique glycan structures and provide insight into adaptation processes to saline environments. *Sci. Rep.* 10, 8232. doi: 10.1038/s41598-020-65135-5
- Ponder, G. R., and Richards, G. N. (1997). Arabinogalactan from Western larch, Part III: alkaline degradation revisited, with novel conclusions on molecular structure. *Carbohydr. Polym.* 34, 251–261. doi: 10.1016/S0144-8617(97)00099-4
- Rabouille, C. (2017). Pathways of unconventional protein secretion. *Trends Cell Biol.* 27, 230–240. doi: 10.1016/j.tcb.2016.11.007
- Rahfeld, P., and Withers, S. G. (2020). Toward universal donor blood: Enzymatic conversion of A and B to O type. *J. Biol. Chem.* 295, 325–334. doi: 10.1074/jbc.REV119.008164
- Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., and Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci.* 106, 1760–1765. doi: 10.1073/pnas.0813167106
- Regente, M., Pinedo, M., San Clemente, H., Balliau, T., Jamet, E., and de la Canal, L. (2017). Plant extracellular vesicles are incorporated by a fungal pathogen and inhibit its growth. *J. Exp. Bot.* 68, 5485–5495. doi: 10.1093/jxb/erx355
- Ren, Y., He, J., and Wang, X. (2007). Changes in activities of three enzymes degrading galactomannan during and following rice seed germination. *Rice Sci.* 14, 295–301. doi: 10.1016/S1672-6308(08)60008-6
- Roe, D. R., and Cheatham, T. E. (2013). PTRAJ and CPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* 9, 3084–3095. doi: 10.1021/ct400341p
- Röhrig, U. F., Goullieux, M., Bugnon, M., and Zoete, V. (2023). Attracting cavities 2.0: improving the flexibility and robustness for small-molecule docking. *J. Chem. Inf. Model.* 63, 3925–3940. doi: 10.1021/acs.jcim.3c00054
- Rose, J. K. C., and Lee, S.-J. (2010). Straying off the highway: trafficking of secreted plant proteins and complexity in the plant cell wall proteome. *Plant Physiol.* 153, 433–436. doi: 10.1104/pp.110.154872
- Rozen, S., and Skaletsky, H. (1999). "Primer3 on the WWW for general users and for biologist programmers," in *Bioinformatics Methods and Protocols* (Humana Press, New Jersey), 365–386. doi: 10.1385/1-59259-192-2.365
- Sakamoto, T., Tsujitani, Y., Fukamachi, K., Taniguchi, Y., and Ihara, H. (2010). Identification of two GH27 bifunctional proteins with β -L-arabinopyranosidase/ α -D-galactopyranosidase activities from *Fusarium oxysporum*. *Appl. Microbiol. Biotechnol.* 86, 1115–1124. doi: 10.1007/s00253-009-2344-6
- Sampedro, J., and Cosgrove, D. J. (2005). The expansin superfamily. *Genome Biol.* 6, 242. doi: 10.1186/gb-2005-6-12-242
- San Clemente, H., Kolkas, H., Canut, H., and Jamet, E. (2022). Plant cell wall proteomes: the core of conserved protein families and the case of non-canonical proteins. *Int. J. Mol. Sci.* 23, 4273. doi: 10.3390/ijms23084273
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019
- Schutze, M. P., Peterson, P. A., and Jackson, M. R. (1994). An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum. *EMBO J.* 13, 1696–1705. doi: 10.1002/j.1460-2075.1994.tb06434.x
- Seibel, N. M., Eljouni, J., Nalaskowski, M. M., and Hampe, W. (2007). Nuclear localization of enhanced green fluorescent protein homomultimers. *Anal. Biochem.* 368, 95–99. doi: 10.1016/j.ab.2007.05.025

- Seifert, G. J. (2020). On the potential function of type II arabinogalactan O-glycosylation in regulating the fate of plant secretory proteins. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.563735
- Serna, L. (2005). A simple method for discriminating between cell membrane and cytosolic proteins. *New Phytol.* 165, 947–952. doi: 10.1111/j.1469-8137.2004.01278.x
- Sharma, P., Sharma, S., Ramakrishna, G., Srivastava, H., and Gaikwad, K. (2022). A comprehensive review on leguminous galactomannans: structural analysis, functional properties, biosynthesis process and industrial applications. *Crit. Rev. Food Sci. Nutr.* 62, 443–465. doi: 10.1080/10408398.2020.1819196
- Shi, Y., Mowery, R. A., Ashley, J., Hentz, M., Ramirez, A. J., Bilgic, B., et al. (2012). Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. *Protein Sci.* 21, 1197–1209. doi: 10.1002/pro.2107
- Shindyalov, I. N., and Bourne, P. E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng. Des. Sel.* 11, 739–747. doi: 10.1093/protein/11.9.739
- Showalter, A. M. (2001). Arabinogalactan-proteins: structure, expression and function. *Cell. Mol. Life Sci.* 58, 1399–1417. doi: 10.1007/PL00000784
- Silva, J., Ferraz, R., Dupree, P., Showalter, A. M., and Coimbra, S. (2020). Three decades of advances in arabinogalactan-protein biosynthesis. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.610377
- Sinnett, M. L. (1990). Catalytic mechanisms of enzymic glycosyl transfer. *Chem. Rev.* 90, 1171–1202. doi: 10.1021/cr00105a006
- Smith, C. C., and Fretwell, S. D. (1974). The optimal balance between size and number of offspring. *Am. Nat.* 108, 499–506. doi: 10.1086/282929
- Snapp, E. L., McCaul, N., Quandt, M., Cabartova, Z., Bontjer, I., Källgren, C., et al. (2017). Structure and topology around the cleavage site regulate post-translational cleavage of the HIV-1 gp160 signal peptide. *eLife* 6, e26067. doi: 10.7554/eLife.26067
- Soh, C.-P., Ali, Z. M., and Lazan, H. (2006). Characterisation of an α -galactosidase with potential relevance to ripening related texture changes. *Phytochemistry* 67, 242–254. doi: 10.1016/j.phytochem.2005.09.032
- Sparkes, I. A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* 1, 2019–2025. doi: 10.1038/nprot.2006.286
- Spiwok, V. (2017). CH/ π Interactions in carbohydrate recognition. *Molecules* 22, 1038. doi: 10.3390/molecules22071038
- Srivastava, R., Deng, Y., and Howell, S. H. (2014). Stress sensing in plants by an ER stress sensor/transducer, bZIP28. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00059
- Steenfot, C., Vakhrushev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T.-B. G., et al. (2013). Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* 32, 1478–1488. doi: 10.1038/emboj.2013.79
- Steeves, R. M., Denton, M. E., Barnard, F. C., Henry, A., and Lambert, J. M. (1999). Identification of three oligosaccharide binding sites in ricin. *Biochemistry* 38, 11677–11685. doi: 10.1021/bi990493o
- Stellmach, H., Hose, R., Råde, A., Marillonnet, S., and Hause, B. (2022). A new set of golden-gate-based organelle marker plasmids for colocalization studies in plants. *Plants* 11, 2620. doi: 10.3390/plants11192620
- Stephenson, E., Estrada, S., Meng, X., Ourada, J., Muszynski, M. G., Habben, J. E., et al. (2019). Over-expression of the photoperiod response regulator ZmCCT10 modifies plant architecture, flowering time and inflorescence morphology in maize. *PLoS One* 14, e0203728. doi: 10.1371/journal.pone.0203728
- Strasser, R., Seifert, G., Doblin, M. S., Johnson, K. L., Ruprecht, C., Pfrengle, F., et al. (2021). Cracking the “Sugar code”: A snapshot of N- and O-glycosylation pathways and functions in plants cells. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.640919
- Studer, G., Rempfer, C., Waterhouse, A. M., Gumienny, R., Haas, J., and Schwede, T. (2020). QMEANDisCo—distance constraints applied on model quality estimation. *Bioinformatics* 36, 1765–1771. doi: 10.1093/bioinformatics/btz828
- Sturgeon, R. J. (1988). “L-arabinose,” in *Methods of Enzymatic Analysis* (VCH Publishers Ltd, Cambridge, UK), 427–431.
- Tan, L., Cheng, J., Zhang, L., Backe, J., Urbanowicz, B., Heiss, C., et al. (2024). Pectic-AGP is a major form of Arabidopsis AGPs. *Carbohydr. Polym.* 330, 121838. doi: 10.1016/j.carbpol.2024.121838
- Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., et al. (2013). An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an Arabinogalactan protein. *Plant Cell* 25, 270–287. doi: 10.1105/tpc.112.107334
- Tan, L., Zhang, L., Black, I., Glushka, J., Urbanowicz, B., Heiss, C., et al. (2023). Most of the rhamnogalacturonan-I from cultured *Arabidopsis* cell walls is covalently linked to arabinogalactan-protein. *Carbohydr. Polym.* 301, 120340. doi: 10.1016/j.carbpol.2022.120340
- Tapernoux-Lüthi, E. M., Böhm, A., and Keller, F. (2004). Cloning, functional expression, and characterization of the raffinose oligosaccharide chain elongation enzyme, galactan: galactan galactosyltransferase, from common bugle leaves. *Plant Physiol.* 134, 1377–1387. doi: 10.1104/pp.103.036210
- Teh, O. K., Singh, P., Ren, J., Huang, L. T., Ariyaratne, M., Salamon, B. P., et al. (2022). Surface-localized glycoproteins act through class C ARFs to fine-tune gametophore initiation in *Physcomitrium patens*. *Development* 149, dev200370. doi: 10.1242/dev.200370
- Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gislason, M. H., Pihl, S. I., Tsirigos, K. D., et al. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* 40, 1023–1025. doi: 10.1038/s41587-021-01156-3
- Tian, C., Kasavajhala, K., Belfon, K. A. A., Raguette, L., Huang, H., Migues, A. N., et al. (2020). fl9SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J. Chem. Theory Comput.* 16, 528–552. doi: 10.1021/acs.jctc.9b00591
- Tryfona, T., Liang, H.-C., Kotake, T., Tsumuraya, Y., Stephens, E., and Dupree, P. (2012). Structural characterization of Arabidopsis leaf arabinogalactan polysaccharides. *Plant Physiol.* 160, 653–666. doi: 10.1104/pp.112.202309
- Tsaniklidis, G., Benovias, A., Delis, C., and Aivalakis, G. (2016). Acidic alpha galactosidase during the maturation and cold storage of cherry tomatoes. *Acta Physiol. Plant* 38, 57. doi: 10.1007/s11738-016-2075-0
- Umer, M., and Kalra, D. K. (2023). Treatment of fabry disease: established and emerging therapies. *Pharmaceuticals* 16, 320. doi: 10.3390/ph16020320
- Van den Ende, W. (2013). Multifunctional fructans and raffinose family oligosaccharides. *Front. Plant Sci.* 4. doi: 10.3389/fpls.2013.00247
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paep, A., et al. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, research0034.1. doi: 10.1186/gb-2002-3-7-research0034
- Van Holle, S., De Schutter, K., Eggermont, L., Tsaneva, M., Dang, L., and Van Damme, E. J. M. (2017). Comparative study of lectin domains in model species: new insights into evolutionary dynamics. *Int. J. Mol. Sci.* 18, 1136. doi: 10.3390/ijms18061136
- Van Holle, S., and Van Damme, E. J. M. (2019). Messages from the past: new insights in plant lectin evolution. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00036
- Van Hove, J., Fouquaert, E., Smith, D. F., Proost, P., and Van Damme, E. J. M. (2011). Lectin activity of the nucleocytoplasmic EUL protein from *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* 414, 101–105. doi: 10.1016/j.bbrc.2011.09.031
- Van Leene, J., Eeckhout, D., Persiau, G., Van De Slijke, E., Geerinck, J., Van Isterdael, G., et al. (2011). “Isolation of transcription factor complexes from arabidopsis cell suspension cultures by tandem affinity purification,” in *Plant Transcription Factors*. Eds. L. Yuan and S. E. Perry (Humana Press, Totowa, NJ), 195–218. doi: 10.1007/978-1-61779-154-3_11
- Vega, M. C., Lorentzen, E., Linden, A., and Wilmanns, M. (2003). Evolutionary markers in the (β / α)8-barrel fold. *Curr. Opin. Chem. Biol.* 7, 694–701. doi: 10.1016/j.cbpa.2003.10.004
- Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., and Taylor, R. D. (2003). Improved protein–ligand docking using GOLD. *Proteins Struct. Funct. Bioinforma.* 52, 609–623. doi: 10.1002/prot.10465
- Vocadlo, D. J., Davies, G. J., Laine, R., and Withers, S. G. (2001). Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* 412, 835–838. doi: 10.1038/35090602
- Voinicic, C. (2022). Modern mannan: a hemicellulose’s journey. *New Phytol.* 234, 1175–1184. doi: 10.1111/nph.18091
- Wagner, P., and Hall, M. N. (1993). Nuclear protein transport is functionally conserved between yeast and higher eukaryotes. *FEBS Lett.* 321, 261–266. doi: 10.1016/0014-5793(93)80121-a
- Wang, R., and Brattain, M. G. (2007). The maximal size of protein to diffuse through the nuclear pore is larger than 60 kDa. *FEBS Lett.* 581, 3164–3170. doi: 10.1016/j.febslet.2007.05.082
- Wang, X., Chung, K. P., Lin, W., and Jiang, L. (2018). Protein secretion in plants: conventional and unconventional pathways and new techniques. *J. Exp. Bot.* 69, 21–37. doi: 10.1093/jxb/erx262
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427
- White, E. R., Reed, T. M., Ma, Z., and Hartman, M. C. T. (2013). Replacing amino acids in translation: Expanding chemical diversity with non-natural variants. *Methods* 60, 70–74. doi: 10.1016/j.jymeth.2012.03.015
- Wilbers, R. H. P., Westerhof, L. B., Van Raaij, D. R., Van Adrichem, M., Prakasa, A. D., Lozano-Torres, J. L., et al. (2016). Co-expression of the protease furin in *Nicotiana benthamiana* leads to efficient processing of latent transforming growth factor- β 1 into a biologically active protein. *Plant Biotechnol. J.* 14, 1695–1704. doi: 10.1111/pbi.12530
- Windels, A., Francaeus, J., Pleiss, J., and Desmet, T. (2024). CANDY: Automated analysis of domain architectures in carbohydrate-active enzymes. *PLoS One* 19, e0306410. doi: 10.1371/journal.pone.0306410
- Wolny, E., Betekhtin, A., Rojek, M., Braszewska-Zalewska, A., Lusinska, J., and Hasterok, R. (2018). Germination and the early stages of seedling development in brachypodium distachyon. *Int. J. Mol. Sci.* 19, 2916. doi: 10.3390/ijms19102916
- Xu, F., Gonneau, M., Faucher, E., Habrylo, O., Lefebvre, V., Domon, J.-M., et al. (2022). Biochemical characterization of Pectin Methyltransferase Inhibitor 3 from *Arabidopsis thaliana*. *Cell Surf.* 8, 100080. doi: 10.1016/j.cesw.2022.100080
- Yadav, R. K., Perales, M., Gruel, J., Girke, T., Jönsson, H., and Reddy, G. V. (2011). WUSCHEL protein movement mediates stem cell homeostasis in the *Arabidopsis* shoot apex. *Genes Dev.* 25, 2025–2030. doi: 10.1101/gad.17258511

- Yang, J., Su, L., Li, D., Luo, L., Sun, K., Yang, M., et al. (2020). Dynamic transcriptome and metabolome analyses of two types of rice during the seed germination and young seedling growth stages. *BMC Genomics* 21, 603. doi: 10.1186/s12864-020-07024-9
- Yu, C.-S., Cheng, C.-W., Su, W.-C., Chang, K.-C., Huang, S.-W., Hwang, J.-K., et al. (2014). CELLO2GO: A web server for protein subCELLular LOcalization prediction with functional gene ontology annotation. *PLoS One* 9, e99368. doi: 10.1371/journal.pone.0099368
- Yu, Q., Tang, C., and Kuo, J. (2000). A critical review on methods to measure apoplastic pH in plants. *Plant Soil* 219, 29–40. doi: 10.1023/A:1004724610550
- Yuan, J., Chen, D., Ren, Y., Zhang, X., and Zhao, J. (2008). Characteristic and expression analysis of a metallothionein gene, *osMT2b*, down-regulated by cytokinin suggests functions in root development and seed embryo germination of rice. *Plant Physiol.* 146, 1637–1650. doi: 10.1104/pp.107.110304
- Zemella, A., Thoring, L., Hoffmeister, C., and Kubick, S. (2015). Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. *ChemBioChem* 16, 2420–2431. doi: 10.1002/cbic.201500340
- Zhang, Q., Hrmova, M., Shirley, N. J., Lahnstein, J., and Fincher, G. B. (2006). Gene expression patterns and catalytic properties of UDP-D-glucose 4-epimerases from barley (*Hordeum vulgare* L.). *Biochem. J.* 394, 115–124. doi: 10.1042/BJ20051329
- Zhang, Y., Li, D., Dirk, L. M. A., Downie, A. B., and Zhao, T. (2021a). ZmAGA1 Hydrolyzes RFOs Late during the Lag Phase of Seed Germination, Shifting Sugar Metabolism toward Seed Germination Over Seed Aging Tolerance. *J. Agric. Food Chem.* 69, 11606–11615. doi: 10.1021/acs.jafc.1c03677
- Zhang, Z., Liu, Y., Dai, H., and Miao, M. (2021b). Characteristics and expression patterns of six α -galactosidases in cucumber (*Cucumis sativus* L.). *PLoS One* 16, e0244714. doi: 10.1371/journal.pone.0244714
- Zhao, T.-Y., Corum Iii, J. W., Mullen, J., Meeley, R. B., Helentjaris, T., Martin, D., et al. (2006). An alkaline α -galactosidase transcript is present in maize seeds and cultured embryo cells, and accumulates during stress. *Seed Sci. Res.* 16, 107–121. doi: 10.1079/SSR2006243
- Zhu, A., Wang, Z.-K., and Goldstein, J. (1995). Identification of tyrosine 108 in coffee bean α -galactosidase as an essential residue for the enzyme activity. *Biochim. Biophys. Acta* 1247, 260–264. doi: 10.1016/0167-4838(94)00228-9
- Zoete, V., Schuepbach, T., Bovigny, C., Chaskar, P., Daina, A., Röhrig, U. F., et al. (2016). Attracting cavities for docking. Replacing the rough energy landscape of the protein by a smooth attracting landscape. *J. Comput. Chem.* 37, 437–447. doi: 10.1002/jcc.24249