Characterization of new fungal carbohydrate esterase family 1 proteins leads to the discovery of two novel dual feruloyl/acetyl xylan esterases

Adiphol Dilokpimol
Westerdijk Fungal Biodiversity Institute https://orcid.org/0000-0002-7208-507X
Bart Verkerk
Westerdijk Fungal Biodiversity Institute
Annie Bellemare
Concordia University
Mathieu Lavallee
Concordia University
Matthias Frommhagen
Wageningen Universiteit en Research
Emilie Nørmølle Underlin
Danmarks Tekniske Universitet Institut for Kemiteknik
Mirjam Anna Kabel
Wageningen Universiteit en Research
Justin Powlowski
Concordia University
Adrian Tsang
Concordia University
Ronald Peter de Vries (r.devries@wi.knaw.nl)
https://orcid.org/0000-0002-4363-1123

Research

Keywords: Feruloyl esterase, Acetyl xylan esterase, hydroxycinnamic acid, Carbohydrate Esterase family 1, Fungi, plant biomass, xylan

Posted Date: March 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-17222/v1

Abstract

Background Feruloyl esterases (FAEs) and acetyl xylan esterases (AXEs) are important accessory enzymes in the deconstruction of plant biomass. Carbohydrate Esterase family 1 (CE1) of the Carbohydrate-Active enZymes database contains both fungal FAEs and AXEs, sharing a high amino acid sequence similarity, even though they target different structural molecules on plant cell wall polysaccharides.

Results We recently classified fungal CE1 into five subfamilies (CE1_SF1-5). In this study, ten novel fungal CE1 enzymes from different subfamilies were heterologously produced in Aspergillus niger and characterized to gain insight on relationships among these esterases. The enzymes from CE1_SF1 possess AXE activity, as they hydrolyzed p NP-acetate and released acetic acid from wheat arabinoxylan, but were not active towards FAE substrates. CE1_SF5 showed FAE activity as they hydrolyzed methyl ferulate and other FAE related substrates, and release ferulic acid from wheat arabinoxylan. These FAEs preferred feruloylated arabinoxylan over pectin. Two CE1_SF2, sharing over 70% amino acid sequence identity, possessed the opposite activity. Interestingly, one enzyme from CE1_SF1 and one from CE1_SF5 possess dual feruloyl/acetyl xylan esterase (FXE) activity. These dual activity enzymes showed expansion of substrate specificity.

Conclusions The new FXEs from CE1 can efficiently release both ferulic acid and acetic acid from feruloylated xylan, making them particularly interesting novel components of industrial enzyme cocktails for plant biomass degradation.

Background

Many by-products from agricultural processing industries contain arabinoxylan as a major component. Arabinoxylan comprises of a β-D-(1→4)-linked xylose backbone, which is substituted with α-L-(1→2)- and/or α-L-(1→3)-linked arabinose residues. The xylan backbone can be acetylated at the *O*-2 and/or *O*-3 positions depending on the types of plants [1]. In commelinid monocots (e.g. wheat, rice and barley), these can be further substituted at the *O*-5 position by ferulic acid (4-hydroxy-3-methoxycinnamic acid) and other hydroxycinnamic acids (e.g. *p*-coumaric acid) [2-5]. 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 rhamnose residue can be substituted with α-L-(1→5)-arabinan or β-D-(1→4)-galactan chains, which both can contain terminal ferulic acid residues [6-8]. Ferulic acid plays a role in defense mechanism against pathogens, due to its antimicrobial property and the di-ferulic acid 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 the enzymatic saccharification for biomass valorization, e.g. for bioethanol production and biorefinery processes.

Acetyl xylan esterases (AXEs) [EC 3.1.1.72] catalyze 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) [EC 3.1.1.73] catalyze the de-esterification of agro-industrial byproducts which liberates ferulic 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 potential 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 part of the fungal FAEs are grouped in CE1, most FAEs are not classified CAZymes [10, 19]. CE1 FAEs are of interest to industry because of their broad substrate range and high synthetic property for potential antioxidants through transferuloylation [20]. 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_SF1-SF5) based on phylogenetic analysis [21]. CE1_SF1 and CE1_SF2 are closely related and split from the same node. However, CE1_SF1 contains characterized AXEs, while CE1_SF2 contains characterized FAEs. CE1_SF4 and CE1_SF5 are also split from a common node, and CE1_SF5 contains characterized FAEs, whereas CE1_SF4 has only uncharacterized members. CE1_SF3 is the smallest branch, which contains only sequences from basidiomycetes with no predicted secreted signal peptide. It 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 [22]. 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 [23], which is distantly related to bacterial enzymes belonging to Antigen85c family (<20% sequence identity) [24]. Thus far, two crystal structures of fungal CE1 enzymes have been reported, i.e. *Anaeromyces mucronatus* AmCE1 (PDB: 5CXU [25]) and *Aspergillus luchuensis Al*AXEA (formerly *Aspergillus awamori*) (PDB: 5X6S [24]).

In this study, we characterized ten fungal CE1 proteins covering CE1_SF1, SF2 and SF5 (Fig. 1, Table 1). These enzymes were heterologously produced in an *Aspergillus niger* strain derived from N400 (CBS 120.49, ATCC 9029, FGSC A1143) and biochemically characterized using both simple model and complex plant biomass substrates to validate their activity and substrate preference. Our results provide new insights into the differences in activity of fungal CE1 enzymes, and shed light on the relationship between these CE1 FAEs and other fungal FAEs.

Results And Discussion

Model substrates do not provide conclusive functional assignment for CE1 enzymes

In this study, we selected ten previously uncharacterized fungal CE1 enzymes covering three CE1 subfamilies for characterization. Four of these enzymes belonged to CE1_SF1, two to CE1_SF2 and four to CE1_SF5 (Table 1). We also included two characterized fungal CE1 enzymes from *Aspergillus niger* (AxeA) [26, 27] and *Chaetomium thermophilum* (AxeA) [28], 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 [29]. The protein concentration was estimated using a densitometric method, showing that *A. niger* AxeA was produced at the highest level (Table 1). Most CE1_SF1 enzymes efficiently hydrolyzed *p*NP-acetate, but not *p*NP-ferulate and four methyl hydroxycinnamates (methyl caffeate, methyl ferulate, methyl *p*-coumarate, and methyl sinapate), indicating that they possess AXE activity, except for the one from *Thermothelomyces thermophilus* which could also hydrolyze *p*NP-ferulate (Table 2).

Despite sharing over 70% amino acid sequence identity, two CE1_SF2 possess the opposite activity. The enzyme from *Corynascus sepedonium* (Fae1) in CE1_SF2 hydrolyzed the four methyl hydroxycinnamates tested, whereas the one from *C. thermophilum* only hydrolyzed *p*NP-acetate. Their different activities were confirmed using plant biomass substrates (see below). Three CE1_SF5 enzymes hydrolyzed all four methyl hydroxycinnamate substrates, except for *C. sepedonium* Fae2 which only hydrolyzed methyl ferulate and methyl sinapate. This is quite unique because the members from this subfamily usually show a broad substrate specificity, in particular those from *T. thermophilus* (formerly *Myceliophthora thermophila* and *Chrysosporium lucknowense*) [30, 31]. For the non-CE1 member, *C. sepedonium* Fae4 could also hydrolyze all tested hydroxycinnamate substrates showing high substrate preference for methyl ferulate. In addition, *C. sepedonium* Fae4 did not possess tannase activity as it did not show activity towards methyl gallate (data not shown). Based on our results all enzymes could hydrolyze *p*NP-acetate, hence it is not possible to use *p*NP-acetate alone for the identification of true AXE activity.

The three subfamilies show different activity towards plant biomass substrates

Acetylated corn fiber 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. All CE1_SF1 enzymes released acetic acid from insoluble wheat arabinoxylan and most of them also released acetic acid from corn fiber oligosaccharides (Fig. 2A, B). *Acremonium thermophilum* Axe1 released 100% acetic acid from corn oligosaccharides, whereas *A. niger* AxeA released up to 70% acetic acid from wheat arabinoxylan, which indicate that they were the most active enzymes on these substrates. Surprisingly, an enzyme from *Chaetomium globosum* in CE1_SF5 also released more than 50% acetic acid from wheat arabinoxylan. This is the first evidence confirming the dual FAE/AXE activity (FXE) towards plant biomass by an enzyme from this family.

The CE1_SF5 enzymes released between 35% and 60% ferulic acid from wheat arabinoxylan, but could not release ferulic acid from sugar beet pectin (Fig. 2C, D). In contrast, the tannase-related *C. sepedonium* Fae4, which does not group with any of the CE1 subfamilies, released almost 100% ferulic acid from sugar beet pectin, whereas it only released 15% from wheat arabinoxylan. This clearly showed a substrate specificity of CE1_SF5 for arabinoxylan. The tannase-related FAEs have higher specificity for pectin, which confirms previous reports [32, 33]. Surprisingly, an enzyme from *T. thermophilus* in CE1_SF1 (Fxe3) also efficiently released ferulic acid from wheat arabinoxylan. For CE1_SF2, *C. sepedonium* Fae1 showed FAE activity towards wheat arabinoxylan and to a lesser extent towards pectin, while the one from *C. thermophilum* showed almost no activity towards the plant cell wall-derived substrates. We did not analyze further the diferulic acid release because wheat arabinoxylan contains only low amount of diferulic acid (Underlin et al., unpublished data).

Our study showed that both *C. globosum* Fxe1 from CE1_SF5 and *T. thermophilus* Fxe3 from CE1_SF1 possess FXE activity, since both efficiently release ferulic acid and acetic acid from arabinoxylan to the same extent. These dual activity enzymes are not only interesting in their removal of side residues for the valorization of plant cell wall polysaccharide, but may also be particularly interesting for enzymatic synthesis for valuable compounds such as novel antioxidants through transesterification [20], because of their expanded substrate specificity which can accommodate different chemical structures. Identifying the amino acids responsible for the expanded substrate specificity will allow us to unlock the full potential of these enzymes.

Homology modelling did not result in amino acid candidate that could explain the substrate specificity

To verify the catalytic and substrate binding sites of fungal CE1 members, we created homology models of the selected CE1 enzymes (Fig 3, 4, Additional file 1) using 5X6S (*A. luchuensis* AXEA) as a template [24]. Based on the Protein Quality Predictor [34], the homology models for CE1_SF1 and SF2 are of very good quality (>80% sequence similarity). However, the models from CE1_SF5 are in 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 Additional file 2.

Recently, it has been shown that Trp160 in *A. luchuensis* AXEA determines the substrate specificity, and replacing it with Ala, Gly, or Pro could expand the substrate specificity towards FAE model substrates [24]. The corresponding amino acid in *T. thermophilus* Fxe3 from CE1_SF1 is also a Trp, whereas *C.*

sepedonium Fae1 and *C. thermophilum* AxeB from CE1_SF2, has an Ala and a Pro instead of Trp160. This indicates that additional amino acids also influence the substrate specificity expansion. We aimed to identify these residues using homology models of the enzymes and their amino acid sequence alignment, but no obvious candidate(s) was identified due to the high sequence similarity of the enzymes within CE1_SF1 (Fig 3, Additional file 2).

Conclusions

Fungal CE1 members are highly diverse and a few amino acid changes could lead to different substrate specificity. Both CE1_SF2 and SF5 possess FAE activity with broad substrate specificity, however only the CE1_SF2 members may have evolved from AXE in CE1_SF1 or *vice versa*. Our new insights and the newly characterized CE1 enzymes will advance their industrial applications, especially for the enzymes with FXE activity which efficiently release both ferulic acid and acetic acid from the arabionxylan.

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 fiber oligosaccharides was provided by Dr. Mirjam Kabel, Wageningen University [2]. Wheat bran was obtained from Wageningen Mill (Wageningen, the Netherlands). 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 (Wicklow, 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/) [35]. 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/) [36] and visualized using Easy Sequencing in Postscript (http://espript.ibcp.fr/ESPript/ESPript/) [37]. Phylogenetic analysis was performed using the Molecular Evolutionary Genetic Analysis program (MEGA7) [38] with 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) [39].

The homology protein models were created by using Homology detection and structure prediction by HMM-HMM comparison (HHPRED, https://toolkit.tuebingen.mpg.de/#/) [40]. The models were validated with different online tools, i.e. 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/) [41] and Protein Quality Predictor (ProQ, https://proq.bioinfo.se/ProQ/ProQ.html) [34]. Protein models were visualized using a molecular visualization system PyMOL (DeLano scientific, Schrödinger, LLC). Autodock (http://autodock.scripps.edu/) [42] was used for 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 [43] and cDNA was obtained following the previously described method [44]. The genes were cloned using an in-house variation of the Ligation-Independent Cloning (LIC) method [45].

All genes were cloned into the ANIp7 shuttle vector [46], which uses the *Aspergillus niger* glucoamylase promoter to drive recombinant gene expression. Positive plasmids were selected based on their resistance to ampicillin and maintained in *E. coli* strain DH5a using standard protocols. Protoplasts of *A. niger* strain CSFG_6005 (N593 *glaA*::*hisG*) were transformed [47] with the recombinant plasmids carrying CE1 genes. Transformants were selected on minimal medium without uracil and uridine [48]. Supernatants from transformants were screened for recombinant protein production after growth in liquid minimal MMJ medium [29] for induction of protein production.

Spores from positive transformants were inoculated in 200 ml MMJ media at a concentration of $2x10^6$ conidia/ml. Supernatants were harvested after incubating stationary cultures for 5 days at 30°C. They were then desalted and concentrated using Vivaflow® cassettes based on the manufacturer protocol (Sartorius, Göttingen, Germany). Buffer exchange was done using 10 mM citrate buffer at pH 5.0.

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 [49]. The reactions (225 µl) contained 0.8 mM *p*-nitrophenyl acetate or ferulate (dissolved in dimethyl sulfoxide), 100 mM sodium phosphate buffer, pH 6 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 [49]. The culture supernatant of *A. niger* harboring 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 [33] at 30°C for 5-30 min. Detection of substrates 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* harboring 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 hydrolyzed 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 fiber 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 rpm (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 fiber oligosaccharides, which were terminated by addition of 50 μL 2 M HCl prior to HPLC analysis.

Ferulic acid content analysis

Release of ferulic and *p*-coumaric acids was monitored by HPLC system (Dionex ICS-5000+ chromatography system; Thermo Scientific, Sunnyvale, CA) equipped with an Acclaim Mixed-Mode WAX-

1 column (3 x 150 mm; Thermo Scientific) and a UV detector (310 nm; Thermo Scientific). The chromatographic separation was carried out as previously described [33] 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/min 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 x 7.8 mm; Bio-Rad, Hercules, CA), a refractive index detector and an UV detector (UV 210 nm, Bio-Rad). An isocratic elution comprising of 5.0 mM sulfuric acid in MillQ water was applied. The flow rate was 0.6 mL/min at 40°C was used. 0.01-2.0 mg/mL of acetic acid was used as a standard for identification and quantification.

Declarations

List of abbreviations

AXE: Acetyl xylan esterase

CE1: Carbohydrate Esterase family 1

FAE: Feruloyl esterase

SF: Subfamily

Acknowledgements

Not applicable.

Funding

This work was supported by Genome Canada and Génome Québec.

Author information

Affiliations

Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

Adiphol Dilokpimol, Bart Verkerk & Ronald Peter de Vries

Centre for Functional and Structural Genomics, Concordia University, 7141 Sherbrooke St. W. Montreal, QC, H4B 1R6, Canada.

Annie Bellemare, Mathieu Lavallee, Justin Powlowski & Adrian Tsang

Laboratory of Food Chemistry, Wageningen University and Research, Bornse Weilanden 9, 6708 WG Wageningen, The Netherlands.

Matthias Frommhagen, Emilie Nørmølle Underlin & Mirjam Anna Kabel

Department of Chemistry, Technical University of Denmark, Building 207, Kemitorvet, DK-2800 Kgs. Lyngby, Denmark.

Emilie Nørmølle Underlin

Contributions

AD, BV performed the main experiments, analyzed the data, and wrote the manuscript. AB, ML, JP prepared recombinant proteins. MF, ENU, MAK performed substrate composition and acetic release analyses. AD, RPdV, AT designed the experiments and revise the manuscript. All authors reviewed and approved the manuscript.

Corresponding author

Correspondence to Adrian Tsang (adrian.tsang@concordia.ca), Ronald Peter de Vries (r.devries@wi.knaw.nl).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional files.

Ethics declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional files.

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. *Carbohydr 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. *Carbohydr Res*. 2013; 381:33-42.
3. de O. Buanafina MM. Feruloylation in grasses: current and future perspectives. *Mol Plant*. 2009; 2:861-872.
4. de Oliveira DM, Finger-Teixeira A, Rodrigues Mota T, Salvador VH, Moreira-Vilar FC, Correa Molinari HB, Craig Mitchell RA, Marchiosi R, Ferrarese-Filho O, Dantas dos Santos W. Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. *Plant Biotechnol J*. 2015; 13:1224-1232.

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. *Carbohydr Res*. 2004; 339:2315-2319.
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; 134:1173-1180.
8. Ralet MC, Thibault JF, Faulds CB, Williamson G. Isolation and purification of feruloylated oligosaccharides from cell walls of sugar-beet pulp. *Carbohydr Res*. 1994; 263:227-241.
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. *Carbohydr Res*. 2000; 327:401-410.
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; 51:72-78.
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 Microb Technol*. 2006; 39:897-902.
15. Razeq FM, Jurak E, Stogios PJ, Yan R, Tenkanen M, Kabel MA, Wang W, Master ER. 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-D495.
18. Adesioye FA, Makhalanyane TP, Biely P, Cowan DA. Phylogeny, classification and metagenomic bioprospecting of microbial acetyl xylan esterases. *Enzyme Microb 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. Antonopoulou I, Dilokpimol A, Iancu L, Mäkelä RM, Varriale S, Cerullo G, Hüttner S, Uthoff S, Jütten P, Piechot A *et al*. The Synthetic Potential of Fungal Feruloyl Esterases: A Correlation with Current Classification Systems and Predicted Structural Properties. *Catalysts*. 2018; 8.
21. 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; 11:869-880.
22. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang H-Y, El-Gebali S, Fraser MI *et al*. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res*. 2018; 47:D351-D360.
23. 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-D429.
24. Komiya D, Hori A, Ishida T, Igarashi K, Samejima M, Koseki T, Fushinobu S. Crystal structure and substrate specificity modification of acetyl xylan esterase from *Aspergillus luchuensis*. *Appl Environ Microbiol*. 2017; 83:e01251-01217.
25. Gruninger Robert J, Cote C, McAllister Tim A, Abbott DW. Contributions of a unique β-clamp to substrate recognition illuminates the molecular basis of exolysis in ferulic acid esterases. *Biochem J*. 2016; 473:839-849.
26. Kormelink FJM, Lefebvre B, Strozyk F, Voragen AGJ. Purification and characterization of an acetyl xylan esterase from *Aspergillus niger*. *J Biotechnol*. 1993; 27:267-282.
27. Koutaniemi S, van Gool MP, Juvonen M, Jokela J, Hinz SW, Schols HA, Tenkanen M. Distinct roles of carbohydrate esterase family CE16 acetyl esterases and polymer-acting acetyl xylan esterases in xylan deacetylation. *J Biotechnol*. 2013; 168:684-692.
28. Berka RM, Grigoriev IV, Otillar R, Salamov A, Grimwood J, Reid I, Ishmael N, John T, Darmond C, Moisan M-C *et al*. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. *Nat Biotechnol*. 2011; 29:922-927.
29. Master Emma R, Zheng Y, Storms R, Tsang A, Powlowski J. A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: recombinant expression, purification and characterization. *Biochem J* 2008; 411:161-170.
30. 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 Microb Technol*. 2012; 50:77-85.
31. 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.
32. 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. *Biochem J*. 2002; 363:377-386.

33. Dilokpimol A, Mäkelä MR, Mansouri S, Belova O, Waterstraat M, Bunzel M, de Vries RP, Hildén KS. Expanding the feruloyl esterase gene family of *Aspergillus niger* by characterization of a feruloyl esterase, FaeC. *N Biotechnol*. 2017; 37, Part B:200-209.
34. Wallner B, Elofsson A. Can correct protein models be identified? *Protein Sci*. 2003; 12:1073-1086.
35. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Meth*. 2011; 8:785-786.
36. Notredame C, Higgins DG, Heringa J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J Mol Biol*. 2000; 302:205-217.
37. Gouet P, Robert X, Courcelle E. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res*. 2003; 31:3320-3323.
38. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016; 33:1870-1874.
39. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy Server. In: Walker JM, editors. *The Proteomics Protocols Handbook*. Totowa, NJ: Humana Press; 2005. p. 571-607.
40. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res*. 2005; 33:W244-W248.
41. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000; 16:404-405.
42. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*. 2009; 30:2785-2791.
43. Bellemare A, John T, Marqueteau S. Fungal Genomic DNA Extraction Methods for Rapid Genotyping and Genome Sequencing. In: de Vries R, Tsang A, Grigoriev I, editors. *Fungal Genomics Methods in Molecular Biology.* vol. 1775. New York: Humana Press; 2018. p. 11-20.
44. Semova N, Storms R, John T, Gaudet P, Ulycznyj P, Min XJ, Sun J, Butler G, Tsang A. Generation, annotation, and analysis of an extensive *Aspergillus niger* EST collection. *BMC Microbiol*. 2006; 6:7-7.
45. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res*. 1990; 18:6069-6074.
46. 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.
47. Debets AJM, Bos CJ. Isolation of small protoplasts from *Aspergillus niger*. *Fungal Genet Rep*. 1986; 33:1575.
48. Käfer E. Meiotic and mitotic recombination in *Aspergillus* and its chromosomal aberrations. *Adv Genet*. 1977; 19:33-131.

49. Dilokpimol A, Mäkelä MR, Varriale S, Zhou M, Cerullo G, Gidijala L, Hinkka H, Brás JLA, Jütten P, Piechot A *et al*. Fungal feruloyl esterases: Functional validation of genome mining based enzyme discovery including uncharacterized subfamilies. *N Biotechnol*. 2018; 41:9-14.

50. Pouvreau L, Jonathan MC, Kabel MA, Hinz SWA, Gruppen H, Schols HA. Characterization and mode of action of two acetyl xylan esterases from *Chrysosporium lucknowense* C1 active towards acetylated xylans. *Enzyme Microb Technol*. 2011; 49:312-320.

Tables

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

Fungal species	CE1[1] subfamily	FAE[2] subfamily	UniProt[3] entry name	Enzyme name	Calculated molecular mass (kDa)	Apparent molecular mass (kDa)	Calculated p*I*	Production (mg/mL)[4]	Reference
Aspergillus niger	CE1_SF1	-	AXE1A_ASPNG	AxeA[7]	30.1	30	4.82	1.42	[26, 27]
Rasamsonia emersonii	CE1_SF1	-	AXE1A_RASEM	Axe1	30.1	30	4.70	0.60	This study
Thermothelomyces thermophilus[5]	CE1_SF1	-	FXE1A_MYCTH[f]	Fxe3[8]	31.6	30	6.21	0.68	This study
Chaetomium thermophilum	CE1_SF1	-	AXE1A_CHATH[g]	AxeA	37.6	39	7.62	0.40	[28]
Corynascus sepedonium[6]	CE1_SF1	-	AXE1B_MYCSE	Axe1	32.5	35	4.57	0.40	This study
Acremonium thermophilum	CE1_SF1	-	AXE1A_ACRTH	Axe1	29.8	40	6.02	0.13	This study
Corynascus sepedonium[f]	CE1_SF2	SF6	FAE1B_MYCSE	Fae1	29.3	30	4.53	0.07	This study
Chaetomium thermophilum	CE1_SF2	SF6	AXE1B_CHATH	AxeB	29.7	60	7.27	0.13	This study
Chaetomium globosum	CE1_SF5	SF5	FXE1A_CHAGL	Fxe1	28.1	30	6.22	0.13	This study
Corynascus sepedonium[6]	CE1_SF5	SF5	FAE1C_MYCSE	Fae2	26.4	30	5.39	0.13	This study
Corynascus sepedonium[6]	CE1_SF5	SF5	FAE1E_MYCSE	Fae3	32.5	40	7.15	0.25	This study
Chaetomium thermophilum	CE1_SF5	SF5	FAE1B_CHATH	FaeA	34.0	45	7.66	0.25	This study
Corynascus sepedonium[6]	Non-CE1	SF2	FAE1F_MYCSE	Fae4	55.1	60	4.97	0.03	This study

[1] Based on Fig. 1

[2] Based on [10]

[3] Based on UniProtKB/Swiss-Prot entry name recommendation (https://www.uniprot.org/help/entry_name)

[4] Based on Additional file 3

[5] (Formerly *Myceliophthora thermophila*)

[6] (Formerly *Myceliophthora sepedonium*)

[7] Also referred to as Axe; *An*AXE; *An*AxeA; *An*AXE1

[8] Trp160 in *Al*AXEA (PDB: 5X6S [24]) 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 [50].

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

Fungi	Enzyme name	Specific activity (mU/mg) [1,2,3]					
		*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	Fxe3	107	34	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	Fae1	232	N.A.	336 [77 %]	200 [46 %]	439 [100 %]	128 [29 %]
Chaetomium thermophilum	AxeB	1205	N.A.	N.A.	N.A.	N.A.	N.A.
Chaetomium globosum	Fxe1	1331	47	135 [100 %]	75 [55 %]	93 [69 %]	63 [47 %]
Corynascus sepedonium	Fae2	512	N.A.	N.A. [0 %]	N.A. [0 %]	50 [100 %]	44 [87 %]
Corynascus sepedonium	Fae3	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	Fae4	764	3133	1,790 [10 %]	734 [4 %]	18,771 [100 %]	488 [3 %]

[1] N.A., no activity detected

[2] 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.

[3] 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%.

Figures

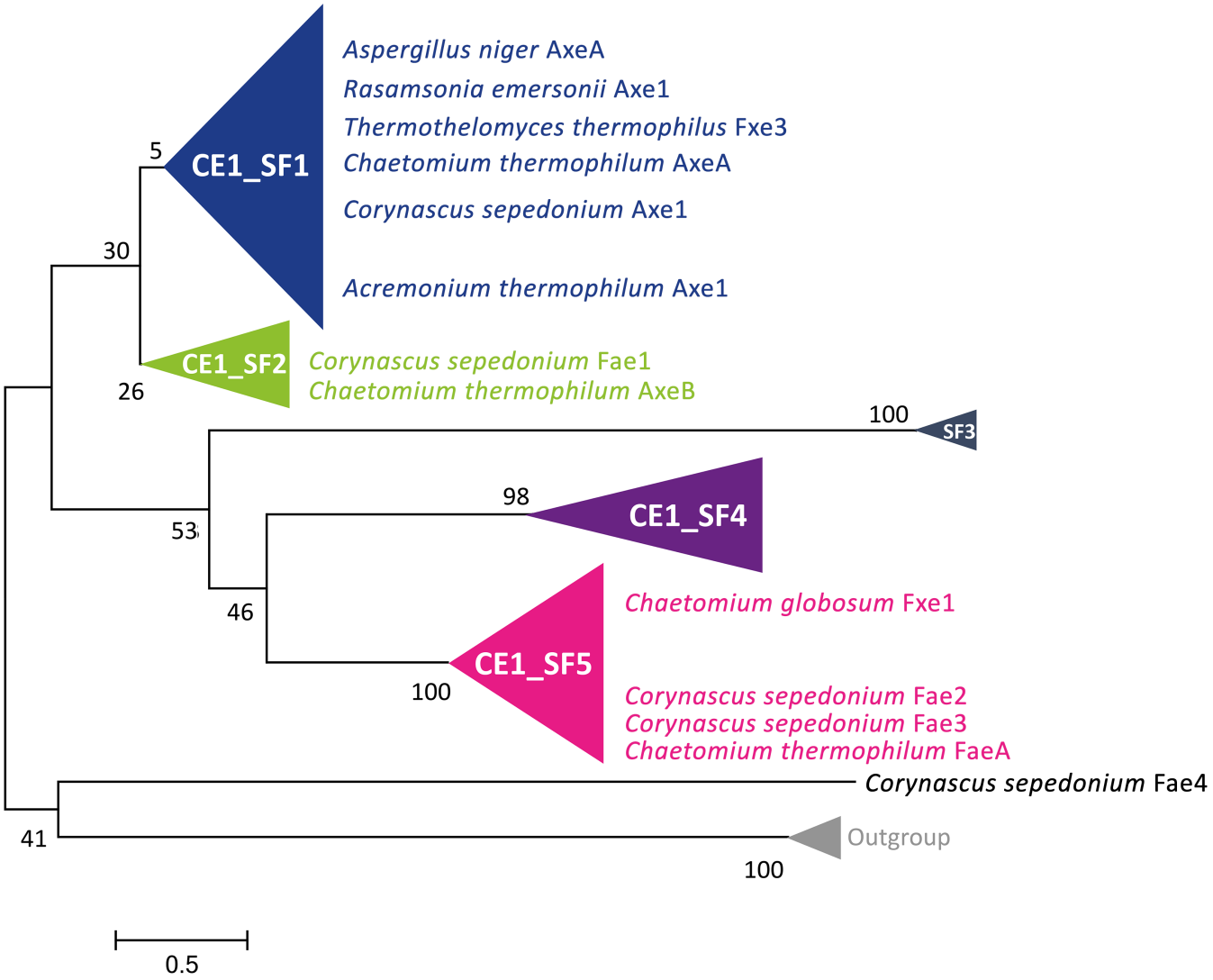

Figure 1

Phylogenetic relationships among the fungal CE1 members. The phylogenetic analysis was based on amino acid sequences (modified from [21]). The evolutionary history was inferred using Neighbour-Joining method with a bootstrap value of 500. Eight non-CE1 FAEs from SF7 [10] were used as an outgroup.

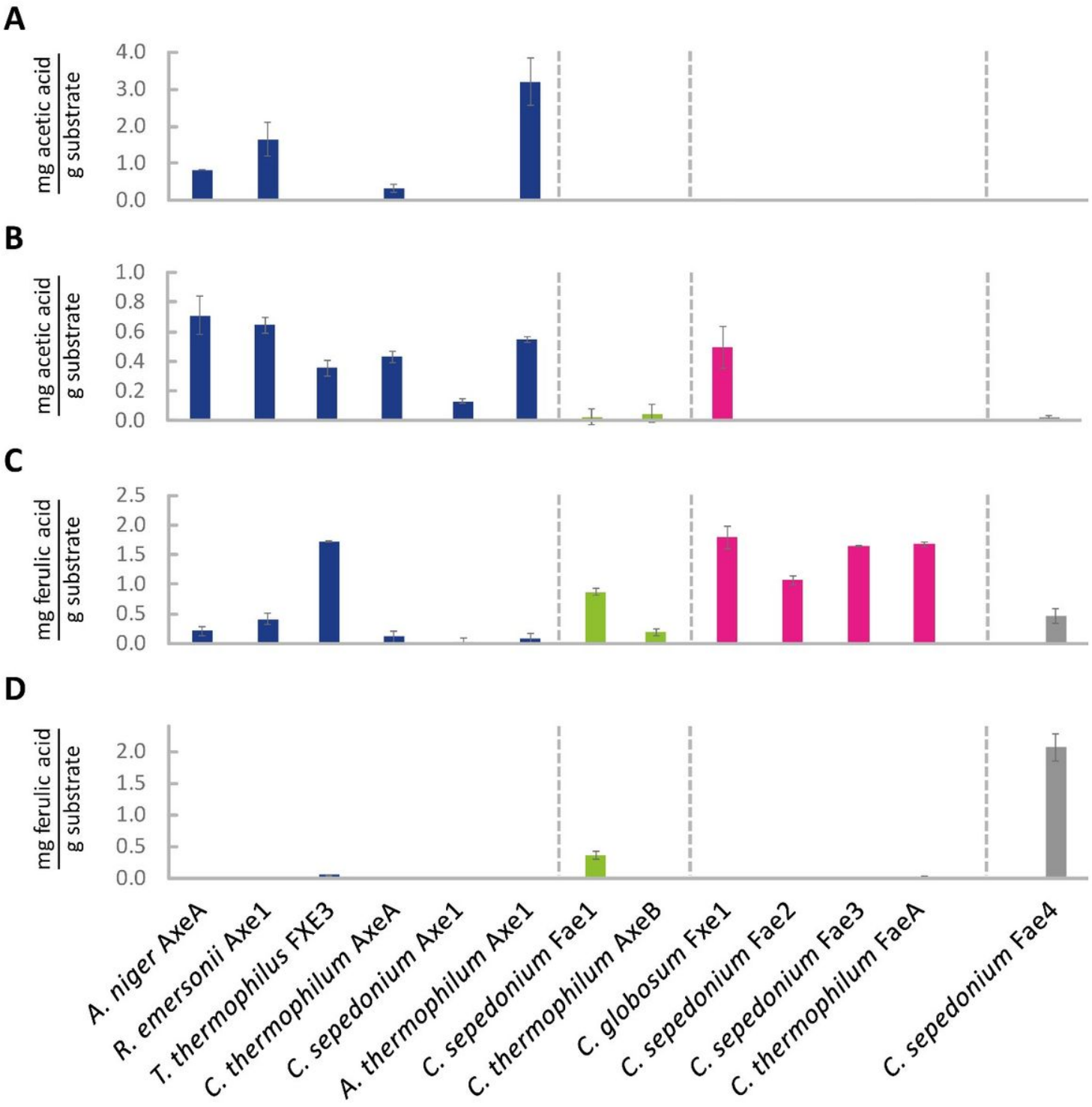

Figure 2

Release of acetic and ferulic acids from plant biomass substrates by CE1 enzymes. (A) acetic acid release from corn fiber 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 pretreated 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.

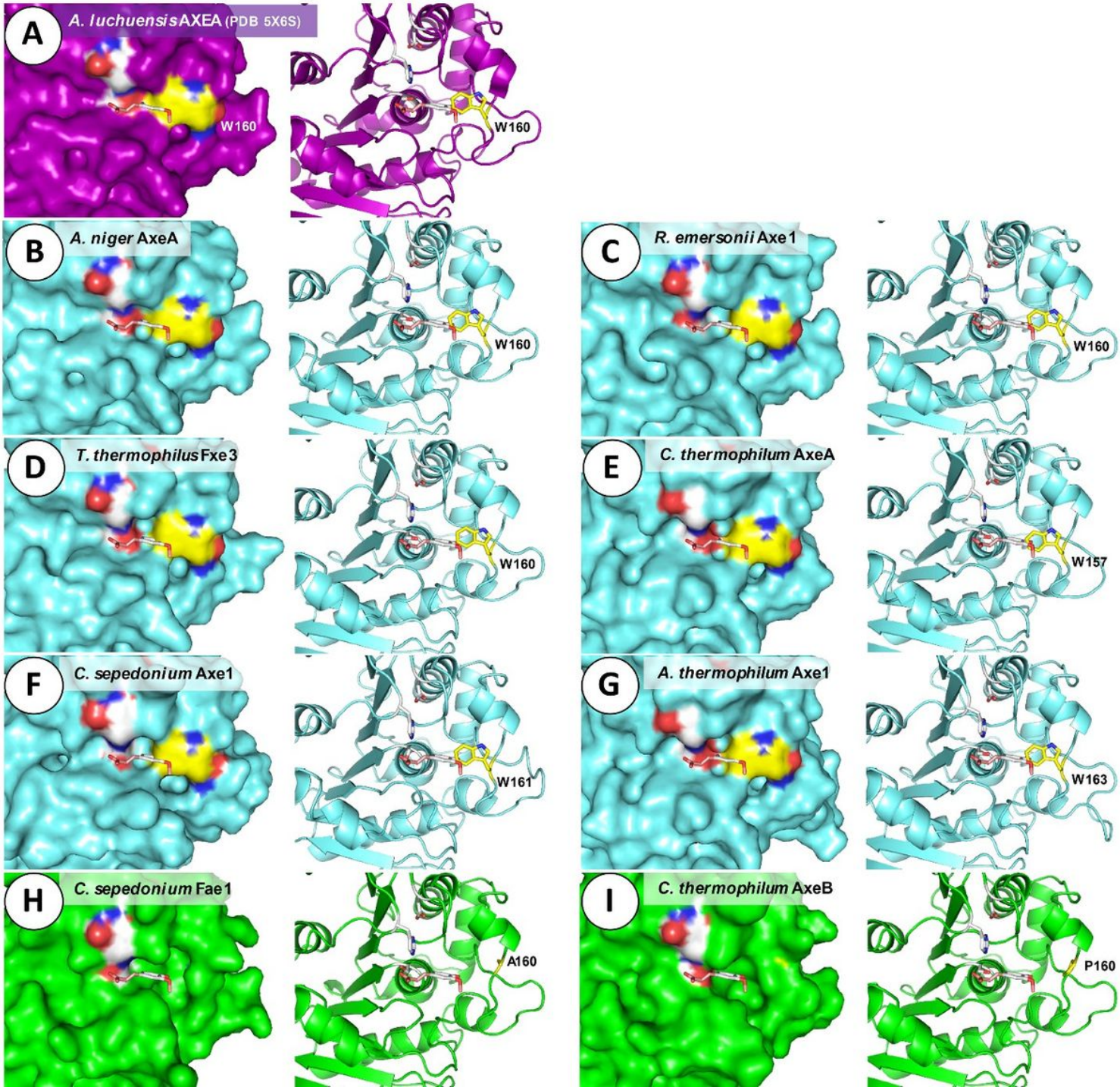

Figure 3

3D diagram representing structure of A. luchuensis AXEA and models of CE1_SF1 and SF2 enzymes. The left panel shows the molecular surface, while the right panel shows a ribbon diagram. (A) Crystal structure of AlAXEA (PDB: 5X6S [24]), (B-G) CE1_SF1 enzymes, (H-I) CE1_SF2 enzymes. The catalytic triad is shown in white, Trp160 in AlAXEA 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 [42].

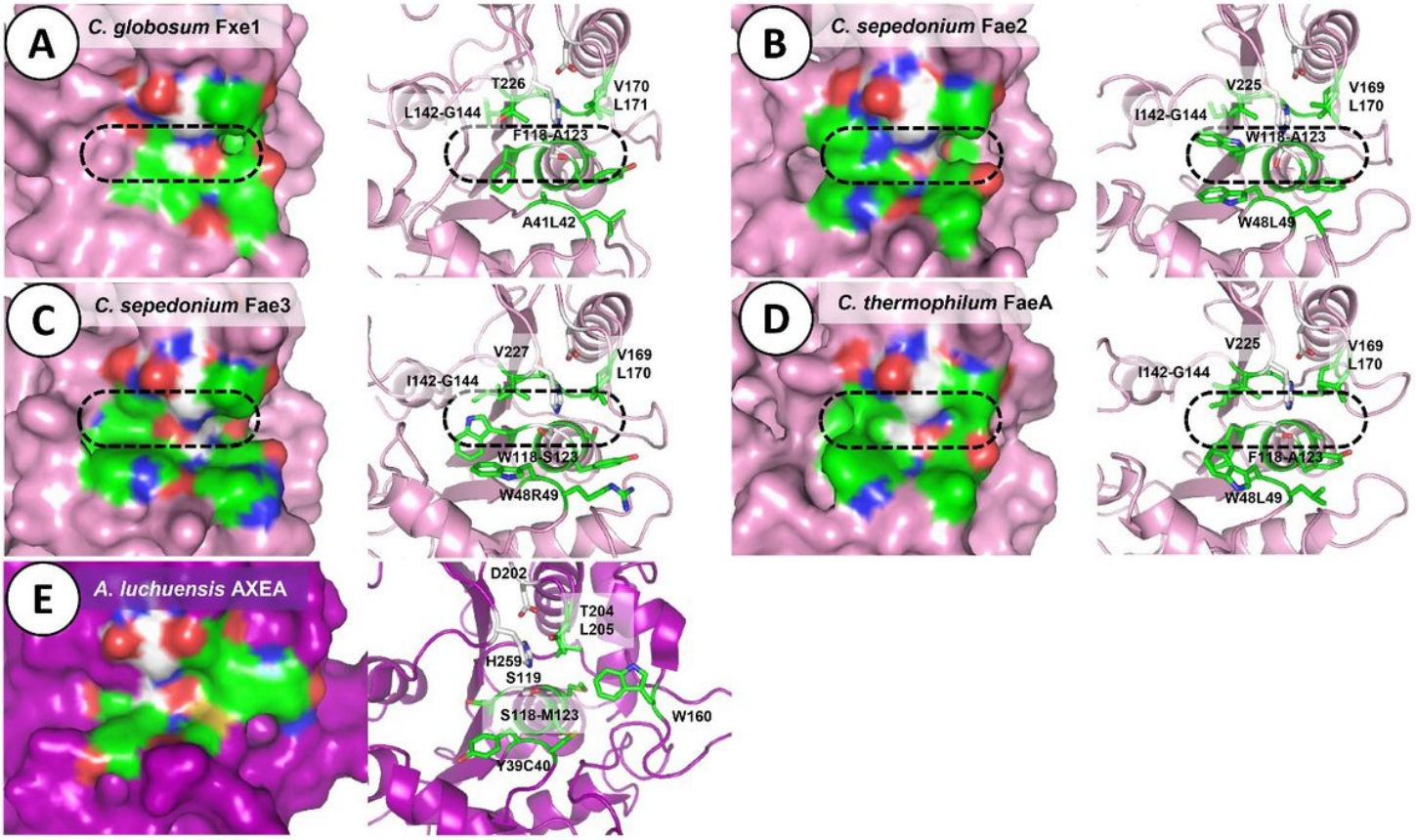

Figure 4

3D diagram representing structure of A. luchuensis AXEA and models of CE1_SF5 enzymes. The left panel shows the molecular surface, while the right panel shows ribbon diagram. (A-D) CE1_SF5 enzymes, (E) crystal structure of AlAXEA (PDB: 5X6S [24]). 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.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS2.pdf
- TableS1.pdf
- FigureS1.pdf