short communications

Molecular determinants of substrate specificity in the feruloyl esterase module of xylanase 10B from *Clostridium thermocellum*

Nicolas Tarbouriech,[a]**‡
José A. M. Prates,**[b]
Carlos M. G. A. Fontes[b] **and
Gideon J. Davies**[a]*****

[a]Structural Biology Laboratory, Department of
Chemistry, University of York, Heslington,
York YO10 5YW, England, and [b]Centro
Interdisciplinar de Investigação em Sanidade
Animal, Faculdade de Medicina Veterinária,
Pólo Univeristário do Alto da Ajuda,
1300-477 Lisboa, Portugal

‡ Present address: EMBL-Grenoble c/o ILL,
F-38042 Grenoble CEDEX, France.

Correspondence e-mail: davies@ysbl.york.ac.uk

Feruloyl esterases play a key role in the degradation of the intricate structure of the plant cell wall by hydrolysing the ferulate ester groups involved in the cross-linking between hemicelluloses and between hemicellulose and lignin. The structure of the feruloyl esterase module of *Clostridium thermocellum* cellulosomal xylanase 10B has been reported previously. It displays the α/β hydrolase fold with a classical Ser-His-Asp catalytic triad. Here, the structures of a Ser-Ala mutant of this feruloyl esterase in complexes with methyl syringate, methyl sinapinate and methyl vanillate are described. Substrate binding is accompanied by subtle conformational changes at amino acids Trp982, Met955, Asn1023 and Ile1019 in the ligand-binding cavity. The structural determinants, particularly the *m*-methoxy substituent, governing the substrate specificity of Xyn10B feruloyl esterase are rationalized.

1. Introduction

The plant cell wall is a complex conglomerate of several polymers including cellulose, hemicellulose, lignin and pectin. Phenolic cross-links using ferulic or 4-coumaric acid contribute to the structural integrity of the cell wall by mediating direct covalent interactions between hemicellulose chains and between hemicellulose and lignin. Ferulate is usually esterified to the arabinofuranosyl side chains of arabinoxylans, one of the most abundant hemicelluloses, comprising a backbone of β-1,4-linked xylose units that in addition to arabinose can also be 'decorated' with covalently linked acetate and methyl glucuronate substituents. Phenolic residues can link directly to lignin or form diferulic acid dimers, which result in covalent cross-linking between hemicelluloses.

The degradation of the plant cell wall by bacteria and fungi involves the cooperative action of a large consortium of enzymes. These enzymes frequently display a very complex modular architecture. In *Clostridia*, the plant cell-wall-degrading enzyme repertoire are arranged in a supramolecular assembly termed the 'cellulosome' (Bayer *et al.*, 1998). Xylanase 10B (Xyn10B) is a modular cellulosomal xylanase from *Clostridium thermocellum* that contains, in addition to a family 10 xylanase domain, two carbohydrate-binding modules and a C-terminal domain, which displays feruloyl esterase activity (Blum *et al.*, 2000). The structure of the Xyn10B feruloyl esterase domain (FAE) has been solved by X-ray crystallography, revealing an α/β-hydrolase fold with a catalytic triad composed of Ser954, His1018 and Asp1058 (Prates *et al.*, 2001). In recombinant *Escherichia coli*-derived protein, the catalytic serine has been shown to be post-translationally modified and was thus modelled as phosphoserine. The structure of a complex of a mutant protein in which the catalytic serine was mutated to alanine (FAE-S954A) confirmed the position of the active site and revealed the ferulate-recognition apparatus (Prates *et al.*, 2001).

Here, we describe the structure of three novel structures of the S954A mutant of the *C. thermocellum* Xyn10B feruloyl esterase module in complexes with methyl syringate, methyl sinapinate and methyl vanillate (Fig. 1). Interestingly, we were unable to obtain complexes with methyl coumarate or caffeic acid. Taken together, the data suggest that the presence of methoxy groups in the *meta* position of the phenolic structures is crucial for the recognition of Xyn10B FAE target substrates.

2. Materials and methods

2.1. Materials

Methyl sinapinate, methyl syringate, methyl vanillate and methyl p-coumarate were purchased from Apin Chemicals (UK) and caffeic acid was acquired from Sigma–Aldrich (UK).

2.2. Methods

Crystals of FAE-S954A were grown by vapour-phase diffusion using the hanging-drop method from 15 mg ml^{-1} protein in 0.1 M Na HEPES buffer pH 7.5 with 50 mM CdSO$_4$, 5%(v/v) glycerol and 1 M sodium acetate as precipitant as described previously (Prates et al., 2001). The unliganded crystals (space group $P2_12_12_1$, with approximate unit-cell parameters a = 65.4, b = 108.8, c = 113.9 Å) were harvested into mother liquor consisting of the crystallization conditions supplemented with 10 mM of the desired ligand for 2–3 h before freezing. All data sets were collected at the European Synchrotron Radiation Facility on beamline ID14-EH2. Data were processed with DENZO and SCALEPACK (Otwinowski & Minor, 1997) and all further computing used programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994) unless otherwise stated. Rigid-body refinement was performed with REFMAC (Murshudov et al., 1997) using the un-liganded unsolvated S954A structure. Initial rigid-body refinement was followed by cycles of solvent building using the solvent option of ARP/wARP (Perrakis et al., 2001). The ligands and heteroatoms were incorporated and the resulting models refined with maximum-likelihood methods using REFMAC (Murshudov et al., 1997). All manual rebuilding was performed with QUANTA (Accelrys, San Diego, CA, USA).

3. Results and discussion

3.1. Side-chain movements in the active site of FAE-S954A complexes

FAE displays the typical α/β-hydrolase fold previously observed in various esterases, lipases and acetylases. Ser954, Asp1018 and His1058 comprise a classical catalytic triad, as first described for serine proteases in 1967 (Matthews et al., 1967). The structures of FAE-S954A in complex with three bound methyl esters (Figs. 1 and 2) were obtained using data collected at the ESRF and refinement statistics are given in Table 1. The overall folds of the complexed and free FAE-S954A structures did not differ significantly (r.m.s.d. lower than 0.15 Å for 270 C$^\alpha$ atoms); however, differences were observed in the conformations of the residues in the active site. The most significant atomic changes detected in the three structures are illustrated in Fig. 3. The side chain of Trp982 rotates around the C$^\alpha$—C$^\beta$ bond such that the CZ3 atom is

moved up to 2 Å towards the m-methoxy group of the substrates. In addition, Met955 is locked in a single conformation close to the substrate instead of being in two conformations as observed previously. Together, these two side chains reduce the size of the hydrophobic pocket and stabilize the methoxy group of the substrates. Another movement observed in all complex structures is a shift of the side chain of Asn1023 towards the p-hydroxy group and towards the second methoxy group in the case of syringate and sinapinate. This conformational change allows the amide group to hydrogen bond to either the hydroxy or the second methoxy group of the substrate at a distance of approximately 3 Å. Finally, Ile1019, which is poorly ordered in the case of the native structure, is stabilized in a conformation in which the side chain points towards the aliphatic chain of the ligands.

3.2. Substrate accommodation

Under the same conditions, soaks with methyl 4-coumarate (similar to methyl ferulate but lacking the m-methoxy substituent)

Table 1
Data sets and model statistics.

Outer resolution shell statistics are given in parentheses.

Complex	Methyl sinapinate	Methyl syringate	Methyl vanillate
Resolution	20–1.4 (1.42–1.40)	20–1.40 (1.42–1.40)	20–1.40 (1.42–1.40)
Completeness	99.0 (93.3)	96.4 (90.6)	99.3 (94.7)
R_{merge} (%)	4.2 (23.3)	5.1 (18.0)	5.6 (33.4)
Multiplicity	5.45 (4.26)	4.22 (4.05)	4.38 (3.40)
$I/\sigma(I)$	35.2 (7.2)	22.4 (6.4)	26.3 (3.1)
R_{cryst}	11.1	11.8	11.7
R_{free}	13.3	13.8	14.0
R.m.s. to target bond distances (Å)	0.021	0.021	0.021
R.m.s. to target angles (°)	1.927	1.925	1.925
Heteroatoms			
Cadmium	12	17	20
Glycerol	6	4	7
Acetate	2	2	2
Substrate	2 methyl sinapinate	2 methyl syringate	2 methyl vanillate
Mean B, protein (Å^2)	13	12	13
Mean B, solvent (Å^2)	27	26	29
Mean B, 'substrate' (Å^2)	24	17	19
PDB code	1wb4	1wb5	1wb6
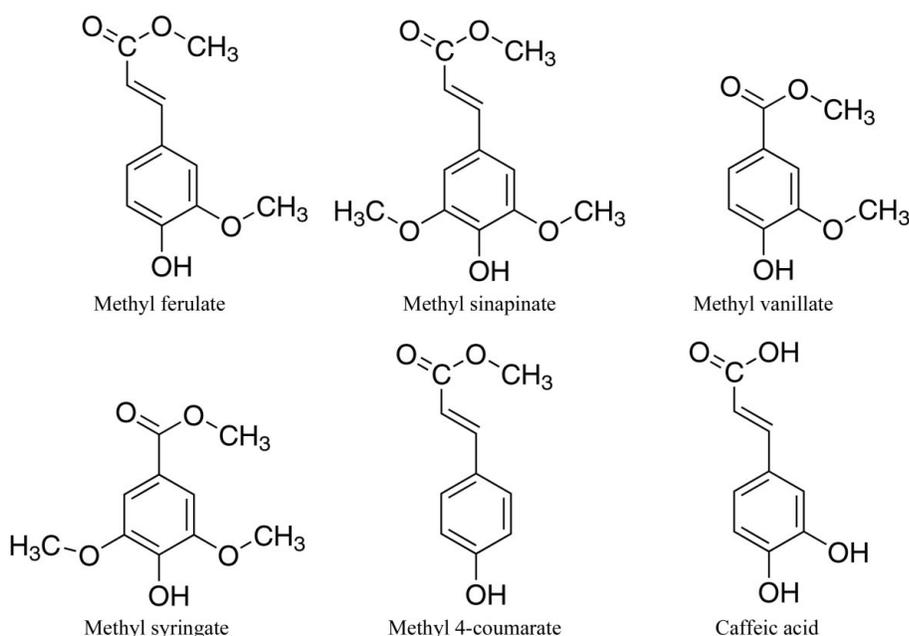
Figure 1
Structure of the ferulate esterase ligands described in this article.

and caffeic acid (similar to ferulic acid but with a hydroxy group in the *meta* position instead of a methoxy group) did not show electron density indicating bound ligand. This suggests that these compounds lack a relevant chemical group that is required for protein recognition

(Fig. 1). The structural data thus support the importance of the *m*-methyl group of the ferulic acid ring for binding, but unfortunately the post-translational modification of the catalytic serine in *E. coli*-derived non-mutant material does not allow us to test these ideas in solution. Interestingly, on the 'opposite' side of the ring the second methoxy group does not prevent binding in the crystal, as both methyl sinapinate and methyl syringate complexes show clear density in the active-site pocket. This would suggest that in addition to the ferulate-based substrates, sinapinate derivatives may also be cleaved by this enzyme. Such activity has recently been shown in solution enzyme assays of the *Sporotrichum thermophile* ferulate esterase (Topakas *et al.*, 2003). The fact that both methyl vanillate and methyl syringate also bind in this pocket also indicates that the length of the separation between the phenyl group and methyl ester in these molecules is not crucial, even if it may contribute to the correct orientation for catalysis; such ideas are further confirmed by observation of the terminal methyl moiety in the shorter substrates as opposed to the constant failure to observe this moiety in complexes with longer substrates.

Figure 2
Electron densities around the active site for the different complexes. The electron densities are maximum-likelihood-weighted $2F_{obs} - F_{calc}$ maps contoured at the 1σ level; electron densities are highlighted in brown for the substrates and in green for the protein. (*a*) Native active site showing the modified serine, (*b*) ferulic acid complex, (*c*) methyl sinapinate complex, (*d*) methyl syringate complex, (*e*) methyl vanillate complex. The figure was prepared using *BOBSCRIPT* (Esnouf, 1999).

4. Conclusion

In general, feruloyl esterases display broad substrate specificities, although the structural motifs that discriminate the various enzymatic targets remain to be elucidated. Here, we provide a first glimpse of the molecular determinants controlling the substrate specificity of the cellulosomal feruloyl esterase module of Xyn10B. It is suggested that the presence of at least one *m*-methoxy group in the substrate is required for binding, which explains why the enzyme cannot recognize methyl coumarate or caffeic acid. Considering the remarkable impact of feruloyl esterases in various industrial and medical applications, it is anticipated that these structural studies will provide an initial framework for the rational design of novel enzymes with improved biotechnological potential.

Figure 3
Superimposition of the active-site environment and substrates. Residues involved in movements are labelled as well as the catalytic triad. The methyl sinapinate complex is in orange, methyl vanillate in pink, methyl syringate in blue and the native uncomplexed structure in white. The figure was prepared using *PyMOL* (DeLano Scientific).

References

Bayer, E. A., Chanzy, H., Lamed, R. & Shoham, Y. (1998). *Curr. Opin. Struct. Biol.* **8**, 548–557.

Blum, D. L., Kataeva, I. A., Li, X. L. & Ljungdahl, L. G. (2000). *J. Bacteriol.* **182**, 1346–1351.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Esnouf, R. M. (1999). *Acta Cryst.* D**55**, 938–940.

Matthews, B. W., Sigler, P. B., Henderson, R. & Blow, D. M. (1967). *Nature (London)*, **214**, 652–656.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Perrakis, A., Harkiolaki, M., Wilson, K. S. & Lamzin, V. S. (2001). *Acta Cryst.* D**57**, 1445–1450.

Prates, J. A., Tarbouriech, N., Charnock, S. J., Fontes, C. M., Ferreira, L. M. & Davies, G. J. (2001). *Structure*, **9**, 1183–1190.

Topakas, E., Stamatis, H., Biely, P., Kekos, D., Macris, B. J. & Christakopoulos, P. (2003). *J. Biotechnol.* **102**, 33–44.