Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The structure and mode of action of *Caldicellulosiruptor bescii* family 3 pectate lyase in biomass deconstruction

Markus Alahuhta,[a] Roman Brunecky,[a] Puja Chandrayan,[b] Irina Kataeva,[b] Michael W. W. Adams,[b] Michael E. Himmel[a] and Vladimir V. Lunin[a]*

[a]Biosciences Center, National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401-3305, USA, and [b]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602-7229, USA

Correspondence e-mail: vladimir.lunin@nrel.gov

Received 25 September 2012
Accepted 11 December 2012

PDB Reference: PL3-cat, 4ew9

The unique active site of the *Caldicellulosiruptor bescii* family 3 pectate lyase catalytic module (PL3-cat) has been structurally described and synergistic digestion studies with *C. bescii* cellulase A have been performed on unpretreated biomass. The X-ray structure of PL3-cat was determined at 1.6 Å resolution (PDB entry 4ew9) in complex with the products of trigalacturonic acid. Comparison with family 1 pectate lyase (PL1) structures shows that the active site of the PL3 catalytic module is considerably different. However, on superimposing the identical sugar rings at the −2 subsites conserved interactions could be identified. Interestingly, only one catalytic residue, the lysine that donates the proton to the carboxylate group in the β-elimination reaction of PL1 (Lys108 in PL3-cat), is conserved in PL3 and there is no arginine to abstract the proton from the C5 carbon of the galactouronate ring. This suggests that the reaction mechanism of PL3 requires different catalytic residues. Most interestingly, comparison with other proton-abstraction reactions reveals that in PL3 the α-proton is abstracted by a lysine, in a striking similarity to enolases. These observations led us to propose that in PL3-cat Lys108 is the catalytic base, Glu84 is the catalytic acid and an acidified water molecule completes the *anti* β-elimination reaction by protonating the O4 atom of the substrate. Also, our digestion experiments with unpretreated switchgrass show that the loadings of *C. bescii* cellobiohydrolase A (CelA) can be lowered by the addition of PL3 to the reaction mixture. This result suggests that PL3 can significantly improve the deconstruction of unpretreated biomass by allowing other enzymes to better access their preferred substrates.

1. Introduction

Pectate lyases are an important part of the plant cell-wall deconstruction mechanism of bacterial and fungal microorganisms. Pectic polysaccharides consist of α-1,4-linked polygalactosyluronic acid residues with regions of alternating galactosyluronic acid and rhamnosyl residues (Willats *et al.*, 2001). Pectate lyases catalyze the degradation of pectins, which are a major component of the primary cell walls of higher plants, in the presence of calcium ions (Marín-Rodríguez *et al.*, 2002). They were first discovered in plant pathogenic bacteria that cause diseases involving the maceration of the parenchymatous tissues of various dicot plants (Collmer & Keen, 1986).

Their ability to break down pectic polysaccharides, which are major structural components of higher plants (Carpita & Gibeaut, 1993), makes pectate lyases necessary for the efficient deconstruction of plant cell walls by microbes. This makes them critical candidates for enzyme cocktails that digest unpretreated biomass from which pectins have not been

Table 1
X-ray data-collection and refinement statistics.

Values in parentheses are for the highest resolution bin.

Data collection	
Space group	*C*2
Unit-cell parameters (Å, °)	a = 136.946, b = 35.898, c = 98.687, β = 132.0
Wavelength (Å)	1.54178
Temperature (K)	100
Resolution (Å)	50–1.6 (1.7–1.6)
Unique reflections	47595 (7847)
Observed reflections	229883 (23619)
$R_{\rm int}$†	0.0584 (0.3498)
Average multiplicity	4.83 (3.01)
$\langle I\rangle/\langle\sigma(I)\rangle$	13.62 (2.74)
Completeness (%)	99.9 (100.0)
Refinement	
$R/R_{\rm free}$	0.164 (0.257)/0.204 (0.298)
Protein atoms	3135
Water molecules	383
Other atoms	136
R.m.s.d. from ideal bond lengths‡ (Å)	0.023
R.m.s.d. from ideal bond angles‡ (°)	2.175
Wilson B factor (Å^2)	18.1
Average B factor (Å^2)	
Protein atoms	21.0
Water molecules	32.3
Ramachandran plot statistics§ (%)	
Allowed	100.0
Favored	95.1
Outliers	0

† $R_{\rm int} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection and $\langle I(hkl)\rangle$ is the mean intensity of a group of equivalents. ‡ Engh & Huber (1991). § Chen *et al.* (2010).

removed by chemical pretreatment. Recently, pectate lyases have been shown to be present in the genomes of eight globally diverse *Caldicellulosiruptor* species (Blumer-Schuette *et al.*, 2012). These enzymes are thus an important target for detailed studies of their structure–function relationships and biochemical roles. *C. bescii* family 3 pectate lyase is especially important because *C. bescii* is the most thermophilic cellulose-degrading organism known (Yang *et al.*, 2010).

The binding site of the unliganded *Bacillus* sp. KSM-P15 family 3 pectate lyase (*Bacillus* PL3) has been characterized previously (Akita *et al.*, 2001); however, owing to the lack of a liganded structure Akita and coworkers were not able to reliably describe the active site and concluded that the conformation of the substrate may be different in *Bacillus* PL3 compared with family 1 pectate lyases. In the *B. subtilis* family 1 pectate lyase (*Bacillus* PL1) *anti* β-elimination reaction the α-proton is abstracted from the C5 carbon of the substrate by an arginine, while the carboxylic acid is protonated by a lysine, generating an enol intermediate which is stabilized by two catalytically important Ca atoms; this is followed by elimination of the β-leaving group (Seyedarabi *et al.*, 2010; Scavetta *et al.*, 1999).

Here, we report the X-ray structure of the PL3 catalytic module from the hyperthermophilic cellulolytic (Yang *et al.*, 2010; Madigan & Marrs, 1997) bacterium *C. bescii* in complex with its products; using synergistic activity measurements, we also studied the importance of PL3 in the deconstruction of unpretreated biomass.

2. Materials and methods

2.1. Cloning, expression and purification

The PL3 catalytic module (PL3-cat) and PL3 containing the carbohydrate-binding module (PL3) encoded by *C. bescii* (Cbes_1854; 193 amino acids) were cloned into pET-45b (Novagen) between *Kpn*I and *Xho*I sites, so that the construct contained an N-terminal 6×His affinity tag. The forward and reverse primers were GGGGTACCAATACGGGGTGGTGTTTTAGTTATTACAG and CCGCTCGAGTTAGTATTGATGTATCTGTGATTGGG, respectively. The 20.9 kDa protein was expressed in *Escherichia coli* BL21 (DE3) cells (10 l culture) induced with 1 m*M* IPTG for 16 h at 310 K. The cell pellet was collected and resuspended in 20 m*M* sodium phosphate pH 7.0 containing 0.3 *M* NaCl (buffer *A*) and the cells were disrupted using a French press. Cell debris was removed by centrifugation at 10 000 rev min^{-1} for 30 min. The supernatant was heated at 343 K for 20 min; after removing denatured protein by centrifugation at 10 000 rev min^{-1} for 30 min, the supernatant was applied onto a HisTrap FF affinity column (5 ml, GE Healthcare) equilibrated with buffer *A*. After washing the column, the protein sample was eluted using an imidazole gradient (0–0.5 *M*) in buffer *A*. Based on SDS–PAGE analysis, fractions containing the pectate lyase costruct were combined, concentrated using ultrafiltration (Amicon, 10 kDa cutoff) and dialyzed against 20 m*M* sodium phosphate pH 7.0. Finally, the remaining impurities were removed by size-exclusion chromatography using a HiLoad Superdex 75 (26/60) column (GE Healthcare, Piscataway, New Jersey, USA) in 20 m*M* acetic acid pH 5.0 containing 100 m*M* NaCl and 5 m*M* $CaCl_2$. The purified protein was concentrated using a Vivaspin 20 concentrator (5 kDa cutoff; GE Healthcare, Little Chalfont, Buckinghamshire, England) and the protein concentration was measured using the Pierce BCA Protein Assay kit (Pierce Biotechnology, Rockford, Illinois, USA).

2.2. Crystallization

PL3-cat crystals were obtained by sitting-drop vapor diffusion in a 96-well plate using Crystal Screen HT from Hampton Research (Aliso Viejo, California, USA). 50 µl well solution was added to the reservoirs and drops were made with 0.2 µl well solution and 0.2 µl protein solution using a Phoenix crystallization robot (Art Robbins Instruments, Sunnyvale, California, USA). The crystals grew at 293 K with 0.1 *M* HEPES pH 7.5, 70%(*v*/*v*) (±)-2-methyl-2,4-pentanediol as the well solution. The protein solution consisted of 15 mg ml^{-1} protein, 20 m*M* acetic acid pH 5, 100 m*M* NaCl, 5 m*M* $CaCl_2$, 5 m*M* trigalacturonic acid.

2.3. Data collection and processing

The PL3-cat crystal was flash-cooled in a nitrogen-gas stream at 100 K before data collection using an in-house Bruker X8 MicroStar X-ray generator with Helios mirrors and a Bruker PLATINUM^{135} CCD detector. Data were indexed

and processed using the Bruker suite of programs v.2011.2-0 (Bruker AXS, Madison, Wisconsin, USA).

2.4. Structure solution and refinement

Intensities were converted into structure factors and 5% of the reflections were flagged for R_{free} calculations using the programs *F2MTZ*, *TRUNCATE*, *CAD* and *UNIQUE* from the *CCP*4 package of programs (Winn *et al.*, 2011). The program *MOLREP* v.10.2.23 (Vagin & Teplyakov, 2010) was used for molecular replacement, using the unliganded structure of a family 3 pectate lyase from *C. bescii* (PDB entry 3t9g; Alahuhta *et al.*, 2011) as the search model. Refinement and manual correction was performed using *REFMAC*5 v.5.5.0109 (Murshudov *et al.*, 2011) and *Coot* v.0.6 (Emsley *et al.*, 2010). The *MolProbity* method (Chen *et al.*, 2010) was used to analyze the Ramachandran plot, and the root-mean-square deviations (r.m.s.d.s) of bond lengths and angles from the ideal values of the Engh and Huber stereochemical parameters (Engh & Huber, 1991) were calculated. The Wilson *B* factor was calculated using *CTRUNCATE* v.1.0.11 (Winn *et al.*, 2011). Average *B* factors were calculated using *ICM* v.3.7-2c (Molsoft LLC, La Jolla, California, USA). The data-collection and refinement statistics are shown in Table 1.

2.5. Structure analysis

Coot, *PyMOL* (http://www.pymol.org) and *ICM* were used for comparing and analyzing structures. C^{α} root-mean-square deviations (C^{α}-r.m.s.d.s) for structure comparisons were calculated using *ICM*. Figs. 1, 2 and 3 were created using *PyMOL*. Fig. 4 was created using *Chemtool* (Brüstle, 2001).

The superimposition of PL3-cat and *Bacillus* PL1 structures (Seyedarabi *et al.*, 2010) was an iterative process in which we initially tried to superimpose the active-site calciums by moving one of the structures manually when normal superimposition with *Coot*, *ICM* and *PyMOL* failed to bring the active sites together and then realised that the sugar rings at the −2 subsites were almost identical. Software superimposition was not used, as by manually moving the structures we were better able to look for other 'imprecise' features of the active site, not just the sugar ring, to find the similar parts of the two structures. A small difference in the sugar rings can cause a large change in the superimposition of the nearby residues.

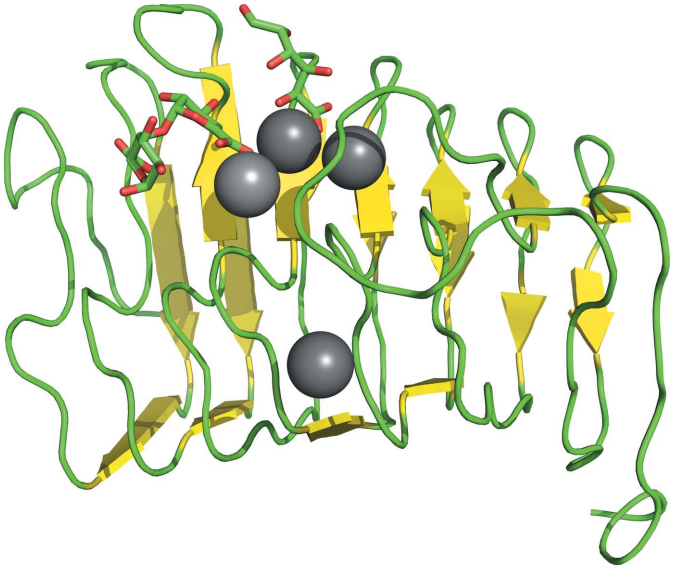

Figure 1
The overall fold of PL3-cat with the contents of the active site. The 4,5-unsaturated digalacturonic acid and the linear form of D-galacturonic acid are shown as sticks with green C atoms and red O atoms. Ca atoms are shown as gray balls.

2.6. Model biomass substrate

Native switchgrass samples were used to evaluate the enhancement of cellulolytic efficiency by PL3 in combination with the multifunctional CelA enzyme from *C. bescii*. To provide a basis for the maximum theoretical sugar yield achievable from each substrate during enzymatic hydrolysis, portions of each of the pretreated solid samples were dried and subjected to the standard two-stage sulfuric acid hydrolysis method for determining structural carbohydrates in lignocelluloses as described by Sluiter *et al.* (2008). In this method, the carbohydrate content of each pretreated sample is calculated from the carbohydrates released.

2.7. Enzymatic synergy assays

Purified PL3 and CelA from *C. bescii* were utilized to determine the synergistic enhancement of biomass. Conversion of native switchgrass was determined at 348 K at an enzyme loading of 15 mg CelA or of 10 mg CelA and 5 mg PL3 protein per gram of cellulose. Digestion assays were performed in 20 m*M* acetate buffer pH 5.5 containing 10 m*M* $CaCl_2$ and 100 m*M* NaCl. Digestions were run continuously for 7 d and sugar release was monitored. Samples were taken at various time points and enzymes were inactivated by boiling for 15 min. Samples were then filtered through 0.45 µm Acrodisc syringe filters and analyzed for glucose and cellobiose by HPLC. 20 µl samples were injected and run on an Agilent 1100 HPLC system equipped with a Bio-Rad Aminex HPX-87H 300 × 7.8 mm column heated to 328 K. A constant flow of 0.6 ml min^{-1} was used with 0.1 *M* H_2SO_4 in water as the mobile phase to give optimal sugar separation. Glucose, xylose and cellobiose were quantified against independent standard curves. All experiments were performed in triplicate and the resulting extents of conversion are shown as a percentage of the maximum theoretical glucan converted.

3. Results and discussion

3.1. The structure of the *C. bescii* PL3 catalytic module

We have solved a 1.6 Å resolution X-ray structure of the *C. bescii* family 3 pectate lyase catalytic module in complex with two molecules: 4-*O*-(4-deoxy-β-L-*threo*-hex-4-enopyranuronosyl)-α-D-galactopyranuronic acid (4,5-unsaturated digalacturonic acid) and the linear form of D-galacturonic acid (Fig. 1). Most importantly, for the first time, this structure shows in high detail the interaction between this enzyme and

the products of trigalacturonic acid bound in the active site (Fig. 2). Three Ca atoms were also bound in the active site. The 4,5-unsaturated digalacturonic acid occupied subsites −2 and −1 and interacted with Ca atoms 1 and 2 (Fig. 2). The linear form of D-galacturonic acid was bound at subsite +1 in contact with Ca atoms 2 and 3 (Fig. 2). The overall fold is a parallel β-helix, which is typical for pectate lyases (Marín-Rodríguez *et al.*, 2002). The resulting structure was refined to an R factor of 0.164 and an $R_{\rm free}$ of 0.204 with two molecules in the asymmetric unit. The structure has been deposited in the Protein Data Bank (PDB) as entry 4ew9.

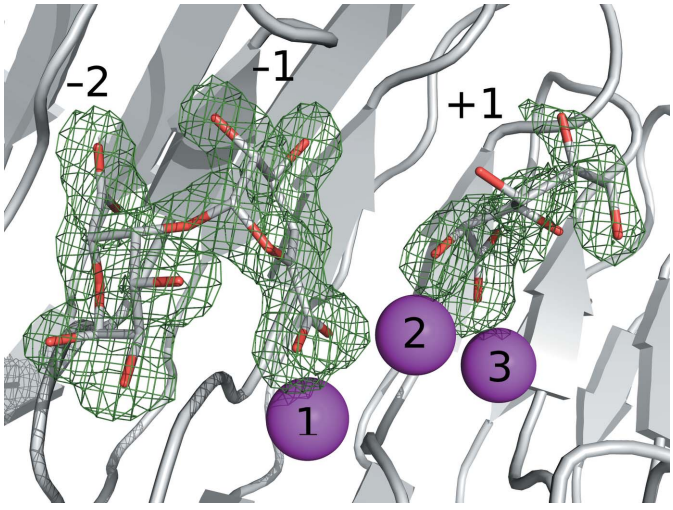

Figure 2
The OMIT electron-density map of PL3-cat with the contents of the active site. The 4,5-unsaturated digalacturonic acid (subsites −2 and −1) and the linear form of D-galacturonic acid (subsite +1) are shown as sticks with gray C atoms and red O atoms. Ca atoms are shown as magenta balls and are numbered 1, 2 and 3 from left to right. This $F_{\rm o} - F_{\rm c}$ map was calculated at 3.0σ after ten cycles of refinement without the ligands.

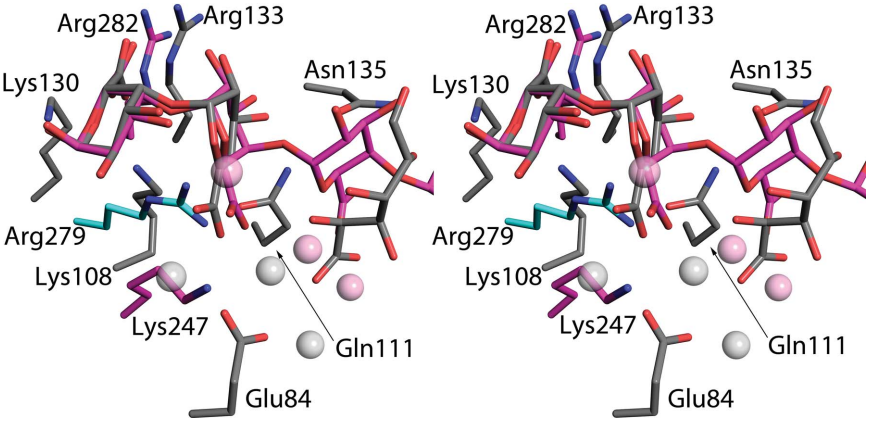

Figure 3
Details of the superimposed active sites of PL3-cat and *Bacillus* PL1 in stereo. The ligands and discussed residues are shown as sticks and Ca atoms are shown as spheres. The PL3-cat C atoms are shown in gray and Ca atoms are shown in light gray. *Bacillus* PL1 C atoms are colored magenta and Ca atoms are colored pink. The catalytically important Arg279 of *Bacillus* PL1 was mutated to an alanine in the liganded structure (PDB entry 2nzm; Seyedarabi *et al.*, 2010), so the unliganded and unmutated native structure (PDB entry 1bn8; Pickersgill *et al.*, 1994) was instead used to show this residue (with C atoms in cyan).

3.2. Comparison with other pectate lyases

In our recent report on the unliganded structure of PL3-cat we concluded that the most similar structure available is that of *Bacillus* PL3, while other structures were only similar in fold (Alahuhta *et al.*, 2011). This structure also confirms the position of the Ca atom that was replaced by an Na atom in our previous structure (Ca atom 1 in Fig. 2). The active site contains a total of three Ca atoms, similar to family 1 pectate lyases (Seyedarabi *et al.*, 2010). Closer inspection of PL3-cat shows that its active site and the overall fold (C^{α}-r.m.s.d. of 1.0 Å) are indeed practically identical to those of *Bacillus* PL3, with most differences causing no changes to the fold and only minor changes to the substrate-binding area outside subsites −2, −1 and +1. Comparison with *Bacillus* PL1 structures (Seyedarabi *et al.*, 2010) reveals that while the overall fold can be superimposed fairly well (C^{α}-r.m.s.d. of 1.9 Å), the active site of PL3-cat is not in the same position. However, by superimposing the nearly identical ligand atoms at the −2 subsites of PL3-cat and *Bacillus* PL1 (PDB entry 2nzm; Seyedarabi *et al.*, 2010) we were able to identify conserved interactions (Fig. 3). There are significant differences in the calcium coordination, but other conserved features superimpose surprisingly well. The interaction between Arg133 and the substrate carboxylate group at the −2 subsite, which seems to be needed for correct positioning of the substrate, is identical in both structures.

Interestingly, only the lysine that donates a proton to the carboxylate group in the β-elimination reaction of *Bacillus* PL1 (Lys108 in PL3-cat; Fig. 3) is conserved in *C. bescii* PL3, and in place of the arginine that abstracts the α-proton from the C5 atom of the galactouronate ring in *Bacillus* PL1 a lysine is present in both *C. bescii* PL3 and *Bacillus* PL3 (Lys130 in PL3-cat; Fig. 3). The function of this lysine seems to be non-catalytic as it is clearly hydrogen bonded to the O5 atom of the galactouronate ring bound at the −2 subsite. This indicates that the reaction mechanism of family 3 pectate lyases requires a different catalytic residue to abstract the α-proton from the C5 atom. The only nearby basic side chain in both *C. bescii* PL3 and *Bacillus* PL3 is a glutamine (Gln111 in PL3-cat; Fig. 3) that is roughly at the position of Ile250 of *Bacillus* PL1. However, a neutral amino acid such as glutamine cannot function as a catalytic base for proton abstraction and would require a significantly different transition-state conformation of the substrate compared with family 1 pectate lyases as this residue points to the C5 α-proton from a significantly different angle.

3.3. The catalytic mechanism of family 3 pectate lyases

We were intrigued by these results and turned to the literature explaining the enzyme-catalyzed proton abstraction in

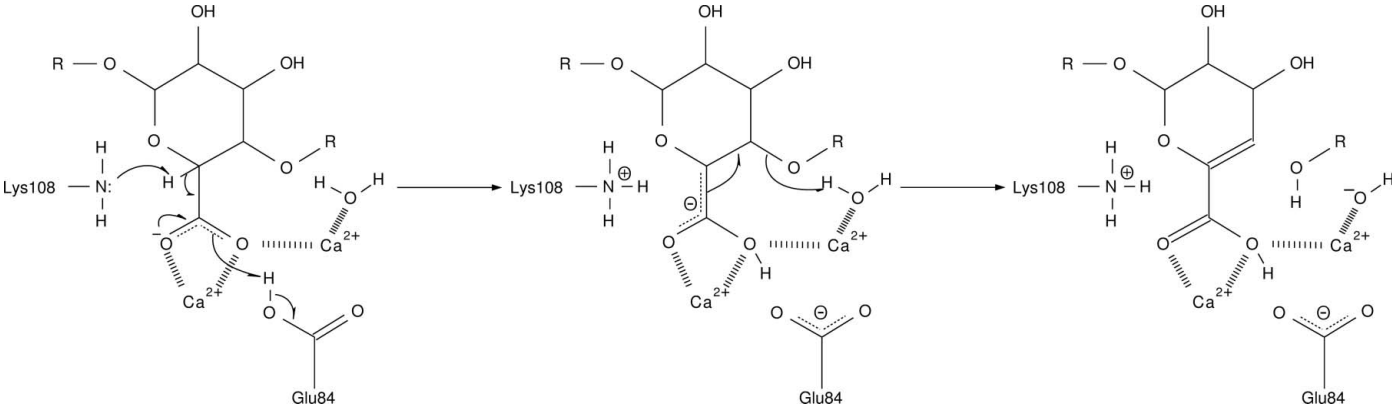

Figure 4
The proposed *anti* β-elimination reaction of family 3 pectate lyase.

detail. The pectate lyase *anti* β-elimination reaction requires a basic residue to abstract the α-proton from the C5 atom of the substrate that has been acidified (by the two catalytically important Ca^{2+} ions) and the concerted action of an acidic residue that donates a proton to the carbonyl group, creating an enol intermediate (Gerlt *et al.*, 1991; Gerlt & Gassman, 1992, 1993). This is followed by elimination of the β-leaving group together with protonation of the O4 atom by another acidic residue or a water molecule acidified by the two catalytically important Ca^{2+} ions, while the enol intermediate is stabilized by the Ca atoms (Scavetta *et al.*, 1999; Seyedarabi *et al.*, 2010). Closer inspection of Fig. 3 indeed shows that a basic residue (Lys108) is present in a position to abstract the α-proton from the C5 atom and an acidic residue (Glu84) is present to donate the proton to the carbonyl group of the enol intermediate. The second acidic residue is missing, as only the neutral Asn135 is nearby, but an acidified water molecule could also protonate the O4 atom of the substrate. Both of the possible catalytic residues, Lys108 and Glu84, have the correct position and distance from the groups that they need to modify: no conformational changes would be required in the substrate and they are present in both PL3 structures. Owing to the absence of acidic residues within range it is likely that an acidified water molecule performs the last protonation step. Also, although lysine is an unlikely residue to abstract a proton, it is not unheard of. In enolases a lysine which has been acidified by two Mg^{2+} ions abstracts the α-proton, in a striking similarity to pectate lyases (Dinovo & Boyer, 1971; Pancholi, 2001). Encouraged by these observations, we propose that Lys108 is the catalytic base, Glu84 is the catalytic acid and an acidified water completes the *anti* β-elimination reaction by protonating the O4 atom of the substrate (Fig. 4).

3.4. Synergy with *C. bescii* CelA in degrading unpretreated switchgrass

The ability of pectate lyases to efficiently degrade pectin, which is a major structural component of plant cell walls (Carpita & Gibeaut, 1993), makes it a candidate for future enzyme mixtures aimed at deconstructing unpretreated biomass. PL3 on its own would not be sufficiently active, but a sufficient extent of degradation can be attained using a mixture with CelA from *C. bescii*. Our premise for this experiment was to determine whether or not the addition of PL3 can improve the overall rate of hydrolysis and thus the economics of unpretreated switchgrass digestion by CelA. We chose *C. bescii* CelA as the accompanying enzyme because it has both endoglucanase and exoglucanase activities (Zverlov *et al.*, 1998) and has been shown to be the main cellulase in *C. bescii* (Lochner *et al.*, 2011; Blumer-Schuette *et al.*, 2012). Also, *C. bescii* has been shown to grow well on crystalline cellulose, as well as unpretreated substrates such as switchgrass and poplar, at 348 K (Yang *et al.*, 2009).

Experimental digestions of unpretreated switchgrass substrates showed that the addition of PL3 to purified CelA cellulase with a 33% decreased loading, giving an overall similar total protein loading overall, can achieve the same level of digestion as CelA digestion of unpretreated switchgrass alone (Fig. 5). While we did not observe an improvement in the rate or extent of digestion, this early result suggests that PL3 may improve access to the cellulose substrate or otherwise improve the overall function of CelA. Clearly, *C. bescii* PL3 can be used to supplement the more carbon-intensive

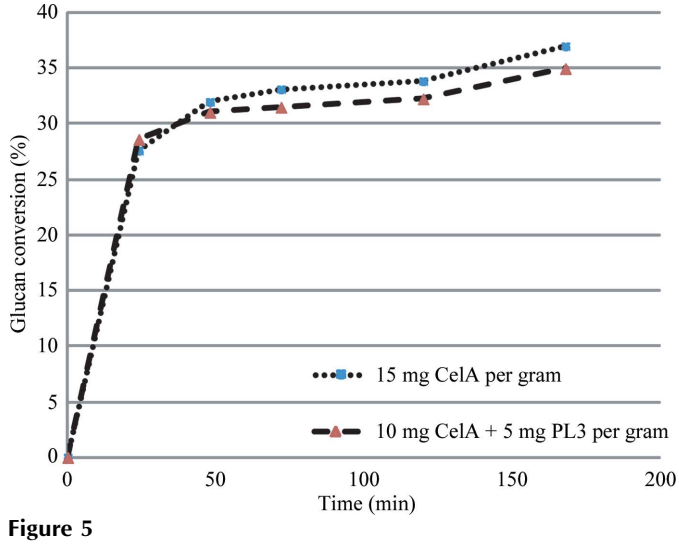

Figure 5
Unpretreated switchgrass conversion by *C. bescii* CelA and PL3.

CelA to improve the economics of unpretreated biomass deconstruction.

4. Conclusions

We have determined the X-ray structure of PL3-cat in complex with its products and assayed its synergistic effect in degrading unpretreated switchgrass together with *C. bescii* CelA. Our results show that the active site of PL3-cat is located differently compared with family 1 pectate lyases. Through careful comparison of available structures and similar proton-abstraction reactions, we have been able to propose a reaction mechanism for family 3 pectate lyases that uses different catalytic residues compared with family 1 pectate lyases. In this reaction Lys108 is the catalytic base, Glu84 is the catalytic acid and an acidified water completes the *anti* β-elimination reaction by protonating the O4 atom of the substrate. We have also investigated the importance of PL3 in the deconstruction of unpretreated biomass with the premise that pectate lyase will improve the efficiency of other enzymes by removing pectins and making it possible for them to access their preferred substrate. In agreement with this hypothesis, our digestion experiments with unpretreated switchgrass show that the amount of *C. bescii* CelA needed can be reduced by adding PL3 to the reaction mixture, thus improving the economics of unpretreated biomass deconstruction.

This work was funded by the US DOE Office of Science, Biological and Environmental Research Program, Bioenergy Research Center (BioEnergy Science Center, BESC) managed by Oak Ridge National Laboratory.

References

Akita, M., Suzuki, A., Kobayashi, T., Ito, S. & Yamane, T. (2001). *Acta Cryst.* D**57**, 1786–1792.
Alahuhta, M., Chandrayan, P., Kataeva, I., Adams, M. W. W., Himmel, M. E. & Lunin, V. V. (2011). *Acta Cryst.* F**67**, 1498–1500.
Blumer-Schuette, S. E. *et al.* (2012). *J. Bacteriol.* **194**, 4015–4028.
Brüstle, M. (2001). *Nachr. Chem.* **49**, 1310–1314.
Carpita, N. C. & Gibeaut, D. M. (1993). *Plant J.* **3**, 1–30.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Collmer, A. & Keen, N. T. (1986). *Annu. Rev. Phytopathol.* **24**, 383–409.
Dinovo, E. C. & Boyer, P. D. (1971). *J. Biol. Chem.* **246**, 4586–4593.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Gerlt, J. A. & Gassman, P. G. (1992). *J. Am. Chem. Soc.* **114**, 5928–5934.
Gerlt, J. A. & Gassman, P. G. (1993). *J. Am. Chem. Soc.* **115**, 11552–11568.
Gerlt, J. A., Kozarich, J. W., Kenyon, G. L. & Gassman, P. G. (1991). *J. Am. Chem. Soc.* **113**, 9667–9669.
Lochner, A., Giannone, R. J., Rodriguez, M., Shah, M. B., Mielenz, J. R., Keller, M., Antranikian, G., Graham, D. E. & Hettich, R. L. (2011). *Appl. Environ. Microbiol.* **77**, 4042–4054.
Madigan, M. T. & Marrs, B. L. (1997). *Sci. Am.* **276**, 82–87.
Marín-Rodríguez, M. C., Orchard, J. & Seymour, G. B. (2002). *J. Exp. Bot.* **53**, 2115–2119.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Pancholi, V. (2001). *Cell. Mol. Life Sci.* **58**, 902–920.
Pickersgill, R., Jenkins, J., Harris, G., Nasser, W. & Robert-Baudouy, J. (1994). *Nature Struct. Biol.* **1**, 717–723.
Scavetta, R. D., Herron, S. R., Hotchkiss, A. T., Kita, N., Keen, N. T., Benen, J. A., Kester, H. C., Visser, J. & Jurnak, F. (1999). *Plant Cell*, **11**, 1081–1092.
Seyedarabi, A., To, T. T., Ali, S., Hussain, S., Fries, M., Madsen, R., Clausen, M. H., Teixteira, S., Brocklehurst, K. & Pickersgill, R. W. (2010). *Biochemistry*, **49**, 539–546.
Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D. & Crocker, D. (2008). *Determination of Structural Carbohydrates and Lignin in Biomass.* http://www.nrel.gov/biomass/pdfs/42618.pdf.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Willats, W. G., McCartney, L., Mackie, W. & Knox, J. P. (2001). *Plant Mol. Biol.* **47**, 9–27.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yang, S.-J., Kataeva, I., Hamilton-Brehm, S. D., Engle, N. L., Tschaplinski, T. J., Doeppke, C., Davis, M., Westpheling, J. & Adams, M. W. W. (2009). *Appl. Environ. Microbiol.* **75**, 4762–4769.
Yang, S.-J., Kataeva, I., Wiegel, J., Yin, Y., Dam, P., Xu, Y., Westpheling, J. & Adams, M. W. W. (2010). *Int. J. Syst. Evol. Microbiol.* **60**, 2011–2015.
Zverlov, V., Mahr, S., Riedel, K. & Bronnenmeier, K. (1998). *Microbiology*, **144**, 457–465.