Computational and Structural Biotechnology Journal 20 (2022) 6023–6032

COMPUTATIONAL ANDSTRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Biochemical characterization and molecular insights into substrate recognition of pectin methylesterase from *Phytophthora infestans*

Rajender Kumar [a], Sanjiv Kumar [a,1], Vincent Bulone [a,b], Vaibhav Srivastava [a,*]

[a] Division of Glycoscience, Department of Chemistry, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, AlbaNova University Center, 106 91 Stockholm, Sweden
[b] College of Medicine and Public Health, Flinders University, Bedford Park Campus, Sturt Road, South Australia 5042, Australia

ARTICLE INFO

Article history:
Received 26 September 2022
Received in revised form 1 November 2022
Accepted 1 November 2022
Available online 4 November 2022

Keywords:
Pectin methylesterases
Phytophthora infestans
Oomycete
Potato late blight
Molecular simulation

ABSTRACT

Pectin methylesterases (PMEs) are a class of carbohydrate-active enzymes that act on the O6-methyl ester groups of the homogalacturonan component of pectins, resulting in de-esterification of the substrate polymers and formation of pectate and methanol. PMEs occur in higher plants and microorganisms, including fungi, oomycetes, bacteria, and archaea. Microbial PMEs play a crucial role in pathogens' invasion of plant tissues. Here, we have determined the structural and functional properties of *Pi*-PME, a PME from the oomycete plant pathogen *Phytophthora infestans*. This enzyme exhibits maximum activity at alkaline pH (8.5) and is active over a wide temperature range (25–50 °C). *In silico* determination of the structure of *Pi*-PME reveals that the protein consists essentially of three parallel β-sheets interconnected by loops that adopt an overall β-helix organization. The loop regions in the vicinity of the active site are extended compared to plant and fungal PMEs, but they are shorter than the corresponding bacterial and insect regions. Molecular dynamic simulations revealed that *Pi*-PME interacts most strongly with partially de-methylated homogalacturonans, suggesting that it preferentially uses this type of substrates. The results are compared and discussed with other known PMEs from different organisms, highlighting the specific features of *Pi*-PME.

1. Introduction

The oomycete *Phytophthora infestans* is responsible for the late blight disease in potato and tomato plants [1,2]. Despite numerous efforts made in the past to control this pathogen, the disease remains a threat to global food security. Microbial pathogens of plants typically secrete cell-wall degrading enzymes that facilitate penetration of the host tissues, e.g. pectin- and cellulose-active enzymes [3]. The breakdown of esterified forms of pectins involves the action of pectin methylesterases (PMEs), a class of enzymes grouped in carbohydrate esterase family 8 in the carbohydrate-active enzymes database (CAZy, https://www.cazy.org/). In addition to their occurrence in insects and microorganisms such as pathogenic fungi, yeast and bacteria, PMEs are ubiquitously present in higher plants [4]. Interestingly, plant PMEs have been shown to play essential roles in plant–microbe interactions and stress responses [5,6]. Del Corpo et al. (2020) showed that a putative *Arabidopsis thaliana* PME, AtPME17, was highly induced in response to the necrotrophic fungus *Botrytis cinerea* [7].

PMEs catalyze the specific hydrolysis of the methyl ester bond at C-6 of galacturonosyl residues in pectins to release the corresponding carboxylic form of the polysaccharide (pectic acid) and methanol [8]. Pectins are a class of heteropolysaccharides abundant in plant primary cell walls where they contribute to cell and tissue integrity and rigidity [9]. The pectins comprise three main polymers, namely homogalacturonan (HG), rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II). In addition, the less abundant xylogalacturonans (XGA) and apiogalacturonans (AGA) are also considered a type of pectins [4]. HG consists of α-1,4-linked D-galacturonic acid residues (GalA), some of which are esterified with methanol and/or acetyl groups [10]. The degree and pattern of methyl esterification in HG have gained much attention in pectin research because these features contribute to the functional properties of pectins, both in plant cell walls and in pectin-based food products. PMEs affect the fine structure of pectins by removing methyl esters from the HG backbone, which in turn facilitates pectin degradation by polygalacturonases [11].

* Corresponding author.
E-mail address: vasri@kth.se (V. Srivastava).
[1] Current address: Faculty of Medical Sciences, Örebro University, 703 62 Örebro, Sweden.

https://doi.org/10.1016/j.csbj.2022.11.003

PME sequences arise from a multigene family, and several PME isoforms are identified in cell walls [12–15]. The isoforms catalyze the same reaction, but they differ in their isoelectric point (pI), molar mass, thermostability, substrate specificity, action pattern and catalytic properties [8,11,16]. PMEs are typically active as monomers, and their molecular mass varies from 25 to 54 kDa. In some instances, PMEs from eukaryotic origin have been shown to be N- or O- glycosylate. This is the case for instance of isoforms from kiwi fruit, orange, jelly fig achenes and Aspergillus species [16]. The isoelectric point of PMEs varies from as low as 3.1 for some fungal PMEs to up to 11 for some tomato PMEs [16,17]. Most of the purified plant PMEs have neutral or alkaline pIs, explaining their affinity with the slightly acidic cell wall. Only a few studies have revealed the occurrence of acidic plant PMEs, for example, in flax, mung bean, jelly fig, aspen, and chicory root [18]. Compared to plant and most bacterial PMEs (e.g., *Erwinia chrysanthemi*), fungal PMEs generally have more acidic pIs [16]. To date, three-dimensional structures of several bacterial (*E. chrysanthemi, Dickeya chrysanthemi,* and *Yersinia enterocolitica*), plant (carrot and tomato), and fungal (*A. niger*) PMEs have been resolved [16,19–21]. Based on these structures, it was found that PMEs are highly similar in their overall folding topology, characterized by a right-handed parallel β-helix structure consisting of three parallel β-sheets with interconnecting loops protruding from the helix core [16,19].

Despite their assumed role in pathogenesis, there are no reports available that describe the biochemical and structural properties of PMEs from oomycete species. In this report, we first explore the sequence and structural diversity of the PME family and present the key conserved motifs of this class of enzymes. We then describe the biochemical and structural properties of *Pi*-PME, a PME from the oomycete *P. infestans*, and compare its specific properties with that of well-characterized PMEs from other organisms.

2. Methods

2.1. Evolutionary sequence conservation and phylogenetic analysis of PMEs

Ten putative PME sequences from *P. infestans* (strain T30-4) are reported in the NCBI database. They present >58 % sequence identity and, out of these, four have a truncated N-terminal end or are shorter in length. In addition, the C-terminal region is found to be more conserved than the N-terminal part of the proteins. Based on preliminary sequence and phylogenetic analysis, one of the representative full-length PMEs, XP 002907462.1 (PITG_02545), was selected for further studies (Fig. S1). This sequence (referred to as *Pi*-PME) was used as a query input in the ConSurf Program [22,23] to calculate sequence conservation and find structurally and functionally important residues across the entire PME family. The Context-Specific Iterated-Basic Local Alignment Search Tool (CSI-BLAST) was employed for homology searches against the UNIREF-90 protein database using an E-value cut-off of 0.00001, a maximum of 95 % identity and a minimal identity of 35 % between sequences. The other parameters were kept at default values for sequence conservation analysis. For phylogenetic analysis of the PME family, homologous protein sequences were obtained from the Uniprot database, including Oomycetes, Fungi, Plant, Bacteria and Archaea domains with an E-value threshold of 0.0001. Sequences that showed >30 % sequence identity and >50 % sequence query coverage were considered for further analysis. Sequence redundancy was removed using the CD-Hit clustering program at 60 % sequence identity [24]. All sequences were manually checked to exclude uncharacterized protein fragments and very short sequences from the analysis. The curated sequences were used for phylogenetic analysis and multiple sequence alignment using the inbuilt MUSCLE program in MEGA11 [25]. This analysis allowed the identification of the key evolutionary conserved residues and highly conserved sequence motifs across the PME family. The evolutionary phylogenetic analysis was inferred using the Maximum Likelihood method and JTT matrix-based model in MEGA11 [25]. Initial tree(s) for the heuristic search were generated automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model and then selecting the topology with a superior log-likelihood value. The final phylogenetic tree was visualized using iTOl (http://itol.embl.de).

2.2. Gene synthesis, protein expression and purification

A truncated version of the *Pi*-PME sequence not containing the signal peptide of the full-length protein (18 amino acid N-terminal segment) was codon optimized for expression in *Escherichia coli*. Synthesis of the corresponding DNA sequence was outsourced (GeneArt, Invitrogen). The construct was cloned between the *Bam*HI and *Xho*I restriction sites of the pET28a vector (Invitrogen) with an N-terminal hexa-histidine tag and the plasmid was transformed into *E. coli* Rosetta 2(DE3) (Novagen) cells for protein expression. An overnight 30 ml culture was inoculated into one litre of LB broth containing ampicillin (100 μg/mL) and chloramphenicol (34 μg/ml), and incubated at 37 °C until an OD_{600} of 0.4–0.6 was reached. The culture was cooled at 4 °C for one hour and protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG; final concentration 0.3 mM). After 18 h induction at 18 °C, the cells were harvested by centrifugation (15 min at 6000 × *g* and 4 °C) and re-suspended in lysis buffer (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 10 mM imidazole). Lysozyme was added to a final concentration of 1 mg/ml and the cell suspension was incubated on ice for 30 min. The cells were subsequently disrupted by sonication (three pulses of 30 s each interspersed by 30 s rest on ice) and the soluble fraction containing the recombinant protein was recovered by centrifugation at 17,000 × *g* for 30 min at 4 °C. The supernatant was added to a Ni-NTA agarose matrix pre-equilibrated with lysis buffer, and the mixture was kept at 4 °C under constant agitation for one h. The matrix was then washed with 5 volumes of lysis buffer, followed by 10 volumes of lysis buffer containing 20 mM imidazole. The recombinant *Pi*-PME protein bound to the matrix was consecutively eluted three times (Elution 1–3; Fig. S2a) with 1 ml of lysis buffer containing 500 mM imidazole and further concentrated using Amicon Ultra centrifugal filters (10 kDa, Merck Millipore, Cork, Ireland). Protein purity and molecular weight were estimated by SDS-PAGE. The concentration of the purified protein was determined using the Bradford assay [26]. The band corresponding to the over-expressed protein was excised from an SDS-PAGE gel and subjected to in-gel trypsin digestion and LC-MS/MS analysis to confirm the identity of the recombinant *Pi*-PME [27]. The protein was concentrated and dialyzed in 1× PBS, pH 7.4 with 10 % glycerol until further use.

2.3. Assay of PME activity

PME activity was assayed in 96-well ELISA plates, essentially as described previously, except for the following minor modifications [28]. Various parameters were first tested to determine the optimal assay conditions. These included the screen of a range of concentrations of purified *Pi*-PME (1, 10, 100, and 1000 nM), different pH (3.0, 4.0, 5.0, 6.0, 6.5, 7.0, 7.5, 8.0, and 8.5) and temperature (25, 30, 37, 40, 45, 50, 55, 60, 70, and 80 °C) conditions. The standard assay mixture designed from the assay optimization work consisted of 89.4 μl of 0.4 mM NAD in 50 mM phosphate buffer

pH 7.5, 8.0 µl of 5 % (w/v) citrus pectin (Sigma P 9135) in water, 0.035 U formaldehyde dehydrogenase (F1879-25UN, Sigma), and 1.0 U alcohol oxidase (A2404-250UN, Sigma). Reactions were initiated by adding the purified *Pi*-PME to a final concentration of 1.9 µg/ml in a 100 µl total reaction volume and the mixtures were incubated at 25 °C. For enzyme kinetics, reactions were performed using different concentrations of pectin (final concentrations of 0.003125, 0.00625, 0.0125, 0.025, 0.05, 0.1, 0.2, or 0.4 %) and NAD (6.25, 12.5, 25, 50, 100, 200, 400 or 800 µM). All reactions were continuously monitored at 340 nm over one h, using a CLARIOstar® UV/Vis spectrophotometer. All assays were accompanied by a negative control performed in the absence of *Pi*-PME.

2.4. Structural modelling and molecular dynamic simulations of the Pi-PME enzyme free and as a complex with pectin substrates

The 3D structure of *Pi*-PME was modeled using the SWISS-MODEL homology-modelling pipeline and the structure of the *A. niger* PME as a template. The latter protein was selected as it exhibits the highest sequence similarity with *Pi*-PME. The modelled *Pi*-PME structure was further minimized using the structure editing module of the USCF Chimera software [29], and used for comparative structural analysis with other reported PME structures. *Pi*-PME was also modelled as a complex with three different pentamers of α-(1–4)-linked D-galacturonosyl residues, namely: (i) a fully de-esterified pentasaccharide; (ii) a partially methyl-esterified pentasaccharide at the C-6 position with alternating methyl groups; and (iii) a fully methyl-esterified pentasaccharide (Fig. 1). The 3D structures of these oligosaccharides were built using the Carb-Builder online server [30] and methyl-esterification modifications were introduced using the xleap and tleap Amber modules. The reducing ends of the saccharides were protected by a β-linked methyl group to avoid any potential impact of unfavorable H-bonding with the free hemiacetal group at C1. Each minimized and optimized structure was used for initial enzyme-substrate complex formation. The 3D coordinates of each HG substrate were docked in the substrate-binding site of *Pi*-PME using the previously reported 3D structure of *D. dadantii* PME-substrate complex (PDB ID: 2NTP) [20]. The structure of each enzyme-substrate complex was minimized using the minimization module of the Chimera software.

The unbound *Pi*-PME and the three enzyme-substrate complexes were simulated using the AMBER 18.0 software package by applying forcefield parameters ff14SB for the protein and GLYCAM_06j-1 for the carbohydrate substrates [31–33]. The leap module of the Amber tool was used for applying all required parameters, including default protonation states of ionizable amino acids at pH 7. Each complex was neutralized by adding Na^+/Cl^- counterions and solvated using TIP3P water model molecules in a truncated octahedron box with a margin distance of 10 Å. All systems were independently equilibrated as described previously [34]. Each complex was computed for up to 200 ns, and trajectory files were recorded every 50 ps and analyzed using the visualizer software Chimera and VMD [35].

3. Results and discussion

3.1. Evolutionary sequence conservation and phylogenetic analyses of PMEs

The ConSurf program was used to investigate the degree of conservation of substrate-binding residues as well as structural and other functionally important residues among the PME homologues. The analysis revealed that the N-terminal segment corresponding to the first 50 amino acids, including the signal peptide (amino acids 1–18), is variable, whereas the C-terminal region of PMEs encompassing the catalytic and substrate-binding residues is evolutionary conserved. The conserved residues N123, K157, Q174, Y178, F199, A222, and R256 are involved in substrate binding, whereas residues Q152, D175, D196, R253, and W255 participate in the catalytic process (Fig. 2). The nonpolar G and A residues located at various positions are also highly conserved and structurally important (Fig. 2). In addition, two conserved sequence motifs, one pentapeptide [L-G-R-P-W] and a tetrapeptide [G-X-X-D], are also found in PMEs (green boxes, Fig. 2). Several conserved residues present in the *Pi*-PME enzyme are also present in the PME of *D. dadantii* [20], a bacterium that causes necrotic blight and soft rot diseases in plants.

Phylogenetic analysis of PME sequences was performed by first identifying 1600 proteins from oomycete, fungal, plant, bacterial, and archaea species that exhibit significant sequence similarity (>30 % sequence identity and >50 % sequence query coverage) with *Pi*-PME. After clustering based on >60 % sequence identity, 99 representative sequences were selected, of which 32 are from fungi, 28 from bacteria, 20 from plants, 12 from oomycetes and one from an algal species. These sequences vary greatly in length, from 283 (oomycetes, e.g. *Phytophthora sojae*) to 811 residues (fungi, e.g. *Talaromyces cellulolyticus*), and their N-terminal regions showed higher sequence diversity.

Phylogenetic analysis revealed that the selected representative PME sequences are divided into three main clusters (Fig. 3). The larger cluster A comprises 46 sequences from fungal and oomycete species, one from a plant and one from an algal species. Cluster A can be further divided into two sub-clusters. Sub-cluster-AI consists essentially of fungal PMEs, while sub-cluster-AII is dominated by oomycete enzymes, including *Pi*-PME (Fig. 3). Most fungal PME sequences showed higher sequence similarity with each other. All sequences from cluster A also contain the highly conserved pen-

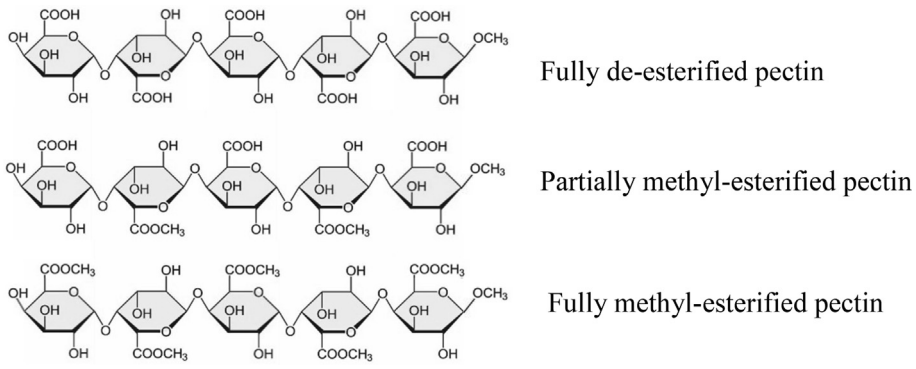

Fig. 1. 2D representation of the three different pentamers of α-(1–4)-linked D-galacturonosyl residues used for modelling and simulation studies.

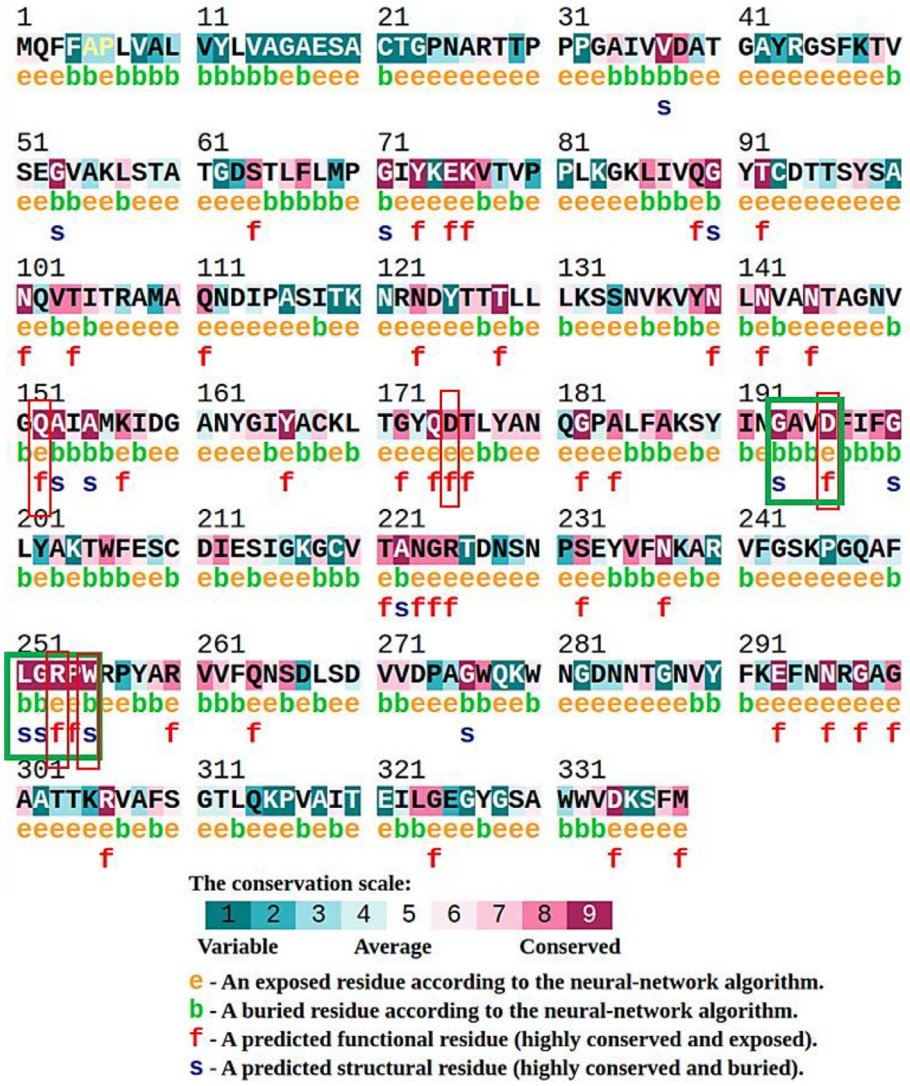

Fig. 2. Evolutionary sequence conservation analysis of the PME family. The red rectangular boxes highlight the conserved key active site residues and the green boxes indicate conserved sequence motifs.

tapeptide motif [L-G-R-P-W] which is located in the active site region. Cluster B comprises 34 PME sequences from Bacteria and Archaea.

Similarly to cluster A, cluster B sequences also contain the conserved pentapeptide, albeit with an additional arginine residue at the C-terminal end of the motif [L-G-R-P-W-R] for the bacterial sequences, and a tyrosine residue at the N-terminal end of the archaea sequences [Y-L-G-R-P-W]. In addition, the archaea sequences are clustered in two distinct subgroups within cluster B.

Cluster C consists of 16 plant PMEs, however, these sequences are less conserved than their counterpart from other taxa. In some plant PMEs, the conserved pentapeptide contains an alanine residue instead of a proline, and the fifth residue (tryptophan) is sometimes replaced by a tyrosine [L-G-R-P/A-W/Y]. The three outlier plant PMEs in the phylogenetic tree (Fig. 3) group closely with the oomycete enzymes from cluster A. In addition, a PME from the plant *Vitis riparia* groups within the common fungal clade in cluster A, which was also reported by Markovic and Janecek in 2004 [8]. In this report [8], it was shown that plant PMEs are more diverse than their fungal and bacterial counterparts. Nonetheless, all conserved residues are located close to the catalytic site and may play an essential structural and functional role in the PME family.

Moreover, a systemic genome sequencing study confirmed that PME genes encode pre-pro-proteins that comprise a signal peptide or signature pattern for this enzyme family [12–14]. The genes can be divided into two groups based on their pro region, with the first group containing only two to three introns and an extended pro region, and the second group comprising five to six introns and a short or nonexistent pro region [12,13]. Our data suggest that PMEs from oomycetes, bacterial and fungal phytopathogenic organisms are closer to the latter group.

3.2. Enzyme activity, optimal temperature and pH

As expected, the purified recombinant *Pi*-PME protein was characterized by an apparent molecular weight of around 38 kDa on SDS-PAGE gels. LC-MS/MS analysis further confirmed the identity of the purified protein, with a sequence coverage of 90 % (Fig. S2). Various concentrations of the recombinant enzyme (1 to 1000 nM) were tested to optimize the enzymatic reaction conditions and demonstrate that the level of activity measured is concentration dependent. For further studies, 100 nM of *Pi*-PME in the presence of 4 mg/ml pectin substrate was considered optimum as no significant increase of activity was observed with higher concentrations of the protein in the conditions tested (data not

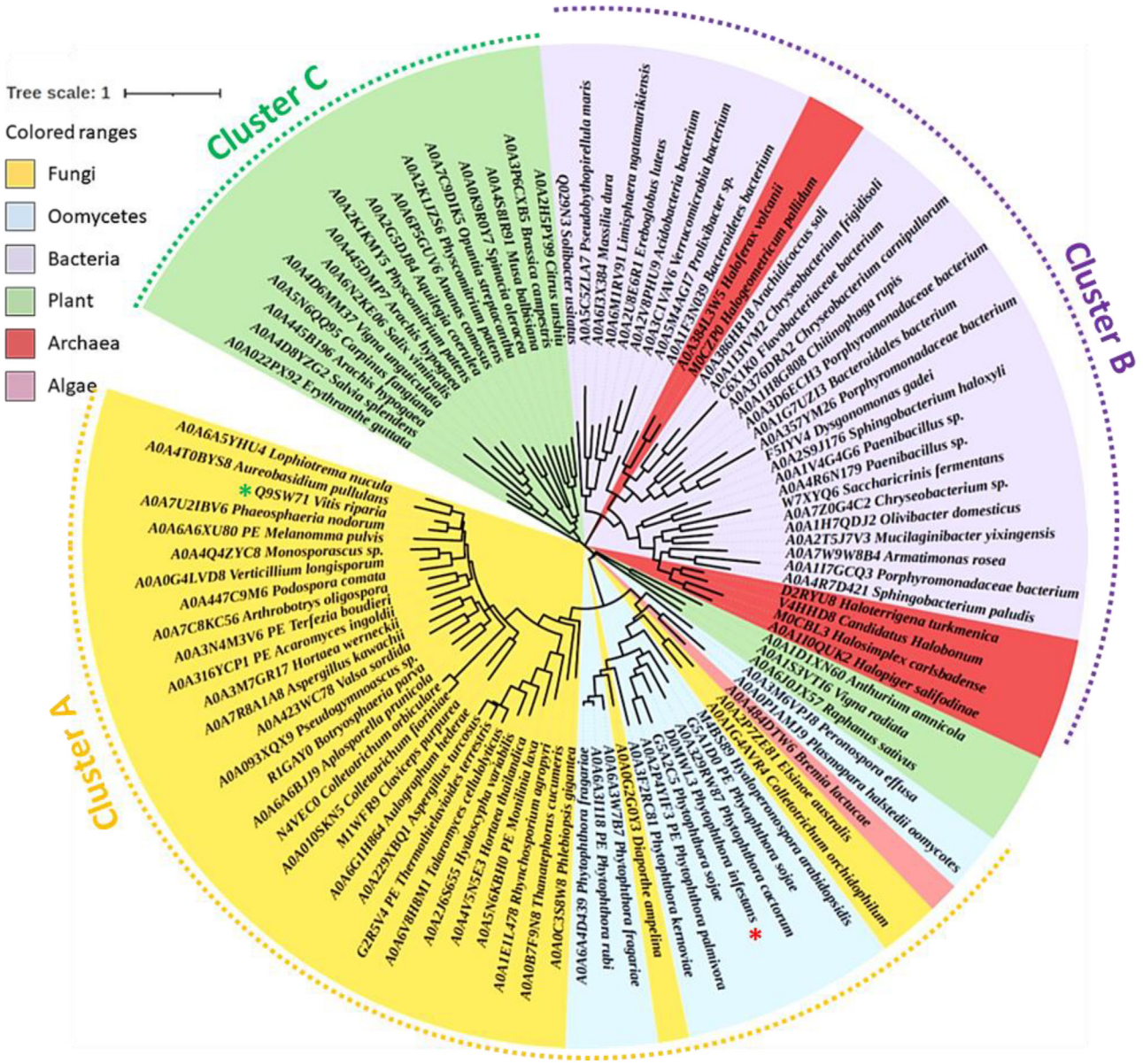

Fig. 3. Phylogenetic analysis of 99 representative PME sequences from different taxa. Note: red * represents *Pi*-PME; green * represents PME from the plant *Vitis riparia.*

shown). Enzyme activity exceeded 250 mol of methanol released per mol of enzyme across a temperature range of 25–45 °C, with the maximum activity observed at 45 °C and a rapid decrease of activity above 50 °C (Fig. 4a). These data are comparable to previously published results on plant PMEs, which shows that this class of enzymes are typically most active at temperatures higher than 40 °C. For example, optimal temperatures of 55 °C have been reported for PMEs from black carrot and tomato [36,37], and 65 °C for a PME from plum [38]. In fewer cases, thermophilic PMEs have been reported, such as in the case of an enzyme from guava which exhibits an optimal temperature of 75–80 °C [39]. *Pi*-PME is moderately alkaline, with an optimal pH of 8.5 and much lower activity below pH 6 (Fig. 4b). The optimal pH of other PMEs has been shown to vary to a large extent depending on the enzyme source, i.e. from below 7 to up to 9 [37,38,40–42]. Enzyme kinetics performed in the presence of a pectin substrate from citrus revealed that *Pi*-PME is characterized by a K_m of 0.007 % and a V_{max} of 10 μmol methanol released/min at a concentration of 50 nM *Pi*-PME.

3.3. Structural modelling of Pi-PME

The *Pi*-PME sequence presents 24–35 % identity with sequences of PME proteins for which 3D structures are available, including enzymes from *A. niger*, *Daucus carota*, *Solanum lycopersicum*, *Y. enterocolitica* and *D. chrysanthemi* [16,19–21]. Modelling of the *Pi*-PME 3D structure was performed using the structure of the *A. niger* enzyme (PDB ID: 5C1C) as a template as this showed the highest percentage of sequence identity (35 %) with *Pi*-PME. The quality of the all-atom modelled structure was evaluated and validated using the MolProbity server [43]. The refined modelled structure was superimposed to the structure of the *A. niger* PME. The root mean square deviation (RMSD) of the protein backbone atoms and heavy atoms were of 1.80 Å and 2.66 Å, respectively. Structural alignment and RMSD analysis showed that the overall structural folding topology of the modelled *Pi*-PME structure is similar to that of *A. niger* and other bacterial and plant PME structures. *Pi*-PME consists mostly of a β-helix arrangement made of three parallel β-sheets interconnected by short loops.

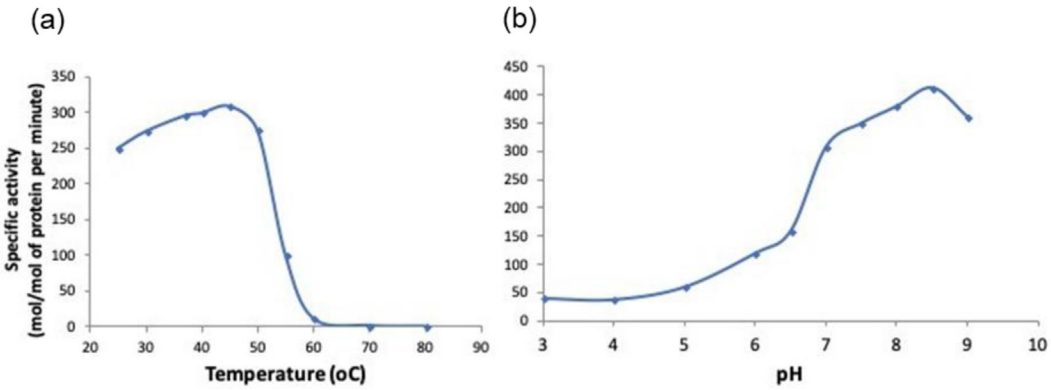

Fig. 4. Determination of the optimal temperature (a) and optimal pH (b), of activity of *Pi*-PME.

The N-terminal region of the protein consists of a β-strand followed by a short α-helix connected to the core β-helix structure. Other studies have reported β-helix structures characterized by secondary elements of parallel β sheets coiled into a large right-handed cylinder [14,16,19]. The core of the *Pi*-PME structure is structurally close to that of these previously reported PME structures. These enzymes also share structural topology with other pectin degrading enzymes, such as polygalacturonases (PG) and pectin lyases (PL). This structural similarity suggests a common ancestral fold for the different classes of pectin degrading enzymes that have evolved to acquire different catalytic properties.

In contrast, the C-terminal end of *Pi*-PME consists of an extended tail and three short and distorted α-helices. The *Pi*-PME substrate-binding site adopts a long shallow-shaped structure surrounded by two extended loops (referred to as loop I and II) that are located on each side of the substrate-binding cleft (Fig. 5a). Loop I has a longer extension than loop II in *Pi*-PME (Fig. 5a). Both loops are separated by an average distance of approximately 23 ± 5 Å and the variance in this distance is mainly due to the flexibility of loop I. The central part of the substrate-binding site is enriched in aromatic amino acids (Tyr, Phe and Trp) which are conserved among members of the PME family.

The 3D structures of PMEs from plants, fungi and oomycetes are entirely superimposable, as illustrated in Fig. 5a, with an example of a PME from each class of organism. The core β-helix structure is highly identical and rigid, except for both loop regions mentioned above (Fig. 5a). Similarly, compared to *Pi*-PME, the core β-helix folding topologies are also similar in bacterial and insect PMEs (Fig. 5b). However, in bacterial PMEs, the region that corresponds to loop II in *Pi*-PME contains additional β-strands and is more extended, while loop I and its surrounding regions are longer than in *Pi*-PME (Fig. 5b). In the case of the insect PME shown in Fig. 5b, the region that corresponds to loop I in *Pi*-PME has a longer extension (around 30 amino acids in length) that forms two α-helices, while loop II and its proximal region are also extended. This structural topology makes the substrate-binding site narrow and deeper in insect PME (Fig. 5b). In summary, we observe that PMEs are characterized by rather low similarity at the sequence level while presenting a similar structural topology.

Compared to *Pi*-PME, the RMSD values between corresponding atoms of fungal, plant, and bacterial PMEs fell within the range 1.8–3.0 Å, whereas a large variation of 7 Å was observed for the insect PME. Loop I showed large conformational changes to accommodate proper fitting of the substrate and facilitate catalysis and product release after demethylation (see further discussion below).

The proportion of positively and negatively charged amino acids is significantly different between PMEs from different organisms [16,20]. Structurally most of the charged residues are lining on the protein surface. The highly conserved and positively charged Arg residue in the conserved pentapeptide motif is located in the substrate-binding cleft of *Pi*-PME and other structurally characterized PMEs [16,20]. A total of 32 positively charged residues (Arg and Lys) and 23 negatively charged residues (Asp and Glu) are present in *Pi*-PME. As for the plant and bacterial enzymes, *Pi*-PME has more basic and less acidic amino acids than fungal PMEs. *Pi*-PME does not contain His residues, whereas up to six His residues are present in PMEs from plant, bacterial, and fungal species [16]. In the substrate-binding cleft of *Pi*-PME, the side chains of six positively charged Arg and Lys residues favor electrostatic binding to the negatively charged groups of pectin. Similar electrostatic interactions have also been described in bacterial, plant and fungal species [16]; however, the number of positively charged amino acids (Arg and Lys) varies. In the case of the PME from the bacterium *E. chrysanthemi*, eight positively charged residues are present, whereas the PMEs from the plant *D. carota* and fungus *A. niger* have three positive residues in their substrate-binding site [16,19,20].

3.4. MD simulation of enzyme-substrate complexes

To better understand the binding mode of homo-oligogalacturonan substrates in the active site of *Pi*-PME, molecular simulation dynamics were performed. The data show that the core part of the *Pi*-PME β-helix structure was rigid, whereas some regions experienced conformational changes during the MD simulation. Region (D113-N121) of loop I showed a dynamic behavior, with most conformations being close to the catalytic center. These conformations take place between the two extreme conformations shown in the average structure presented in Fig. 6. In addition, some conformational changes were also observed in the N- and C-terminal regions of the protein, whereas limited changes occurred in loop II.

The average structure of *Pi*-PME obtained from the MD simulation was used for modelling enzyme-substrate complexes. For this purpose, three homogalacturonan penta-oligosaccharides with different degrees of methylesterification were used. The demethylated and fully methylated substrates did not interact strongly with the substrate-binding site as both were released promptly during the simulation (Fig. 7).

It was also observed that the fully demethylated substrate is released from the binding site through the loop I region. Loop I residues K120 and R122 appear to play a key role in the exclusion mechanism of this substrate. Loop I showed significant dynamic behavior and the MD simulation revealed that substrate binding involves electrostatic interactions between the positively charged functional groups of the K120 and R122 residues and hydroxyl groups of the substrates. In the initial stage of the simulation (0–

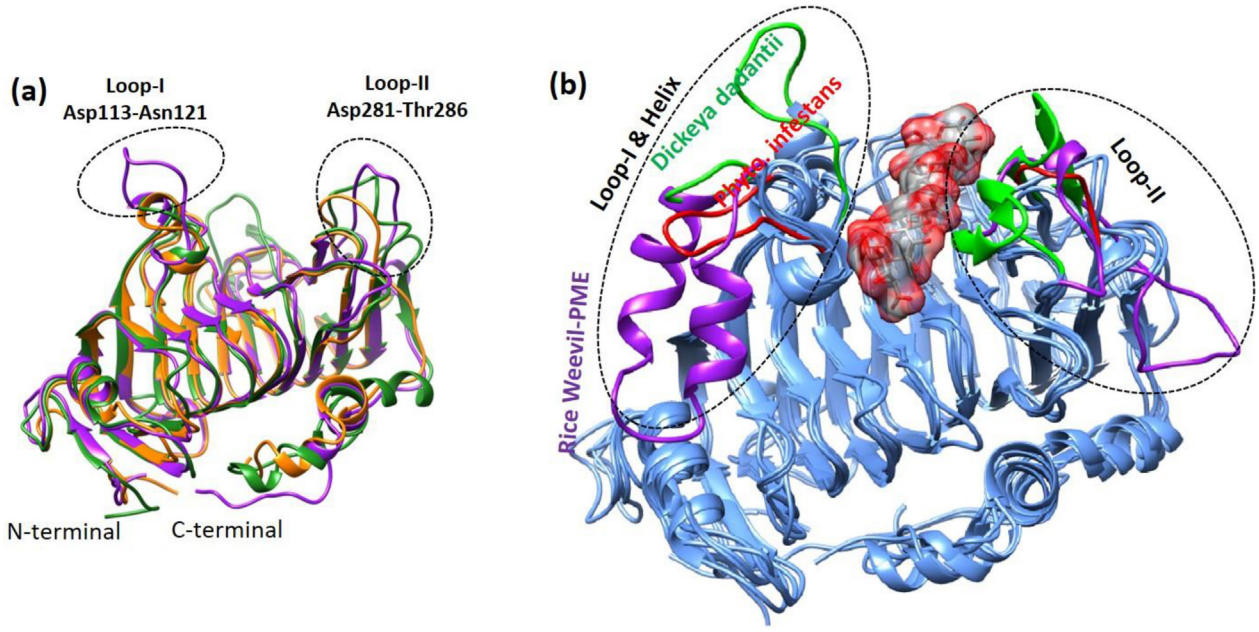

Fig. 5. (a) Model of the 3D structure of *Pi*-PME (violet) superimposed on the template structure of *A. niger* (orange) and *D. carota* (green), (b) Structural superposition of other PMEs structure; loops region variations are highlighted in voilet (Rice Weevil PME); green (PME from *D. dadantii*); and red (*Pi*-PME).

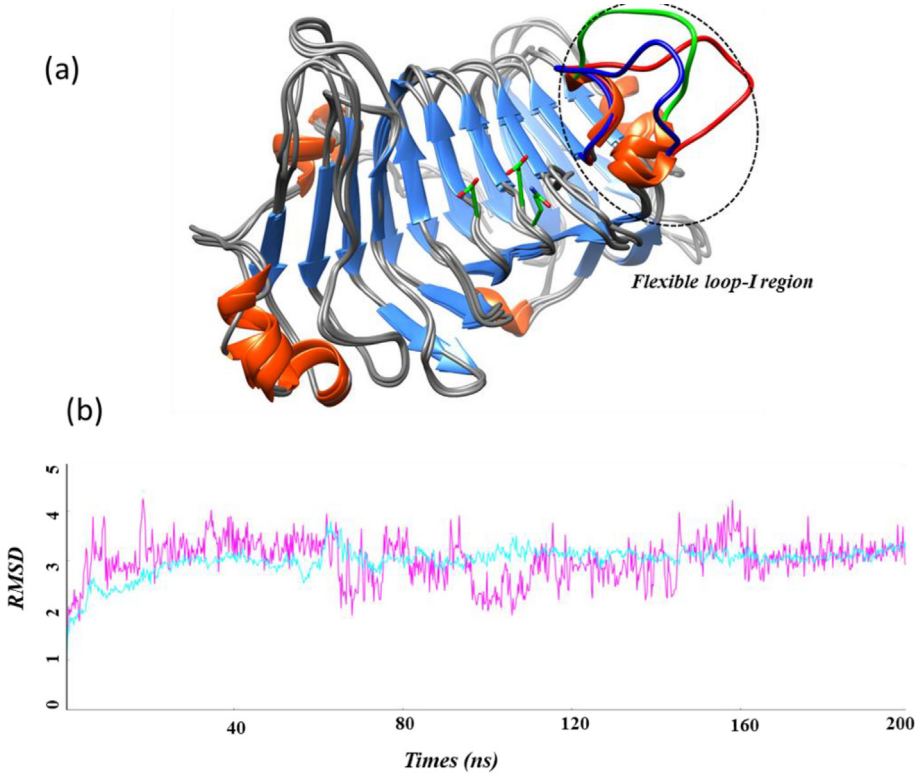

Fig. 6. An average structure of *Pi*-PME obtained after MD simulation showing two extreme conformations of the flexible loop I region (red and blue), (b) RMSD vs time evolution of the entire *Pi*-PME protein (light blue) and loop I region (magenta).

1 ns) following binding of the substrate in the enzyme active site, the loop I region bent toward the catalytic site, with R122 interacting with the hydroxyl group of pectin sugar ring. As the simulation steps progressed, the substrate moved toward loop I to interact with K120. At about 10 ns, a large-scale movement is observed in the loop I region and between 10 and 90 ns a further large conformational change led to the exclusion of the substrate (Fig. 7). The release of the demethylated substrate from the binding site of *Pi*-PME is a progressive process driven by the loop I region and its dynamic behavior. In contrast, the fully methylesterified substrate was readily excluded from the binding site right at the beginning of the simulation run (Fig. 7). This fast-expelling movement was accompanied by weak interactions with the loop II region.

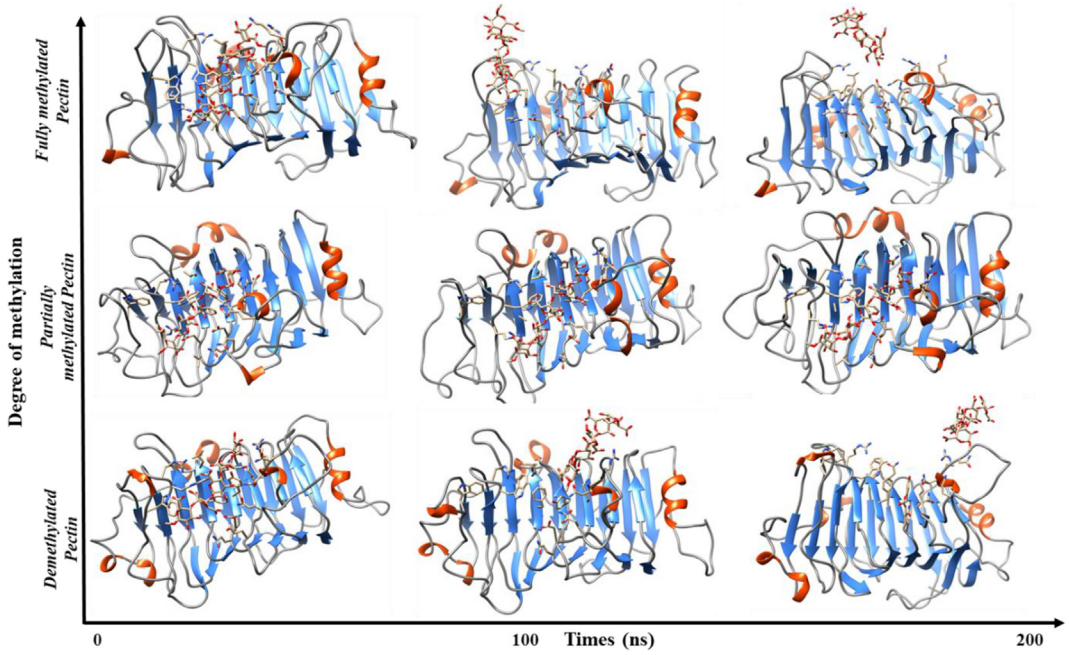

Fig. 7. MD simulation of complexes of *Pi*-PME and three different homo-oligogalacturonan substrates.

The partially methylated homo-oligogalacturonan substrate consistently bound into the binding site of *Pi*-PME during the entire MD simulation calculations (Fig. 8a). Residues R256, Y178, K157, R122 and N123, interacted with the substrate largely through hydrogen bonding (Fig. 8b). In some conformational stages, D196, W255, D124 and K120 also showed hydrogen bonding with the substrate, but with lower occupancies. We observed that the conformations of the partially methylated substrate obtained from different time intervals during the entire simulation were superimposable, with an RSMD value of > 0.5 Å. In this Michaelis complex, the leaving methyl group is positioned near the conserved residues Q152, D175 and D196, which are involved in the formation of the intermediate product and further release of the demethylated product. Residues R253 and Y255 are conserved amongst all PMEs and, although not directly involved in the catalytic process, are predicted to assist substrate holding and protonation of D175. These data show that *Pi*-PME uses the same reaction mechanism as proposed earlier for other PMEs [19,20].

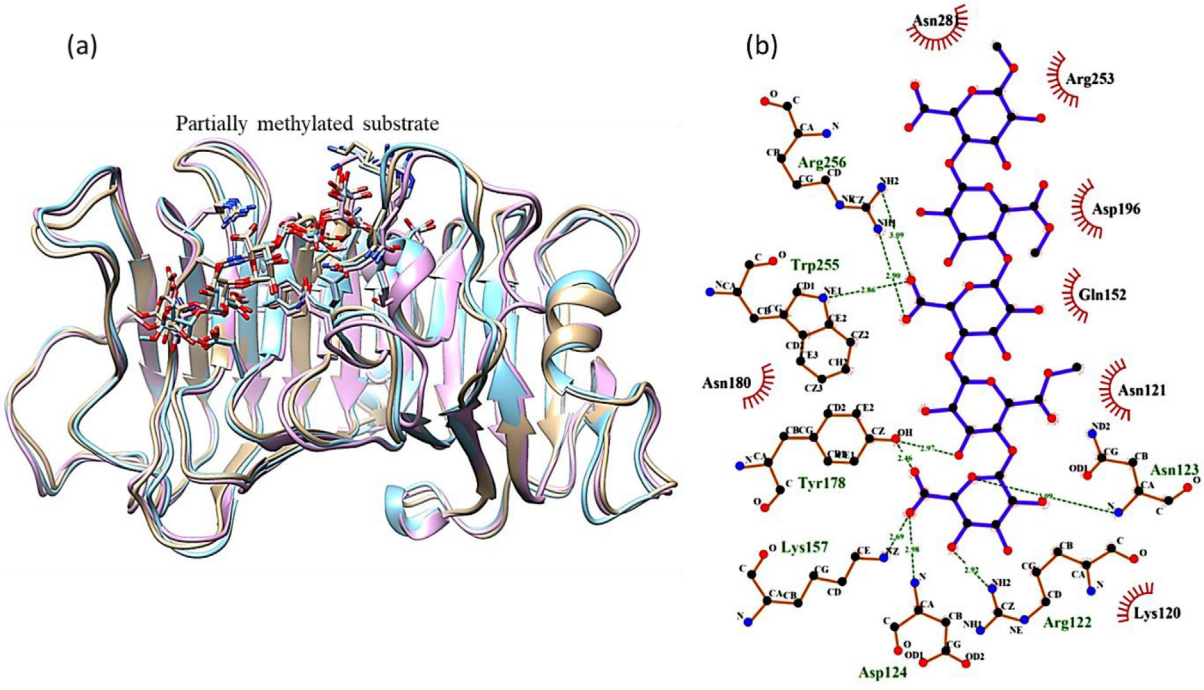

Fig. 8. (a) Conformations of the substrate obtained from different time intervals (0–200 ns) during the entire simulation, (b) 2D representation of the substrate-binding mode in the active site of *Pi*-PME.

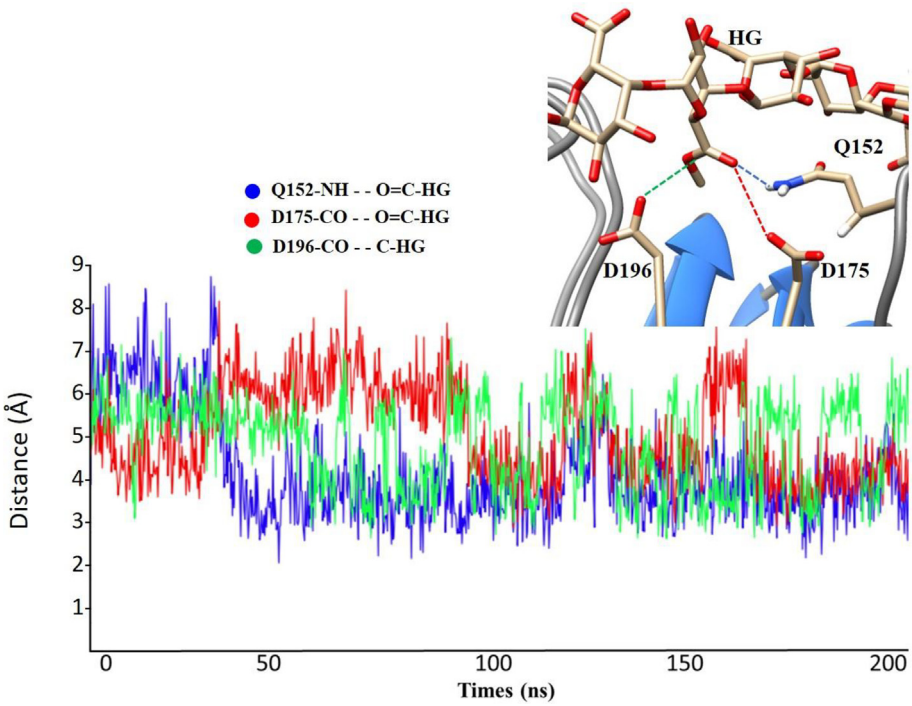

Fig. 9. Key active site residues involved in the catalytic mechanism and their distance from the methylester group of the substrate during MD simulation.

The MD simulation result of the enzyme-substrate complex provided detailed molecular-level information on the substrate binding mode and interaction with the catalytic residues.

The simulation trajectory of the enzyme-substrate complex indicates that the active site residue D196 leads the nucleophilic attack on the carbonyl carbon of the C6 ester. In few conformations, the side-chain carboxylic group of D196 comes closer to the substrate C6 ester, with a typical distance close to 3 Å (Fig. 9). The deprotonated state of D196 can be sustained by the hydrogen bonding interaction between R253 and D196. Another active site residue, D175, is situated on the same plane as D196 but acts as a general acid-base catalyst in the PME reaction mechanism [20]. In addition, the conserved Q152 residue forms a hydrogen bond with the carbonyl group of the ester (Fig. 9). Johansson et al. (2002) suggested that a pair of conserved catalytic residues form an oxyanion hole for the negatively charged acyl intermediate that provides stability to the complex [19]. In the case of *Pi*-PME, these residues are Q152 and D175, and we speculate the same role for this pair. It is worth mentioning that the dynamic behaviour of the loop I region and the electrostatic potential of the *Pi*-PME binding groove could play a key role in the enzyme's processivity, resulting in sequential de-methylesterification of the sugar units [44].

4. Conclusion

We have explored the sequence diversity of PMEs, identified conserved sequence motifs, and conducted a first-time *in silico* structural and *in vitro* biochemical characterization of an oomycete PME. The *Pi*-PME enzyme was active over a broad range of pH and temperature, with maximum activity at an alkaline pH of 8.5 and temperature of 45 °C. Furthermore, the 3D structure of *Pi*-PME, its substrate binding specificity and dynamic behavior were determined and compared with that of similar enzymes from other taxonomic groups. This fundamental knowledge can potentially be exploited to better understand the role of oomycete pectin-degrading enzymes during plant infection.

CRediT authorship contribution statement

Rajender Kumar: Data curation, Visualization, Software, Investigation, Validation, Formal analysis, Methodology, Writing - original draft. **Sanjiv Kumar:** Methodology, Formal analysis, Investigation, Visualization. **Vincent Bulone:** Conceptualization, Supervision, Funding acquisition, Writing – review & editing. **Vaibhav Srivastava:** Conceptualization, Supervision, Funding acquisition, Resources, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by a grant to V.B and V.S. from the European Commission (Horizon 2020 FET-OPEN Grant Agreement No. 828940). The authors also acknowledge the Swedish Infrastructure Committee (SNIC) resources utilized for molecular simulation studies under the project SNIC 2021/5-471.

References

[1] Akino S, Takemoto D, Hosaka K. *Phytophthora infestans*: a review of past and current studies on potato late blight. J Gen Plant Pathol 2014;80:24–37.
[2] Nowicki M, Foolad MR, Nowakowska M, Kozik EU. Potato and tomato late blight caused by *Phytophthora infestans*: an overview of pathology and resistance breeding. Plant Dis 2012;96:4–17.
[3] Kotoujansky A. Molecular genetics of pathogenesis by soft-rot Erwinias. Ann Rev Phytopathol 1987;25:405–30.
[4] Mohnen D. Pectin structure and biosynthesis. Curr Opin Plant Biol 2008;11:266–77.
[5] Coculo D, Lionetti V. The Plant Invertase/Pectin Methylesterase Inhibitor Superfamily. Front Plant Sci 2022;13:863892.

[6] Lionetti V, Cervone F, Bellincampi D. Methyl esterification of pectin plays a role during plant-pathogen interactions and affects plant resistance to diseases. J Plant Physiol 2012;169:1623–30.
[7] Del Corpo D, Fullone MR, Miele R, Lafond M, Pontiggia D, Grisel S, et al. AtPME17 is a functional *Arabidopsis thaliana* pectin methylesterase regulated by its PRO region that triggers PME activity in the resistance to *Botrytis cinerea*. Mol Plant Pathol 2020;21:1620–33.
[8] Markovic O, Janecek S. Pectin methylesterases: sequence-structural features and phylogenetic relationships. Carbohydr Res 2004;339:2281–95.
[9] Wu HC, Bulgakov VP, Jinn TL. Pectin Methylesterases: Cell Wall Remodeling Proteins Are Required for Plant Response to Heat Stress. Front Plant Sci 2018;9:1612.
[10] Voragen AG, Coenen G-J, Verhoef RP, Schols HA. Pectin, a versatile polysaccharide present in plant cell walls. Struct Chem 2009;20:263–75.
[11] Bonnin E & Pelloux J (2020) Pectin degrading enzymes. Pectin: technological and physiological properties: 37-60, ISBN : 978-3-030-53420-2.
[12] Richard L, Qin LX, Goldberg R. Clustered genes within the genome of *Arabidopsis thaliana* encoding pectin methylesterase-like enzymes. Gene 1996;170:207–11.
[13] Arabidopsis Genome I. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 2000;408:796–815.
[14] Micheli F. Pectin methylesterases: cell wall enzymes with important roles in plant physiology. Trends Plant Sci 2001;6:414–9.
[15] Wen B, Zhang F, Wu X, Li H. Characterization of the tomato (*Solanum lycopersicum*) pectin methylesterases: evolution, activity of isoforms and expression during fruit ripening. Front Plant Sci 2020;11:238.
[16] Kent LM, Loo TS, Melton LD, Mercadante D, Williams MA, Jameson GB. Structure and properties of a non-processive, salt-requiring, and acidophilic pectin methylesterase from *Aspergillus niger* provide insights into the key determinants of processivity control. J Biol Chem 2016;291: 1289–306.
[17] de Freitas ST, Handa AK, Wu Q, Park S, Mitcham EJ. Role of pectin methylesterases in cellular calcium distribution and blossom-end rot development in tomato fruit. Plant J 2012;71:824–35.
[18] Sénéchal F, Wattier C, Rustérucci C, Pelloux J. Homogalacturonan-modifying enzymes: structure, expression, and roles in plants. J Exp Bot 2014;65:5125–60.
[19] Johansson K, El-Ahmad M, Friemann R, Jornvall H, Markovic O, Eklund H. Crystal structure of plant pectin methylesterase. FEBS Lett 2002;514:243–9.
[20] Fries M, Ihrig J, Brocklehurst K, Shevchik VE, Pickersgill RW. Molecular basis of the activity of the phytopathogen pectin methylesterase. EMBO J 2007;26:3879–87.
[21] Boraston AB, Abbott D. Structure of a pectin methylesterase from *Yersinia enterocolitica*. Acta Crystallogra F: Struct Biol Commun 2012;68: 129–33.
[22] Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 2010;38:W529–33.
[23] Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 2016;44:W344–50.
[24] Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 2006;22:1658–9.
[25] Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis version 11. Mol Biol Evol 2021;38:3022–7.
[26] Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976;72:248–54.
[27] Srivastava V, Malm E, Sundqvist G, Bulone V. Quantitative proteomics reveals that plasma membrane microdomains from poplar cell suspension cultures are enriched in markers of signal transduction, molecular transport, and callose biosynthesis. Mol Cell Proteom 2013;12:3874–85.
[28] Grsic-Rausch S, Rausch T. A coupled spectrophotometric enzyme assay for the determination of pectin methylesterase activity and its inhibition by proteinaceous inhibitors. Anal Biochem 2004;333:14–8.
[29] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera–a visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–12.
[30] Kuttel MM, Stahle J, Widmalm G. CarbBuilder: Software for building molecular models of complex oligo- and polysaccharide structures. J Comput Chem 2016;37:2098–105.
[31] Case DA, Cheatham 3rd TE, Darden T, Gohlke H, Luo R, Merz Jr KM, et al. The Amber biomolecular simulation programs. J Comput Chem 2005;26:1668–88.
[32] Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, et al. GLYCAM06: a generalizable biomolecular force field. Carbohydrates J Comput Chem 2008;29:622–55.
[33] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput 2015;11:3696–713.
[34] Kumar R, Henrissat B, Coutinho PM. Intrinsic dynamic behavior of enzyme:substrate complexes govern the catalytic action of beta-galactosidases across clan GH-A. Sci Rep 2019;9:10346.
[35] Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14(33–8):27–8.
[36] Van den Broeck I, Ludikhuyze LR, Van Loey AM, Hendrickx ME. Effect of temperature and/or pressure on tomato pectinesterase activity. J Agri Food Chem 2000;48:551–8.
[37] Ünal MÜ, Bellur E. Extraction and characterisation of pectin methylesterase from black carrot (*Daucus carota L.*). Food Chem 2009;116:836–40.
[38] Nunes CS, SnM C, Saraiva JA, Coimbra MA, Hendrickx ME, Van Loey AM. Thermal and high-pressure stability of purified pectin methylesterase from plums (*Prunus domestica*). J Food Biochem 2006;30:138–54.
[39] Leite KMSC, Tadiotti AC, Baldochi D, Oliveira OMMF. Partial purification, heat stability and kinetic characterization of the pectinmethylesterase from Brazilian guava, *Paluma* cultivars. Food Chem 2006;94:565–72.
[40] Carvalho AB, De Assis SA, Cerqueira Leite KM, Bach EE, de Faria Oliveira OM. Pectin methylesterase activity and ascorbic acid content from guava fruit, cv. Predilecta, in different phases of development. Int J Food Sci Nutr 2009;60:255–65.
[41] Do Amaral SH, De Assis SA, De Faria Oliveira OMM. Partial purification and characterization of pectin methylesterase from orange (*Citrus sinensis*) cv. pera-rio. J Food Biochem 2005;29:367–80.
[42] Ciardiello MA, Tamburrini M, Tuppo L, Carratore V, Giovane A, Mattei B, et al. Pectin methylesterase from kiwi and kaki fruits: purification, characterization, and role of pH in the enzyme regulation and interaction with the kiwi proteinaceous inhibitor. J Agri Food Chem 2004;52:7700–3.
[43] Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, et al. MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci 2018;27:293–315.
[44] Mercadante D, Melton LD, Jameson GB, Williams MA. Processive pectin methylesterases: the role of electrostatic potential, breathing motions and bond cleavage in the rectification of Brownian motions. PLoS ONE 2014;9: e87581.