The crystal structure of a hyperthermoactive exopolygalacturonase from *Thermotoga maritima* reveals a unique tetramer

Tjaard Pijning a, Gertie van Pouderoyen a, Leon Kluskens b, John van der Oost b, Bauke W. Dijkstra a,*

a Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
b Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands

**Abstract**

The exopolygalacturonase from *Thermotoga maritima* is the most thermoactive and thermostable pectinase known to date. Here we present its crystal structure at 2.05 Å resolution. High structural homology around the active site allowed us to propose a model for substrate binding, explaining the exo-cleavage activity and specificity for non-methylated saturated galacturonate at the non-reducing end. Furthermore, the structure reveals unique features that contribute to the formation of stable tetramers in solution. Such an oligomerization has not been observed before for polygalacturonases.

© 2009 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

**1. Introduction**

The cell wall of higher plants consists of a complex matrix of polysaccharides such as pectin, cellulose and hemicellulose [1,2]. Pectin contains homogalacturonan and rhamnogalacturonan regions, which may be methylated, acetylated or glycosylated by various carbohydrates. Because of the complexity of pectin, its degradation requires an array of depolymerases and esterases [3,4]. Polygalacturonases (PGs) are depolymerases that cleave the backbone glycosidic linkages of pectin using a hydrolytic reaction type. They have been classified in glycosyl hydrolase family 28 (GH28) [5], a family that also includes rhamnogalacturonases [6], and are either endo-acting or exo-acting enzymes.

Structural information of pectinolytic enzymes comes from fungal and bacterial endo-PGs [7–13], one bacterial exo-PG [14] and rhamnogalacturonase [15]. They all display a right-handed parallel β-helix fold; their active site cleft is open on both sides in endo-PGs or occluded on one side in the exo-PG [14]. Experimental structures of substrate- or product-bound complexes [10,14,16,17], as well as modelling studies [7,10], have led to understanding of how PGs bind and process their substrates [3]. The galacturonate polymer substrate is bound in a mixed $2\frac{1}{3}$ helical conformation and interacts with a set of conserved residues. Its non-reducing end sugar unit is bound towards the N-terminal side of the β-helix; its carboxylate group is anchored by a conserved lysine and a conserved glycine-serine cis-peptide. As a result, this sugar unit is distorted to a half-chair conformation. Furthermore, three aspartates approach the substrate from the same side; one of these acts as a general acid, while the other two activate a conserved water molecule proposed to attack the anomeric carbon atom. After cleavage of the glycosidic bond, the product has an inversed anomeric configuration.

Pectin-degrading enzymes have received broad attention from the food industry, e.g. for the clearing of fruit juices [2,4,18–20]. Because of the elevated temperatures used during these processes, there is high interest in thermostable pectinases. Previously, Kluskens et al. [19] and also Parisot et al. [21] have identified two pectinases in the hyperthermophile *Thermotoga maritima* (*Tm*), an extracellular pectate lyase and an intracellular polygalacturonase. The lyase produces oligogalacturonates with a D$_4$,5-unsaturated unit at their non-reducing ends. The polygalacturonase is a 448-residue exo-acting enzyme releasing mono-galacturonate from non-reducing ends. It only accepts saturated and non-methylated substrates [19]; it does not accept products of the lyase, which are non-saturated at their non-reducing ends. In gel filtration
experiments an apparent molecular weight (MW) of 212 kDa has been observed for Tm ExoPG, indicating that the enzyme is tetrameric in solution. With a melting temperature of 105 °C and maximum activity at 80 °C, it is the most thermostable and thermoactive pectinase known to date [19].

Here we report the 2.05 Å crystal structure of Tm ExoPG, the first structure of a polygalacturonase from a hyperthermophilic bacterium. A model of a bound tetragalacturonate substrate, allows us to explain the substrate specificity (exo-activity and acceptance of non-methylated, saturated polygalacturonate only), as well as the product specificity (release of mono-galacturonate). The crystal structure indicates a unique tetrameric arrangement with a head-to-head intermolecular β-helix, in which hydrophobic patches are shielded from the solvent, and suggests that tetramerization, which is also observed in solution, may contribute to the thermostability of the enzyme.

2. Materials and methods

2.1. Crystallization, data collection and data processing

Cloning, expression and purification of Tm ExoPG have been described by Kluskens et al. [19]. Crystallization experiments were performed using the hanging drop vapour diffusion method at 293 K; drops were prepared by mixing 1.5 μl protein solution (12 mg/ml in 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)–NaOH, pH 7.5, 3 mM NaN3), which we still had available, but was not needed for phasing because of the successful molecular replacement solution. Results were processed and analyzed with Dynamics software. Intermolecular interactions in the crystal structure and the stability of possible protein oligomers were analyzed using the PISA web server [36].

2.2. Structure determination and refinement

Using the FFSAS03 server [24] and SCWRL [25], homology models for Tm ExoPG were generated. The three structures with highest identity were used as templates: Verrinia enterococilitica (Ye) ExoPG (β-helical domain [14]), Aspergillus niger EndoPG II [12] and A. niger EndoPG I [11], having sequence identities of 18%, 18% and 16%, respectively.

Molecular replacement was then performed with PHASER [26]. The highest scoring solution was subjected to a round of automatic building with ARP/wARP [27] and further refined with REFMAC [28], alternated with cycles of map inspection and manual rebuilding with COOT [29]. TLS parameters were included in the later stages of refinement [30]. The quality of the model was analyzed with Molproby [31]; secondary structure elements were assigned with DSSP [32]. The coordinates and structure factor amplitudes have been deposited with the PDB (accession code 3JUR).

2.3. Substrate modelling

The (empty) native structure of Tm ExoPG was taken as a starting point; to model a bound tetragalacturonate we proceeded as follows. For the –1 and +1 subsites, a superposition of the three catalytic aspartates was made with the galactofuranuronate/galactopyranuronate bound structure of Stereum purpureum (Sp) EndoPG I (PDB entry 1KCD [10]). The galactopyranuronate (GalpA) at the +1 subsite was left unchanged, while the galactofuranuronate (GalpA) bound at the –1 subsite was remodelled in the same way as described by Shimizu et al. [10]. The resulting GalpA–GalpA digalacturonate was extended at the +2 and +3 subsites to a tetragalacturonate by superposition of the experimentally observed tetragalacturonate bound to Erwinia chrysanthemi pectate lyase Pell (PDB entry 3BBY [16]), using the PRODRG server [33]. The glycosidic torsion angles between the +1/+2 and +2/+3 units were adjusted to values described by Braccini et al. [34] for the B-conformer of digalacturonate (ϕ = 99°, ψ = 161°). To model an unsaturated Δ4,5-galacturonate, which is identical to 4,5-dehydroglucuronate, the 4,5-dehydroglucuronate unit from PDB entry 1HMW [35] was superposed on the substrate –1 GalpA from our tetragalacturonate model, using the C1, C2, C5 and O5 atoms.

2.4. Oligomeric state

Dynamic light scattering (DLS) experiments were performed using a DynaPro MS800TC instrument (Wyatt Technology Corporation, Santa Barbara, CA, USA) at 20–60 °C. The sample was a 0.02 μM filtered, SeMet-labelled Tm ExoPG solution (12 mg/ml in 10 mM HEPES–NaOH, pH 7.5, 3 mM NaN3), which we still had available, but was not needed for phasing because of the successful molecular replacement solution. Results were processed and analyzed with Dynamics software. Intermolecular interactions in the crystal structure and the stability of possible protein oligomers were analyzed using the PISA web server [36].

3. Results and discussion

3.1. Overall structure

The crystal structure of Tm ExoPG was solved by molecular replacement at 2.05 Å resolution using a composite model (see Section 2 and Table 1 for details). The asymmetric unit contains four molecules, which can be superimposed with an root mean square deviation (R.M.S.D.) between 0.19 and 0.32 Å for 444 Ca’s. One sequence correction was made based on the electron density; residue 216 was changed from valine to isoleucine.

The overall fold of Tm ExoPG resembles that of other GH28 enzymes: its core consists of a right-handed parallel β-helix, which is decorated with loops and secondary structure elements on one side ("top"), and a 26-residue N-terminal extension comprising an α-helix and a loop at the “bottom” (Fig. 1). The β-helix is com-

Table 1

<table>
<thead>
<tr>
<th>Data collection</th>
<th>P2_1,2,3_1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell (Å)</td>
<td>a = 63.3, b = 155.9, c = 200.7</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>34.3–2.05 (2.16–2.05)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.098 (0.523)</td>
</tr>
<tr>
<td>Rfree</td>
<td>0.057 (0.410)</td>
</tr>
<tr>
<td>Number of measured reflections</td>
<td>423 763 (27195)</td>
</tr>
<tr>
<td>Number of unique reflections</td>
<td>114 618 (11 297)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>91.6 (63.6)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Crystallographic refinement</th>
<th>Asymmetric unit content (4 monomers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of protein atoms (non-hydrogen)</td>
<td>14 138</td>
</tr>
<tr>
<td>Number of water molecules</td>
<td>1161</td>
</tr>
<tr>
<td>Overall B-factor (Å²)</td>
<td>28.1</td>
</tr>
<tr>
<td>Rmerge/Rfree</td>
<td>0.175/0.220</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stereochemical quality of the model</th>
<th>R.M.S.D. bond lengths (Å)</th>
<th>0.008</th>
</tr>
</thead>
<tbody>
<tr>
<td>R.M.S.D. bond angles (°)</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot (favoured)</td>
<td>1670 (94.4%)</td>
<td></td>
</tr>
<tr>
<td>Ramachandran plot (allowed)</td>
<td>102 (5.4%)</td>
<td></td>
</tr>
</tbody>
</table>
posed of three sheets, PB1, PB2 and PB3. The β-strands in the sheets are connected by turns or loops (T1, T2 and T3), some of which are long inserts. The β-helix is unique in several ways. First, it is longer than in any other GH28 structure. It comprises 2 incomplete N-terminal coils contributing to PB2 and PB3, and 12 complete coils up to the very C-terminus of the enzyme. So far, GH28 enzymes have been found displaying 7–10 complete coils [4]. Second, in the last three coils of Tm ExoPG, the β-strands contribute to more than one sheet (green coils in Fig. 1, β35–β37 in Fig. S1 [40]). A third unique feature of the β-helix in Tm ExoPG is that the very C-terminal coil (residues 434–447) is antiparallel to the other strands, and pairs antiparallel with the C-terminal strand of a neighboring molecule, thus extending the β-helix of the first molecule with the β-helix of the neighboring molecule (Fig. 2, left panel). The result of this is that the hydrophobic interior of the β-helix is shielded from the solvent. This is in contrast to other GH28 structures, where often a C-terminal capping helix, a disulphide-constrained loop, or a less hydrophobic exposed loop is present that shields the interior of the β-helix from the solvent. The N-terminal side of the β-helix of Tm ExoPG, is capped by an α-helix, a feature observed in most other β-helix folds.

3.2. Active site cleft

In Tm ExoPG, a cleft is found on the “top” of the PB1 sheet, formed by the T3 loops or inserting following coils 3–7, and the T1 loops from coils 8 and 11 (Fig. 1b). The position of this cleft corresponds to that of other GH28 enzymes. At the N-terminal end, the cleft in Tm ExoPG is deeper and forms a pocket, with the T3 loops after coil 3 and especially coil 4 forming a high-rising wall behind it. Consequently, the cleft is only accessible from the C-terminal side of the β-helix, explaining the exo-activity. A similar situation is observed in Ye ExoPG [14]; however, the pocket of Tm ExoPG is smaller and provides only one subsite (see Section 3.3), which explains the product specificity of the enzyme.

In PGs, 8 residues are strictly conserved [37]; in Tm ExoPG these residues (N237, D239, D260, D261, H296, G297, R327 and K329) are located at equivalent positions (Fig. S1). Among them, the three proposed catalytic aspartates are positioned at the bottom of the pocket. The structural conservation extends to residues Y362 [6] and the G302-S303 cis-peptide motif, positioned at the opposite side of the cleft.

3.3. Model for substrate binding

Kluskenes et al. [19] have shown that the enzyme cleaves its substrate from the non-reducing end and has maximum activity on (GalpA)4, suggesting that the enzyme has four sugar-binding subsites. Therefore, to analyze substrate binding we modelled a tetragalacturonate substrate in our structure (see Section 2 for details). Since the pocket is only large enough to provide one subsite (−1), the other GalpA residues must occupy subsites +1 to +3. The validity of the resulting model (Fig. 3a) is supported by several observations. First, the residues that surround the −1 and +1 subsites are structurally highly conserved. A superposition of the eight conserved residues, the G302-S303 cis-peptide and Y362 with those of Sp EndoPG I yields an R.M.S.D. of 0.58 Å. Second, the modelled galacturonate units do not clash with any protein residues. Third, the positions of the carboxylate groups and most of the hydroxyls of the modelled substrate coincide with the observed positions of water molecules in the native structure (Fig. 3b).

Kluskenes et al. [19] have shown that the enzyme cleaves its substrate from the non-reducing end and has maximum activity on (GalpA)4, suggesting that the enzyme has four sugar-binding subsites. Therefore, to analyze substrate binding we modelled a tetragalacturonate substrate in our structure (see Section 2 for details). Since the pocket is only large enough to provide one subsite (−1), the other GalpA residues must occupy subsites +1 to +3. The validity of the resulting model (Fig. 3a) is supported by several observations. First, the residues that surround the −1 and +1 subsites are structurally highly conserved. A superposition of the eight conserved residues, the G302-S303 cis-peptide and Y362 with those of Sp EndoPG I yields an R.M.S.D. of 0.58 Å. Second, the modelled galacturonate units do not clash with any protein residues. Third, the positions of the carboxylate groups and most of the hydroxyls of the modelled substrate coincide with the observed positions of water molecules in the native structure (Fig. 3b).

Accordingly, like in Sp EndoPG I, the galacturonate bound in the −1 subsite was modelled in a 4H3 half-chair conformation: C2, C1, O5 and C5 are coplanar, a feature that is characteristic of the postulated oxocarbenium ion-like transition state in GH reactions. Furthermore, this galacturonate unit is oriented such that its carboxylate is anchored by the conserved G302-S303 cis-peptide, and by the side chains of K329, Y362 and E304. Consequently, the glycosidic torsion angles between the −1 and +1 subsites (ψ = 8°, ψ = −150°) deviate considerably from those found for 2 or 3, helical conformations [7,34]. At the +2 and +3 subsites, the glycosidic torsion angles are representative of a 21 helical polygalacturonate conformation.

Inspection of the interactions between protein and modelled substrate reveals a conserved network of hydrogen bonds in the −1 and +1 subsites (Fig. 3a and b), even though some of the interacting residues are not conserved (K146/N91, W214/Q120, K148/H150, Tm ExoPG/Sp EndoPG I numbering). For the +2 and +3 subsites, residues that may interact with the substrate (either directly or via water molecules) are H296, E323, R324, R327 and E354. Further towards the C-terminal end of the enzyme, a patch of positively charged side chains (R357, K387 and R391) is observed at the surface that may interact with the carboxylate groups of longer substrates.

The three proposed catalytic aspartates (D239, D260 and D261) are positioned below the −1 and +1 subsites, approaching the substrate at its β-face (Fig. 3a and b). Their carboxylate groups are positioned at 4.8–5.7 Å from each other, similar to what has been
observed in other PGs. The aspartate closest to the glycosidic bond between the −1 and +1 units is D260 (distance O\(_d\)− glycosidic oxygen: 2.7 Å). It is appropriately oriented to act as the catalytic acid, protonating the oxygen of the scissile glycosidic linkage. The other two aspartates bind a conserved water molecule, activating it for its function as a nucleophile. This water molecule is oriented in-line with the C1 atom of the −1 galacturonate unit (distance: 3.7 Å) and the glycosidic oxygen, and is therefore positioned appropriately to attack the anomeric C1 carbon atom.

Our model also confirms the substrate specificity of the enzyme: in the −1 and +1 subsites, there is no room for substrates with methylated carboxylate groups or acetylated C3 hydroxyl groups (Fig. 3b). For a Δ4,5-unsaturated substrate in the −1 subsite (Fig. 3c), hydrogen bond interactions with the enzyme are
only partially maintained because of its half-chair conformation, and the missing C4 hydroxyl which creates an unfilled space. Moreover, the glycosidic oxygen between the −1 and +1 subsites likely takes a different position, which will lead to conformational strain and hence a lower affinity for the enzyme. Furthermore, the position of the glycosidic oxygen may be less favourable for protonation.

3.4. Quaternary structure: the tetramer

To support the observation of a tetramer in gel filtration experiments [19] and to examine the effect of temperature on the oligomeric state, DLS experiments were performed at 20–60 °C (Table 2). The results clearly indicate that Tm ExoPG is multimeric in solution; the MW values are close to the theoretical MW of a tetramer (204 kDa). A second observation is that with increasing temperature the overall polydispersity decreases. Thus, at high temperatures the enzyme is tetrameric and is somewhat less prone to aspecific aggregation.

In the crystal structure of Tm ExoPG, the asymmetric unit contains four molecules (A–D), which interact extensively. Analysis of the interactions by PISA [36] suggested that these molecules could form a stable tetramer, represented in Fig. 4a. There are two types of interactions. Molecule pairs A/B and C/D are engaged in a head-to-head interaction via their C-terminal ends, with an average buried surface area of 711 Å². The second type of interaction involves pairs A/C and B/D, interacting mainly with the C-terminal half of their PB3 sheets, and an average buried surface area of 719 Å². The total buried surface area upon tetramer formation is ~5790 Å², corresponding to 9.4% of the total surface area of all four molecules. To our knowledge, tetramerization of pectin-degrading enzymes or other β-helical proteins has not been reported before. The only known example of oligomerization is tomato pectin methylesterase which has been shown to form dimers at high concentration [38].

In the tetramer (overall dimensions 139 × 103 × 57 Å), the active site clefts of molecules A and D are on the same side; those of molecules B and C are on the opposite side (Fig. 4a). The distance between two clefts on the same side is about 67 Å, suggesting that substrate binding and processing likely are not influenced by tetramerization.

A more detailed look at the interactions within the tetramer reveals remarkable features. The head-to-head arrangement of the A/B and C/D pairs is achieved by the antiparallel pairing of their C-terminal β-strands (residues 434–447) (Fig. 4b). In fact, 15 main chain H-bonds are observed, thus combining the two molecules to form a continuous β-helix. A few other H-bonds and four salt bridges also contribute to this interface. The formation of this interface shields a large hydrophobic surface at the C-terminal end of the β-helix (interior) from solvent. A survey of the PDB for similar intermolecular β-helices yielded only one example, Helicobacter pylori vacuolating toxin; the crystal structure of its C-terminal P55 domain reveals a head-to-head dimer with intermolecular β-helix formation [39]. However the functional significance of this interaction is not clear. For the side-by-side pairs (A/C and B/D), the interacting residues are mainly hydrophobic in nature, and formation of these pairs buries a second hydrophobic patch on the surface of the enzyme. Only three H-bonds and one or two salt bridge interactions are observed (Fig. 4c). Also for this side-by-side arrangement of β-helical proteins, no other examples were found in the PDB. Both types of intermolecular contacts thus shield hydrophobic patches, possibly enhancing tetramer stability at high temperatures.

### Table 2
Dynamic light scattering results.

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Estimated MW (kDa)</th>
<th>Overall polydispersity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>189</td>
<td>60</td>
</tr>
<tr>
<td>30</td>
<td>176</td>
<td>48</td>
</tr>
<tr>
<td>40</td>
<td>186</td>
<td>45</td>
</tr>
<tr>
<td>50</td>
<td>177</td>
<td>43</td>
</tr>
<tr>
<td>60</td>
<td>193</td>
<td>42</td>
</tr>
</tbody>
</table>
The unique interactions between the β-helices and the hyper-thermal stability of the enzyme suggest that the protomer structure is “tailor-made” to form stable tetramers at the high temperature growing conditions of Tm. Especially, the unique C-terminal anti-parallel single β-strand, responsible for strong intermolecular circular β-sheet formation, is remarkable in this respect. The use of thermostable pectinolytic enzymes such as Tm ExoPG could be successful in the clarification process of fruit juices, especially for fruits with high pectin content which are poorly soluble and viscous at normal processing temperatures.

Acknowledgement

We thank the staff of beamline ID14-2 at the ESRF, Grenoble for data collection facilities and assistance.

Appendix A. Supplementary data


References