



**HAL**  
open science

## Plant polygalacturonase structures specify enzyme dynamics and processivities to fine-tune cell wall pectins

Josip Safran, Wafae Tabi, Vanessa Ung, Adrien Lemaire, Olivier Habrylo, Julie Bouckaert, Maxime Rouffle, Aline Voxeur, Paula Pongrac, Solène Bassard, et al.

### ► To cite this version:

Josip Safran, Wafae Tabi, Vanessa Ung, Adrien Lemaire, Olivier Habrylo, et al.. Plant polygalacturonase structures specify enzyme dynamics and processivities to fine-tune cell wall pectins. *The Plant cell*, 2023, pp.1-19. 10.1093/plcell/koad134 . hal-04106623

**HAL Id: hal-04106623**

**<https://u-picardie.hal.science/hal-04106623v1>**

Submitted on 2 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## RESEARCH ARTICLE

# Plant polygalacturonase structures specify enzyme dynamics and processivities to fine-tune cell wall pectins

Josip Safran<sup>1#</sup>, Wafae Tabi<sup>1#</sup>, Vanessa Ung<sup>2#</sup>, Adrien Lemaire<sup>1</sup>, Olivier Habrylo<sup>1</sup>, Julie Bouckaert<sup>3</sup>, Maxime Rouffle<sup>1</sup>, Aline Voxeur<sup>4</sup>, Paula Pongrac<sup>1</sup>, Solène Bassard<sup>1</sup>, Roland Molinié<sup>1</sup>, Jean-Xavier Fontaine<sup>1</sup>, Serge Pilard<sup>5</sup>, Corinne Pau-Roblot<sup>1</sup>, Estelle Bonnin<sup>6</sup>, Danaé Sonja Larsen<sup>2</sup>, Mélanie Morel-Rouhier<sup>7</sup>, Jean-Michel Girardet<sup>7</sup>, Valérie Lefebvre<sup>1</sup>, Fabien Sénéchal<sup>1§</sup>, Davide Mercadante<sup>2§</sup>, Jérôme Pelloux<sup>1§</sup>

<sup>1</sup>UMRT INRAE 1158 BioEcoAgro – BIOPI Biologie des Plantes et Innovation, Université de Picardie, 33 Rue St Leu, 80039 Amiens, France. <sup>2</sup>School of Chemical Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand. <sup>3</sup>UMR 8576 Unité de Glycobiologie Structurale et Fonctionnelle (UGSF), 50 Avenue de Halley, 59658 Villeneuve d'Ascq, France. <sup>4</sup>Université Paris-Saclay, INRAE, AgroParisTech, Institut Jean-Pierre Bourgin (IJPB), 78000, Versailles, France <sup>5</sup>Plateforme Analytique, Université de Picardie, 33, Rue St Leu, 80039 Amiens, France. <sup>6</sup>INRAE, UR 1268 Biopolymers, Interactions Assemblies, CS 71627, 44316 Nantes Cedex 3, France. <sup>7</sup>Université de Lorraine, INRAE, IAM, F-54000 Nancy, France.

#: Contributed equally as first authors, §: Contributed equally as last authors

**Short title:** Plant polygalacturonases processivities differ

**One-sentence summary:** Combined experimental and computational approaches show that the fine crystal structures of plant polygalacturonases differ, which has consequences for their processivities and effects on plants.

**Corresponding authors:** Jérôme Pelloux (jerome.pelloux@u-picardie.fr), Davide Mercadante ([davide.mercadante@auckland.ac.nz](mailto:davide.mercadante@auckland.ac.nz)), and Fabien Sénéchal (fabien.senechal@u-picardie.fr)

The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plcell/pages/General-Instructions>) are: Jérôme Pelloux (jerome.pelloux@u-picardie.fr) and Davide Mercadante (davide.mercadante@auckland.ac.nz).

## Abstract

Polygalacturonases (PGs) fine-tune pectins to modulate cell wall chemistry and mechanics, impacting plant development. The large number of PGs encoded in plant genomes leads to questions on the diversity and specificity of distinct isozymes. Herein, we report the crystal structures of two *Arabidopsis thaliana* polygalacturonases, POLYGALACTURONASE LATERAL ROOT (PGLR) and ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE2 (ADPG2), which are co-expressed during root development. We first determined the amino acid variations and steric clashes that explain the absence of inhibition of the plant PGs by endogenous PG-Inhibiting Proteins (PGIPs). Although their beta helix folds are highly similar, PGLR and ADPG2 subsites in the substrate-binding groove are occupied by divergent amino acids. By combining molecular dynamic simulations, analysis of enzyme kinetics and hydrolysis products, we showed that these structural differences translated into distinct enzyme-substrate dynamics and enzyme processivities: ADPG2 showed greater substrate fluctuations with hydrolysis products, oligogalacturonides (OGs), with a degree of polymerization (DP) of  $\leq 4$ , while the DP of OGs generated by PGLR was between 5 and 9. Using the *Arabidopsis* root as a developmental model, exogenous application of purified enzymes showed that the highly processive ADPG2 had major effects on both root cell elongation and cell adhesion. This work highlights the importance of PG processivity on pectin degradation regulating plant development.

## Introduction

The plant primary cell wall is composed of an intricate network of polysaccharides and proteins that is constantly being remodelled. Cell wall remodelling involves changes in its mechanical properties, which ultimately affect the extent of cell growth or the response to environmental stress (Bidhendi and Geitmann, 2016). Pectin, the major polysaccharide of the primary cell wall of dicotyledonous species such as *Arabidopsis* (*Arabidopsis thaliana*), is composed of homogalacturonan (HG): a homopolymer of  $\alpha$ -1,4-linked-D-galacturonic acid (GalA) units, that can be substituted with methylester and/or acetyl groups (Mohnen, 2008). The control of the degree of polymerization (DP) of HG by polygalacturonases (PGs) impacts diverse developmental processes such as root and hypocotyl growth, stomata functioning, cell separation during pollen formation and pollen tube elongation (Rhee et al., 2003; Ogawa et al., 2009; Xiao et al., 2014, 2017; Rui et al., 2017; Hocq et al., 2020).

Importantly, phytopathogenic organisms, including parasitic plants, also produce PGs, thus contributing to host colonization by degrading the physical barrier of the plant cell wall (Mutuku et al., 2021). Although all perform the hydrolysis of the  $\alpha$ -(1–4) glycosidic bond between two adjacent non-methylesterified GalA units, PGs can differ in their mode of action and are referred to as endo-PGs or exo-PGs if they either hydrolyse in the middle of the HG chain or attack from its non-reducing end (Park et al., 2010; Markovič and Janeček, 2001).

Resolved structures of PGs, all from microorganisms, fold into a right-handed parallel beta-helix and harbour four conserved amino acids (AA) stretches in their active site: namely NTD, DD, GHG and RIK (Markovič and Janeček, 2001). In a typical endo-PG, such as that from *Aspergillus aculeatus* PG1 (AaPG1), the active site is organized in a tunnel-like binding cleft, allowing the enzyme to bind the polysaccharide and produce pectic fragments called oligogalacturonides (OGs) of various DP and with different methyl and acetyl substitutions (Cho et al., 2001; Kohorn, 2016; Davidsson et al., 2017). In contrast, the structure of exo-PGs differs, loop extension turns the open-ended channel into a closed pocket, restricting the attack to the non-reducing end of the substrate, and releasing non-methylesterified GalA monomers or dimers (Abbott and Boraston, 2007).

It has been reported that pathogenic PGs are inhibited by Polygalacturonase Inhibiting Proteins (PGIPs), expressed by plants upon infection, either through competitive or non-competitive interactions, in a strategic attempt by plants to limit pectin degradation and pathogenic invasion (Benedetti et al., 2011; Kalunke et al., 2015). However, the nature of this inhibition is PG-specific as certain PGIPs were ineffective in mediating PG inhibition (Benedetti et al., 2013; Kalunke et al., 2015). In contrast a number of plant PGs are not inhibited by plant PGIPs, which suggests yet unidentified and specific structural features among this class of enzymes.

The PG-mediated degradation of HG can have two distinct consequences: i) it can impact polysaccharide rheology, decreasing cell wall stiffness and promoting cell growth (or infection by pathogens) and/or ii) it can produce OGs, which can act as signalling molecules (Ferrari et al., 2013; Davidsson et al., 2017). It seems likely that the fine composition of OG arrays produced by a myriad of differentially expressed PG isoforms can modulate the oligosaccharide interactions with cell wall integrity receptors, triggering distinct downstream signalling events (Kohorn, 2016).

In plants, PGs are encoded by large multigenic families (68 genes in *Arabidopsis thaliana*). The rationale for such an abundance of PGs in the context of the cell wall remains unclear. Considering such a large number of genes, and potential compensation mechanisms mediated by partial functional redundancy between isoforms, the use of reverse or forward genetic mutants can only bring partial clues to sample the diversity of the plant PG landscape.

Here we report the biochemical and structural characterization of two plant PGs, POLYGALACTURONASE LATERAL ROOT (PGLR) and ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE2 (ADPG2), whose expression overlaps in Arabidopsis roots during lateral root formation, in particular in cells above emerging primordia (González-Carranza et al., 2007; Hocq et al., 2020; Kumpf et al., 2013; Swarup et al., 2008). We first determined the structural features that explain the absence of inhibition of plant PG by endogenous PGIPs. We next found that, although having an overall conserved structure and overlapping functional profiles, enzymes have key and noticeable differences in their processivities. The investigation of PGLR and ADPG2 crystal structures, together with enzyme-substrate complexes, via combined experimental and computational approaches, including binding kinetics, Molecular Dynamics (MD) simulations, LC-MS/MS profiling of digestion products, indeed highlighted the existence of a link between enzyme-substrate interactions and

dynamics, enzyme activities and processivities. Using exogenous application of purified enzymes as a tool, we further showed that these distinct modes of action can translate into peculiar effects on root development. This overall shows that, despite apparent gene redundancy, plant PGs have distinct biochemical activities leading to peculiar consequences on plant development, which could be a key for the fine spatial and temporal tuning of cell wall chemistry and mechanics.

## Results

### Crystal structures of *A. thaliana* PGLR and ADPG2 reveal a conserved $\beta$ -fold

ADPG2 was produced as an active recombinant protein in the yeast *Pichia pastoris* and subsequently purified (Supplemental Figure S1A and Hocq et al., 2020). ADPG2 was biochemically characterized with regards to substrate, pH and temperature dependence (Supplemental Figure S1, B-D). Using polygalacturonic acid (PGA) as a substrate, at 25°C, PGLR and ADPG2 differ in their  $K_m$  (14.57 versus 3.0 mg.mL<sup>-1</sup>) and  $V_{max}$  (30.8 versus 11.0 nmol of GalA.min<sup>-1</sup>. $\mu$ g<sup>-1</sup>). Protein structures were determined by X-ray crystallography with the final models' geometry, processing and refinement statistics summarized in Table 1. We solved the crystal structure of PGLR (AA 429, AA1-18 and 409-429 not modelled,) at a resolution of 1.3 Å using molecular replacement (Figure 1A and Supplemental Figure S2A). PGLR crystallised as a single molecule in a P1 asymmetric unit. The crystal structure of ADPG2 (AA 420, AA 1-41 and 406-420 not modelled,) was resolved at a resolution of 2.0 Å (Figure 1A and Supplemental Figure S2B). ADPG2 crystals belonged to the orthorhombic space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> with chains A and B having a C $\alpha$  root mean square deviation (rmsd) of 0.924 Å.

PGLR and ADPG2 fold in right-handed parallel  $\beta$ -helical structure, which is common to pectinases (Figure 1A, Cho et al., 2001). This  $\beta$ -helix is formed by three repeating parallel  $\beta$ -sheets - PB1, PB2 and PB3, which contain 11, 12, and 11 parallel  $\beta$ -strands, respectively, as well as a small  $\beta$ -sheet, PB1a, having only 3  $\beta$ -strands (Supplemental Figure S3, A and B). T1, T1a, T2 and T3-turns connect the PB1-PB2, PB1-PB1a, PB2-PB3 and PB3-PB1  $\beta$ -sheets, respectively (Supplemental Figure S3, C and D, Yoder and Journak, 1995).

PGLR and ADPG2 show a  $\alpha$ -helix at the N-terminus, interacting with the T1 turn through the establishment of a disulphide bridge (PGLR, C46-C76; ADPG2, C71-C98), which shields the hydrophobic core of the enzyme (van Santen et al., 1999). Superimposition of PGLR and ADPG2 structures resulted in a rmsd of 2.299 Å, predominantly due to a deviation in the region surrounding the active site, in particular N130–P142 (T3 turn, PGLR numbering) and Y304-V318 (T1a turn, PGLR numbering). Between these loops, a large cleft (10.29 Å wide for PGLR and 14.46 Å for ADPG2), open at both sides is present, exposing PB1 for accommodating the substrate and identifying PGLR and ADPG2 as putative endo-PGs (Hocq et al., 2020; Abbott and Boraston, 2007; Cho et al., 2001).

### **Structural differences underlie the lack of plant PG–plant PGIP interactions**

While PGLR and ADPG2 show low sequence identity with fungal polygalacturonase enzymes (sequence identity: 19%-25% with *Aspergillus aculeatus* (AaPG1), *Aspergillus niger* (AnPGI and AnPGII), *Fusarium phyllophilum* (FpPG1), *Pectobacterium carotovorum* (PcPG1), *Chondrostereum purpureum* (CpPG1)), they show high structural similarity with a rmsd of 4.753 to 7.761 Å between all atoms (Supplemental Figure S4, A and B). Still, PGLR does not interact with plant PGIPs, as shown by the lack of inhibition of PGLR activity by *Phaseolus vulgaris* PGIP2 (PvPGIP2), while this interaction exists with fungal PGs (Benedetti et al., 2013; Hocq et al., 2020).

To understand the structural basis of this absence of inhibition of plant PG activity by PGIP we superimposed the resolved structures of PGLR and ADPG2 onto the *Fusarium phyllophilum* PG (FpPG1) - PvPGIP2 complex (Figure 1B, Benedetti et al., 2011, 2013). In FpPG1, a S120-N121-S122-N123 stretch, within the protein's N-terminal loop, plays a key role in the PG-PGIP interaction (N121 notably interacting with H110 of PvPGIP2). PGLR and ADPG2 N-terminal loops are, on the other hand, rich in bulkier and chemically different residues, including and K161, K164 and K166 for ADPG2 (Supplemental Figure S5). At the C-terminus, A274, the AA that contributes to hydrophobic-stabilizing interactions for the FpPG1-PvPGIP is replaced by G277/G278 and G303/G304 in PGLR and ADPG2, respectively (Figure 1C, Benedetti et al., 2013). Moreover, plant PGs have a specific H to P (P190/P216) substitution together with W275/Y301 insertion which can hinder the PG-PGIP interaction (Federici et al., 2001).

We next modelled Arabidopsis POLYGALACTURONASE INHIBITING PROTEIN1 and 2 (AtPGIP1 and AtPGIP2), which superimpose to PvPGIP2 with a rmsd of 1.194 and 1.201 Å, respectively (Figure 1D). The analysis of the models for PGLR/ADPG2-AtPGIP1/AtPGIP2 complexes showed that multiple AA are involved in steric clashes (between 81 and 275 atom contacts depending of the PG-PGIP pair, Supplemental Data Set S1), which, together with the above-mentioned structural features, can explain the absence of the interaction between PGs and PGIPs from Arabidopsis, and lack of protein-mediated inhibition of PG activity in planta (Figure 1E and Supplemental Figure S6, A and B).

To assess whether above-mentioned steric clashes determined for the interactions between PGLR/ADPG2 and AtPGIP1/AtPGIP2 can be expanded to other plant PGs, we modelled the complexes between AtPGIP1/AtPGIP2 and POLYGALACTURONASE INVOLVED IN EXPANSION1 (PGX1), PGX2, and PGX3, that were previously characterized (Xiao et al., 2014, 2017; Rui et al., 2017). This analysis showed that, depending on the AtPGIP/PG pairs, between 157 and 333 AA are involved in steric clashes further diminishing the possibility of interaction between plant PGs and plant PGIPs (Supplemental Figure S7 and Supplemental Data Set S2).

### **PGs with conserved active sites show differences alongside the binding groove subsites known to be important for substrate interaction and processivity**

Comparison of PGLR and ADPG2 sequences and structures with PGs from bacteria and fungi reveal that the active site is formed by four conserved structural motifs NTD, DD, GHG, RIK positioned at subsites -1 and +1 of the PB1 (André-Leroux et al., 2009; Shimizu et al., 2002; Pagès et al., 2000). Eight of these AA, N191/N217, D193/D219, H196/H222 D214/D240, D215/D241, H237/H263, R271/R297, and K273/K299 (PGLR/ADPG2 numbering) are strictly conserved, with the three aspartates responsible for the hydrolysis of the substrate (Figure 2A, Park et al., 2008; Markovič and Janeček, 2001; Shimizu et al., 2002; van Santen et al., 1999). To determine the importance of specific AA, five site-directed mutations were designed for PGLR: D215A occurring in the active site, R271Q (subsite +1), and the histidine mutants H196K, (subsite -1), H237K (subsite +1) and H196K/H237K (Supplemental Figure S8). Histidine residues could potentially modulate the activity of the enzyme by controlling the protonation state of residues placed in subsites flanking the hydrolysis site (Figure 2A).

Their activities on PGA and dissociation binding constants ( $K_d$ ) on a substrate of DP12 and degree of methylesterification 5 (DM5) (represented by a mix of OGs of mean DP12 and DM5 on which PGLR shows activity, Supplemental Figure S9) were determined. D215A and R271Q mutations resulted in a loss of activity (total loss for D215A) with a two- to four-time reduction in binding affinity, in particular for R271Q ( $K_d$  of 2567 nM and 4840 nM for D215A and R271Q, respectively compared to 1246 nM for wild type (WT), Figure 2B). While binding affinities of all histidine mutants were not significantly different, the H237K and H196K/H237K mutants showed slight residual activity while the H196K mutant had featured only 48% residual activity compared to that of the WT.

Although having conserved active sites, sequence and structure analyses showed that twelve AA positioned alongside the binding groove (subsites from -5 to +5), previously shown to be of importance for substrate interaction and processivity, differ between PGLR and ADPG2 (Figure 2, C and D) and the fungal AaPG1 (Supplemental Figure S10, Cho et al., 2001; André-Leroux et al., 2009; Pagès et al., 2000). For instance, at subsite -5, PGLR harbours R146, that can be responsible for the interaction with a carboxylate group of GalA, while ADPG2 harbours T172. Similarly, at subsite -4, Q198 in PGLR is replaced by T224 in ADPG2. At subsite -4, -3 and -2 a patch formed by Q198, Q220 and the positively charged K246 in PGLR is mutated into T224, E246 and D272 in ADPG2. At subsite -1 S269 in ADPG2, that can form hydrogen bonds with the substrate is mutated into G243 in PGLR. Finally, at subsite +2 and +3 D293 and K322 in ADPG2 are replaced by T267 and A296 in PGLR. While active and binding sites in AA are important for enzyme-substrate interactions and activity, subtle differences in AA composition might play a role in enzyme dynamics and binding affinities.

### **Molecular dynamic (MD) simulations reveal distinct substrate-dependent dynamics of PGLR and ADPG2**

The large number and chemical diversity of interactions across the binding groove make structural comparisons between different PG isoforms poorly informative. Such a diversity can result in different dynamic behaviours of enzymes and/or substrates, which could translate into different functional profiles. We performed MD simulations on PGLR and ADPG2 in complex with either a fully demethylesterified (pattern 1) or 60% methylesterified (pattern 2) decasaccharides

(Figure 2E), able to occupy the entire binding groove (subsites from -5 to +5). We first simulated PGLR, as well as H196K and H237K mutants in complex with fully demethylated deca-saccharides, and the analysis of substrate dynamics, through the quantification of subsite-specific root mean square fluctuations (RMSF), revealed a trend between enzymatic activity (Figure 2B), substrate dynamics (Figure 2, F-H) and the total number of contacts between the substrate and enzymes (Supplemental Figure S11, A and B).

MD simulations of the PGLR mutants (H196K and H237K) revealed how substrate dynamics is affected all along the binding groove, even with a single histidine mutation occurring in subsites either towards the non-reducing end (H196K – subsite -1) or the reducing end of the sugar (H237K – subsite +1). Overall, a rigidification of the substrate coincides with the loss of activity observed in experiments (Figure 2B), with the H237K mutant (strong loss of activity) showing the lowest RMSF in subsites -1 to +5 compared to the H196K (48% residual activity) and the WT (highest substrate dynamics, Figure 2, G and H).

The substrate dynamics can be also seen when comparing the RMSF of ADPG2 and PGLR when in complex with either de-methylated or methylated deca-saccharides. For both enzymes, de-methylated oligomers are overall less dynamic, hence more tightly bound in the binding groove (Figure 3, A and B). Quantitative differences in the RMSF of the two complexes suggest that, for the same substrate either being de-methylated or partially methylated, the binding to PGLR is tighter. Moreover, for each of the substrates, that vary in their degrees of methylation, ADPG2 has a higher activity compared to PGLR (Supplemental Figure S1E), which again corroborates the observation that methylated substrates are overall less dynamic in complex with PGLR when compared to ADPG2 (Figure 3, A and B).

The observed substrate dynamics is linked to the total number of contacts with the enzyme, with some noticeable differences between the two isoforms. When in complex with de-methylated substrates, both enzymes establish a larger number of contacts with the oligosaccharides. PGLR has however the ability to make a larger number of contacts, which is especially relevant for salt-bridges and hydrogen-bonds (Figure 3C). The reduced substrate dynamics when bound to histidine PGLR mutants corresponds with a higher number of contacts (Supplemental Figure S11A). A

comparison of the enzymatic motions revealed that PGLR and ADPG2, while engaged to the same decasaccharide substrate, explore separate conformational states, which are especially related to the fluctuations of unstructured regions flanking the binding groove. While for PGLR these are the regions flanking the substrate's non-reducing end (residues K108, R146, K169), in the case of ADPG2 they flank the binding cleft and in proximity of the substrate's reducing end (Supplemental Figure S12, A and B).

Relevant differences can also be observed between the electrostatic potentials of the two enzymes, calculated by solving the Poisson-Boltzmann equation in implicit solvent (Supplemental Figure S12, C and D). Compared to ADPG2, PGLR shows a much more positively charged electrostatic potential within the substrate binding cleft, in line with pronouncedly reduced dynamics for a negatively charged (demethylesterified) substrate, which would undergo much stronger electrostatically dominated interactions with the enzyme. Overall, subtle differences within the amino acid composition of certain enzyme subsites can convey specifically different activity profiles from a seemingly identical fold, which is likely to generate distinct substrate binding affinities, and end-products.

### **The differences in PGLR and ADPG2 binding's kinetics leads to specific pools of pectin-derived fragments**

The calculated RMSF shows differences in enzyme-substrate dynamics once the substrate is bound, which could reflect differences in the binding affinities of the enzymes towards specific substrates. Using the fluorescence-based switchSENSE® aptasensor, we determined binding kinetics for enzyme-substrate interactions for both PGLR and ADPG2, by quantifying substrate association ( $k_{on}$ ) and dissociation ( $k_{off}$ ) rate constants, as well as equilibrium dissociation constant ( $K_D$ ) using substrates with various DPs and DMs (PGA, pectins DM 20-34%, OGs of DP12DM5, DP12DM30, DP12DM60, Table 2).

ADPG2 displayed affinities much higher for low-DM substrates (i.e. PGA and DP12DM5) than those determined with the high-DM pectins ( $K_D$  ca. 10 to 60 times lower; Table 2) and comparable to those of PGLR. Considering the kinetics constants, PGLR and ADPG2 show no difference for  $k_{on}$  for pectins of low DM, including PGA and DP12DM5 (1320/1120 and 953/833  $M^{-1}s^{-1}$ , respectively). In contrast, when the DM of the substrate increases (DM 20-34%, DP12DM30 and DP12DM60), the  $k_{on}$  is always

roughly 3 to 16 times higher for PGLR compared to ADPG2. This suggests that for methylesterified pectins, PGLR, in line with the MD simulations and lower RMSF compared to ADPG2, associates much tighter with the substrate. This is as well reflected by the lower  $K_D$  determined for PGLR compared to ADPG2. No such drastic differences are measured for  $k_{off}$ , as values for PGLR and ADPG2 are in the same range for most substrates.

To determine whether the differences in subsites structure, enzyme dynamics and binding affinities can translate into differences in the processivity of PGLR and ADPG2, we assessed the products generated by either of the enzymes. Using PGA as a substrate, PGLR or ADPG2 maximum activities were reached after 1-hour digestion, generating products that cannot be further hydrolysed. ADPG2 total activity was higher than that measured for PGLR. Furthermore, the addition of ADPG2 following a first hour substrate incubation with PGLR, led to an increase in total PG activity, confirming putative differences in processivity between the two enzymes, ADPG2 being able to hydrolyse PGLR's end-products (Figure 4A).

We then used a recently developed LC-MS/MS oligoprofiling approach (Voxeur et al., 2019) to analyse the reaction products and confirmed, using PGA as a substrate, that both enzymes have endo activities, as suggested by the structural features of the binding cleft, and that ADPG2 releases higher proportion of short-sized OGs ( $\leq DP4$ ) compared to PGLR (Figure 4B). On pectic substrates of DM 20-34%, the pool of OGs produced by PGLR differed to that of ADPG2 (Figure 4C). In particular, PGLR released de-methylesterified OGs of DP5 to DP9, as well as specifically methylesterified forms of more than 6 GalA units that were either poorly represented or absent in the pool of end-products produced by ADPG2. The main products of ADPG2 were indeed de-methylesterified OGs of DP2 to DP4, as well as large amount of GalA4Me (Figure 4C and Figure 4C, Inset). When comparing the OGs produced by PGLR, ADPG2 and AaPG1 upon enzymatic activity on pectins with DM between 20 and 34% using principal component analysis (PCA), PGLR and AaPG1 were separated according to the first dimension (Dim1 54.6% of the variance) while ADPG2 clustered according to second dimension (Dim2 40.4% of the variance), with main loadings being, as an example, GalA2, GalA3, GalA4Me2, GalA9Me3 (Supplemental Figure S13, A and B).

Overall, ADPG2 and PGLR have nearly identical folds that, through distinct subsite structure and enzymes' dynamics, could translate into different enzymatic processivities. Indeed, PGLR and ADPG2 differ in their intrinsic processivities,  $P^{Intr}$ ,

being described as the average number of consecutive catalytic acts before enzyme-substrate dissociation.  $P^{Intr}$  is dependent on the dissociation probability,  $P_d$ , calculated using the turnover number ( $k_{cat}$ ) and rate constant of dissociation ( $k_{off}$ , Horn et al., 2012).  $P_d$  values were  $4.8 \times 10^{-4}$  and  $5.1 \times 10^{-5}$ , and  $P^{Intr}$  values were 2081 and 19777, for PGLR and ADPG, respectively (Supplemental Figure S1D). This data shows that, albeit acting both as processive enzymes ( $P_d \ll 1$ ), PGLR and ADPG2 differ in the extent by which they act on the substrate, with ADPG2 being much more processive than PGLR, as reflected by the lower size of the released products detected with LC-MS/MS.

### **The exogenous application of PGs with different processivities have distinct effects on root development**

Considering the localization of the expression of *PGLR* and *ADPG2* during root development, we tested the activity of both enzymes on root cell walls, whose pectins can be both methylesterified and acetylated (Willats et al., 2001; Kumpf et al., 2013; Swarup et al., 2008). Noticeably, PGLR released a higher proportion of acetylated OGs (including GalA5Ac, GalA6Ac, GalA6Ac2) compared to ADPG2, in addition to longer oligomers on average (Figure 5A and Figure 5A, Inset). Similar to what was observed on methylesterified pectins, the main OGs produced by ADPG2 were of lower DPs as compared to what produced by PGLR, corresponding mainly to unsubstituted GalA2 and GalA3.

As a read-out, and to determine how distinct processivities of PGLR and ADPG2 on HG can translate into distinct phenotypes *in muro*, we assessed the effects of exogenously-applied purified enzymes on developing Arabidopsis roots. Iso-activities of PGLR and ADPG2 were added in the culture medium of 6-day-old seedlings, for either one or three days, and phenotypic changes were examined. If one day's application of either of the enzymes did not affect root length, ADPG2 significantly impaired root elongation when applied for three days (Figure 5B). In contrast, in the latter condition, a slight effect was measured for PGLR albeit non-significant. (Figure 5B).

To determine more precisely if cells are differentially affected upon enzyme application depending on their spatial positioning (meristematic, elongating or fully elongated), we then measured the length of the firsts 50 cells from the root tip after

three days of enzymes' application, using EGFP-LTI6b reporter line that specifically labels plasma membrane (Figure 5C, Kurup et al., 2005). Cell length was not affected by the application of either of the enzymes up to the 40<sup>th</sup> cell. In contrast, the application of ADPG2 drastically reduced the length of the cells in the elongation zone as early as cell 40; while the effects measured for PGLR were from cell 46 onwards and were lower compared to that of ADPG2 (Figure 5D).

Further differences between the enzymes can be highlighted by analysing their effects on the morphology of the root cap, the structure at the tip of the root which supports growth and protects the root meristem. The application of ADPG2 for three days had much drastic effects on root cap detachment as compared to that of PGLR suggesting that it has more drastic effects on cell-to cell adhesion (Figure 5E). Altogether, this shows that the biochemical specificities/processivities of the two enzymes will ultimately translate into distinct effects on development, when applied exogenously in the culture media.

## Discussion

PGs play an important role in the control of pectin chemistry, contributing to changes in the cell wall mechanics, with important consequences on plant development (Ogawa et al., 2009; Xiao et al., 2014, 2017; Rui et al., 2017). In Arabidopsis, PGs are encoded by 68 genes: an abundance which is hard to rationalize within the context of the plant cell wall. Here we elucidated the structure-to-function relationships for two plant PGs, PGLR and ADPG2, whose gene expression patterns overlap in Arabidopsis roots. Both enzymes have nearly identical triple  $\beta$ -helix folds commonly found in other pectinases, including fungal endo-PGs (Shimizu et al., 2002; Cho et al., 2001; van Santen et al., 1999), pectin/pectate lyases (Vitali et al., 1998; Yoder and Journak, 1995; Lietzke et al., 1996) and rhamnogalacturonases (Petersen et al., 1997), with a large cleft opened at both sides that accommodates oligomeric substrates and confirms that PGLR and ADPG2 are endo-PGs (van Santen et al., 1999).

The resolution of the crystal structure for plant PGs first rationalized the structural determinants of the absence of inhibition of plant enzymes by plant PGIPs, as PGLR activity was indeed not inhibited by *P. vulgaris* PGIP2 (PvPGIP2, Hocq et al., 2020). Structurally, the key AA of *F. phyllophilum* FpPG1 (S120-N121-S122-N123) needed for determining the interaction of this pathogenic PG with PvPGIP2, are absent

in the T3 loop of PGLR and ADPG2. The homology modelling of Arabidopsis AtPGIP1 and AtPGIP2 further highlighted the absence of PGIP-mediated regulation of endogenous PG activity in plants as, albeit having highly conserved structure with that of PvPGIP2, they are lacking H110 and Q224 residues, required for inhibition (Ferrari et al., 2003). In addition, analysis of structural/homology models of AtPGIP1/AtPGIP2 in complex with a number of previously characterized PGs, including PGX1, PGX2 and PGX3, shows a high number of steric clashes between the different complexes (Xiao et al., 2014; Rui et al., 2017; Xiao et al., 2017). Based on this representative selection of PGs, this further demonstrates that plant PGIP are highly unlikely to interact, and inhibit, plant PGs. This suggests that cellular regulation of plant PG is mediated by other means at the cell wall, one of which being, as demonstrated in this study, the potential differential processivities of the enzymes.

The main challenge in understanding subtle differences between isoforms of PGs and other carbohydrate binding enzymes (CBEs) are mostly related to the large binding interface that characterizes the interaction between CBEs and oligomeric substrates. We tackled this challenge by designing strategic mutations across the binding cleft of the structurally characterized PGLR and functionally analysing the enzymes with combined computational and experimental methodologies. Our findings confirmed the importance of D215 for substrate hydrolysis, as well as R271 in binding and positioning the substrate at the catalytic subsite +1, as previously reported for fungal PGs (van Santen et al., 1999; Park et al., 2008).

Besides residues actively important in stabilizing the substrate, we find that other interactions in subsites flanking the catalytic subsite, crucially regulate substrate dynamics and corresponds with enzymatic activity. Histidine to lysine mutants in PGLR (H196K, H237K and H196K/H237K), that might generally be important in controlling the observed pH-dependent activity of other PGs, show how the distribution of charges affects substrate dynamics. Most interestingly, substrate rigidification reported by MD upon the insertion of a positive charges, increases the number of contacts with the substrate across the substrate binding interface and negatively impacts enzymatic activity as reported by the experimental biochemical characterization of the mutants. The importance of substrate dynamics in the activity of other CBEs has been also previously reported and it might be a key factor in regulating the processive activity of CBEs more generally, with processivity being limited by substrate dissociation (Mercadante et al., 2013, 2014).

We next investigated whether the processivities of PGLR and ADPG2 differ, which could be related to their different subsite's composition affecting enzymes' dynamics. For instance, D293 and K322 in ADPG2 are replaced by T267 and A296 in PGLR, which could modify the enzyme-substrate interaction and the enzyme specificity. The determination of the dynamics, measured as the RMSF, of the enzymes in complex with a decasaccharide of GalA showed that i) for a given enzyme, the enzyme's dynamics differs with the DM of the substrate and ii) ADPG2 was overall more dynamic, with a higher RMSF, as compared to PGLR. Together with these simulations, the determination of the binding kinetics of the enzyme-substrate interactions led to hypothesizing distinct processivities for the two enzymes. When considering pectins of high DM (DM 20-34%, DP12DM30, DP12DM60), the affinities of both enzymes are  $k_{on}$ -dominated, with PGLR associating much tighter with the substrate. Interestingly, the affinity of ADPG2 for the low-DM substrates is higher than that towards the high-DM pectins and is comparable to the affinities determined for PGLR. Considering the lubricating hypothesis, inferred from the studies on pectin methylesterases, and intrinsic processivity calculations, ADPG2 acts more processively on the HG chain than PGLR, and that would occur more favourably with low-DM substrates (Figure 6, Vitali et al., 1998; Mercadante et al., 2013).

Altogether, despite overall higher  $K_m$  value, these results are in accordance with both the lower RMSF, the substrate being more tightly bound inside the active site, and  $k_{on}$ , substrate association rate constant for the active site measured for PGLR. This would impair the sliding of the enzyme onto the chain, leading to enzyme-substrate dissociation and reiteration of enzyme attack onto the chain (Figure 6A). Such distinct processivities effectively translated into different end-products, with ADPG2 releasing OGs of short DP (methylesterified or not) from either commercial substrates or root cell wall extracts, while PGLR released a high proportion of non-methylesterified OGs of higher DP (Figure 6B). As highlighted by the fact that ADPG2 can hydrolyse PGLR-generated OGs, one could envisage a cooperative action of both enzymes in the cell wall to finely-tune HG structure during root development.

A number of studies previously showed the impact of the changes in PG activity, either through the study of loss-of-function mutants or over-expressing lines in *Arabidopsis*, on developmental processes as diverse as dark-grown hypocotyl development, stomata formation and root development (Rhee et al., 2003; Ogawa et al., 2009; Xiao et al., 2014, 2017; Rui et al., 2017; Hocq et al., 2020). However,

considering the size of the *PG* gene family, functional genomics approaches (mutants and/or overexpressing lines) can lead to counter intuitive results where changes in expression of one given *PG* gene leads to either increase, or decrease of total PG activity, owing to compensation mechanisms among the gene family. For instance, in *pglr* mutants, an increase in roots' total PG activity was measured, related to the upregulation of the expression of a subset of PG-encoding genes (Hocq et al., 2020). These results somehow show the challenge in assessing the precise role of given isoforms in planta by using mutants and specific complementation approaches, including promoter and enzyme swaps.

By using exogenous application of purified enzymes as a tool, one could expect visualizing more direct effects, which can allow linking the enzymes' processivities to their impact on pectins' structure, on their interaction with other cell wall components, on cell wall integrity, and ultimately on plant development. We however have to bear in mind that the plant's response to exogenously added proteins might differ to that of enzymes secreted in planta, in part related, but not exclusively, to the accessibility to their substrates, to the extent of protein glycosylation, or enzymes' concentrations and pH range in the apoplastic space (both largely unknown). However, we showed that the exogenous application of the highly processive ADPG2 had indeed different consequences on root development as compared to that observed with PGLR: HG remodelling upon ADPG2 application leads to strong defects in root elongation and in cell adhesion at the root cap. The root cap phenotype of ADPG2-treated roots is similar to that reported for *ROOT CAP PG1 (RCPG1)* overexpressing lines, known to be involved in root cap removal, suggesting that enzymes might share common biochemical specificities and/or processivities (Kamiya et al., 2016).

Altogether, in planta, and considering the overlapping expression patterns of *PGLR* and *ADPG*, their joint action could be required for proper cell wall hydrolysis leading to primordia emergence (González-Carranza et al., 2002; Hocq et al., 2020). Another approach, which might ease the conclusions that could be drawn from exogenous application experiments would be to use alternative plant models (*Physcomitrium patens* or *Marchantia polymorpha*) for which the number of PG-encoding genes is reduced. Using such organisms, gene compensation may be limited and phenotypes quantifiable.

Our work demonstrates that PGLR and ADPG2, albeit having a highly conserved structural fold, show subtle differences in their amino-acids composition at

the binding groove. This can translate into differences in enzymes' dynamics, substrate specificities, binding kinetics, leading to distinct processivities that have specific impacts on plant development. This shows the extent by which, among the multigenic family, each of the isoforms has distinct specificities that would be required, at the cell wall, to temporally and spatially control pectin structure. This further highlights that, for this class of enzymes, the gene redundancy at the genome level is unlikely to reflect redundant biochemical specificities. Our study now paves the way for a better understanding of how PG's processivities can control polysaccharides chemistry and mechanical properties *in muro*.

## **Materials and Methods**

### **Sequence analysis**

The presence of putative signal peptide in PGLR and ADPG2 was predicted using SignalP-5.0 Server (<http://www.cbs.dtu.dk/services/SignalP/>). Glycosylation sites were predicted using NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>). Sequence alignments were performed using MEGA and Clustal Omega multiple sequence alignment programs (Kumar et al., 2018).

### **Cloning, heterologous expression and purification of PGLR and ADPG2**

PGLR was previously expressed in the yeast *Pichia pastoris* and biochemically characterized (Hocq et al., 2020). Cloning and protein expression was done as previously described (Safran et al., 2021; Hocq et al., 2020). PGLR mutants were created using *Arabidopsis thaliana* cDNA and specific primers carrying mutations (Supplemental Table S1) in a two-step PCR reaction. Firstly, full gene primer and primer carrying mutations were used to create the two PCR amplicon corresponding to the full gene. Secondly, two parts of the gene were fused in a second PCR reaction. PGLR full gene primers were synthesised with *EcoRI* and *NotI* restriction sites. After PCR amplification gene carrying mutations were digested with *EcoRI* and *NotI* (New England Biolabs, Hitchin, UK) during 1h 30 min at 37°C. Gene were ligated into pPICZ $\alpha$ B vector (Invitrogen, Carlsbad, California, United States) using T4 ligase (New England Biolabs, Hitchin, UK) during a 12h ligation at 4°C. The pPICZ $\alpha$ B vector was

subsequently used for *Pichia pastoris* transformation following manufacturer's protocol (Invitrogen, Carlsbad, California, United States) .

At2g41850 (ADPG2) coding sequence was synthesized as a codon-optimized (Eurofins Luxembourg City, Luxembourg) version in pPICZ $\alpha$ B vector for *Pichia pastoris* expression.

### **PGLR and ADPG2 enzyme analysis**

Bradford method was used to determine the protein concentration, with bovine serum albumin (A7906, Sigma) as a standard. Deglycosylation was performed using Peptide-N-Glycosidase F (PNGase F) at 37 °C for one hour according to the supplier's protocol (New England Biolabs, Hitchin, UK). Enzyme purity and molecular weight were estimated by 12% (v/v) SDS-PAGE using mini-PROTEAN 3 system (BioRad, Hercules, California, United States). Gels were stained using PageBlue Protein Staining Solution (Thermo Fisher Scientific) according to the manufacturer's protocol.

### **PGLR and ADPG2 biochemical characterization**

The substrate specificity of PGLR and ADPG2 were determined with the DNS method as previously described (Hocq et al., 2020; Safran et al., 2021). PolyGalacturonic Acid (PGA, 81325, Sigma); Citrus pectin with DM 20-34% (P9311, Sigma), DM 55-70% (P9436, Sigma) were used as substrates. Results were expressed as nmol of GalA.min<sup>-1</sup>. $\mu$ g<sup>-1</sup> of proteins. The optimum temperature was determined by incubating the enzymatic reaction between 25 and 60°C during 60 min using PGA (0.4%, w/v) at pH5. The pH optimum was determined between pH 4 and 7 using 100 mM sodium acetate buffer (pH 3 to 5) and phosphate citrate buffer (pH 6 to 8) and 0.4% (w/v) PGA as a substrate. The PGLR and ADPG2 kinetic parameters were calculated using GraphPad Prism8 (version 8.4.2.) with PGA as a substrate. The reactions were performed using 1 to 8 mg.mL<sup>-1</sup> PGA concentrations during 10 min at 25°C in 50 mM sodium acetate (pH5). All experiments were realized in triplicate.

### **Digestion of cell wall pectins and released OGs profiling**

OGs released after digestions by recombinant PGLR and ADPG2 were identified as described (Voxeur et al., 2019). Briefly, PGA (81325, Sigma) or citrus pectin with DM 24-30% (P9311, Sigma) or OGs DP12DM5 (degree of polymerization centred on 12 and average DM of 5%) were prepared at 0.4% (w/v) final concentration

diluted in 100 mM ammonium acetate buffer (pH 5) and incubated with either PGLR and ADPG2 at  $0.03 \mu\text{g} \cdot \mu\text{L}^{-1}$ . Non-digested pectins were pelleted by centrifugation and the supernatant dried in speed vacuum concentrator (Concentrator plus, Eppendorf, Hamburg, Germany). The same procedure was applied for pectins from roots of *Arabidopsis thaliana* Col-0. Roots were cut, incubated in ethanol 100% (w/v) for 24 h, washed two times 5 min with acetone 100% (w/v) and left to dry 24 h. Thirty roots per replicate were rehydrated in 150  $\mu\text{L}$  100 mM ammonium acetate pH 5 during 2 h at room temperature and digested with PGLR or ADPG2 at  $0.02 \mu\text{g} \cdot \mu\text{L}^{-1}$  on average, using the above-mentioned protocol. Separation of OGs was achieved using an ACQUITY UPLC Protein BEH SEC column (125Å, 1.7  $\mu\text{m}$ , 4.6 mm x 300 mm), and the analysis was done as described (Safran et al., 2021). The data represent a minimum of three replicates.

### **Microscale thermophoresis**

Molecular interactions between PGLRs (WT and mutants) and DP12DM5 was done using Microscale thermophoresis (MST) approach as described with some modifications (Sénéchal et al., 2017). Briefly, PGLRs were labelled with monolith protein labelling kit blue NHS amine reactive (Lys, NanoTemper, catalog no. MO-L003) and conserved in MST buffer (50 mM Tris pH 7.4, 150 mM NaCl, 10 mM MgCl<sub>2</sub>, 0.05 % v/v tween-20). For all experiments, constant final concentration of labelled PGLRs was 1650 nM. Mix of OGs centered on DP12DM5 was prepared at 14028 nM concentration in MST buffer/dH<sub>2</sub>O in 1:1 ratio. For all experiments, a constant concentration of labelled PGLRs was titrated with decreasing concentrations of non-labelled DP12DM5 from 7014 to 0.214 nM. The resulting mixtures were loaded into a Monolith NT.115 series standard capillaries (NanoTemper, catalogue no. MO-K002). Thermophoresis experiments were performed with 40% of MST power and 20% of LED power for fluorescence acquisition. The data represents minimum three replicates.

### **Time-resolved molecular dynamics measurements**

PGLR and ADPG2 (used as ligands) were immobilized on an electro-switchable DNA biochip MPC-48-2-R1-S placed into a biosensor analyzer switchSENSE® DRX (Dynamic Biosensors GmbH, Planegg, Germany). For that, a covalent conjugate between PGLR or ADPG2 and a 48mer ssDNA was first prepared with the amine

coupling kit supplied by Dynamic Biosensors and purified by anion-exchange chromatography onto a proFIRE® system (Dynamic Biosensors), then hybridized with a complementary ssDNA attached on the surface of the biochip and carrying a Cy5 fluorescent probe at its free extremity. When analytes injected in the microfluidic system bind to the oscillating dsDNA nanolevers, the nanolever movement is altered by the additional friction imposed. Kinetic measurements for 2 min (association) and for 5 min (dissociation) were performed in 5 mM sodium acetate buffer, pH 5.5, with a flow rate of 100  $\mu\text{l}\cdot\text{min}^{-1}$  at 25°C with different concentrations of various analytes: PGA (81325, Sigma), citrus pectin with DM 24-30% (P9311, Sigma) and pool of OGs centred on DP12DM5, DP12DM30 and DP12DM60 at 25, 50 and 100  $\mu\text{M}$ . The fluorescence traces were analysed with the switchANALYSIS® software (V1.9.0.33, Dynamic Biosensors). The association and dissociation rates ( $k_{\text{on}}$  and  $k_{\text{off}}$ ), dissociation constant ( $K_{\text{D}} = k_{\text{off}}/k_{\text{on}}$ ) and the error values were derived from a global single exponential fit model, upon double referencing correction (blank and real-time, Müller-Landau and Varela, 2021). The experiments were performed in three replicates.

### Intrinsic processivity calculations

The intrinsic processivity potential ( $P^{\text{Intr}}$ ), a parameter corresponding to the number of consecutive catalytic steps before dissociation from the substrate was used as a measure of the processivities of PGLR and ADPG2 as described in Horn et al. (2012). The calculation of  $P^{\text{Intr}}$  is given in the Eq. 1.

$$\text{(Eq. 1)} \quad P^{\text{Intr}} = -\frac{1}{\ln(1 - P_d)}$$

The dissociation probability ( $P_d$ ) is expressed as a rate constant for two processes, (i) the turnover number ( $k_{\text{cat}}$ ) and (ii) the enzyme–substrate complex dissociation constant ( $k_{\text{off}}$ ).  $P_d$  is related to  $k_{\text{cat}}$  and  $k_{\text{off}}$  according to Eq. 2. In the case of processive enzymes  $P_d \ll 1$ .

$$\text{(Eq. 2)} \quad P_d = \frac{k_{\text{off}}}{k_{\text{off}} + k_{\text{cat}}}$$

The turnover number ( $k_{\text{cat}}$ ) was calculated using GraphPad Prism8 (version 8.4.2.) by fitting the non-linear regression curve following the Eq. 3, where Y is enzyme velocity, X is the substrate concentration,  $K_m$  is the Michaelis-Menten constant in the

same units as  $X$  and  $Et$  is the concentration of enzyme catalytic sites, 0.02307 and 0.001944 nM for PGLR and ADPG2, respectively.

$$\text{(Eq. 3) } k_{cat} = \frac{Y * (Km + X)}{Et * X}$$

### **Crystallization of proteins**

PGLR and ADPG2 were concentrated at 10 mg.mL<sup>-1</sup>. Crystallization was performed using the sitting-drop vapour-diffusion method at 18°C. Crystallisation conditions were screened using a mosquito robot (SPT Labtech) and the PACT premier™ plate (Molecular dimensions, Sheffield, UK). PGLR and ADPG2 (100 nL) were mixed with an equal volume of precipitant (1:1). The crystals that resulted in best diffraction data were obtained with 0.2 M sodium fluoride, 0.1 M bis-tris propane pH 8.5, 20% (w/v) PEG 3350 (H1 condition, PACT premier™ plate) for PGLR and 0.2 M sodium malonate dibasic monohydrate, 20% (w/v) PEG 3350 (E12 condition, PACT premier™ plate) for ADPG2. Crystals for PGLR and ADPG2 formed after 6 and 2 months, respectively. Scale-up of the best condition was realized by mixing 1 µl of the best precipitant condition with 1 µl of the enzyme in the hanging drop vapor-diffusion method.

### **X-ray data collection and processing**

Crystals were mixed with precipitation solution and PEG 3350 (35% w/v) before mounting in a loop and flash cooling in liquid nitrogen. The diffraction data were collected at PROXIMA-1 beamline (Synchrotron Soleil, Saint Aubin, France), at a temperature of -100.15°C using a PILATUS 6M end EIGER 16M detector (Dectris). Data were collected using X-rays with wavelength of 0.978564 Å. For PGLR, three data sets were collected from the same crystal to 1.3 Å resolution. Intensities were integrated, scaled and merged using XDS (Kabsch, 2010a) and XSCALE (Kabsch, 2010b). For ADPG2, one data set was collected to 2.0 Å resolution. Intensities were processed using XDS (Kabsch, 2010a). PGLR crystal belonged to triclinic space group P1 with one molecule in asymmetric unit, while ADPG2 belongs to orthorhombic space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> with two molecules in asymmetric units.

### **Structure solution and refinement and analysis**

For PGLR and ADPG2 structure and function prediction I-TASSER prediction software was used (Zhang, 2008). ColabFold and AlphaFold were used for AtPGIP1, AtPGIP2, PGX1, PGX2 and PGX3 modelling (Mirdita et al., 2022). The structure of PGLR was solved by molecular replacement using *Phaser* (McCoy et al., 2007). Model was built using *Autobuild* and refined using *Refine* from PHENIX (1.20.1-4487) suite (Liebschner et al., 2019). The model was iteratively improved with *Coot* (Emsley et al., 2010) and *Refine*. ADPG2 structure was solved by molecular replacement using ADPG2 I-tasser starting model and the above-mentioned iterative procedure. UCSF Chimera was used for creation of graphics (Pettersen et al., 2004).

### **Modelling and molecular dynamics simulations**

Molecular dynamics (MD) simulations were carried out on both the WT PGLR and ADPG2 protein structures in complex with fully de-methylesterified decasaccharides, as well as partially methylesterified decasaccharides (as described in Mercadante et al., 2014). Additionally, PGLR mutants H196K and H237K, modelled from the resolved X-ray crystal structures using PyMOL, were also simulated, in complex with fully de-methylesterified decasaccharides (The PyMOL Molecular Graphics System. Version 2.0 Schrödinger, LLC.).

Molecular topologies of the complexes were created according to the parameters of the AMBER14SB\_parmbsc1 forcefield (Lindorff-Larsen et al., 2010). The complexes were placed in cubic boxes with a solute-box distance of 1.0 nm and solvated with water molecules parameterised according to the TIP3P water model (Jorgensen et al., 1983). To neutralise the system's net charge and reach a salt concentration of 0.165 M, Na<sup>+</sup> and Cl<sup>-</sup> ions were added before energy-minimisation was performed.

The systems were then energy minimized, to resolve clashes between particles using a steep-descent algorithm with a step size of 0.01, considering convergence when the particle-particle force was 1000 kJ mol<sup>-1</sup> nm<sup>-1</sup>. Particle-particle forces were computing considering van der Waals and electrostatic interactions occurring up to 1.0 nm, treating long-range electrostatics in the Fourier space using the Particle Mesh Ewald (PME) summation method.

After minimization, solvent equilibration was achieved in two stages to reach constant temperature and pressure. The first stage was performed in the nVT ensemble while the second in the nPT ensemble. Solvent equilibration through the

nVT ensemble was carried out for 1 ns, with the equation of motion integrated with a time step of 2 fs, targeting a reference temperature of 310.15 K coupled every 0.1 ps using the V-rescale thermostat (Berendsen et al., 1984).

In this step, each particle in the system was assigned random velocities based on the Maxwell-Boltzmann distribution (Rowlinson, 2005) obtained at 310.15 K. Equilibration of the solvent through the nPT ensemble was then carried out for 1 ns starting from the last step (coordinates and velocities) of the previous equilibration, at a reference temperature of 310.15 K, coupled every 0.1 ps using the V-rescale thermostat (Berendsen et al., 1984). In this step, pressure coupling was conducted at 1 bar, with pressure coupled isotropically every 2.0 ps using the Parrinello-Rahman barostat (Parrinello and Rahman, 1981). Particle-particle interactions were calculated by building pair lists using the Verlet scheme. A cutoff of 1.0 nm was used to compute short-range van der Waals and electrostatic interactions sampled via a Coulomb potential. The Particle Mesh Ewald (PME) algorithm (Darden et al., 1993), with a Fourier grid spacing of 0.16 and a cubic B-spline interpolation level of 4, was used to compute, in the Fourier space, long-range electrostatic interactions past the cutoff.

Simulations were then performed on both in-house machines and on NeSI's (New Zealand eScience Infrastructure) high performance cluster, Mahuika, using GROMACS (Groningen MACHine for Chemical Simulation) version 2020.5 (Van Der Spoel et al., 2005). For each of the 6 complexes, simulations were run for 200 ns using a time step of 2 fs and replicated 5 times for a total simulation time of 1  $\mu$ s per complex. Each replicate differed in terms of the random sets of particle velocities generated through the nVT ensemble. Molecular dynamics trajectories were recorded every 10 ps. For analysis, the first 50 ns of each production run were considered equilibration time and discarded.

Analyses were conducted using in-house Python 3 scripts implemented Jupyter notebooks (Kluyver et al., 2016). Porcupine plots were created using data from a normalised principal component analysis calculated using GROMACS. Figures were created and rendered with Matplotlib (Hunter, 2007), VMD (Visual Molecular Dynamics, Humphrey et al., 1996) and UCSF Chimera (Pettersen et al., 2004).

### **Poisson-Boltzmann calculations of electrostatic potentials**

The protonation states of each amino acid were assigned according to the pKa curves calculated at pH = 4 for PGLR and pH = 5 for ADPG2, using the PROPKA software (Søndergaard et al., 2011). Atomic charges and radii for the protein atoms were assigned using the PDB2PQR software (Dolinsky et al., 2004) according to the parameters of the AMBER14SB\_parmbsc1 forcefield (Lindorff-Larsen et al., 2010), while atomic charges and radii for the sugar atoms were obtained from our previous work (Irani et al., 2018). The surface electrostatic potentials for WT PGLR and ADPG2 were then calculated solving the non-linearized form of the Poisson-Boltzmann equation through the APBS (Adaptive Poisson–Boltzmann Solver) software on a cubic grid composed of 193 grid points across the x-, y- and z- directions (Jurrus et al., 2018).

These calculations followed a stepwise approach where the Poisson-Boltzmann equation is first solved on a coarse mesh grid with a length of 155 Å and a spacing of 0.8 Å; then on a fine mesh grid with a length of 125 Å and a spacing of 0.64 Å. Calculations were solved considering a temperature of 218.15 K with a mobile ionic charge of +/- 1  $e_c$ , an ionic concentration of 0.165 M and an ionic radius of 2.0 Å. The protein dielectric constant was set at 4.0 and the solvent dielectric constant was set to 78.54. The protein surface electrostatic potentials were then visualised and coloured on the protein's molecular surface using VMD (Humphrey et al., 1996).

### **Plant growth conditions**

Sterile seeds of *Arabidopsis thaliana* Col-0 and EGFP-LTI6b (Kurup et al., 2005) plasma membrane marker-lines were sowed and grown in 400 µl liquid *Arabidopsis* ½ Murashige and Skoog medium (sucrose 10 g.L<sup>-1</sup>, MES monohydrate 0.5 g.L<sup>-1</sup> (Duchefa), in 24 well-plates (Murashige and Skoog, 1962). After 48 hours stratification, plates were placed in growth chamber under long day conditions (16 hours light/8 hours dark, 120 µmol.m<sup>-2</sup>.s<sup>-1</sup>, 21°C, spectra 400 to 700 nm).

### **Exogenous application of enzymes on Arabidopsis seedlings**

After 6 days, *Arabidopsis thaliana* seedlings were supplemented with 0.051 µg/µl and 0.015 µg/µl filter-sterilized PGLR and ADPG2, respectively using 0.2 µm PES filter (Whatman TM Puradisc TM 13 mm) in a volume of liquid MS medium of 200 µl to reach iso-activity. Plantlets were allowed to grown for another 1 day (T1) or 3 days (T3). Negative controls correspond to 6-, 7- or 9-days cultures with buffer only

(T0 ØEnz, T1 ØEnz and T3 ØEnz, respectively). For each of these conditions, measurements of primary root lengths were done using ImageJ software with NeuronJ plugin. For each condition, 30-40 plants were measured. For cell lengths determination, approximately 1 mm from the tip of the root of 3 to 7 plants were photographed under UV light using a stereomicroscope (ZEISS SteREO Discovery.V20). Images were assembled using MosaicJ plugin from Image J.

The length of the firsts 50 rhizodermal cells, starting from the first cell of the columella, were measured using image J software with NeuronJ plugin. Phenotypical observations where performed following ruthenium red staining (0.05% (w/v) in water, Sigma-Aldrich R-2751) under binocular microscope (Leica EZ4). The data represents minimum three replicates.

### **Statistical analysis**

All experiments show representative set of biological replicates. Biological replicates were performed with at least three technical repeats which are independent assays with the same biological materials. Sample sizes and number of technical replicates, used in each experiment are specified in the figure legends. Data were analysed by two-way ANOVA with Sidask's multiple comparison test (if not stated otherwise). Data are presented in Supplemental Data Set S3. Bar graphs and dot plots were used to show mean  $\pm$  SD and individual data points and were generated using Prism8. All statistical analysis was performed using GraphPad Prism8 (version 8.4.2.).

### **Accession Numbers**

Sequence data from this article can be found in the TAIR data libraries under gene number At5g14650 (*POLYGALACTURONASE LATERAL ROOT, PGLR*), At2g41850 (*ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE2, ADPG2*), At3g26610 (*POLYGALACTURONASE INVOLVED IN EXPANSION1, PGX1*), At1g78400 (*POLYGALACTURONASE INVOLVED IN EXPANSION2, PGX2*), At1g48100 (*POLYGALACTURONASE INVOLVED IN EXPANSION3, PGX3*), AT5G06860 (*POLYGALACTURONASE INHIBITING PROTEIN1, AtPGIP1*) and AT5G06870 (*POLYGALACTURONASE INHIBITING PROTEIN2, AtPGIP2*).

Fungal PGs sequences correspond to PDB codes 1BHE (*Pectobacterium carotovorum* PG1, PcPG1), 1NHC (*Aspergillus niger* PGI, AnPGI), and 1CZF

(AnPGII), 1HG8 (*Fusarium phyllophilum* PG1, FpPG1) 1IB4 (*Aspergillus aculeatus*, AaPG1) and 1KCD (*Chondrostereum purpureum*, CpPG1) and for PGIP, 1OGQ (*Phaseolus vulgaris*, PvPGIP2).

The final structure of PGLR and ADPG2 have been deposited in the PDB as entries 7B7A and 7B8B, respectively.

### **Supplemental Data**

Supplemental Figure S1. Purification and biochemical characterization of ADPG2.

Supplemental Figure S2. Crystallised PGLR and ADPG2 in asymmetric unit and glycosylation sites.

Supplemental Figure S3. PGLR and ADPG2 represent right-handed parallel  $\beta$ -helical structure.

Supplemental Figure S4. PGLR and ADPG2 sequence and structure identity with selected fungal enzymes.

Supplemental Figure S5. PGLR and ADPG2 N-terminal loops.

Supplemental Figure S6. Structural determinants of the absence of interaction between AtPGIP2 and PGLR-ADPG2.

Supplemental Figure S7. Structural determinants of the potential absence of interaction between AtPGIP1 and AtPGIP2 with PGX1, PGX2 and PGX3

Supplemental Figure S8. SDS-PAGE representing the purified wild type and mutants proteins of PGLR.

Supplemental Figure S9. Oligogalacturonides produced by PGLR and ADPG2 from pectins of DP12DM5.

Supplemental Figure S10. Structure of subsites of AaPG1.

Supplemental Figure S11. PGLR H196K and H237K mutants contact calculations.

Supplemental Figure S12. Porcupine plots and surface electrostatic potential of PGLR and ADPG2.

Supplemental Figure S13. PCA of OGs produced by PGLR, ADPG2 and AaPG1.

Supplemental Table S1. Primers for cloning mutated forms of PGLR and ADPG2 into pPICzαB expression vectors.

Supplemental Data Set S1. AtPGIP1 and AtPGIP2 contact analysis with PGLR and ADPG2.

Supplemental Data Set S2. AtPGIP1 and AtPGIP2 contact analysis with PGX1, PGX2, PGX3.

Supplemental Data Set S3. Statistical analysis

### **Funding Information**

This work was supported by a grant from the Agence Nationale de la Recherche (ANR-17-CE20-0023) and by the Conseil Regional Hauts-de-France and the FEDER (Fonds Européen de Développement Régional) through a PhD grant awarded to J.S.

### **Acknowledgements**

We wish to thank Pierre Legrand and all the staff at Proxima1 beamline (Synchrotron SOLEIL, Gif sur Yvette, France) for X-ray diffraction and data collection. The technical assistance of Maša Boras, a former master student is gratefully acknowledged.

### **Author Contributions**

J.P., F.S., V.L. and D.M. designed research; J.S., W.T., V.U., A.L., O.H.; J.B., A.V., S.B., M.R., P.P., E.B., S.P., D.S.L, M.M.R., J-M.G., D.M., C. P-R., and V.L. performed research; J.S., W.T., V.U., A.L., R.M., J-X.F., J-M.G., D.M., V.L., F.S. and J.P. analyzed data; J.S., W.T., F.S., V.L., D.M., and J.P. wrote the paper with input from J.B., E.B., J-M.G, M.M.R.

### **Conflict of interest statement**

The authors declare no competing interests.

## Tables

**Table 1. Data collection, processing and refinement statistics for PGLR and ADPG2.**

Statistics for the highest-resolution shell are shown in parentheses.

Characteristic	PGLR	ADPG2
Data collection		
Diffraction source	PROXIMA1A	PROXIMA1A
Wavelength (Å)	0.978	0.978
Temperature (°C)	-100.15	-100.15
Detector	PILATUS3 6M	EIGER 16M
Crystal to detector distance (mm)	190.0	279.3
Rotation range per image (°)	0.1	0.1
Total rotation range (°)	360	360
Crystal data		
Space group	P1	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
<i>a</i> , <i>b</i> , <i>c</i> (Å)	38.97, 41.83, 63.33	71.78, 88.56, 113.87
$\alpha$ , $\beta$ , $\gamma$ (°)	93.25, 99.86, 114.95	90.00, 90.00, 90.00
Subunits per asymmetric unit	1	2
Data statistics		
Resolution range (Å)	33.33 - 1.3	44.61 - 2.0
Total No. of reflection	761821 (55201)	645204 (65696)
No. of unique reflection	83668 (8043)	47381 (4654)
No. of reflections, test set	4182 (401)	2368 (233)
$R_{\text{merge}}$ (%)	7.64 (77.7)	8.9 (97)
Completeness (%)	96.1 (92.0)	99.9 (99.3)
$\langle I/\sigma(I) \rangle$	16.24 (2.78)	16.91 (2.56)
Multiplicity	9.1(6.9)	13.6 (14.0)
CC <sub>1/2</sub> (%)	99 (86.3)	99 (92.1)
Refinement		
$R_{\text{crys}}/R_{\text{free}}$ (%)	14.2 / 17.7	18.9 / 23.0
Average B - factor (Å <sup>2</sup> )	29.1	27.89
No. of non-H atoms		
Protein	3085	5563
Ion	-	5
Ligand	100	-
Water	609	999
Total	3794	6567
R.m.s. deviations		
Bonds (Å)	0.015	0.006
Angles (°)	1.59	1.06
Ramachandran plot		
Most favoured (%)	94.6	93.58
Allowed (%)	5.4	6.28
Outlier (%)	-	0.14

**Table 2. kon, koff and kD measurements for PGLR and ADPG using substrates of various degrees of polymerization.**

PGA: Polygalacturonic acid, pectins DM 20-34%: Commercial pectins of DM30%, DP12DM5/DP12DM30/DP12DM60: Pool of OG centered on DP12 with increasing DM (5%, 30%, 60%). Values correspond to means  $\pm$  SE of three replicates.

Substrate	PGLR			ADPG2		
	$K_D$ ( $\mu\text{M}$ )	$k_{on}$ ( $\text{M}^{-1}\text{s}^{-1}$ )	$k_{off}$ ( $\text{ms}^{-1}$ )	$K_D$ ( $\mu\text{M}$ )	$k_{on}$ ( $\text{M}^{-1}\text{s}^{-1}$ )	$k_{off}$ ( $\text{ms}^{-1}$ )
PGA	8.12 $\pm$ 0.7	1320 $\pm$ 110	10.7 $\pm$ 0.2	4.3 $\pm$ 1.1	1120 $\pm$ 230	4.8 $\pm$ 0.8
DM 20-34%	12.1 $\pm$ 2.2	1010 $\pm$ 180	12.2 $\pm$ 0.4	194 $\pm$ 89	62.8 $\pm$ 28.6	12.2 $\pm$ 0.3
DP12DM5	12.6 $\pm$ 1.0	953 $\pm$ 71	12 $\pm$ 0.3	10.7 $\pm$ 0.9	833 $\pm$ 64	8.9 $\pm$ 0.3
DP12DM30	26.8 $\pm$ 3.6	268 $\pm$ 23	7.2 $\pm$ 0.7	267 $\pm$ 79	28.8 $\pm$ 7.8	7.7 $\pm$ 0.9
DP12DM60	38 $\pm$ 7.4	196 $\pm$ 29	7.5 $\pm$ 0.9	155 $\pm$ 47	54.9 $\pm$ 16.0	8.5 $\pm$ 0.7

## REFERENCES

- Abbott, D.W. and Boraston, A.B.** (2007). The Structural Basis for Exopolygalacturonase Activity in a Family 28 Glycoside Hydrolase. *J. Mol. Biol.* **368**: 1215–1222.
- André-Leroux, G., Tessier, D., and Bonnin, E.** (2009). Endopolygalacturonases reveal molecular features for processivity pattern and tolerance towards acetylated pectin. *Biochim. Biophys. Acta - Proteins Proteomics* **1794**: 5–13.
- Benedetti, M., Andreani, F., Leggio, C., Galantini, L., Di Matteo, A., Pavel, N.V., De Lorenzo, G., Cervone, F., Federici, L., and Sicilia, F.** (2013). A single amino-acid substitution allows endo-polygalacturonase of *Fusarium verticillioides* to acquire recognition by PGIP2 from *Phaseolus vulgaris*. *PLoS One* **8**: 1–11.
- Benedetti, M., Leggio, C., Federici, L., de Lorenzo, G., Pavel, N.V., and Cervone, F.** (2011). Structural resolution of the complex between a fungal polygalacturonase and a plant polygalacturonase-inhibiting protein by small-angle x-ray scattering. *Plant Physiol.* **157**: 599–607.
- 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.

- Bidhendi, A.J. and Geitmann, A.** (2016). Relating the mechanics of the primary plant cell wall to morphogenesis. *J. Exp. Bot.* **67**: 449–461.
- Cho, S.W., Lee, S., and Shin, W.** (2001). The X-ray structure of *Aspergillus aculeatus* Polygalacturonase and a Modeled structure of the Polygalacturonase-Octagalacturonate Complex. *J. Mol. Biol.* **311**: 863–878.
- Darden, T., York, D., and Pedersen, L.** (1993). Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**: 10089–10092.
- Davidsson, P., Broberg, M., Kariola, T., Sipari, N., Pirhonen, M., and Palva, E.T.** (2017). Short oligogalacturonides induce pathogen resistance-associated gene expression in *Arabidopsis thaliana*. *BMC Plant Biol.* **17**: 1–17.
- Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A.** (2004). PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**: 665–667.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K.** (2010). Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**: 486–501.
- Federici, L., Caprari, C., Mattei, B., Savino, C., Di Matteo, A., De Lorenzo, G., Cervone, F., and Tsernoglou, D.** (2001). Structural requirements of endopolygalacturonase for the interaction with PGIP (polygalacturonase-inhibiting protein). *Proc. Natl. Acad. Sci. U. S. A.* **98**: 13425–13430.
- Ferrari, S., Savatin, D. V., Sicilia, F., Gramegna, G., Cervone, F., and Lorenzo, G. de** (2013). Oligogalacturonides: plant damage-associated molecular patterns and regulators of growth and development. *Front. Plant Sci.* **4**: 1–9.
- Ferrari, S., Vairo, D., Ausubel, F.M., Cervone, F., and De Lorenzo, G.** (2003). Tandemly duplicated *Arabidopsis* genes that encode polygalacturonase-inhibiting proteins are regulated coordinately by different signal transduction pathways in response to fungal infection. *Plant Cell* **15**: 93–106.
- González-Carranza, Z.H., Elliott, K.A., and Roberts, J.A.** (2007). Expression of polygalacturonases and evidence to support their role during cell separation processes in *Arabidopsis thaliana*. *J. Exp. Bot.* **58**: 3719–3730.
- González-Carranza, Z.H., Whitelaw, C.A., Swarup, R., and Roberts, J.A.** (2002).

Temporal and Spatial Expression of a Polygalacturonase during Leaf and Flower Abscission in Oilseed Rape and Arabidopsis. *Plant Physiol.* **128**: 534–543.

**Hocq, L. et al.** (2020). The exogenous application of AtPGLR, an endo - polygalacturonase, triggers pollen tube burst and repair. *Plant J.* **103**: 617–633.

**Horn, S.J., Sorlie, M., Vårum, K.M., Våljamäe, P., and Eijsink, V.G.H.** (2012). Measuring processivity. *Methods Enzymol.* **510**: 69–95.

**Humphrey, W., Dalke, A., and Schulten, K.** (1996). VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **14**: 33–38.

**Hunter, J.D.** (2007). Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**: 90–95.

**Irani, A.H., Mercadante, D., and Williams, M.A.K.** (2018). On the electrophoretic mobilities of partially charged oligosaccharides as a function of charge patterning and degree of polymerization. *Electrophoresis* **39**: 1497–1503.

**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.

**Jurrus, E. et al.** (2018). Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**: 112–128.

**Kabsch, W.** (2010a). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**: 133–144.

**Kabsch, W.** (2010b). Xds. *Acta Crystallogr. D. Biol. Crystallogr.* **66**: 125–32.

**Kalunke, R.M., Tundo, S., Benedetti, M., Cervone, F., De Lorenzo, G., and D'Ovidio, R.** (2015). An update on polygalacturonase-inhibiting protein (PGIP), aleucine-rich repeat protein that protects crop plants against pathogens. *Front. Plant Sci.* **6**: 1–17.

**Kamiya, M., Higashio, S.Y., Isomoto, A., Kim, J.M., Seki, M., Miyashima, S., and Nakajima, K.** (2016). Control of root cap maturation and cell detachment by BEARSKIN transcription factors in Arabidopsis. *Development* **143**: 4063–4072.

- Kluyver, T. et al.** (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. Position. Power Acad. Publ. Play. Agents Agendas - Proc. 20th Int. Conf. Electron. Publ. ELPUB 2016: 87–90.
- Kohorn, B.D.** (2016). Cell wall-associated kinases and pectin perception. *J. Exp. Bot.* **67**: 489–494.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K.** (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**: 1547–1549.
- Kumpf, R.P., Shi, C.L., Larrieu, A., Stø, I.M., Butenko, M.A., Péret, B., Riiser, E.S., Bennett, M.J., and Aalen, R.B.** (2013). Floral organ abscission peptide IDA and its HAE/HSL2 receptors control cell separation during lateral root emergence. *Proc. Natl. Acad. Sci. U. S. A.* **110**: 5235–5240.
- Kurup, S., Runions, J., Köhler, U., Laplaze, L., Hodge, S., and Haseloff, J.** (2005). Marking cell lineages in living tissues. *Plant J.* **42**: 444–453.
- Liebschner, D. et al.** (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **75**: 861–877.
- Lietzke, S.E., Scavetta, R.D., Yoder, M.D., and Journak, F.** (1996). The Refined Three-Dimensional Structure of Pectate Lyase E from *Erwinia chrysanthemi* at 2.2 Å Resolution. *Plant Physiol.* **111**: 73–92.
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., and Shaw, D.E.** (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma.* **78**: 1950–1958.
- Markovič, O. and Janeček, Š.** (2001). Pectin degrading glycoside hydrolases of family 28: sequence-structural features, specificities and evolution. *Protein Eng. Des. Sel.* **14**: 615–631.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J.** (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**: 658–674.

- Mercadante, D., Melton, L.D., Jameson, G.B., and Williams, M.A.K.** (2014). Processive pectin methylesterases: The role of electrostatic potential, breathing motions and bond cleavage in the rectification of brownian motions. *PLoS One* **9**: 1–11.
- Mercadante, D., Melton, L.D., Jameson, G.B., Williams, M.A.K., and De Simone, A.** (2013). Substrate dynamics in enzyme action: Rotations of monosaccharide subunits in the binding groove are essential for pectin methylesterase processivity. *Biophys. J.* **104**: 1731–1739.
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M.** (2022). ColabFold: making protein folding accessible to all. *Nat. Methods* **19**: 679–682.
- Mohnen, D.** (2008). Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* **11**: 266–277.
- Müller-Landau, H. and Varela, P.F.** (2021). Standard operation procedure for switchSENSE DRX systems. *Eur. Biophys. J.* **50**: 389–400.
- Murashige, T. and Skoog, F.** (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol. Plant.* **15**: 473–497.
- Mutuku, J.M., Cui, S., Yoshida, S., and Shirasu, K.** (2021). Orobanchaceae parasite–host interactions. *New Phytol.* **230**: 46–59.
- Ogawa, M., Kay, P., Wilson, S., and Swain, S.M.** (2009). Arabidopsis Dehiscence Zone Polygalacturonase1 (ADPG1), ADPG2, and Quartet2 are polygalacturonases required for cell separation during reproductive development in Arabidopsis. *Plant Cell* **21**: 216–233.
- Pagès, S., Heijne, W.H.M., Kester, H.C.M., Visser, J., and Benen, J.A.E.** (2000). Subsite mapping of *Aspergillus niger* endopolygalacturonase II by site-directed mutagenesis. *J. Biol. Chem.* **275**: 29348–29353.
- Park, K.-C., Kwon, S.-J., Kim, P.-H., Bureau, T., and Kim, N.-S.** (2008). Gene structure dynamics and divergence of the polygalacturonase gene family of plants and fungus. *Genome* **51**: 30–40.
- Park, K.C., Kwon, S.J., and Kim, N.S.** (2010). Intron loss mediated structural

dynamics and functional differentiation of the polygalacturonase gene family in land plants. *Genes and Genomics* **32**: 570–577.

**Parrinello, M. and Rahman, A.** (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**: 7182–7190.

**Petersen, T.N., Kauppinen, S., and Larsen, S.** (1997). The crystal structure of rhamnogalacturonase a from *Aspergillus aculeatus*: A right-handed parallel  $\beta$  helix. *Structure* **5**: 533–544.

**Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E.** (2004). UCSF Chimera--A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**: 1605–1612.

**Rhee, S.Y., Osborne, E., Poindexter, P.D., and Somerville, C.R.** (2003). Microspore Separation in the quartet 3 Mutants of *Arabidopsis* Is Impaired by a Defect in a Developmentally Regulated Polygalacturonase Required for Pollen Mother Cell Wall Degradation. *Plant Physiol.* **133**: 1170–1180.

**Rowlinson, J.S.** (2005). The Maxwell-boltzmann distribution. *Mol. Phys.* **103**: 2821–2828.

**Rui, Y., Xiao, C., Yi, H., Kandemir, B., Wang, J.Z., Puri, V.M., and Anderson, C.T.** (2017). POLYGALACTURONASE INVOLVED IN EXPANSION3 functions in seedling development, rosette growth, and stomatal dynamics in *Arabidopsis thaliana*. *Plant Cell* **29**: 2413–2432.

**Safran, J., Habrylo, O., Cherkaoui, M., Lecomte, S., Voxeur, A., Pilard, S., Bassard, S., Pau-Roblot, C., Mercadante, D., Pelloux, J., and Sénéchal, F.** (2021). New insights into the specificity and processivity of two novel pectinases from *Verticillium dahliae*. *Int. J. Biol. Macromol.* **176**: 165–176.

**van Santen, Y., Benen, J.A.E., Schroter, K.H., Kalk, K.H., Armand, S., Visser, J., and Dijkstra, B.W.** (1999). 1.68-angstrom crystal structure of endopolygalacturonase II from *Aspergillus niger* and identification of active site residues by site-directed mutagenesis. *J. Biol. Chem.* **274**: 30474–30480.

**Sénéchal, F., Habrylo, O., Hocq, L., Domon, J.M., Marcelo, P., Lefebvre, V., Pelloux, J., and Mercadante, D.** (2017). Structural and dynamical

characterization of the pH-dependence of the pectin methylesterase-pectin methylesterase inhibitor complex. *J. Biol. Chem.* **292**: 21538–21547.

**Shimizu, T., Nakatsu, T., Miyairi, K., Okuno, T., and Kato, H.** (2002). Active-site architecture of endopolygalacturonase I from *Stereum purpureum* revealed by crystal structures in native and ligand-bound forms at atomic resolution. *Biochemistry* **41**: 6651–6659.

**Søndergaard, C.R., Olsson, M.H.M., Rostkowski, M., and Jensen, J.H.** (2011). Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK<sub>a</sub> values. *J. Chem. Theory Comput.* **7**: 2284–2295.

**Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J.C.** (2005). GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **26**: 1701–1718.

**Swarup, K. et al.** (2008). The auxin influx carrier LAX3 promotes lateral root emergence. *Nat. Cell Biol.* **10**: 946–954.

**The PyMOL Molecular Graphics System. Version 2.0 Schrödinger, LLC.**

**Vitali, J., Schick, B., Kester, H.C.M., Visser, J., and Journak, F.** (1998). The Three-Dimensional Structure of *Aspergillus niger* Pectin Lyase B at 1.7-Å Resolution. *Plant Physiol.* **116**: 69–80.

**Voxeur, A. et al.** (2019). Oligogalacturonide production upon *Arabidopsis thaliana*-*Botrytis cinerea* interaction. *Proc. Natl. Acad. Sci. U. S. A.* **116**: 19743–19752.

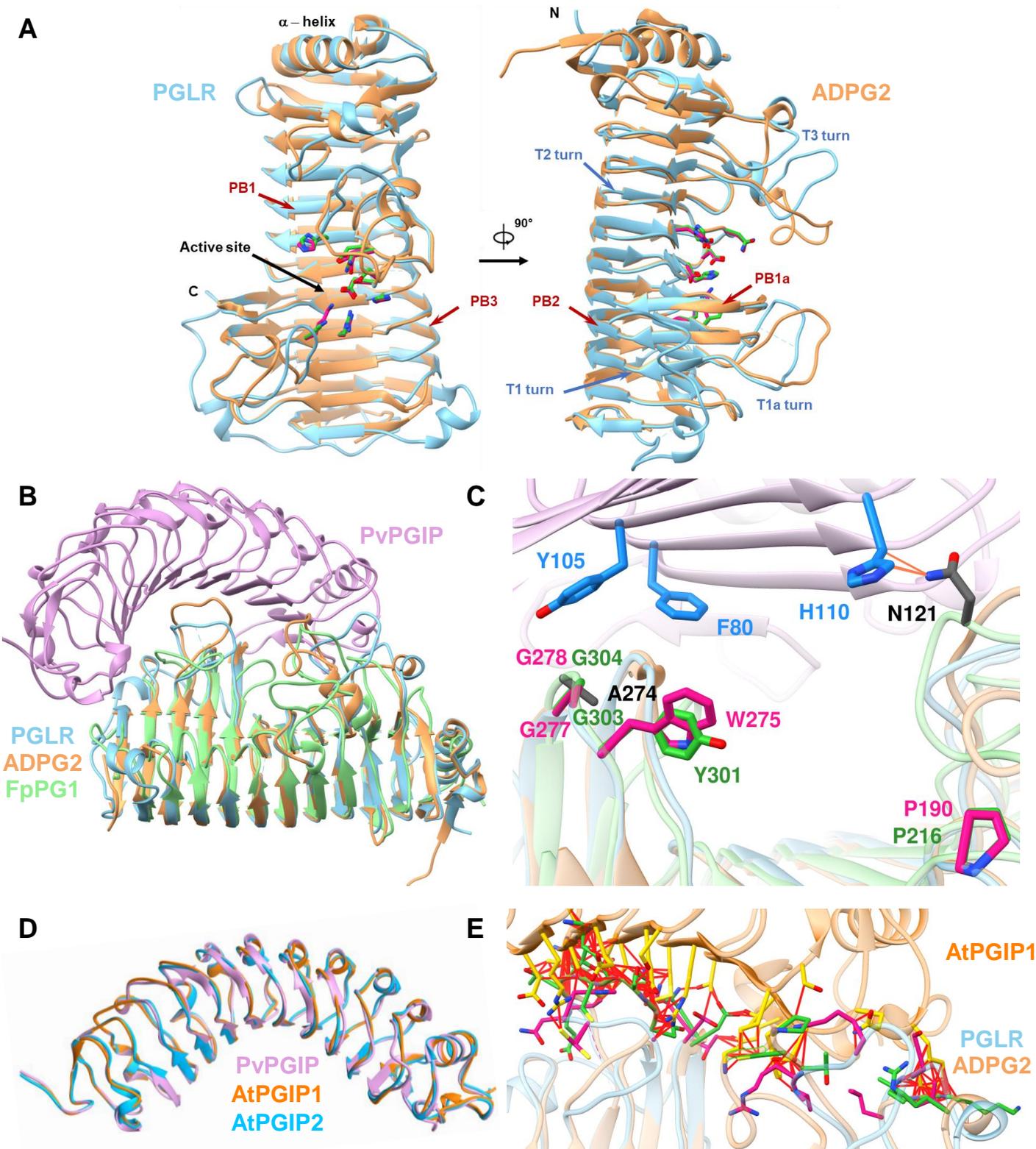
**Willats, W.G.T., McCartney, L., Mackie, W., and Knox, J.P.** (2001). Pectin: cell biology and prospects for functional analysis. *Plant Mol. Biol.* **47**: 9–27.

**Xiao, C., Barnes, W.J., Zamil, M.S., Yi, H., Puri, V.M., and Anderson, C.T.** (2017). Activation tagging of *Arabidopsis* POLYGALACTURONASE INVOLVED IN EXPANSION2 promotes hypocotyl elongation, leaf expansion, stem lignification, mechanical stiffening, and lodging. *Plant J.* **89**: 1159–1173.

**Xiao, C., Somerville, C., and Anderson, C.T.** (2014). POLYGALACTURONASE INVOLVED IN EXPANSION1 functions in cell elongation and flower development in *Arabidopsis*. *Plant Cell* **26**: 1018–1035.

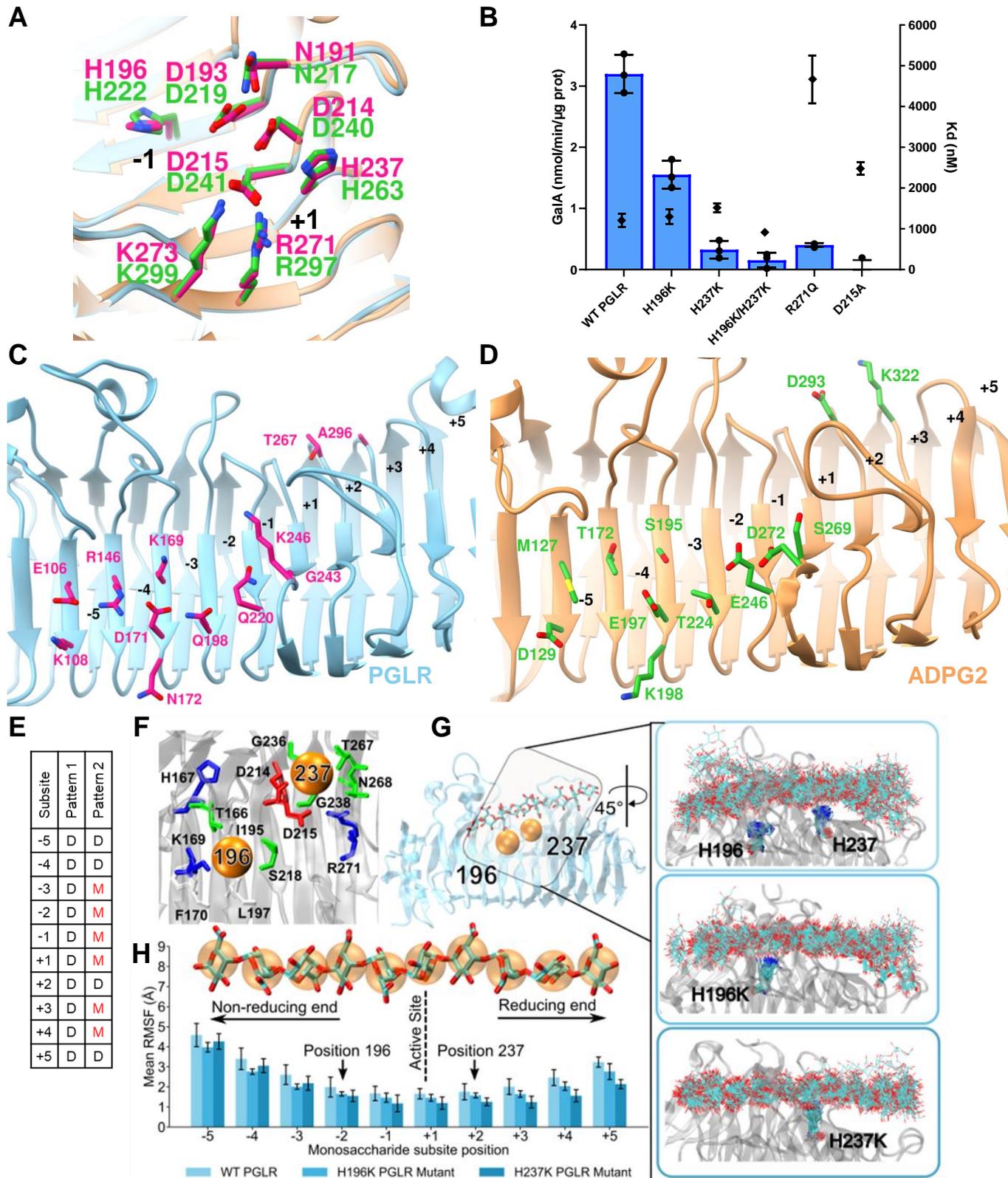
**Yoder, M.D. and Journak, F.** (1995). The Refined Three-Dimensional Structure of Pectate Lyase C Implications for an Enzymatic Mechanism. *Plant Physiol* **107**: 349–364.

**Zhang, Y.** (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**: 1–8.



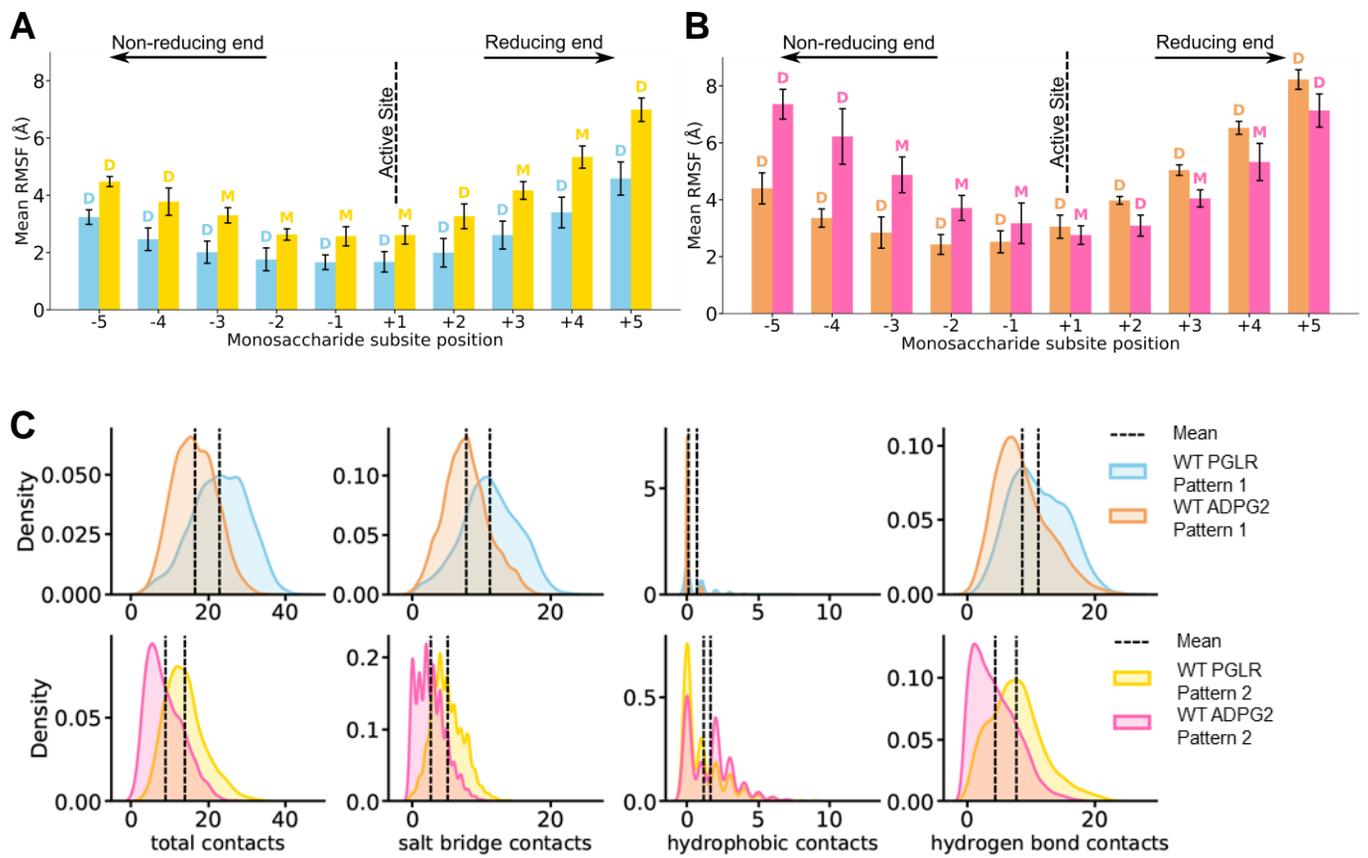
**Figure 1. Structure comparison of PGLR and ADPG2 and identification of novel amino-acids required for activity**

A) Overall structure of PGLR and ADPG2 represented in ribbon diagrams which are colored in blue and brown, respectively. PGLR and ADPG2 active site amino acids are pink and green colored.  $\beta$ -sheets and turns are indicated by red and blue arrows. B) Ribbon representation of *Phaseolus vulgaris* PGIP2 (PvPGIP2, plum), PGLR (blue), ADPG2 (brown), *Fusarium phyllophilum* PG (FpPG1, green). C) Detailed representation of aa involved in PvPGIP2-FpPG1 interaction (PvPGIP2 amino-acids in blue and FpPG1 amino-acids in grey), with orange lines representing van der Waals contacts. Key aa (N121, A274) mediating the interaction in FpPG1 are absent in PGLR and ADPG2. AA that can hinder the PG-PGIP interaction are represented in pink (PGLR) and green (ADPG2). D) Superimposition of crystallised PvPGIP2 with models of Arabidopsis PGIP1 (AtPGIP1, orange) and PGIP2 (AtPGIP2, blue). E) Interactions of AtPGIP1 with PGLR and ADPG2. Amino acids of AtPGIP1 (yellow), PGLR (pink) and ADPG2 (green) included in clashes closer than 0.6 Å are shown. The red lines represent atoms overlap of minimum 0.6 Å.



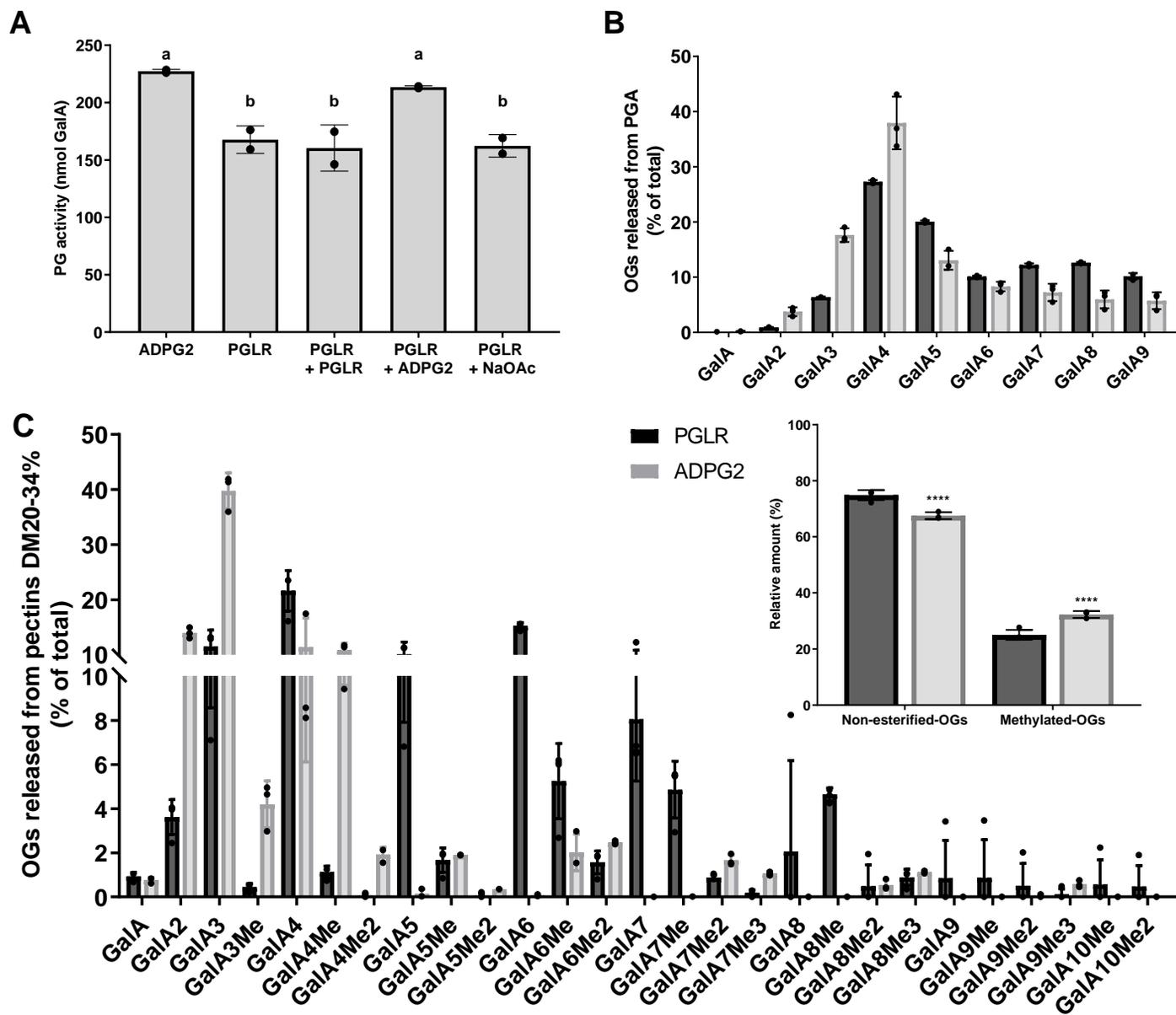
**Figure 2: Structure of the PGLR-ADPG2 active site and binding groove. Role of H196 and 237 for PGLR activity**

A) Active site of PGLR/ADPG2 highlighting absolutely conserved AA. D193/D219, D214/D240 and D215/D241 are AA involved in substrate hydrolysis. Black numbers indicate the subsites. B) Total PG activity of WT and mutated forms of PGLR (H196K, H237K, H196K/H237K, R271Q, D217A) on PGA (blue bars), and MST analysis of the interaction between WT and mutated forms of PGLR using a substrate of DP12 and DM5 (black rhomboids). Values correspond to means  $\pm$  SD of three replicates. C) Structure of PGLR binding groove (subsite -5 to +5). D) Structure of ADPG2 binding groove (subsites -5 to +5). E) Sequence of the fully de-methylesterified (pattern 1) or 60% methylesterified (pattern 2) deca-saccharides simulated in complex with ADPG2 and PGLR. D: de-methylesterified GalA, M: methylesterified GalA. F) cross-section of the substrate binding groove highlighting the positions of H196 and H237, which are represented as orange spheres. Positively and negatively charged residues are shown in blue and red, respectively, while polar residues are shown in green and represented as sticks. G) PGLR in complex with a deca-saccharide substrate, with insets showing the conformational ensembles of the substrate in complex with WT PGLR, H196K and H237K, by reporting conformations obtained every 10 ns. H) Root mean square fluctuations (RMSF) of de-methylesterified deca-saccharide across the binding groove for WT PGLR and PGLR mutants. Figures values correspond to means  $\pm$  SD of three replicates.



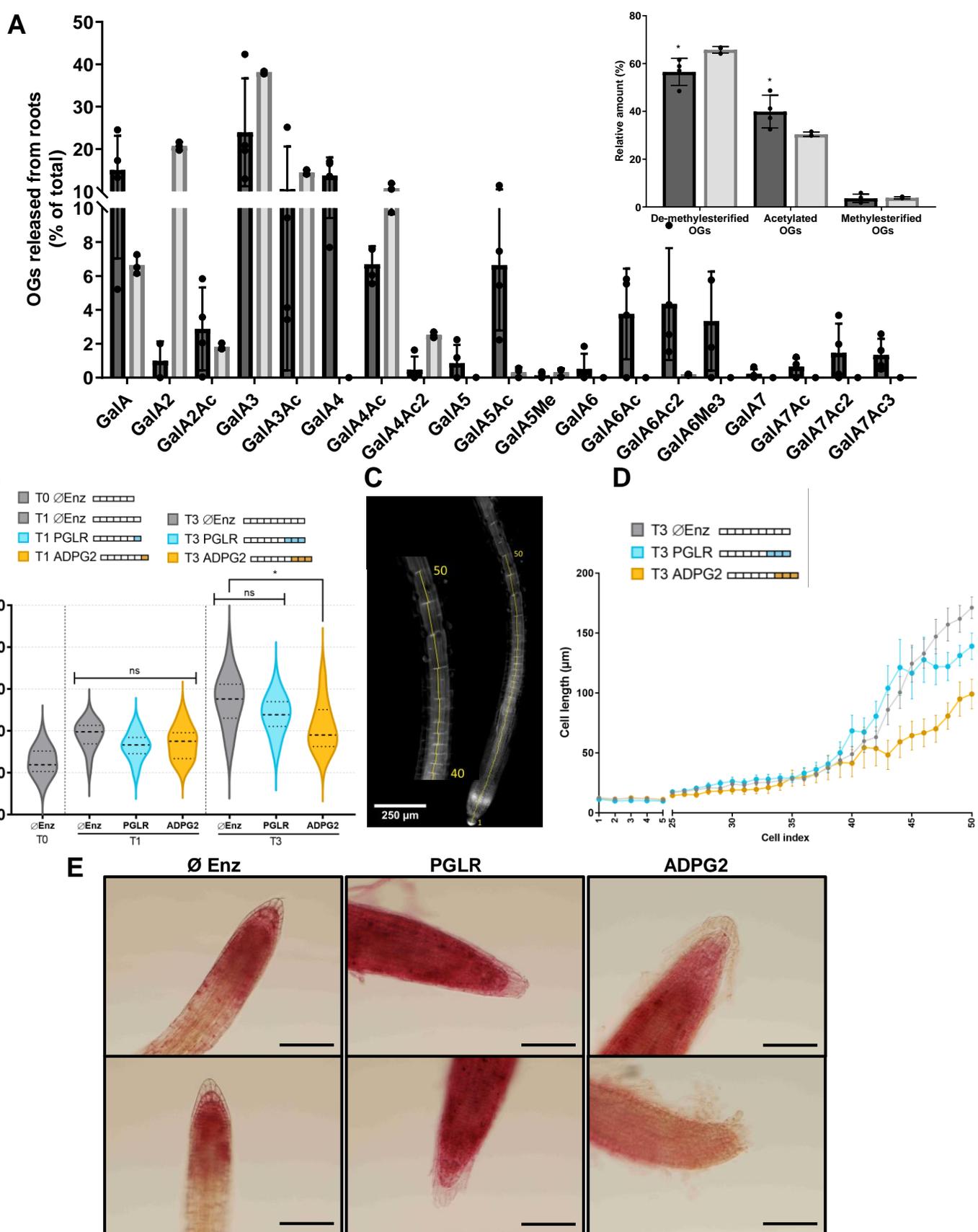
**Figure 3. PGLR and ADPG2 show distinct substrate dynamics.**

A-B) Root mean square fluctuations (RMSF) of each monosaccharide bound across the binding groove of PGLR (A) or ADPG2 (B). In each panel, fully de-methylesterified (pattern 1 – cyan in A and orange in B) or 60% methylesterified deca-saccharides (pattern 2 – yellow in A and pink in B) are shown. C) Analysis of the contacts between PGLR or ADPG2 and substrates either fully de-methylesterified (pattern 1) or characterized by 60% methylesterification (pattern 2).

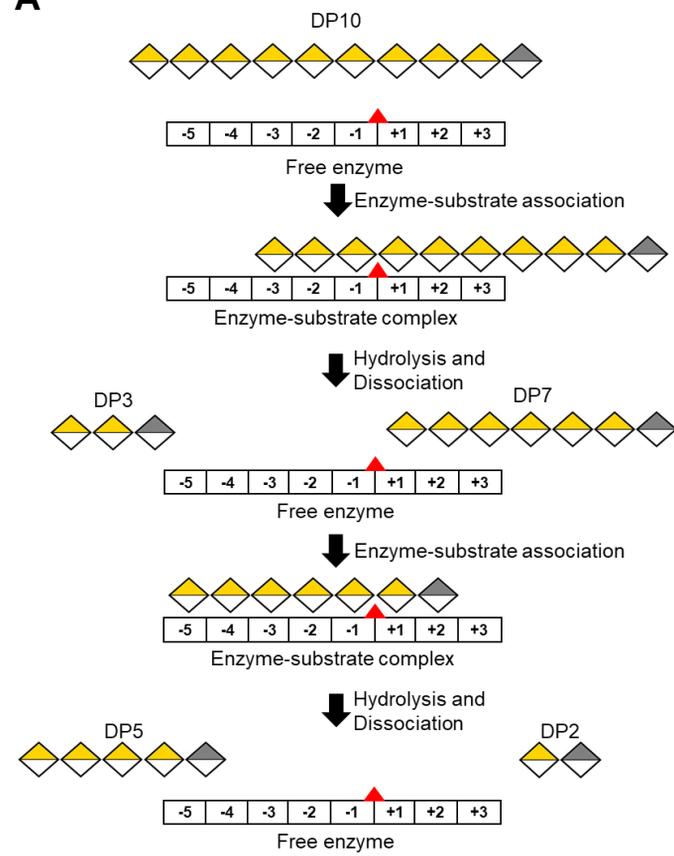
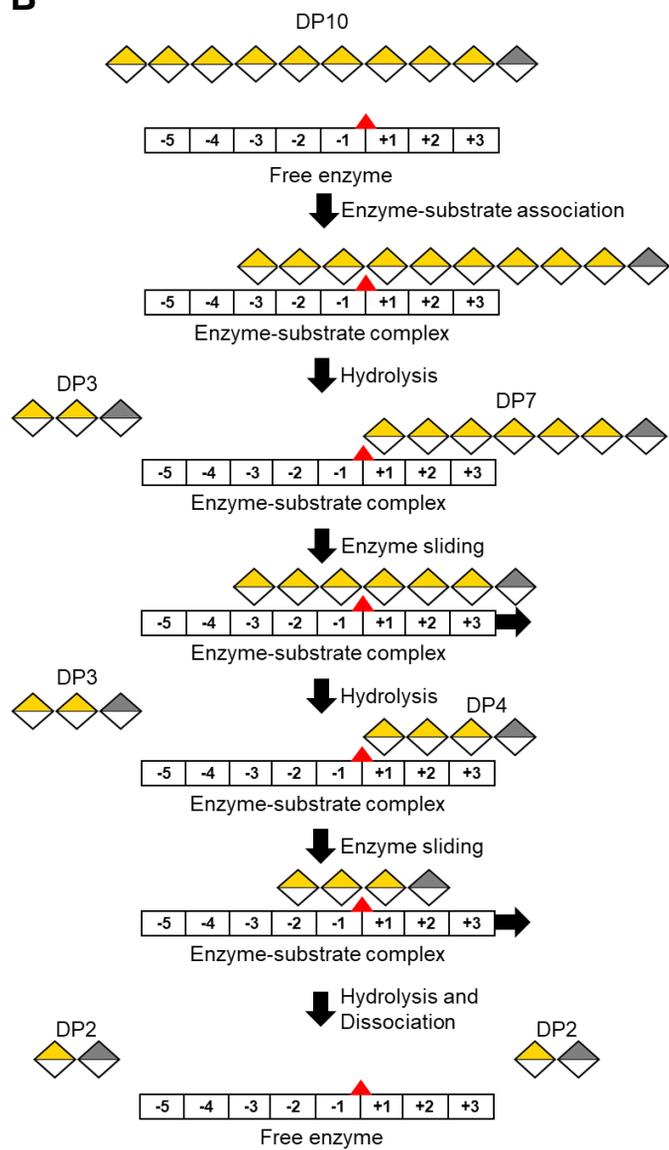


**Figure 4. PGLR and ADPG2 release distinct OGs.**

A) Activity tests performed on PGA (DM 0) after 1 hour digestion by ADPG2, PGLR and by adding PGLR or ADPG2 for 1 hour after a first digestion by PGLR. NaOAc (sodium acetate): negative control. B) Oligoprofiling of OGs released after 1 hour digestion of PGA by PGLR (black) or ADPG2 (grey) at 40°C, pH 5.2. C) Oligoprofiling of OGs after overnight digestion of pectins DM 20-34% by PGLR (black) or ADPG2 (grey) at 40°C, pH 5.2. Inset: Cumulative OGs released by PGLR and ADPG2 after over-night digestion on pectins DM 20-34% at 40°C, pH 5.2. In all figures values correspond to means  $\pm$  SD of three replicates. a, b, \*\*\*\* indicates statistically significant difference,  $P < 0.001$ .



**Figure 5. PGLR and ADPG2 are active on roots pectins and have distinct effects on root length and root cap.**  
 A) Oligoprofiling of OGs after digestion of roots cell wall by PGLR (black) and ADPG2 (grey) at 40°C, pH 5.2 after overnight digestion, (Inset: Cumulative OGs released by PGLR (black) and ADPG2 (grey) after over-night digestion of roots cell walls at 40°C, pH 5.2). \* indicates statistically significant difference,  $P < 0.05$ . B) Effects of the exogenous application of PGLR and ADPG2 on total root length of Arabidopsis seedlings. PGLR and ADPG2 were applied at iso-activities for one or three days on 6-day-old seedlings grown in liquid media. The value marked with \* indicates statistically significant difference between controls and ADPG2 analyzed by the one-way ANOVA with Tukey multiple comparison test  $P < 0.0001$ .  $n \geq 14$ , ns = non significant. C) Root cell numbering using EGFP-LTI6b reporter lines. D) Effects of 3-days exogenous application of PGLR and ADPG2 on the cell length of the firsts 50 root cells of 7-day-old seedlings.  $N > 50$ . E) Effects of 3-days exogenous application of PGLR and ADPG2 on root cap structure of 7-day-old seedlings (2 representative images per condition). Buffer ( $\emptyset$  Enz) was used as negative control. Scale bar represents 100  $\mu\text{m}$ .

**A****B**

**Figure 6. Model of PGLR and ADPG2 processivity**

A) PGLR shows low processive dynamics where enzyme-substrate association is followed by hydrolysis and dissociation of the substrate from the enzyme. This low processivity produces OGs of variable DPs. B) ADPG2 sliding motion after forming enzyme-substrate complex allows multiple substrate hydrolysis while staying attached to the substrate showing highly processive dynamics. Processive enzymes can produce small DP OGs. Galacturonic acid are yellow colored. Galacturonic acid reducing end is grey-colored. PG subsites are indicated by numbers. Red triangle represents the hydrolysis site.