FEBS Letters

FEBS PRESS science publishing by scientists

RESEARCH LETTER

Screening of novel fungal Carbohydrate Esterase family 1 enzymes identifies three novel dual feruloyl/acetyl xylan esterases

Adiphol Dilokpimol[1,*], Bart Verkerk[1], Xinxin Li[1], Annie Bellemare[2], Mathieu Lavallee[2], Matthias Frommhagen[3], Emilie Nørmølle Underlin[3,4], Mirjam A. Kabel[3], Justin Powlowski[2], Adrian Tsang[2] and Ronald P. de Vries[1]

1 Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, The Netherlands
2 Centre for Functional and Structural Genomics, Concordia University, Montreal, Canada
3 Laboratory of Food Chemistry, Wageningen University and Research, The Netherlands
4 Department of Chemistry, Technical University of Denmark, Lyngby, Denmark

Correspondence
R. P. de Vries, Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands
Tel: +31 302122600
E-mail: r.devries@wi.knaw.nl

Present address
*Protein Production Team, VTT Technical Research Center of Finland Ltd., Espoo, Finland

(Received 27 December 2021, revised 9 February 2022, accepted 10 February 2022, available online 28 February 2022)

doi:10.1002/1873-3468.14322

Edited by Ulf-Ingo Flügge

Feruloyl esterases (FAEs) and acetyl xylan esterases (AXEs) are important enzymes for plant biomass degradation and are both present in Carbohydrate Esterase family 1 (CE1) of the Carbohydrate-Active enZymes database. In this study, ten novel fungal CE1 enzymes from different subfamilies were heterologously produced and screened for their activity towards model and complex plant biomass substrates. CE1_1 enzymes possess AXE activity, while CE1_5 enzymes showed FAE activity. Two enzymes from CE1_2 and one from CE1_5 possess dual feruloyl/acetyl xylan esterase (FXE) activity, showing expansion of substrate specificity. The new FXEs from CE1 can efficiently release both feruloyl and acetyl residues from feruloylated xylan, making them particularly interesting novel components of industrial enzyme cocktails for plant biomass degradation.

Keywords: acetyl xylan esterase; carbohydrate esterase family 1; feruloyl esterase; fungi; hydroxycinnamic acid; plant biomass

Many side streams from agricultural processing industries contain arabinoxylan as a major component. Arabinoxylan comprises a β-D-(1 → 4)-linked D-xylopyranosyl backbone, which is substituted with α-L-(1 → 2)- and/or α-L-(1 → 3)-linked L-arabinofuranosyl residues [1]. In commelinid monocots (e.g. wheat, rice and barley), L-arabinofuranosyl residues can be further substituted at the *O*-5 position by feruloyl (4-hydroxy-3-methoxycinnamoyl) and other hydroxycinnamoyl (e.g. *p*-coumaroyl) residues [2–5]. The xylan backbone can also be acetylated at the *O*-2 and/or *O*-3 positions depending on the types of plants.

Ferulic acid is also present in rhamnogalacturonan I (RGI) in pectin, which has a backbone of alternating α-L-(1 → 2)-rhamnose and α-D-(1 → 4)-galacturonic acid residues. The L-rhamnose residue can be substituted with α-L-(1 → 5)-arabinan or β-D-(1 → 4)-galactan chains, which both can contain terminal feruloyl residues [6–8]. Ferulic acid plays a role in the defence mechanism of plants against pathogens, due to its antimicrobial property, and the di-feruloyl cross-links between xylan chains increase the physical strength and integrity of plant cell walls [3,4]. Both acetylation and feruloylation inhibit the action of endo-acting plant cell wall-degrading enzymes and, therefore, hinder enzymatic saccharification for biomass valorization, for example, for bioethanol production and biorefinery processes.

Abbreviations
AXE, acetyl xylan esterase; FA, ferulic acid; FAE, feruloyl esterase; SF, subfamily.

FEBS Letters **596** (2022) 1932–1943 © 2022 The Authors. *FEBS Letters* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies

Acetyl xylan esterases (AXEs) [EC3.1.1.72] catalyse the hydrolysis of ester linkages between acetyl groups (in the form of acetylation) and xylan, which leads to the release of acetic acid. Feruloyl esterases (FAEs) [EC3.1.1.73] catalyse the de-esterification of agro-industrial side streams, which liberates ferulic acid and other plant phenolic acids [9–11]. AXEs and FAEs facilitate the degradation of complex plant cell wall polysaccharides by removing the ester bonds in plant polymers, providing access to glycoside hydrolases and polysaccharide lyases [12–14]. Apart from being used as accessory enzymes in the saccharification process, AXEs and FAEs are also being used as biocatalysts for the synthesis of a broad range of novel bioactive components for the food, cosmetics and pharmaceutical industries [10,15,16].

Based on the Carbohydrate-Active enZymes (CAZy) database (http://www.cazy.org [17]), fungal AXEs are primarily classified in Carbohydrate Esterase (CE) families CE1–CE6 and CE16 [18]. While a part of the fungal FAEs is grouped in CE1, most FAEs are not classified CAZymes [10,19]. Even though they target structurally different substrates, both FAEs and AXEs from CE1 share a high sequence similarity.

We recently classified the fungal members of CE1 into five subfamilies (CE1_1-5, Fig. 1, Table S1) based on phylogenetic analysis [20]. CE1_1 and CE1_2 are closely related and split from the same node. However, CE1_1 contains characterized AXEs, while CE1_2 contains characterized FAEs. CE1_4 and CE1_5 also split from a common node, of which CE1_5 contains characterized FAEs, whereas CE1_4 has only uncharacterized members. CE1_3 is the smallest branch, which contains only sequences from basidiomycetes with no predicted secreted signal peptide. CE1_3 is distantly related to the other fungal CE1 members (esterase, PHB depolymerase, IPR010126), but related to the esterase D/S-formylglutathione hydrolase sequences (IPR014186) according to InterPro classification [21]. The first report of an enzyme having both AXE and FAE activity was from *Aspergillus* [22], but as the gene was not identified, this enzyme was not assigned to a specific CAZy family at the time, although it was later classified to CE1. More recently, CE1 enzymes with dual AXE/FAE activity (in this study referred to as FXE) have been identified in the fungal CE1_2 subfamily [23] and from a bacterial metagenome [24]. In comparison with bacterial CE1 enzymes, fungal CE1 enzymes mainly belong to the Esterase_phb family based on the ESTHER database, a classification for proteins with an α/β-hydrolase fold [25], which is distantly related to bacterial enzymes belonging to Antigen85c family (< 20% sequence identity) [26]. Thus far, two crystal structures of fungal CE1 enzymes have been reported, that is, CE1_1 *Aspergillus luchuensis Al*AXEA (formerly *Aspergillus awamori*) (PDB: 5X6S [26]) and the anaerobic fungus *Anaeromyces mucronatus* AmCE1 (PDB: 5CXU [27]).

In this study, we characterized ten CE1 proteins from thermophilic fungal species covering CE1_1, 2 and 5 (Fig. 1, Table 1, Table S1) to obtain a better biochemical support for the fungal CE subfamilies. These enzymes were heterologously produced in an *Aspergillus niger* strain derived from N400 (CBS 120.49, ATCC 9029, FGSC A1143 and NRRL3) and evaluated using both simple model and complex plant biomass substrates to determine their activity and substrate preference. Interestingly, this not only results in the identification of another FXE in CE1_2, but also in CE1_5, suggesting that the dual activity is not restricted to a single subfamily. Our results provide insights into the differences in activity of fungal CE1 enzymes and shed light on the relationship between these CE1 FAEs and other fungal FAEs.

Materials and methods

Materials

Methyl ferulate, methyl *p*-coumarate, methyl sinapate and methyl caffeate were obtained from Apin chemicals (Abingdon, United Kingdom), while *p*-nitrophenyl ferulate was from Taros Chemicals (Dortmund, Germany). The acetylated oligosaccharide mix from corn fibre oligosaccharides was provided by Dr. Mirjam Kabel, Wageningen University [2,28]. Detailed analysis of this substrate was recently published (referred to as CFoligo) [29]. Insoluble wheat arabinoxylan (P-WAXYI, from wheat flour), endo-(1 → 5)-α-arabinanase (E-EARAB, GH43 from *Aspergillus niger*) and endo-(1 → 4)-β-galactanase (E-EGALN, GH53 from *A. niger*) were from Megazyme (Bray, Ireland). Sugar beet pectin (Pectin Betapec RU301) was from Herbstreith & Fox KG (Neuenbürg, Germany). The endo-β-(1 → 4)-xylanase (GH11 from *Thermomyces lanuginosus*) and other chemicals were from Sigma-Aldrich (Merck KGaA, Darmstadt, Germany).

Bioinformatics

Genome mining and phylogenetic analysis were performed based on [10]. Signal peptides were predicted using SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) [30]. Manual gene model correction of selected sequences was performed based on BlastX to identify and remove putative introns. A multiple sequence alignment was performed using T-COFFEE Multiple Sequence Alignment Server (http://tcoffee.crg.cat/) [31] and visualized using Easy Sequencing

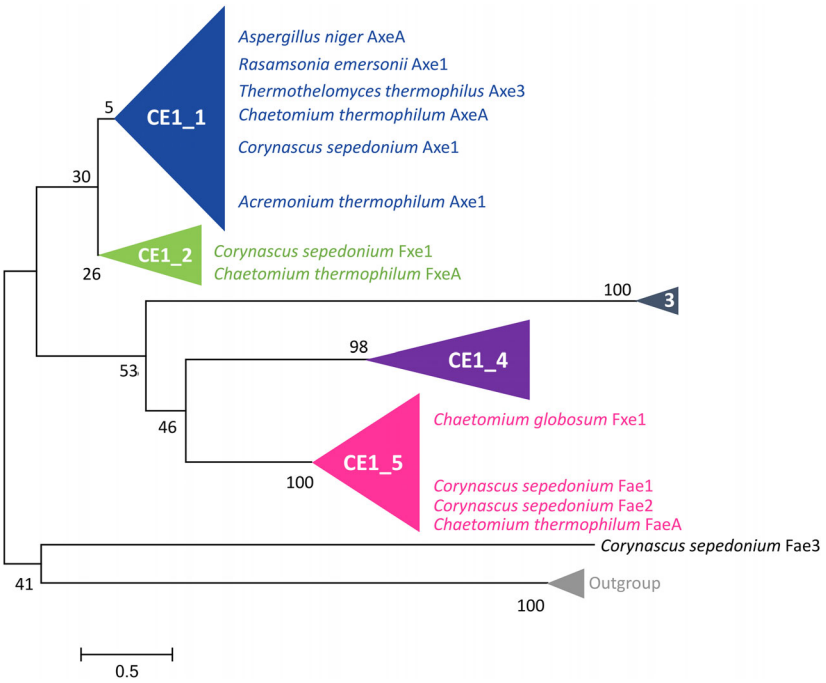

Fig. 1. Phylogenetic relationships among the fungal CE1 members. The phylogenetic analysis was based on amino acid sequences (modified from [20]). The evolutionary history was inferred using the Neighbour-Joining method with a bootstrap value of 500. Eight non-CE1 FAEs from SF7 [10] were used as an outgroup.

in Postscript (http://espript.ibcp.fr/ESPript/ESPript/) [32]. Phylogenetic analysis was performed using the Molecular Evolutionary Genetic Analysis programme (MEGA7) [33] with the Neighbour-Joining method with a bootstrap value of 500. Theoretical molecular masses and p*I* were calculated by the ExPASy–ProtParam tool (http://www.expasy.ch/tools/protparam.html) [34].

The homology protein models were created by using homology detection and structure prediction by HMM–HMM comparison (HHPRED, https://toolkit.tuebingen.mpg.de/#/) [35]. The models were validated with different online tools, that is, protein model check (https://swift.cmbi.umcn.nl/servers/html/modcheck.html), PSI-blast-based secondary structure PREDiction (PSIPRED, http://bioinf.cs.ucl.ac.uk/psipred/) [36] and Protein Quality Predictor (ProQ, https://proq.bioinfo.se/ProQ/ProQ.html) [37]. Protein models were visualized using a molecular visualization system PyMOL (DeLano Scientific, Schrödinger, LLC). Autodock (http://autodock.scripps.edu/) [38] was used for the prediction of the putative substrate-binding sites.

Cloning of CE1 genes and production of CE1 enzymes

Carbohydrate Esterase family 1 sequences were obtained from our genome portal (http://genome.fungalgenomics.ca). Gene sequences were amplified by PCR using genomic DNA or cDNA as the template DNA. Genomic DNA was extracted using the DNeasy® Plant Mini kit (QIAGEN, Hilden, Germany) as previously described [39], and cDNA was obtained following the previously described method [40]. The genes were cloned using an in-house variation of the ligation-independent cloning (LIC) method [41].

All genes were cloned into the ANIp7 shuttle vector [42], which uses the *As. niger* glucoamylase promoter to drive recombinant gene expression. Positive plasmids were selected based on their resistance to ampicillin and maintained in *Escherichia coli* strain DH5α using standard protocols. Protoplasts of *A. niger* strain CSFG_6005 (N593 *glaA*::*hisG*) were transformed [43] with the recombinant plasmids carrying CE1 genes. Transformants were selected on minimal medium without uracil and uridine [44]. Supernatants from transformants were screened for recombinant protein production after growth in liquid minimal MMJ medium [45] for induction of protein production.

For protein production, a standard Petri dish containing 25 mL of MMJ medium was inoculated with positive transformants to a final concentration of 1×10^6 conidia·mL^{-1}, and the culture was grown for 5 days at 30 °C without agitation. Culture supernatant was filtered through a 0.45-μm membrane filter (Whatman, Kent, UK), and the filtrate (17.2 mL) was then concentrated using a 10 kDa cut-off Jumbosep centrifugal filter unit (Pall). The concentrated culture medium was replaced by 10 mM sodium acetate buffer including 0.1 M NaCl (pH 5.0) by repeated dilution and concentration (to 2 mL) using the 10 kDa cut-off centrifugal filter unit and then applied to a SuperdexTM 75 FPLC gel filtration column (1 cm × 30 cm; Amersham Biosciences, Freiburg, Germany) equilibrated with the same buffer. Protein concentrations were assessed from SDS/PAGE gels by densitometric method using IMAGEJ program [46] with bovine serum albumin as a standard), Fig S1.

Table 1. Accession number, subfamily, molecular mass and production level of characterized CE1 enzymes in this study.

Fungal species	CE1[a] subfamily	FAE[b] subfamily	UniProt[c] entry name	Enzyme name	Calculated molecular mass (kDa)	Apparent molecular mass (kDa)	Calculated p*I*	Production ($mg \cdot mL^{-1}$)[d]	References
Aspergillus niger	CE1_1	-	AXE1A_ASPNG	AxeA[g]	30.1	30	4.82	1.42	[47, 48]
Rasamsonia emersonii	CE1_1	-	AXE1A_RASEM	Axe1	30.1	30	4.70	0.60	This study
Thermothelomyces thermophilus[e]	CE1_1	-	FXE1A_MYCTH[f]	Axe3[h]	31.6	30	6.21	0.68	This study
Chaetomium thermophilum	CE1_1	-	AXE1A_CHATH[g]	AxeA	37.6	39	7.62	0.40	[49]
Corynascus sepedonium[f]	CE1_1	-	AXE1B_MYCSE	Axe1	32.5	35	4.57	0.40	This study
Acremonium thermophilum	CE1_1	-	AXE1A_ACRTH	Axe1	29.8	40	6.02	0.13	This study
Corynascus sepedonium[f]	CE1_2	SF6	FAE1B_MYCSE	Fxe1	29.3	30	4.53	0.07	This study
Chaetomium thermophilum	CE1_2	SF6	AXE1B_CHATH	FxeA	29.7	60	7.27	0.13	This study
Chaetomium globosum	CE1_5	SF5	FXE1A_CHAGL	Fxe1	28.1	30	6.22	0.13	This study
Corynascus sepedonium[f]	CE1_5	SF5	FAE1C_MYCSE	Fae1	26.4	30	5.39	0.13	This study
Corynascus sepedonium[f]	CE1_5	SF5	FAE1E_MYCSE	Fae2	32.5	40	7.15	0.25	This study
Chaetomium thermophilum	CE1_5	SF5	FAE1B_CHATH	FaeA	34.0	45	7.66	0.25	This study
Corynascus sepedonium[f]	Non-CE1	SF2	FAE1F_MYCSE	Fae3	55.1	60	4.97	0.03	This study

[a]Based on Fig. 1.; [b]Based on [10].; [c]Based on UniProtKB/Swiss-Prot entry name recommendation (https://www.uniprot.org/help/entry_name).; [d]Based on Figure A2.; [e]Formerly *Myceliophthora thermophila*.; [f]Formerly *Myceliophthora sepedonium*.; [g]Also referred to as Axe; *An*AXE; *An*AxeA; *An*AXE1.; [h]Trp160 in *Al*AXEA (PDB: 5X6S [25]) corresponds to a Trp in FXE3 from *Thermothelomyces thermophilus* ATCC 42464 (this study), but corresponds to an Arg in *Mt*AXE3 from *Thermothelomyces thermophilus* C1 [56].

Reference enzymes were produced in previous studies [10,23].

Enzyme activity assay using p-nitrophenyl substrates

The activity assays of the enzymes towards the *p*-nitrophenyl substrates were performed in 250 µL total reaction volume per sample as described previously [47]. The reactions (225 µl) contained 0.8 mM *p*-nitrophenyl acetate (Sigma-Aldrich, Zwijndrecht, Netherlands) or ferulate (dissolved in dimethyl sulfoxide), 100 mM sodium phosphate buffer, pH 7 containing 2.5% triton-X. In total, 25 µL culture supernatant was added to the reaction mixture and incubated at 30 °C. The release of *p*-nitrophenol (*p*NP) was spectrophotometrically quantified by following the absorbance at 410 nm for 30 min with a 2-min interval and calculated as described previously [47]. pH and temperature profiles were performed as described previously [23]. The culture supernatant of *A. niger* harbouring ANIp7 plasmid without insert was used as a negative control. All assays were performed in triplicate. The amount of released *p*NP was determined using 0.02–2.0 mM *p*NP as a standard curve. One unit of enzyme activity was defined as the amount of enzyme that released 1 µmol of *p*NP from *p*NP-substrate per min under the above-mentioned assay condition.

Enzyme activity assay using methyl hydroxycinnamates

Activity towards methyl hydroxycinnamate substrates (methyl caffeate, methyl ferulate, methyl *p*-coumarate and methyl sinapate) was assayed in 250 µL reaction mixtures as described previously [23] at 30 °C for 5–30 min. Detection of substrate reduction was performed at 340 nm with a 2-min interval. The activity was determined from the standard curves of the substrates (0.001–0.5 mM). The culture supernatant of *A. niger* harbouring ANIp7 plasmid without insert was used as a negative control. All assays were performed in triplicate. The standard curves of 5–250 µM methyl substrates were used to calculate the amount of hydrolysed methyl substrates. One unit of enzyme activity was defined as the amount of enzyme that released 1 µmol of hydroxycinnamic acid from methyl substrates per min under the assay condition.

Hydrolytic activity towards polysaccharide substrates

The activity towards polysaccharides was determined using acetylated corn fibre oligosaccharides [2] and insoluble wheat arabinoxylan for AXE activity and insoluble wheat arabinoxylan and sugar beet pectin for FAE activity. To pre-treat the substrates, 1% (w/v) polysaccharide substrate, 50 mM McIlvian buffer pH 4.5, 0.02% sodium azide and 100 µg endo-β-(1 → 4)-xylanase (for insoluble wheat arabinoxylan), or 100 µg endo-α-(1 → 5)-arabinanase and 100 µg endo-β-(1 → 4)-galactanase (for sugar beet pectin) were incubated at 30 °C for 72 h, followed by heat inactivation at 95 °C for 10 min to terminate the reaction. Then, 1 µg of CE1 enzyme was added to the pre-treated substrates and incubated at 30 °C for 24 h with shaking at 100 r.p.m. (total reaction volume: 600 µL). Reactions were terminated by heating at 95 °C for 10 min followed by centrifugation for 15 min at 4 °C. For ferulic acid content analysis, 200 µL of supernatant was mixed with 600 µL 100% acetonitrile (1:3, v/v). For acetic acid content analysis, the reaction supernatants were used directly for HPLC analysis, except for the reactions containing acetylated corn fibre oligosaccharides, which were terminated by addition of 50 µL 2 M HCl prior to HPLC analysis.

Ferulic acid content analysis

Release of ferulic acid, *p*-coumaric acid and diferulic acid was monitored by HPLC system (Dionex ICS-5000+ chromatography system; Thermo Scientific, Sunnyvale, CA) equipped with an Acclaim Mixed-Mode WAX-1 column (3 × 150 mm; Thermo Scientific) and a UV detector (310 nm; Thermo Scientific). The chromatographic separation was carried out as previously described [23] using an isocratic elution of a 25 mM potassium phosphate buffer, 0.8 mM pyrophosphate pH 6.0 in acetonitrile 50% (v/v) with a flow rate of 0.25 $mL \cdot min^{-1}$ at 30 °C. 0.25–50 µM ferulic and *p*-coumaric acids were used as standards for identification and quantitation.

Acetic acid content analysis

Release of acetic acid was monitored by an HPLC system (Dionex ICS-3000 chromatography system; Thermo Scientific) equipped with an Aminex HPX 87H column with Guard-column (300 × 7.8 mm; Bio-Rad, Hercules, CA), a refractive index detector and an UV detector (UV 210 nm, Bio-Rad). An isocratic elution comprising 5.0 mM sulfuric acid in MillQ water was applied. The flow rate was 0.6 $mL \cdot min^{-1}$ at 40 °C was used. 0.01–2.0 $mg \cdot mL^{-1}$ of acetic acid was used as a standard for identification and quantification.

Results and discussion

Screening of CE1 enzymes using different simple model substrates

In this study, we selected ten previously uncharacterized CE1 enzymes proteins from thermophilic fungal species, that is, *Thermothelomyces thermophilus*, *Chaetomium thermophilum*, *Chaetomium globosum*,

Corynascus sepedonium, *Rasamsonia emersonii* and *Acremonium thermophilum* [23] for characterization to provide additional support for functional diversification of fungal CE1 subfamilies. Four of these enzymes belonged to CE1_1, two to CE1_2 and four to CE1_5 (Table 1, Fig. 1, Table S1). When possible, we used multiple enzymes from the same species to also enable comparison within a fungal species. We also included two characterized fungal CE1 enzymes from *Aspergillus niger* (AxeA) [48,49] and *Chaetomium thermophilum* (AxeA) [50], and an enzyme that did not belong to any CE1 subfamily, but was a tannase-related FAE candidate (FAE_SF2 in [10]) to validate the substrate preference. Heterologous production of CE1 enzymes was performed using *A. niger* strain CSFG_6005 as a host [45]. The protein concentration was estimated using a densitometric method, showing that *A. niger* AxeA was produced at the highest level (Table 1, Fig. S1). Most CE1_1 enzymes efficiently hydrolysed *p*NP-acetate, but not *p*NP-ferulate nor four methyl hydroxycinnamates (methyl caffeate, methyl ferulate, methyl *p*-coumarate and methyl sinapate), indicating that they possess AXE activity (Table 2).

Despite sharing over 70% amino acid sequence identity, the two CE1_2 hydrolysed different substrates. The enzyme from *Co. sepedonium* in CE1_2 hydrolysed *p*NP-acetate and the four methyl hydroxycinnamates, whereas the one from *Ch. thermophilum* hydrolysed *p*NP-acetate and *p*NP-ferulate. Their different substrate specificities were confirmed using plant biomass substrates (see below). Three of the CE1_5 enzymes hydrolysed all four methyl hydroxycinnamate substrates, while *Co. sepedonium* Fae1 only hydrolysed methyl ferulate and methyl sinapate. This is quite distinct because the members from this subfamily usually show a broad substrate specificity, in particular those from *Th. thermophilus* (formerly *Myceliophthora thermophila* and *Chrysosporium lucknowense*) [51,52]. The non-CE1 member, *Co. sepedonium* Fae3, could also hydrolyse all four hydroxycinnamate substrates and showed a high substrate preference for methyl ferulate. In addition, *Co. sepedonium* Fae3 did not possess tannase activity as it did not show activity towards methyl gallate (data not shown). All selected enzymes showed the highest activity at pH 7 with a temperature optimum of 30–40 °C, even though they originated from thermophilic species. As all enzymes could hydrolyse *p*NP-acetate, our results confirm a previous report that the use of *p*NP-acetate alone is insufficient for the identification of true AXE activity [53].

The three subfamilies show different specificity towards plant biomass substrates

Acetylated corn fibre oligosaccharides and insoluble wheat arabinoxylan were used to determine AXE activity, whereas insoluble wheat arabinoxylan and sugar beet pectin were used to determine FAE activity. Because of the limited availability of acetylated corn

Table 2. Activity towards model substrates for AXE and FAE activities.

		Specific activity (mU·mg^{-1}) [a,b,c]					
Fungi	Enzyme name	*p*NP-acetate	*p*NP-ferulate	Methyl *p*-coumarate	Methyl caffeate	Methyl ferulate	Methyl sinapate
Aspergillus niger	AxeA	32395	N.A.	N.A.	N.A.	N.A.	N.A.
Rasamsonia emersonii	Axe1	32157	N.A.	N.A.	N.A.	N.A.	N.A.
Thermothelomyces thermophilus	Axe3	107	N.A.	N.A.	N.A.	N.A.	N.A.
Chaetomium thermophilum	AxeA	5277	N.A.	N.A.	N.A.	N.A.	N.A.
Corynascus sepedonium	Axe1	3511	N.A.	N.A.	N.A.	N.A.	N.A.
Acremonium thermophilum	Axe1	1616	N.A.	N.A.	N.A.	N.A.	N.A.
Corynascus sepedonium	Fxe1	232	N.A.	336 [77%]	200 [46%]	439 [100%]	128 [29%]
Chaetomium thermophilum	FxeA	1205	34	N.A.	N.A.	N.A.	N.A.
Chaetomium globosum	Fxe1	1331	47	135 [100%]	75 [55%]	93 [69%]	63 [47%]
Corynascus sepedonium	Fae1	512	N.A.	N.A. [0%]	N.A. [0%]	50 [100%]	44 [87%]
Corynascus sepedonium	Fae2	230	N.A.	51 [73%]	34 [48%]	70 [100%]	64 [91%]
Chaetomium thermophilum	FaeA	696	5009	150 [16%]	62 [7%]	930 [100%]	140 [15%]
Corynascus sepedonium	Fae3	764	3133	1790 [10%]	734 [4%]	18 771 [100%]	488 [3%]

[a]N.A., no activity detected.; [b]One unit of enzyme activity is defined as the amount of enzyme releasing 1 μmol of *p*-nitrophenol from *p*NP-acetate or *p*NP-ferulate, or releasing 1 μmol of hydroxycinnamic acid from methyl substrates per min under assay conditions. Active indicates the enzyme was active, but the specific activity could not be calculated.; [c]Number in square bracket indicates the preference ratio among four methyl substrates for FAE activity, calculated as a percentage of the highest activity for each enzyme that was set to 100%.

fibre oligosaccharides, these were not used to determine ferulic acid release, even though they also contain feruloyl residues. All CE1_1 enzymes released acetyl residues from insoluble wheat arabinoxylan, and most of them also released acetyl residues from corn fibre oligosaccharides (Fig. 2A, B). *Ac. thermophilum* Axe1 released 100% acetyl residues from corn oligosaccharides, while *As. niger* AxeA released up to 70% acetyl residues from wheat arabinoxylan, which indicate that they were the most active enzymes on these substrates. In addition, the two enzymes from CE1_2 and an enzyme from *Ch. globosum* in CE1_5 also released more than 25% acetyl residues from wheat arabinoxylan. This demonstrates that the occurrence of enzymes with dual activity occurs not just in CE1_2, as reported previously [23], and in the bacterial clade [24], but also in a second fungal subfamily (CE1_5). This may imply that the switch between AXE and FAE activity occurs commonly in fungal CE1 enzymes and may be due to the composition of the substrates in the biotope of different fungal species that pushes the activity to either AXE or FAE or both (FXE).

The CE1_2 and 5 enzymes released between 30% and 60%, respectively, of the feruloyl residues from wheat arabinoxylan. However, they could not release feruloyl residues from sugar beet pectin, except for *Ch. thermophilum* FxeA which could release around 15% of the feruloyl residues in this substrate (Fig. 2C, D). The CE1_1 enzymes mildly released feruloyl residues from wheat arabinoxylan (less than 10%) and did not release feruloyl residues from sugar beet pectin. In contrast, the tannase-related *Co. sepedonium* Fae3, which does not group with any of the CE1 subfamilies, released almost 100% of the feruloyl residues from sugar beet pectin, whereas it only released 15% from wheat arabinoxylan. These results show a clear substrate preference of CE1 enzymes for arabinoxylan. The tannase-related FAE has a higher specificity for pectin, which confirms previous reports on such enzymes [54,55].

Both *Co. sepedonium* Fxe1 and *Ch. thermophilum* FxeA from CE1_2 and *Ch. globosum* Fxe1 from CE1_5 possess dual feruloyl/acetyl xylan esterase (FXE) activity, since they efficiently release feruloyl and acetyl residues from arabinoxylan to the same extent. These dual activity enzymes are not only

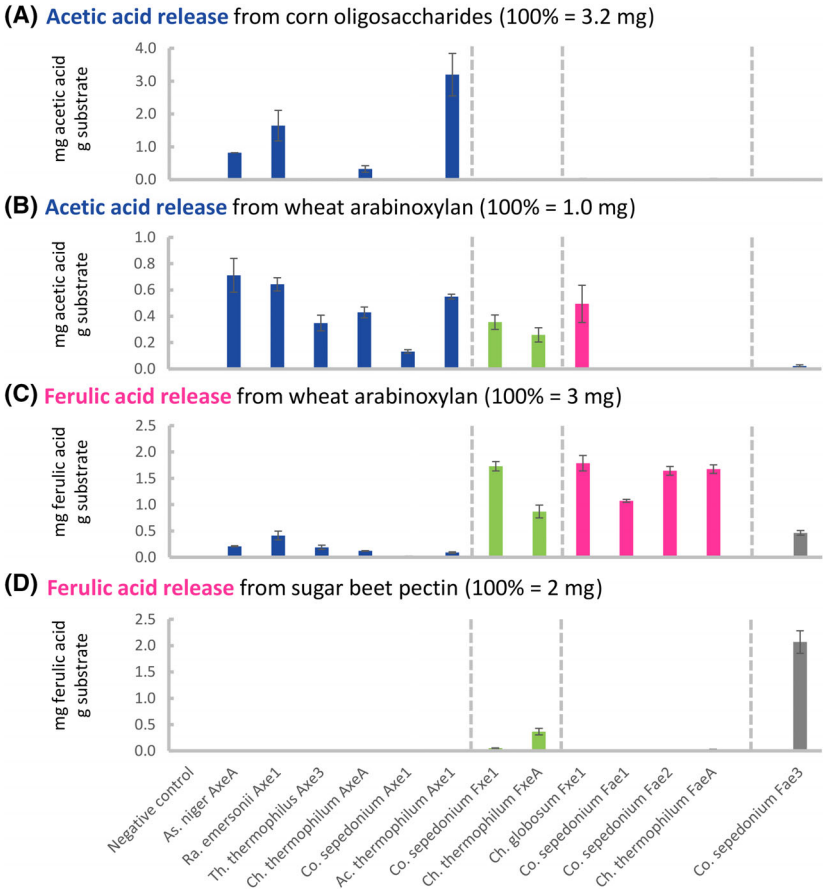

Fig. 2. Release of acetic and ferulic acids from plant biomass substrates by CE1 enzymes. (A) Acetic acid release from corn fibre oligosaccharides (total acetic acid, 3.2 mg). (B) Acetic acid release from insoluble wheat arabinoxylan (total acetic acid, 1.0 mg). (C) Ferulic acid release from insoluble wheat arabinoxylan (total ferulic acid, 3.0 mg). (D) Ferulic acid release from sugar beet pectin (total ferulic acid, 2.0 mg). The substrates were pre-treated with endo-β-(1 → 4)-xylanase (for wheat arabinoxylan) or endo-(1 → 5)-α-arabinanase and endo-(1 → 4)-β-galactanase (for sugar beet pectin). Error bars represent standard deviation of two replicate measurements.

interesting in their removal of side residues for the valorization of plant cell wall polysaccharide, but may also be interesting for enzymatic synthesis for valuable compounds such as novel antioxidants through transesterification [56], because of their expanded substrate specificity which can accommodate different chemical structures. Identifying the amino acids responsible for the expanded substrate specificity would allow us to unlock the full potential of these enzymes.

Homology modelling suggests amino acid candidates influencing the substrate specificity

To verify the catalytic- and substrate-binding sites of fungal CE1 members, we created homology models of the selected CE1 enzymes (Figs. 3 & 4, Table S2) using 5X6S (*As. luchuensis* AXEA) as a template [26]. Based on the Protein Quality Predictor [37], the homology models for CE1_1 and 2 are of very good quality (> 80% sequence similarity). However, the

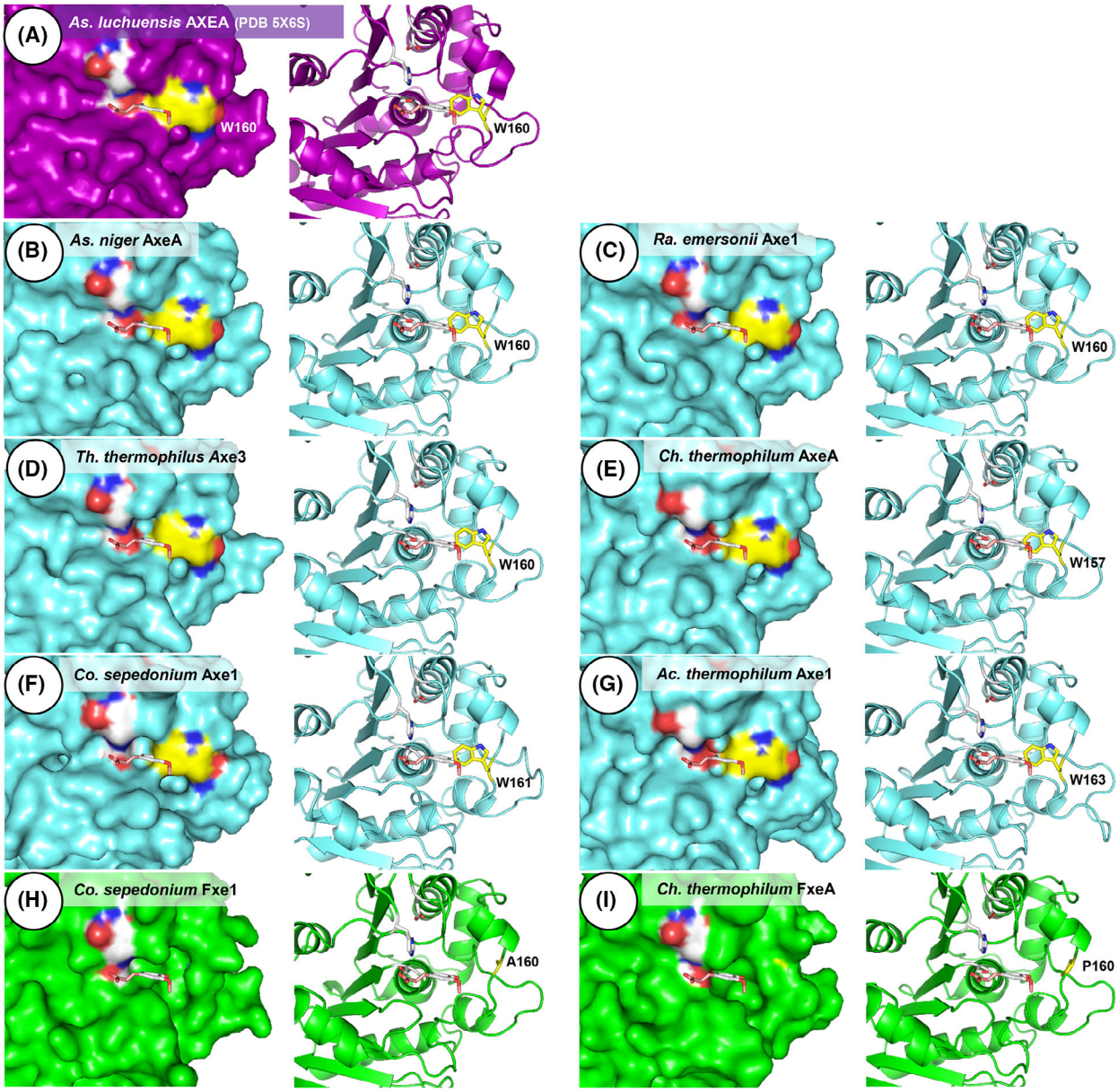

Fig. 3. 3D diagram representing structure of *As. luchuensis* AXEA and models of CE1_1 and 2 enzymes. The left panel shows the molecular surface, and the right panel shows a ribbon diagram. (A) Crystal structure of *Al*AXEA (PDB: 5X6S [26]). (B-G) Models of CE1_1 enzymes. (H-I) CE1_2 enzymes. The catalytic triad is shown in white, and Trp160 in *Al*AXEA and the corresponding amino acid in other enzymes are shown in yellow. Ferulic acid (shown as a stick) binding at the active site was predicted using AutoDock [38].

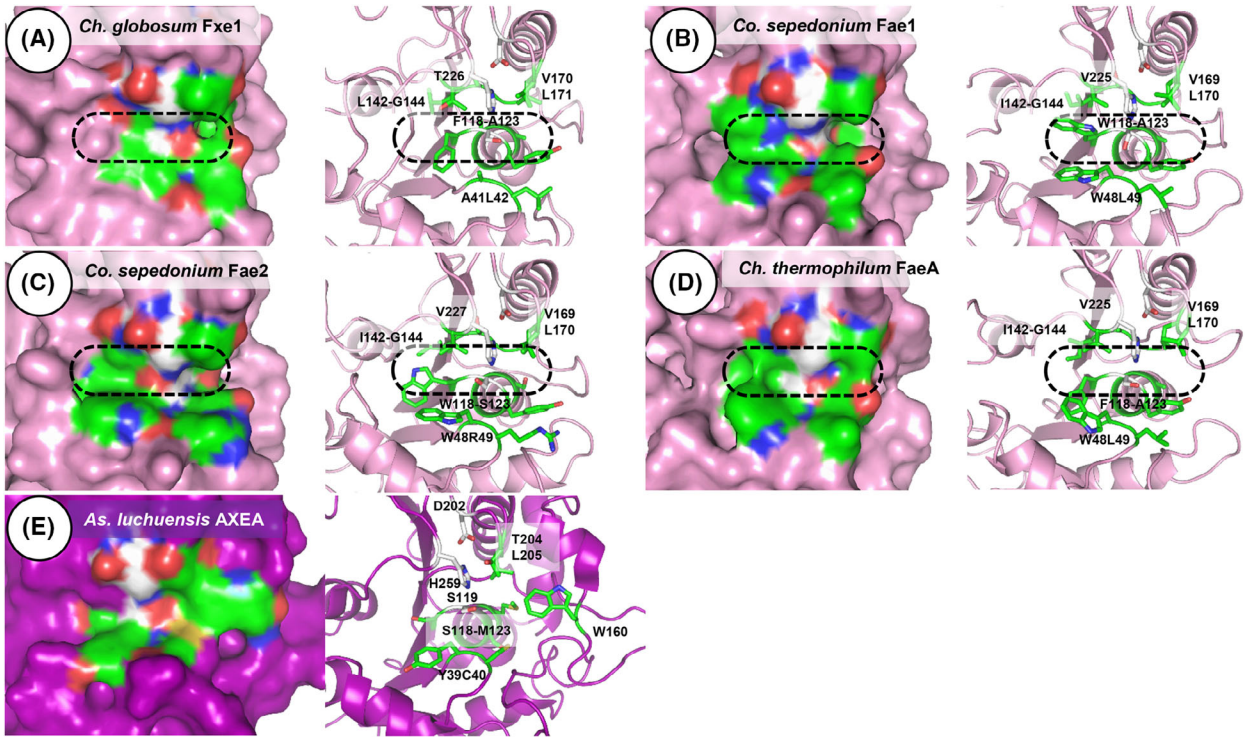

Fig. 4. 3D diagram representing structure of *As. luchuensis* AXEA and models of CE1_5 enzymes. The left panel shows the molecular surface, and the right panel shows ribbon diagram. (A-D) Models of CE1_5 enzymes. (E) Crystal structure of *Al*AXEA (PDB: 5X6S [26]). The catalytic triad is shown in white; amino acid residues within 5 Å of catalytic Ser and His (representing putative substrate-binding site) are shown in green. Black dash indicates possible substrate-binding site.

models from CE1_5 are of lower quality because of the low similarity (< 35% sequence similarity) to the reported structures (5X6S and 5CXU); hence, we only used these to locate the catalytic- and possible substrate-binding sites (Fig. 4). The multiple sequence alignment with conserved and unique regions among the CE1 subfamilies is shown in Fig. S2.

Recently, it has been shown that Trp160 in *As. luchuensis* AXEA (belonging to CE1_1) determines the substrate specificity, and replacing it with Ala, Gly or Pro could expand the substrate specificity towards FAE model substrates [26]. The corresponding amino acids in CE1_2 can be Ala, Asn, Gln, Pro, Ser and Thr instead of Trp160 [57]. These smaller amino acids allow the accommodation of larger molecules, for example, ferulic acid at the catalytic site (Fig. 3H-J). In case of CE1_5, two Trp residues (Trp142 or Trp144) located adjacent to the catalytic Ser143 in *Aspergillus oryzae* (AoFaeD) determine the FAE activity towards the methyl hydroxycinnamate substrates [58]. Replacing Trp with Phe or Tyr (Trp142Phe and Trp144Tyr, respectively) also improved its hydrolytic activity towards arabinoxylan. All our selected CE1_5 enzymes have Tyr instead of Trp144 in AoFaeD, and two also have Phe instead of Trp142 supporting their activity towards methyl hydroxycinnamate substrates. While the other selected CE1_5 enzymes have Val, *Ch. globosum* Fxe1 from CE1_5 has Thr226 which may interact with the acetyl residue contributing to its dual FXE activity (Fig. 4). Future studies in which these residues are mutated will evaluate their involvement in the diverse substrate specificity of these enzymes.

Conclusion

Fungal CE1 members are highly diverse, and a few amino acid changes may lead to different substrate specificity. CE1_1 enzymes possess AXE activity, while CE1_5 enzymes possess FAE activity with broad substrate specificity. The CE1_2 members, which may have evolved from CE1_1 or *vice versa*, possess dual feruloyl/acetyl xylan esterase activity. In addition, one CE1_5 also possesses dual activity, which could be caused by variation in one amino acid close to the catalytic site. Our new CE1 enzymes could be useful for various biotechnological applications, especially for the enzymes with dual activity that efficiently release

both ferulic acid and acetic acid from the arabinoxylan.

Acknowledgements

This work was supported by Genome Canada and Génome Québec. This work was partly supported by the China Scholarship Council to XL (grant no: 201803250066). The funding sources had no role in study design, in the collection, analysis and interpretation of data, in the writing of the manuscript, nor in the decision to submit the article for publication.

Conflict of interest

The authors declare no conflicts of interests.

Author contributions

AD contributed to investigation, formal analysis and writing—original draft. BV, AB, ML, MF and ENU contributed to investigation and formal analysis. XL contributed to formal analysis and writing—review and editing. AT, MAK and JP contributed to conceptualization, supervision, visualization, funding acquisition, project administration, resources and writing—review and editing. RPdV contributed to conceptualization, supervision, funding acquisition, project administration, resources and writing—review and editing.

Data accessibility

Data of this study that are not included in the manuscript are available in the supplemental files. Previously published data used in this study is referred to in the text.

References

1 Kabel MA, de Waard P, Schols HA, Voragen AGJ. Location of *O*-acetyl substituents in xylo-oligosaccharides obtained from hydrothermally treated *Eucalyptus* wood. *Carbohyd Res*. 2003;**338**:69–77.
2 Appeldoorn MM, de Waard P, Kabel MA, Gruppen H, Schols HA. Enzyme resistant feruloylated xylooligomer analogues from thermochemically treated corn fiber contain large side chains, ethyl glycosides and novel sites of acetylation. *Carbohyd Res*. 2013;**381**:33–42.
3 de O. buanafina MM. Feruloylation in grasses: current and future perspectives. *Molecular Plant*. 2009;**2**:861–72.
4 de Oliveira DM. Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. *Plant Biotechnol J*. 2015;**13**:1224–32.
5 Harris PJ, Trethewey JAK. The distribution of ester-linked ferulic acid in the cell walls of angiosperms. *Phytochem Rev*. 2010;**9**:19–33.
6 Levigne S, Ralet MC, Quemener B, Thibault JF. Isolation of diferulic bridges ester-linked to arabinan in sugar beet cell walls. *Carbohyd Res*. 2004;**339**:2315–9.
7 Levigne SV, Ralet MCJ, Quemener BC, Pollet BNL, Lapierre C, Thibault JFJ. Isolation from sugar beet cell walls of arabinan oligosaccharides esterified by two ferulic acid monomers. *Plant Physiol*. 2004;**134**:1173–80.
8 Ralet MC, Thibault JF, Faulds CB, Williamson G. Isolation and purification of feruloylated oligosaccharides from cell walls of sugar-beet pulp. *Carbohyd Res*. 1994;**263**:227–41.
9 de Vries RP, Visser J. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiol Mol Biol Rev*. 2001;**65**:497–522.
10 Dilokpimol A, Mäkelä MR, Aguilar-Pontes MV, Benoit-Gelber I, Hildén KS, de Vries RP. Diversity of fungal feruloyl esterases: updated phylogenetic classification, properties, and industrial applications. *Biotechnol Biofuels*. 2016;**9**:231.
11 Wong DWS. Feruloyl esterase - A key enzyme in biomass degradation. *Appl Biochem Biotechnol*. 2006;**133**:87–112.
12 de Vries RP, Kester HCM, Poulsen CH, Benen JAE, Visser J. Synergy between enzymes from *Aspergillus* involved in the degradation of plant cell wall polysaccharides. *Carbohyd Res*. 2000;**327**:401–10.
13 Gottschalk LMF, Oliveira RA, Bon EPDS. Cellulases, xylanases, β-glucosidase and ferulic acid esterase produced by *Trichoderma* and *Aspergillus* act synergistically in the hydrolysis of sugarcane bagasse. *Biochem Eng J*. 2010;**51**:72–8.
14 Tabka MG, Herpoël-Gimbert I, Monod F, Asther M, Sigoillot JC. Enzymatic saccharification of wheat straw for bioethanol production by a combined cellulase xylanase and feruloyl esterase treatment. *Enzyme Microbial Technol*. 2006;**39**:897–902.
15 Razeq FM, Jurak E, Stogios PJ, Yan R, Tenkanen M, Kabel MA, et al. A novel acetyl xylan esterase enabling complete deacetylation of substituted xylans. *Biotechnol Biofuels*. 2018;**11**:74.
16 Topakas E, Vafiadi C, Christakopoulos P. Microbial production, characterization and applications of feruloyl esterases. *Process Biochem*. 2007;**42**:497–509.
17 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2013;**42**: D490–5.
18 Adesioye FA, Makhalanyane TP, Biely P, Cowan DA. Phylogeny, classification and metagenomic bioprospecting of microbial acetyl xylan esterases. *Enzyme Microbial Technol*. 2016;**93–94**:79–91.

19 Udatha DBRKG, Kouskoumvekaki I, Olsson L, Panagiotou G. The interplay of descriptor-based computational analysis with pharmacophore modeling builds the basis for a novel classification scheme for feruloyl esterases. *Biotechnol Adv*. 2011;**29**:94–110.
20 Mäkelä MR, Dilokpimol A, Koskela SM, Kuuskeri J, de Vries RP, Hildén K. Characterization of a feruloyl esterase from *Aspergillus terreus* facilitates the division of fungal enzymes from Carbohydrate Esterase family 1 of the carbohydrate-active enzymes (CAZy) database. *Microb Biotechnol*. 2018;**11**:869–80.
21 Mitchell AL. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res*. 2018;**47**:D351–60.
22 Tenkanen M, Schuseil J, Puls J, Poutanen K. Production, purification and characterisation of an esterase liberating phenolic acids from lignocellulosics. *J Biotechnol*. 1991;**18**:69–84.
23 Li X, Griffin K, Langeveld S, Frommhagen M, Underlin EN, Kabel MA, et al. Functional validation of two fungal subfamilies in carbohydrate esterase family 1 by biochemical characterization of esterases from uncharacterized branches. *Front Bioeng Biotechnol*. 2020;**8**:694.
24 Hameleers L, Penttinen L, Ikonen M, Jaillot L, Fauré R, Terrapon N, et al. Polysaccharide utilization loci-driven enzyme discovery reveals BD-FAE: a bifunctional feruloyl and acetyl xylan esterase active on complex natural xylans. *Biotechnol Biofuels*. 2021;**14**:127.
25 Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A. ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. *Nucleic Acids Res*. 2013;**41**:D423–9.
26 Komiya D, Hori A, Ishida T, Igarashi K, Samejima M, Koseki T, et al. Crystal structure and substrate specificity modification of acetyl xylan esterase from *Aspergillus luchuensis*. *Appl Environ Microbiol*. 2017;**83**:e01251–17.
27 Gruninger RJ, Cote C, McAllister TA, Abbott DW. Contributions of a unique β-clamp to substrate recognition illuminates the molecular basis of exolysis in ferulic acid esterases. *Biochemical Journal*. 2016;**473**:839–49.
28 Appeldoorn M, Kabel M, Van Eylen D, Gruppen H, Schols H. Characterization of oligomeric xylan structures from corn fiber resistant to pretreatment and simultaneous saccharification and fermentation. *J Agricult Food Chem*. 2010;**58**:11294–301.
29 Underlin EN, Frommhagen M, Dilokpimol A, van Erven G, de Vries RP, Kabel MA. Feruloyl esterases for biorefineries: Subfamily classified specificity for natural substrates. *Front Bioeng Biotechnol*. 2020;**8**:332.
30 Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;**8**:785–6.
31 Notredame C, Higgins DG, Heringa J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol*. 2000;**302**:205–17.
32 Gouet P, Robert X, Courcelle E. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res*. 2003;**31**:3320–3.
33 Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;**33**:1870–4.
34 Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, et al.Protein identification and analysis tools on the ExPASy Server. In The Proteomics Protocols Handbook (Walker JM, ed), 2005; pp. 571–607. Humana Press, Totowa, NJ.
35 Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res*. 2005;**33**:W244–8.
36 McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;**16**:404–5.
37 Wallner B, Elofsson A. Can correct protein models be identified? *Protein Sci*. 2003;**12**:1073–86.
38 Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*. 2009;**30**:2785–91.
39 Bellemare A, John T, Marqueteau S. Fungal genomic DNA extraction methods for rapid genotyping and genome sequencing. In Fungal Genomics. Methods in Molecular Biology (de Vries R, Tsang A, Grigoriev I, eds), 2018; pp. 11–20. Humana Press, New York.
40 Semova N, Storms R, John T, Gaudet P, Ulycznyj P, Min XJ, et al. Generation, annotation, and analysis of an extensive *Aspergillus niger* EST collection. *BMC Microbiol*. 2006;**6**:7.
41 Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res*. 1990;**18**:6069–74.
42 Storms R, Zheng Y, Li H, Sillaots S, Martinez-Perez A, Tsang A. Plasmid vectors for protein production, gene expression and molecular manipulations in *Aspergillus niger*. *Plasmid*. 2005;**53**:191–204.
43 Debets AJM, Bos CJ. Isolation of small protoplasts from *Aspergillus niger*. *Fungal Genet Rep*. 1986;**33**:1575.
44 Käfer E. Meiotic and mitotic recombination in *Aspergillus* and its chromosomal aberrations. *Adv Genet*. 1977;**19**:33–131.
45 Master ER, Zheng Y, Storms R, Tsang A, Powlowski J. A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: recombinant expression,

purification and characterization. *Biochemical J*. 2008;**411**:161–70.
46 Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;**9**:671–5.
47 Dilokpimol A. Fungal feruloyl esterases: Functional validation of genome mining based enzyme discovery including uncharacterized subfamilies. *New Biotechnol*. 2018;**41**:9–14.
48 Kormelink FJM, Lefebvre B, Strozyk F, Voragen AGJ. Purification and characterization of an acetyl xylan esterase from *Aspergillus niger*. *J Biotechnol*. 1993;**27**:267–82.
49 Koutaniemi S, van Gool MP, Juvonen M, Jokela J, Hinz SW, Schols HA, et al. Distinct roles of carbohydrate esterase family CE16 acetyl esterases and polymer-acting acetyl xylan esterases in xylan deacetylation. *J Biotechnol*. 2013;**168**:684–92.
50 Berka RM. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat Biotechnol*. 2011;**29**:922–7.
51 Kühnel S, Pouvreau L, Appeldoorn MM, Hinz SWA, Schols HA, Gruppen H. The ferulic acid esterases of *Chrysosporium lucknowense* C1: Purification, characterization and their potential application in biorefinery. *Enzyme Microbial Technol*. 2012;**50**:77–85.
52 Topakas E, Moukouli M, Dimarogona M, Christakopoulos P. Expression, characterization and structural modelling of a feruloyl esterase from the thermophilic fungus *Myceliophthora thermophila*. *Appl Microbiol Biotechnol*. 2012;**94**:399–411.
53 Puchart V, Vršanská M, Mastihubová M, Topakas E, Vafiadi C, Faulds CB, et al. Substrate and positional specificity of feruloyl esterases for monoferuloylated and monoacetylated 4-nitrophenyl glycosides. *J Biotechnol*. 2007;**127**:235–43.
54 de Vries RP, vanKuyk PA, Kester HCM, Visser J. The *Aspergillus niger* faeB gene encodes a second feruloyl esterase involved in pectin and xylan degradation and is specifically induced in the presence of aromatic compounds. *Biochemical Journal*. 2002;**363**:377–86.
55 Dilokpimol A, Mäkelä MR, Mansouri S, Belova O, Waterstraat M, Bunzel M, et al. Expanding the feruloyl esterase gene family of *Aspergillus niger* by characterization of a feruloyl esterase, FaeC. *New Biotechnol*. 2017;**37**:200–9.
56 Antonopoulou I. The synthetic potential of fungal feruloyl esterases: A correlation with current classification systems and predicted structural properties. *Catalysts*. 2018;**8**:242.
57 Li X, Griffin K, Langeveld S, Frommhagen M, Underlin EN, Kabel MA, et al. Functional validation of two fungal subfamilies in Carbohydrate Esterase family 1 by biochemical characterization of esterases from uncharacterized branches. *Front Bioeng Biotechnol*. 2020;**8**:694.
58 Koseki T, Handa H, Watanabe Y-Y, Ohtsuka M, Shiono Y. An unusual feruloyl esterase from *Aspergillus oryzae*: two tryptophan residues play a crucial role for the activity. *J Mol Catal B Enzym*. 2016;**133**:S560–8.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Multiple sequence alignment of fungal CE1 members.

Fig. S2. Multiple sequence alignment of fungal CE1 members with manual curation. * indicates biochemically characterized fungal enzymes from CE1. Red triangles indicate (putative) catalytic triad (Ser, Asp, His).

Table. S1. Overview of all protein sequences used in this study.

Table. S2. Parameters for homology protein modelling and verification.