Insights into the structure and functionality of feruloyl esterases for the efficient utilization of lignocellulosic biomass

Apisan Phienluphon

Table of contents

Chapter One

General introduction

1.1 Lignocellulosic biomass

Lignocellulosic biomass refers to plant-based material that is not used for food or feed and primarily encompasses agricultural and forestry residues. This sustainable and abundant resource, lignocellulose, has garnered significant attention due to its potential to address pressing global challenges, including climate change and food crises [1,2]. Lignocellulosic biomass holds promise as a renewable resource with multiple applications, such as the production of biofuels, bio-based chemicals, materials, and energy, offering sustainable solutions that mitigate the impact of climate change and help alleviate the food crisis [3].

Lignocellulosic biomass is structurally resilient due to its components of cellulose, hemicellulose, and lignin. Typically, it contains 40–60% cellulose, 20–35% hemicellulose, and 15–40% lignin, varying with agricultural source [4]. Cellulose consists of elongated glucose chains connected through β-1,4-glycosidic bonds **(Fig. 1-1)** [5]. On the other hand, hemicelluloses constitute complex, branched heteropolysaccharides comprising hexoses (glucose, mannose, and galactose), pentoses (xylose, and arabinose), and various modification groups (ferulic, *p*-coumaric, sinapic, caffeic, and acetic acids) [6]. Meanwhile, lignin, which is composed of phenolic compounds, primarily *p*-coumaryl, coniferyl, and sinapyl alcohol units, imparts rigidity, strength, and resistance to microbial attack to plant cell walls **(Fig. 1-1)** [7]. Additionally, the lignin-carbohydrate complex, a crosslink between lignin and carbohydrates, is suggested to play a crucial role in the recalcitrance of woody biomass [8]. These components of plant cell walls can be used in the production of biomass-derived biofuels and functional materials [5–7].

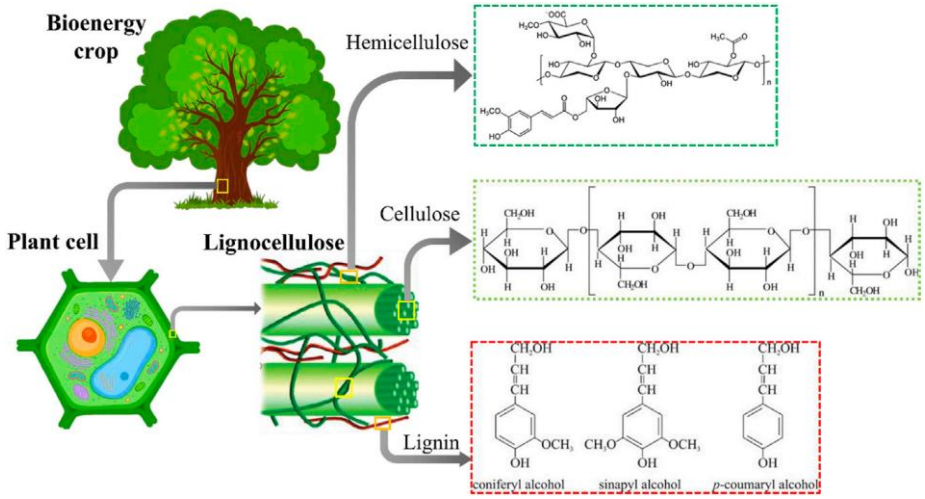

Fig. 1-1. Schematic illustration of the lignocellulosic biomass composition [7].

1.2 Hemicellulases and its accessory enzymes

Hemicellulases are a group of enzymes that play a crucial role in the degradation of hemicellulose, a complex and heterogeneous polysaccharide found in the cell walls of plants [10,11]. Hemicelluloses are polysaccharides that surround cellulose fibers in plant cell walls, contributing to the recalcitrant structure of plant biomass. Hemicellulases include a variety of glycoside hydrolases, polysaccharide lyases, and carbohydrate esterases [10]. These enzymes collaborate to break down the intricate structure of hemicellulose into smaller oligosaccharides and monosaccharides [11]. Some key hemicellulolytic enzymes include xylanases (XYNs), mannanases, β-xylosidases, α-galactosidases, and arabinofuranosidases. XYNs target β-1,4-D-xylose polymeric residues, while mannanases degrade β-1,4-D-mannans, and β-xylosidases cleave β-1,4-xylooligosaccharides. α-galactosidases and arabinofuranosidases are important for the removal of side groups like arabinose and galactose.

Accessory enzymes also play a crucial role in biomass degradation by assisting hemicellulolytic enzymes. These enzymes, including carbohydrate esterases, hydrolyze the ester bonds between hemicellulose and carboxylic acids. Additionally, accessory enzymes like feruloyl esterases (FAEs) and glucuronoyl esterases are necessary to uncouple hemicellulose from the lignin fraction in lignin-carbohydrate complex moieties

[11]. Previously, the coordination of XYN and their accessory counterparts is essential for the effective degradation of lignocellulosic biomass [12–16].

1.3 Feruloyl esterase

Feruloyl esterase or ferulic acid esterase (FAE; E.C. 3.1.1.73) is a hemicellulase accessory enzyme that breakdowns the ester linkage in plant cell walls. The primary function of feruloyl esterase is to hydrolyze the ester bonds between ferulic acid (FA) and arabinose on the arabinoxylan **(Fig. 1-2)**. FA, along with other hydroxycinnamic acids (*p*-coumaric, caffeic, sinapic), are often connected to the hemicellulose component of the biomass through the ester bonds. Notably, FA and its dehydrodimers also play a significant role in the recalcitrance of biomass by crosslinking to other complex structures such as lignin or hemicellulose [17]. Consequently, by cleaving these bonds, feruloyl esterase contributes to the liberation of hemicellulose and other complex carbohydrates, making them more accessible for further enzymatic or microbial degradation [17].

In addition to the hydrolytic activity, FAEs exhibit the (trans)esterification activity, which can be harnessed for producing potential antioxidants **(Fig. 1-3)**. Previous studies demonstrated that the bioactive feruloyl compounds can be synthesized by modifying the reaction conditions or optimizing FAE biocatalysts such as employing water-in-oil microemulsions [18], detergentless microemulsions [19–22], and immobilization techniques [23–27] to facilitate these reactions.

Moreover, the promise of FAE uses extend across various industries. The resulting product, FA, exhibits antioxidant properties, which are valuable in the fields of food, pharmaceuticals, and cosmetics [28–30]. The ability of FA to neutralize harmful free radicals enhances its potential as a natural additive in food products, nutraceuticals, and dietary supplements, thereby aiding in the promotion of human health [29–31].

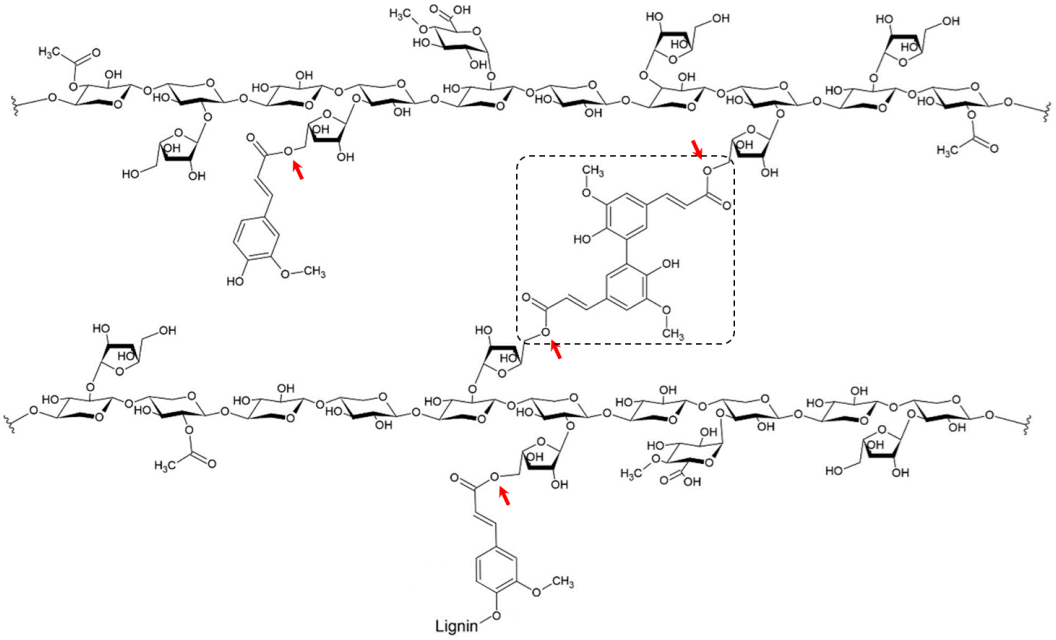

Fig. 1-2. Example of feruloylated arabinoxylan, where FA is esterified to the *O*-5 hydroxyl group on L-arabinofuranose (red arrows) [29]. Crosslink of 5,5'-diferulate between hemicellulose chains is indicated in dotted box. FA linkage with lignin is denoted.

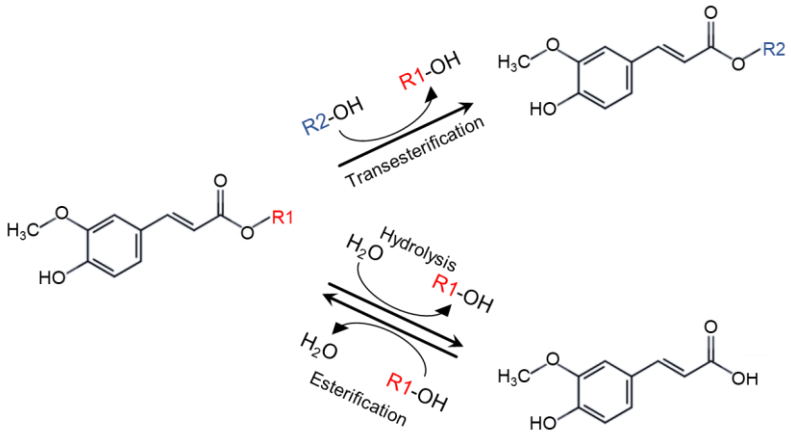

Fig. 1-3. Reaction schemes of FAE hydrolysis, esterification, and transesterification. R1 and R2 are alkyl groups.

1.4 Diversity of fungal FAE

FAEs are a diverse group of enzymes commonly found in lignocellulolytic bacteria and fungi. Primarily, fungal FAE was classified into four functional sub-classes, termed A, B, C, and D, based on the sequence similarity and their specific esterase activities against synthetic methyl hydroxycinnamates and ability to release 5,5'-diferulic acid from natural substrate **(Fig. 1-4)** [32]. However, due to the rapid expansion in the number of sequenced FAEs, particularly from fungal sources, the refined classification based on the phylogenetic analysis was updated to 7 subfamilies (SF1–7) [33], 12 feruloyl esterase families (FEF1–12) [34], and 13 SFs (SF1–13) [35], respectively **(Fig. 1-5)**. The latest revised classification suggested that FAEs evolved from highly divergent esterase families including tannases (SF1–4, SF9–11), acetyl xylan esterases (AXEs; SF5–6), lipases (SF7), and lipase-choline esterases (SF8, SF12–13).

Considering their activity on the ester bond between lignin and carbohydrates, FAEs are classified as members of the carbohydrate esterase family 1 (CE1) in the Carbohydrate-Active enZYme (CAZy) database [36]. Additionally, they are identified as members in the ESTerases and α/β-Hydrolase Enzymes and Relatives (ESTHER) database [37].

Recent phylogenetic analysis of FAEs on fungal enzymes from the CE1 family of CAZy (containing SF5 and SF6 FAEs) revealed five subfamilies (CE1_SF1–5; see in **Fig. 1-6**) [38,39]. Certain members of CE1_SF1, SF2, and SF5 have been previously studied and characterized, such as AXEs belonging to SF1 [38]. In the ESTHER database, fungal CE1 members are classified into the Esterase_phb, Antigen85c, and FaeC families [37]. The Esterase_phb family comprises members of CE1_SF1 and SF2, which share close genetic relationships **(Fig. 1-6)**. The Antigen85c family comprises members of CE1_SF3, while the FaeC family has recently been shown to include members of CE1_SF5. It is noted that there is currently no corresponding ESTHER family for members of CE1_SF4.

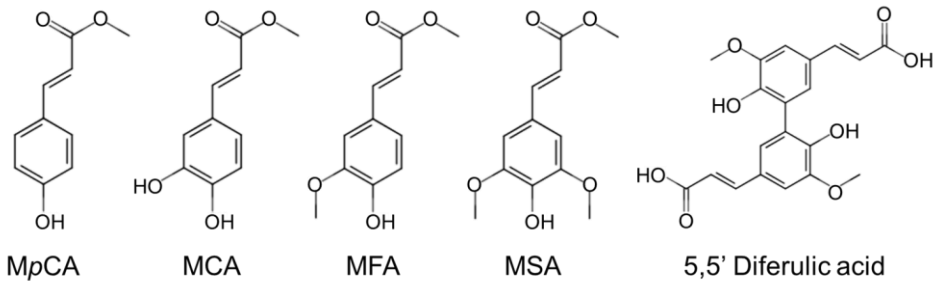

Fig. 1-4. Chemical structures of methyl hydroxycinnamates (methyl *p*-coumarate (M*p*CA), methyl caffeate (MCA), methyl ferulate (MFA), methyl sinapate (MSA)) and 5, 5' diferulate.

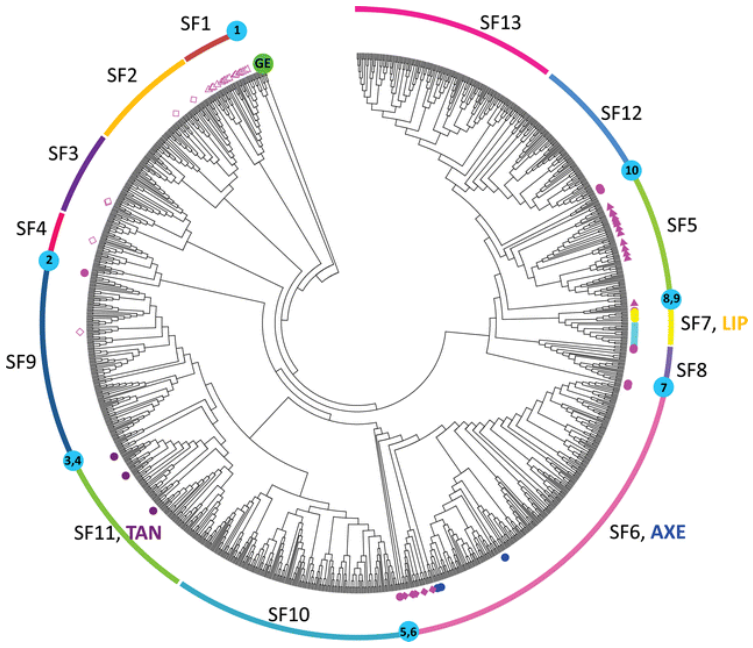

Fig. 1-5. An updated phylogenetic relationship among the fungal FAEs proposed by Dilokpimol, et al. [33].

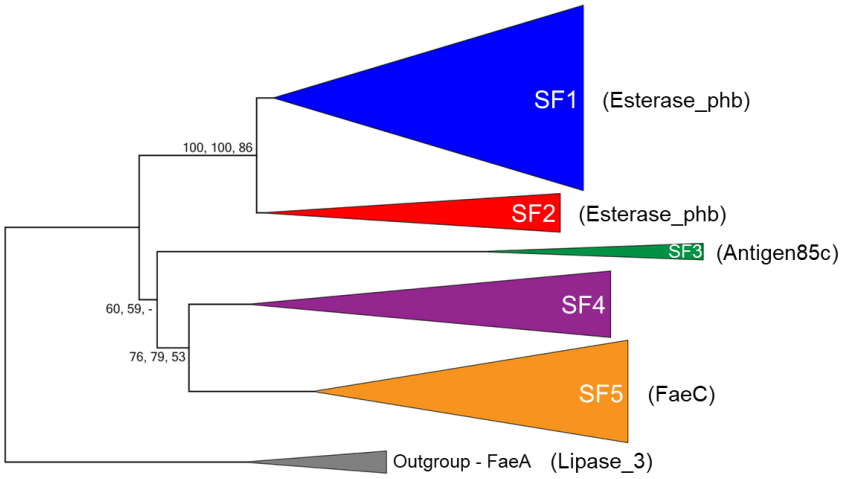

Fig. 1-6. Phylogenetic relationships among the fungal CE1 members proposed by Komiya, et al. [36] and Makela, et al. [37]. The names of the esterase classes according to the ESTHER classification system are shown in parentheses. FAEs from FAE SF7 (Dilokpimol, et al. 2016), belonging to the Lipase_3, was used as an outgroup.

1.5 FAE structural diversity

To gain insight into the functions and biological interactions of FAEs, it is imperative to study structures at the atomic level. Previously, among 13 SFs within the fungal FAE classification, crystal structures have been elucidated for four FAEs belonging to SF1, SF2, SF7, and SF8, as well as one FAE-related enzyme from SF6 **(Fig. 1-7)**. Those crystal structures are tannase-related (SF1) *Ao*FaeB from *Aspergillus oryzae* (PDB ID: 3WMT) [40], tannase-related (SF2) *Fo*FaeC from *Fusarium oxysporum* (related PDB entries: 6FAT, 8GHH) [41,42], lipase-related (SF7) *An*FaeA from *A. niger* (related PDB entries: 1UWC, 1USW, 1UZA, 2BJA, 2IX9, 2HL6) [43–45], CE1-related (SF8) *Am*Fae1A from *Anaeromyces mucronatus* (PDB ID: 5CXU, 5CXX) [46], and SF6 FAE-related *Al*AXE from *A. luchuensis* (PDB: 5X6S) [38].

Additionally, six distinct characteristics of FAE crystal structures from 11 bacterial sources were also observed **(Fig. 1-8)**. These crystal structures include FAE domain of cellulosomal xylanase 10B from *Clostridium thermocellum* (related PDB entries: 1GKK, 1GKL, 1JJF, 1JT2, 1WB4, 1WB5, 1WB6, 4BAG, 4H35, 5FXM) [47–51], Est1E from rumen bacterium *Butyrivibrio proteoclasticus* (PDB ID: 2WTM, 2WTN)

[52], cinnamoyl esterase Lj0536 from *Lactobacillus johnsonii* (PDB ID: 3PF8, 3PF9, 3PFB, 3PFC, 3QM1, 3S2Z) [53], FAE LP_0796 from *L. plantarum* (PDB ID: 7EBO) [54], FAE from *L. buchneri* (PDB ID: 7Z2U, 7Z2V, 7Z2X) [55], FAE from *Bacteroides intestinalis* (related PDB entries: 5VOL, 6MOT, 6MOU) [56,57], FAE from *Streptomyces cinnamoneus* (PDB ID: 5YAE, 5YAL) [58], FAE type-A from *Alistipes shahii* (PDB ID: 7DQ9) [59], GthFAE from *Geobacillus thermoglucosidasius* (PDB ID: 7WWH) [60], EstF27 from uncultured bacteria (PDB ID: 4ZRS) [61], and FAE from the multi-enzyme CE1-GH62-GH10 from uncultured bacteria (PDB ID: 6RZN, 6RZO) [62].

Both bacterial and fungal FAEs share a typical *α/β* hydrolase fold, which is well-known for esterases, lipases, thioesterases, amidases, hydrolases, and lyases [37,63,64]. Furthermore, as serine hydrolases, FAEs contain a conserved catalytic triad (Ser, His, Asp; S-H-D), where serine is a catalytic center with a common consensus nucleophilic elbow (GXSXG pentapeptide; where X is any amino acid residue) [65,66].

This structural diversity allows FAEs to adapt to various ecological niches and substrates, making them integral components of microbial ecosystems. However, due to the diversity of FAEs in terms of sequence homology and the limited number of solved structures, their classification and the prediction of their functions and physicochemical properties can be challenging.

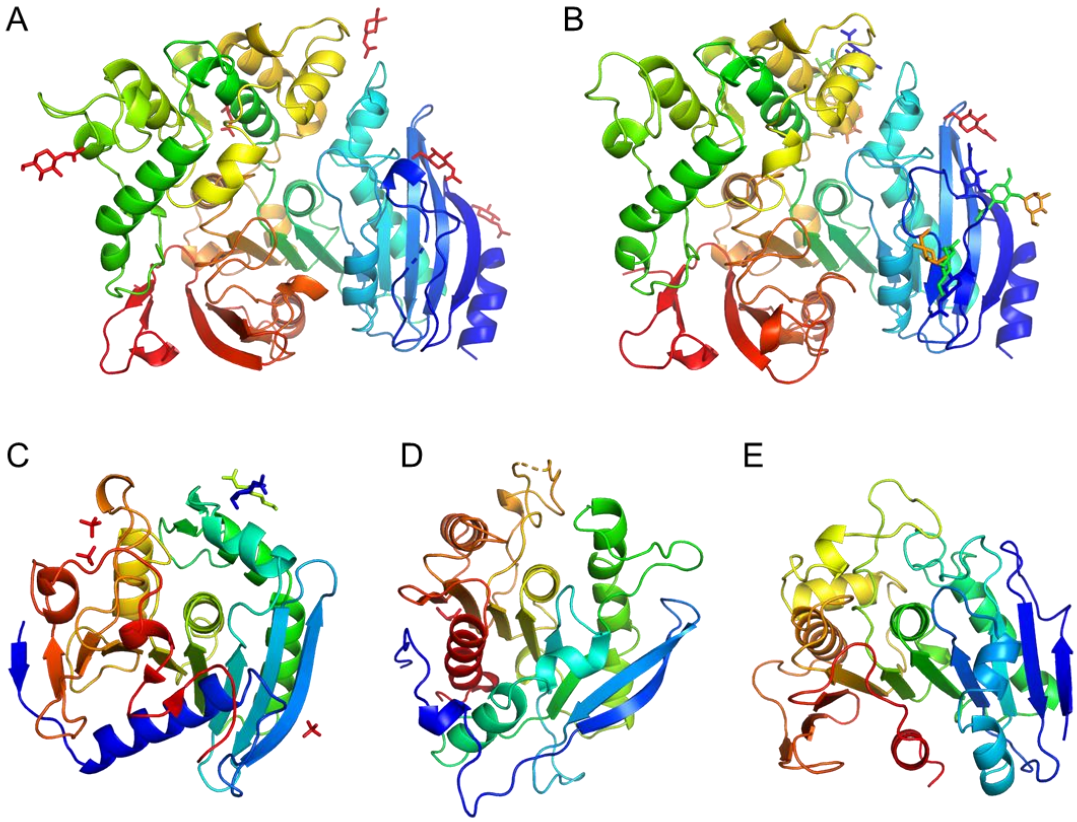

Fig. 1-7. Crystal structures of fungal FAEs in SF1, 2, 7, and 8, and SF6-related AXE. Monomeric structure of *Ao*FaeB from *A. oryzae* (A), *Fo*FaeC from *F. oxysporum* (B), *An*FaeA from *A. niger* (C), *Am*Fae1A from *A. mucronatus* (D), and *Al*AXE from *A. luchuensis* (E).

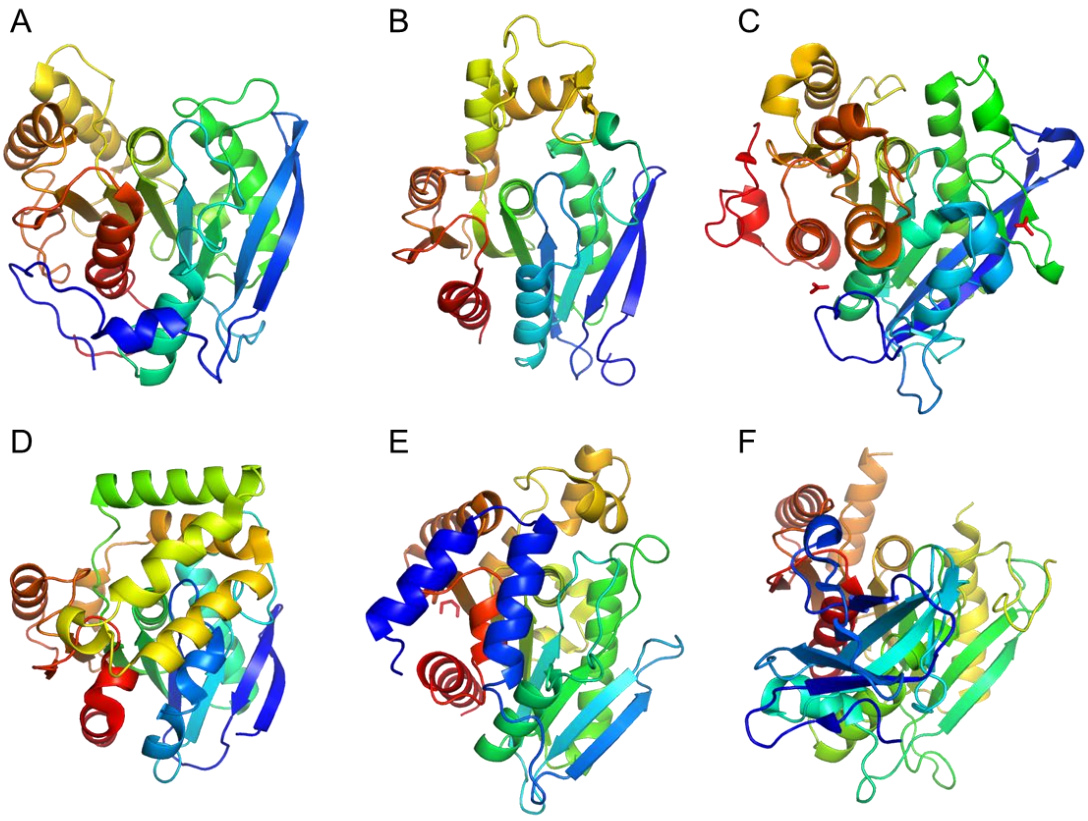

Fig. 1-8. Crystal structures of bacterial FAEs. Monomeric structure of FAE domain of cellulosomal xylanase 10B from *C. thermocellum* (A), cinnamoyl esterase Lj0536 from *L. johnsonii* (B), FAE from *S. cinnamoneus* (C), FAE type-A from *A. shahii* (D), EstF27 from unculture bacteria (E), and FAE from the multi-enzyme CE1-GH62-GH10 from unculture bacteria (F).

1.6 FAE motifs involved in enzyme catalysis

As a member of the *α/β* hydrolase superfamily, FAE structures contain the conserved catalytic core that mainly constitutes the S-H-D triad, the oxyanion hole, and the hydrophobic pocket **(Fig. 1-9)** [64,67,68]. Each component plays a critical role during enzyme catalysis. The S-H-D catalytic triad, crucial for carboxylic ester hydrolysis, forms the hydrogen-bonding network [64,69]. Recent research has investigated the role of the hydrogen-bonding network in the two-step hydrolysis mechanism of *An*FaeA, involving acylation and deacylation, through quantum mechanics/molecular mechanics-based

transition path sampling simulations [70]. During acylation, a covalent bond forms between the O^{γ} of Ser133 in *An*FaeA and the carbonyl carbon of FA, while deacylation corresponds to its cleavage.

Concurrently, the oxyanion hole of FAEs, which is comprised of the two NH groups, is located at the backbone adjacent to the catalytic serine [67]. It plays a role as the hydrogen bond donor to stabilize the negatively charged carbonyl oxygen atom of ferulate [68,70]. Additionally, the hydrophobic pocket in FAEs provides a suitable environment for binding with hydroxycinnamate esters and promotes substrate-binding efficiency. Besides the polar interaction mentioned earlier, the interactions involved in substrate binding are hydrophobic-hydrophobic, CH-π, and/or π-π interactions [40,43].

This structural and mechanistic information serves as the foundation for understanding how FAEs function in the degradation of plant cell wall components, their roles in microbial ecosystems, and their potential applications in various biotechnological processes. Nonetheless, the field of FAE research continues to evolve, presenting exciting opportunities for further exploration, discovery, and innovation in both fundamental and applied science.

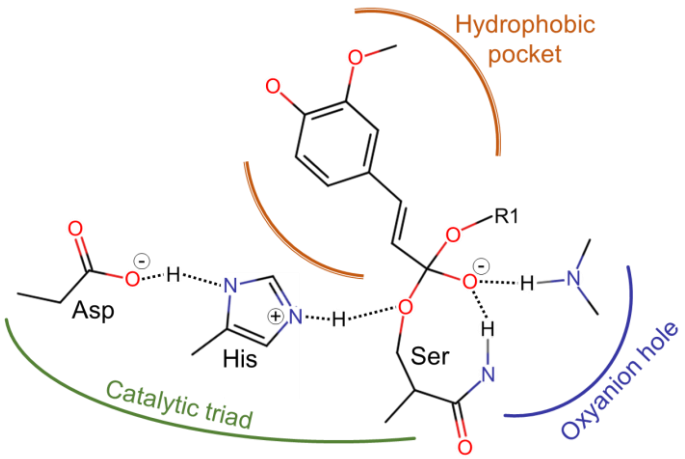

Fig. 1-9. Catalytic triad, oxyanion hole, and hydrophobic pocket in the active site of FAE.

1.7 Yeast surface display

Yeast surface display (YSD) is a powerful molecular biology technique that has revolutionized the field of protein engineering and molecular evolution [71]. It serves as a versatile platform for displaying heterologous proteins, peptides, and antibodies on the surface of yeast cells [72]. Regarding the potential to present various protein onto the yeast cell surface, YSD technique has a wide range of application in biotechnology **(Fig. 1-10)** [71]. It has been shown that this method has advanced the development of biocatalyst and holds tremendous promise for the sustainable production of valuable chemicals and biofuels from lignocellulosic biomass [73].

YSD stands out as a preferred choice in research and biotechnology due to several advantages. Firstly, yeast strains including *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Yarrowia lipolytica*, are considered "generally recognized as safe" (GRAS) by the United States Food and Drug Administration (FDA). Secondly, yeast cells possess the capability to carry out eukaryotic post-translational modifications, which is reported to promote enzyme efficiency and stability [74,75]. Thirdly, the straightforward nature of yeast cell culture and genetic manipulation ensures the correct folding and secretion of intricate and large protein structures. Lastly, YSD is compatible with flow cytometric analysis, facilitating efficient analysis and selection of displayed proteins.

YSD has been developed by continuous advances in genetic engineering strategies that fine-tune its capabilities [73]. These optimizations encompass various aspects of the yeast cell machinery, including promoters [76], signal peptides [77], protein linkers [78], anchoring proteins [79,80], yeast cell wall modification [80,81], secretory pathway enhancements [82,83], and multi-enzyme assembly [84–88].

In essence, YSD has evolved into a transformative technology at the intersection of genetic engineering, enzymatic catalysis, and bioprocessing. Its ability to tailor-make whole-cell biocatalysts for diverse applications, coupled with the innovative concept of enzyme proximity, positions YSD as a key player in the sustainable production of valuable compounds from renewable resources. As advancements in genetic engineering continue to unfold, YSD holds the promise of unlocking new frontiers in biotechnology, paving the way for greener and more efficient industrial processes.

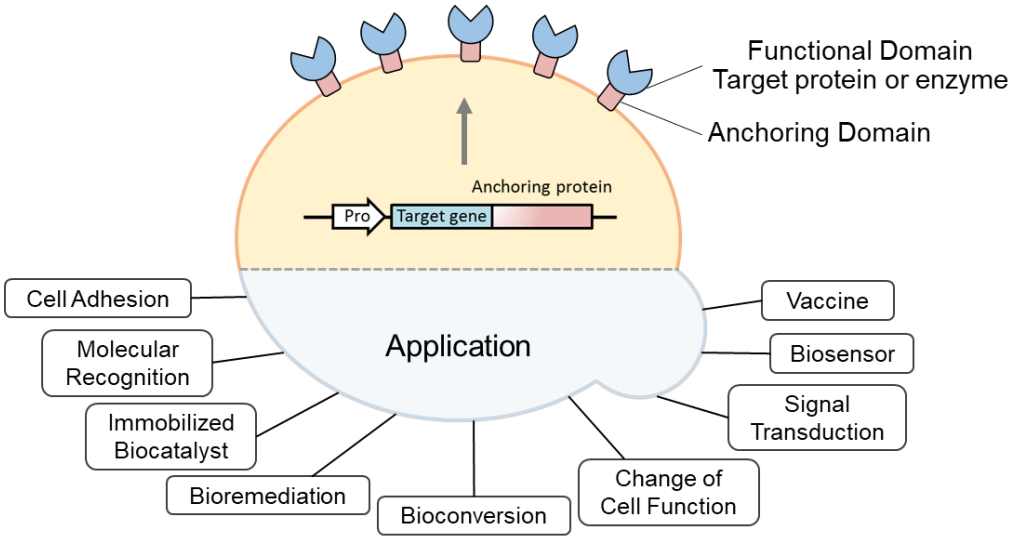

Fig. 1-10. Yeast cell surface display engineering and its application for biotechnology (Modified from Kondo and Ueda [69]).

1.8 Aim of this study

The general objective of this PhD thesis is to explore and advance my understanding of the molecular mechanisms underlying the enzymatic degradation of lignocellulosic biomass by SF6 and SF5 FAEs **(Fig. 1-5)** (alternatively, referred to as CE1_SF2 and CE1_SF5, respectively; see **Fig. 1-6**). Despite their potential for both enzymatic hydrolysis [89,90] and (trans)esterification [19–22,26], structural study of SF6 and SF5 FAEs in both free and complex forms remains unexplored. Therefore, this study aims to provide structural insights into substrate recognition, binding, and hydrolysis by these FAEs, illuminating their catalytic activities and potential for biochemical production. Concurrently, I investigate the synergistic proximity effect between XYN and FAE using YSD technology.

My key objective in the **second chapter** is to uncover the molecular mechanisms governing substrate recognition and catalysis by *As*FaeE. Utilizing X-ray crystallography, functional mutagenesis, and kinetic analysis, I aim to investigate the conformational changes occurring upon substrate binding, particularly in the loop covering the substrate-binding pocket. The study aims to identify critical amino acid residues, such as Leu150, Pro158 and Phe159, and elucidate their roles in substrate binding, substrate selectivity,

and catalytic activity. Additionally, a comparative analysis of substrate-binding mechanisms between *As*FaeE with other FAEs is undertaken to offer insights for designing more efficient FAE variants capable of processing a broader range of substrates.

The main thrust of the **third chapter** is the structural characterization of *Acremonium alcalophilum* FaeD (*Aa*FaeD; SF5). The primary goal is to provide the first structural insights into SF5 FAEs, focusing particularly on the catalytic domain of *Aa*FaeD (*Aa*FaeD-CD). This involves determining the structure of *Aa*FaeD-CD alone and in complex with FA to elucidate the substrate-binding site and the hydrophobic cleavage mechanism. The study also incorporates structure-based functional mutagenesis to understand the roles of key residues in catalytic activity. Notably, the research identifies an F120Y mutant with enhanced catalytic activity and explores the structural distinctions between *Aa*FaeD, *As*FaeE, and other FAEs, providing an explanation for why SF5 FAEs can target a broader range of substrates.

The primary focus of the **fourth chapter** is to investigate the synergistic action of cell-surface-displayed XYN and FAE on *Pichia pastoris* for improved lignocellulosic biomass decomposition. The study aims to assess the impact of the proximity effect between XYN and FAE on the same cell surface. My goal is to engineer *P. pastoris* strains capable of co-displaying both enzymes, validating the cell-surface display through immunofluorescence. The research evaluates the synergistic hydrolysis of sugarcane trash and the production of reducing sugars and FA, comparing different *P. pastoris* strains and additives like 2-mercaptoethanol. Ultimately, this study seeks to advance the efficiency of lignocellulosic biomass utilization and sustainable bioprocessing through synergistic enzyme co-display.

In summary, this PhD thesis aims to contribute significantly to the understanding of how different FAE enzymes recognize, bind to, and hydrolyze lignocellulosic substrates. Additionally, it explores strategies for enhancing the efficiency of biomass decomposition, with potential implications for biofuel and biochemical production from renewable resources.

1.9 References

[1] Panahi, H.K.S., Dehhaghi, M., Kinder, J.E., Ezeji, T.C., A review on green liquid fuels for the transportation sector: A prospect of microbial solutions to climate change, *Biofuel Research Journal*. (2019) 6 995–1024.

[2] Kurowska, K., Marks-Bielska, R., Bielski, S., Kryszk, H., Jasinskas, A., Food security in the context of liquid biofuels production, *Energies*. (2020) 13 6247.

[3] Okolie, J.A., Nanda, S., Dalai, A.K., Kozinski, J.A., Chemistry and specialty industrial applications of lignocellulosic biomass, *Waste and Biomass Valorization*. (2021) 12 2145–2169.

[4] Zoghlami, A., Paës, G., Lignocellulosic biomass: Understanding recalcitrance and predicting hydrolysis, *Frontiers in Chemistry*. (2019) 7 874.

[5] Etale, A., Onyianta, A.J., Turner, S.R., Eichhorn, S.J., Cellulose: A review of water interactions, applications in composites, and water treatment, *Chemical Reviews*. (2023) 123 2016–2048.

[6] Rao, J., Lv, Z., Chen, G., Peng, F., Hemicellulose: Structure, chemical modification, and application, *Progress in Polymer Science*. (2023) 140 101675.

[7] Alzagameem, A., Klein, S.E., Bergs, M., Do, X.T., Korte, I., Dohlen, S., Hüwe, C., Kreyenschmidt, J., Kamm, B., Larkins, M., Schulze, M., Antimicrobial activity of lignin and lignin-derived cellulose and chitosan composites against selected pathogenic and spoilage microorganisms, *Polymers*. (2019) 11 670.

[8] Giummarella, N., Pu, Y., Ragauskas, A.J., Lawoko, M., A critical review on the analysis of lignin carbohydrate bonds, *Green Chemistry*. (2019) 21 1573–1595.

[9] Sankaran, R., Markandan, K., Khoo, K.S., Cheng, C.K., Ashokkumar, V., Deepanraj, B., Show, P.L., The expansion of lignocellulose biomass conversion into bioenergy via nanobiotechnology, *Frontiers in Nanotechnology*. (2021) 3 793528.

[10] Juturu, V., Wu, J.C., Insight into microbial hemicellulases other than xylanases: A review, *Journal of Chemical Technology and Biotechnology*. (2013) 88 353–363.

[11] Østby, H., Várnai, A., Hemicellulolytic enzymes in lignocellulose processing, *Essays in Biochemistry*. (2023) 67 533–550.

[12] Braga, C.M.P., Delabona, P. da S., Lima, D.J. da S., Paixão, D.A.A., Pradella, J.G. da C., Farinas, C.S., Addition of feruloyl esterase and xylanase produced on-site improves sugarcane bagasse hydrolysis, *Bioresource Technology*. (2014) 170 316–324.

[13] Mkabayi, L., Malgas, S., Wilhelmi, B.S., Pletschke, B.I., Evaluating feruloyl esterase-xylanase synergism for hydroxycinnamic acid and xylo-oligosaccharide

production from untreated, hydrothermally pre-treated and dilute-acid pre-treated corn cobs, *Agronomy*. (2020) 10 688.

[14] Mafa, M.S., Malgas, S., Pletschke, B.I., Feruloyl esterase (FAE-1) sourced from a termite hindgut and GH10 xylanases synergy improves degradation of arabinoxylan, *AMB Express*. (2021) 11 21.

[15] Schmitz, E., Leontakianakou, S., Norlander, S., Nordberg Karlsson, E., Adlercreutz, P., Lignocellulose degradation for the bioeconomy: The potential of enzyme synergies between xylanases, ferulic acid esterase and laccase for the production of arabinoxylo-oligosaccharides, *Bioresource Technology*. (2022) 343 126114.

[16] Marasinghe, S.D., Jo, E., Hettiarachchi, S.A., Lee, Y., Eom, T.Y., Gang, Y., Kang, Y.H., Oh, C., Characterization of glycoside hydrolase family 11 xylanase from *Streptomyces* sp. strain J103; its synergetic effect with acetyl xylan esterase and enhancement of enzymatic hydrolysis of lignocellulosic biomass, *Microbial Cell Factories*. (2021) 20 129.

[17] de Oliveira, D.M., Finger-Teixeira, A., Rodrigues Mota, T., Salvador, V.H., Moreira-Vilar, F.C., Correa Molinari, H.B., Craig Mitchell, R.A., Marchiosi, R., Ferrarese-Filho, O., Dantas dos Santos, W., Ferulic acid: A key component in grass lignocellulose recalcitrance to hydrolysis, *Plant Biotechnology Journal*. (2015) 13 1224–1232.

[18] Giuliani, S., Piana, C., Setti, L., Hochkoeppler, A., Pifferi, P.G., Williamson, G., Faulds, C.B., Synthesis of pentylferulate by a feruloyl esterase from *Aspergillus niger* using water-in-oil microemulsions, *Biotechnology Letters*. (2001) 23 325–330.

[19] Antonopoulou, I., Leonov, L., Jütten, P., Cerullo, G., Faraco, V., Papadopoulou, A., Kletsas, D., Ralli, M., Rova, U., Christakopoulos, P., Optimized synthesis of novel prenyl ferulate performed by feruloyl esterases from *Myceliophthora thermophila* in microemulsions, *Applied Microbiology and Biotechnology*. (2017) 101 3213–3226.

[20] Antonopoulou, I., Iancu, L., Jütten, P., Piechot, A., Rova, U., Christakopoulos, P., Optimized enzymatic synthesis of feruloyl derivatives catalyzed by three novel feruloyl esterases from *Talaromyces wortmannii* in detergentless microemulsions, *Computational and Structural Biotechnology Journal*. (2018) 16 361–369.

[21] Antonopoulou, I., Papadopoulou, A., Iancu, L., Cerullo, G., Ralli, M., Jütten, P., Piechot, A., Faraco, V., Kletsas, D., Rova, U., Christakopoulos, P., Optimization of enzymatic synthesis of L-arabinose ferulate catalyzed by feruloyl esterases from *Myceliophthora thermophila* in detergentless microemulsions and assessment of its antioxidant and cytotoxicity activities, *Process Biochemistry*. (2018) 65 100–108.

[22] Antonopoulou, I., Dilokpimol, A., Iancu, L., Mäkelä, M.R., Varriale, S., Cerullo, G., Hüttner, S., Uthoff, S., Jütten, P., Piechot, A., Steinbüchel, A., Olsson, L., Faraco, V., Hildén, K.S., de Vries, R.P., Rova, U., Christakopoulos, P., The synthetic potential of fungal feruloyl esterases: A correlation with current classification systems and predicted structural properties, *Catalysts*. (2018) 8 242.

[23] Thörn, C., Gustafsson, H., Olsson, L., Immobilization of feruloyl esterases in mesoporous materials leads to improved transesterification yield, *Journal of Molecular Catalysis B: Enzymatic*. (2011) 72 57–64.

[24] Bonzom, C., Schild, L., Gustafsson, H., Olsson, L., Feruloyl esterase immobilization in mesoporous silica particles and characterization in hydrolysis and transesterification, *BMC Biochemistry*. (2018) 19 1–12.

[25] Grajales-Hernández, D.A., Velasco-Lozano, S., Armendáriz-Ruiz, M.A., Rodríguez-González, J.A., Camacho-Ruíz, R.M., Asaff-Torres, A., López-Gallego, F., Mateos-Díaz, J.C., Carrier-bound and carrier-free immobilization of type A feruloyl esterase from *Aspergillus niger*: Searching for an operationally stable heterogeneous biocatalyst for the synthesis of butyl hydroxycinnamates, *Journal of Biotechnology*. (2020) 316 6–16.

[26] Zerva, A., Antonopoulou, I., Enman, J., Iancu, L., Jütten, P., Rova, U., Christakopoulos, P., Optimization of transesterification reactions with CLEA-Immobilized feruloyl esterases from *Thermothelomyces thermophila* and *Talaromyces wortmannii*, *Molecules*. (2018) 23 2403.

[27] Grajales-Hernández, D., Armendáriz-Ruiz, M., Velasco-Lozano, S., López-Gallego, F., Mateos-Díaz, J.C., Chitosan-based CLEAs from *Aspergillus niger* type A feruloyl esterase: high-productivity biocatalyst for alkyl ferulate synthesis, *Applied Microbiology and Biotechnology*. (2020) 104 10033–10045.

[28] Mathew, S., Abraham, T.E., Ferulic acid: An antioxidant found naturally in plant cell walls and feruloyl esterases involved in its release and their applications, *Critical Reviews in Biotechnology*. (2004) 24 59–83.

[29] Ou, S., Kwok, K.C., Ferulic acid: Pharmaceutical functions, preparation and applications in foods, *Journal of the Science of Food and Agriculture*. (2004) 84 1261–1269.

[30] Faulds, C.B., What can feruloyl esterases do for us?, *Phytochemistry Reviews*. (2010) 9 121–132.

[31] Antonopoulou, I., Sapountzaki, E., Rova, U., Christakopoulos, P., Ferulic acid from plant biomass: A phytochemical with promising antiviral properties, *Frontiers in Nutrition*. (2022) 8 777576.

[32] Crepin, V.F., Faulds, C.B., Connerton, I.F., Functional classification of the microbial feruloyl esterases, *Applied Microbiology and Biotechnology*. (2004) 63 647–652.

[33] Benoit, I., Danchin, E.G.J., Bleichrodt, R.J., De Vries, R.P., Biotechnological applications and potential of fungal feruloyl esterases based on prevalence, classification and biochemical diversity, *Biotechnology Letters*. (2008) 30 387–396.

[34] Udatha, D.B.R.K.G., Kouskoumvekaki, I., lsson, L., Panagiotou, G., The interplay of descriptor-based computational analysis with pharmacophore modeling builds the basis for a novel classification scheme for feruloyl esterases, *Biotechnology Advances*. (2011) 29 94–110.

[35] Dilokpimol, A., Mäkelä, M.R., Aguilar-Pontes, M.V., Benoit-Gelber, I., Hildén, K.S., De Vries, R.P., Diversity of fungal feruloyl esterases: Updated phylogenetic classification, properties, and industrial applications, *Biotechnology for Biofuels*. (2016) 9 231.

[36] Drula, E., Garron, M.L., Dogan, S., Lombard, V., Henrissat, B., Terrapon, N., The carbohydrate-active enzyme database: Functions and literature, *Nucleic Acids Research*. (2022) 50 D571–D577.

[37] 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 Research*. (2013) 41 D423–D429.

[38] 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*, *Applied and Environmental Microbiology*. (2017) 83 e01251-17.

[39] Mäkelä, M.R., Dilokpimol, A., Koskela, S.M., Kuuskeri, J., de Vries, R.P., 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, *Microbial Biotechnology*. (2018) 11 869–880.

[40] Suzuki, K., Hori, A., Kawamoto, K., Thangudu, R.R., Ishida, T., Igarashi, K., Samejima, M., Yamada, C., Arakawa, T., Wakagi, T., Koseki, T., Fushinobu, S., Crystal structure of a feruloyl esterase belonging to the tannase family: A disulfide bond near a catalytic triad, *Proteins: Structure, Function and Bioinformatics*. (2014) 82 2857–2867.

[41] Dimarogona, M., Topakas, E., Christakopoulos, P., Chrysina, E.D., The crystal structure of a *Fusarium oxysporum* feruloyl esterase that belongs to the tannase family, *FEBS Letters*. (2020) 594 1738–1749.

[42] Ferousi, C., Kosinas, C., Nikolaivits, E., Topakas, E., Dimarogona, M., Crystal structure of the *Fusarium oxysporum* tannase-like feruloyl esterase FaeC in complex with p-coumaric acid provides insight into ligand binding, *FEBS Letters*. (2023) 597 1415–1427.

[43] McAuley, K.E., Svendsen, A., Patkar, S.A., Wilson, K.S., Structure of a feruloyl esterase from *Aspergillus niger*, *Acta Crystallographica Section D: Biological Crystallography*. (2004) 60 878–887.

[44] Hermoso, J.A., Sanz-Aparicio, J., Molina, R., Juge, N., González, R., Faulds, C.B., The crystal structure of feruloyl esterase A from *Aspergillus niger* suggests evolutive functional convergence in feruloyl esterase family, *Journal of Molecular Biology*. (2004) 338 495–506.

[45] Faulds, C.B., Molina, R., Gonzalez, R., Husband, F., Juge, N., Sanz-Aparicio, J., Hermoso, J.A., Probing the determinants of substrate specificity of a feruloyl esterase, *An*FaeA, from *Aspergillus niger*, *FEBS Journal*. (2005) 272 4362–4371.

[46] Gruninger, R.J., Cote, C., McAllister, T.A., Abbott, D.W., Contributions of a unique *β*-clamp to substrate recognition illuminates the molecular basis of exolysis in ferulic acid esterases, *Biochemical Journal*. (2016) 473 839–849.

[47] Prates, J.A.M., Tarbouriech, N., Charnock, S.J., Fontes, C.M.G.A., Ferreira, L.M.A., Davies, G.J., The structure of the feruloyl esterase module of xylanase 10B from *Clostridium thermocellum* provides insights into substrate recognition, *Structure*. (2001) 9 1183–1190.

[48] Schubot, F.D., Kataeva, I.A., Blum, D.L., Shah, A.K., Ljungdahl, L.G., Rose, J.P., Wang, B.C., Structural basis for the substrate specificity of the feruloyl esterase domain of the cellulosomal xylanase Z from *Clostridium thermocellum*, *Biochemistry*. (2001) 40 12524–12532.

[49] Tarbouriech, N., Prates, J.A.M., Fontes, C.M.G.A., Davies, G.J., Molecular determinants of substrate specificity in the feruloyl esterase module of xylanase 10B from *Clostridium thermocellum*, *Acta Crystallographica Section D: Biological Crystallography*. (2005) 61 194–197.

[50] Pereira, P.J.B., Royant, A., Panjikar, S., De Sanctis, D., In-house UV radiation-damage-induced phasing of selenomethionine-labeled protein structures, *Journal of Structural Biology*. (2013) 181 89–94.

[51] Nurizzo, D., Bowler, M.W., Caserotto, H., Dobias, F., Giraud, T., Surr, J., Guichard, N., Papp, G., Guijarro, M., Mueller-Dieckmann, C., Flot, D., McSweeney, S., Cipriani, F., Theveneau, P., Leonard, G.A., RoboDiff: Combining a sample changer and goniometer for highly automated macromolecular crystallography experiments, *Acta Crystallographica Section D: Structural Biology*. (2016) 72 966–975.

[52] Goldstone, D.C., Villas-Bôas, S.G., Till, M., Kelly, W.J., Attwood, G.T., Arcus, V.L., Structural and functional characterization of a promiscuous feruloyl esterase (Est1E) from the rumen bacterium *Butyrivibrio proteoclasticus*, *Proteins: Structure, Function and Bioinformatics*. (2010) 78 1457–1469.

[53] Lai, K.K., Stogios, P.J., Vu, C., Xu, X., Cui, H., Molloy, S., Savchenko, A., Yakunin, A., Gonzalez, C.F., An inserted *α*/*β* subdomain shapes the catalytic pocket of *Lactobacillus johnsonii* cinnamoyl esterase, *PLoS ONE*. (2011) 6 e23269.

[54] Zhang, H., Wen, B., Liu, Y., Du, G., Wei, X., Imam, K.M.S.U., Zhou, H., Fan, S., Wang, F., Wang, Y., Xin, F., A reverse catalytic triad Asp containing loop shaping a wide substrate binding pocket of a feruloyl esterase from *Lactobacillus plantarum*, *International Journal of Biological Macromolecules*. (2021) 184 92–100.

[55] Kasmaei, K.M., Kalyani, D.C., Reichenbach, T., Jiménez-Quero, A., Vilaplana, F., Divne, C., Crystal structure of the feruloyl esterase from *Lentilactobacillus buchneri* reveals a novel homodimeric state, *Frontiers in Microbiology*. (2022) 13 1050160.

[56] Wefers, D., Cavalcante, J.J.V., Schendel, R.R., Deveryshetty, J., Wang, K., Wawrzak, Z., Mackie, R.I., Koropatkin, N.M., Cann, I., Biochemical and structural analyses of two cryptic esterases in bacteroides intestinalis and their synergistic activities with cognate xylanases, *Journal of Molecular Biology*. (2017) 429 2509–2527.

[57] Pereira, G. V., Abdel-Hamid, A.M., Dutta, S., D'Alessandro-Gabazza, C.N., Wefers, D., Farris, J.A., Bajaj, S., Wawrzak, Z., Atomi, H., Mackie, R.I., Gabazza, E.C., Shukla, D., Koropatkin, N.M., Cann, I., Degradation of complex arabinoxylans by human colonic Bacteroidetes, *Nature Communications*. (2021) 12 459.

[58] Uraji, M., Tamura, H., Mizohata, E., Arima, J., Wan, K., Ogawa, K., Inoue, T., Hatanaka, T., Loop of *Streptomyces* feruloyl esterase plays an important role in the enzyme's catalyzing the release of ferulic acid from biomass, *Applied and Environmental Microbiology*. (2018) 84 e02300-17.

[59] Wei, X., Wang, Y.L., Wen, B.T., Liu, S.J., Wang, L., Sun, L., Gu, T.Y., Li, Z., Bao, Y., Fan, S.L., Zhou, H., Wang, F., Xin, F., The *α*-helical cap domain of a novel esterase from gut *Alistipes shahii* shaping the substrate-binding pocket, *Journal of Agricultural and Food Chemistry*. (2021) 69 6064–6072.

[60] Yang, W., Sun, L., Dong, P., Chen, Y., Zhang, H., Huang, X., Wu, L., Chen, L., Jing, D., Wu, Y., Structure-guided rational design of the *Geobacillus thermoglucosidasius* feruloyl esterase GthFAE to improve its thermostability, *Biochemical and Biophysical Research Communications*. (2022) 600 117–122.

[61] Cao, L.C., Chen, R., Xie, W., Liu, Y.H., Enhancing the thermostability of feruloyl esterase EstF27 by directed evolution and the underlying structural basis, *Journal of Agricultural and Food Chemistry*. (2015) 63 8225–8233.

[62] Holck, J., Fredslund, F., Møller, M.S., Brask, J., Krogh, K.B.R.M., Lange, L., Welner, D.H., Svensson, B., Meyer, A.S., Wilkens, C., A carbohydrate-binding family 48 module enables feruloyl esterase action on polymeric arabinoxylan, *Journal of Biological Chemistry*. (2019) 294 17339–17353.

[63] Mindrebo, J.T., Nartey, C.M., Seto, Y., Burkart, M.D., Noel, J.P., Unveiling the functional diversity of the alpha/beta hydrolase superfamily in the plant kingdom, *Current Opinion in Structural Biology*. (2016) 41 233–246.

[64] Denesyuk, A., Dimitriou, P.S., Johnson, M.S., Nakayama, T., Denessiouk, K., The acid-base-nucleophile catalytic triad in ABH-fold enzymes is coordinated by a set of structural elements, *PLoS ONE*. (2020) 15 e0229376.

[65] Liu, X., Zhou, M., Xing, S., Wu, T., He, H., Bielicki, J.K., Chen, J., Identification and biochemical characterization of a novel hormone-sensitive lipase family esterase est19 from the antarctic bacterium *Pseudomonas* sp. E2-15, *Biomolecules*. (2021) 11.

[66] Bai, Q., Cai, Y., Yang, L., Wu, Q., Su, M., Tang, L., Substitution of L133 with methionine in GXSXG domain significantly changed the activity of Penicillium expansum lipase, *Catalysis Letters*. (2022) 152 2047–2055.

[67] Dimitriou, P.S., Denesyuk, A.I., Nakayama, T., Johnson, M.S., Denessiouk, K., Distinctive structural motifs co-ordinate the catalytic nucleophile and the residues of the oxyanion hole in the alpha/beta-hydrolase fold enzymes, *Protein Science*. (2019) 28 344–364.

[68] Simón, L., Goodman, J.M., Enzyme catalysis by hydrogen bonds: The balance between transition state binding and substrate binding in oxyanion holes, *Journal of Organic Chemistry*. (2010) 75 1831–1840.

[69] Topf, M., Várnai, P., Richards, W.G., Ab initio QM/MM dynamics simulation of the tetrahedral intermediate of serine proteases: Insights into the active site hydrogen-bonding network, *Journal of the American Chemical Society*. (2002) 124 14780–14788.

[70] Silveira, R.L., Knott, B.C., Pereira, C.S., Crowley, M.F., Skaf, M.S., Beckham, G.T., Transition path sampling study of the feruloyl esterase mechanism, *Journal of Physical Chemistry B*. (2021) 125 2018–2030.

[71] Kondo, A., Ueda, M., Yeast cell-surface display - Applications of molecular display, *Applied Microbiology and Biotechnology*. (2004) 64 28–40.

[72] Fujita, Y., Takahashi, S., Ueda, M., Tanaka, A., Okada, H., Morikawa, Y., Kawaguchi, T., Arai, M., Fukuda, H., Kondo, A., Direct and efficient production

of ethanol from cellulosic material with a yeast strain displaying cellulolytic enzymes, *Applied and Environmental Microbiology*. (2002) 68 5136–5141.

[73] Teymennet-Ramírez, K. V., Martínez-Morales, F., Trejo-Hernández, M.R., Yeast surface display system: Strategies for improvement and biotechnological applications, *Frontiers in Bioengineering and Biotechnology*. (2022) 9 794742.

[74] Koseki, T., Takahashi, K., Handa, T., Yamane, Y., Fushinobu, S., Hashizume, K., N-linked oligosaccharides of *Aspergillus awamori* feruloyl esterase are important for thermostability and catalysis, *Bioscience, Biotechnology and Biochemistry*. (2006) 70 2476–2480.

[75] Bonzom, C., Hüttner, S., Mirgorodskaya, E., Chong, S.L., Uthoff, S., Steinbüchel, A., Verhaert, R.M.D., Olsson, L., Glycosylation influences activity, stability and immobilization of the feruloyl esterase 1a from *Myceliophthora thermophila*, *AMB Express*. (2019) 9.

[76] Inokuma, K., Hasunuma, T., Kondo, A., Efficient yeast cell-surface display of exo- and endo-cellulase using the SED1 anchoring region and its original promoter, *Biotechnology for Biofuels*. (2014) 7 8.

[77] Kajiwara, K., Aoki, W., Ueda, M., Evaluation of the yeast surface display system for screening of functional nanobodies, *AMB Express*. (2020) 10 51.

[78] Lown, P.S., Cai, J.J., Ritter, S.C., Otolski, J.J., Wong, R., Hackel, B.J., Extended yeast surface display linkers enhance the enrichment of ligands in direct mammalian cell selections, *Protein Engineering, Design and Selection*. (2021) 34 1–9.

[79] Phienluphon, A., Mhuantong, W., Boonyapakron, K., Deenarn, P., Champreda, V., Wichadakul, D., Suwannarangsee, S., Identification and evaluation of novel anchoring proteins for cell surface display on *Saccharomyces cerevisiae*, *Applied Microbiology and Biotechnology*. (2019) 103 3085–3097.

[80] Inokuma, K., Kitada, Y., Bamba, T., Kobayashi, Y., Yukawa, T., Riaan Den Haan, &, Heber Van Zyl, W., Kondo, A., Hasunuma, T., Improving the functionality of surface-engineered yeast cells by altering the cell wall morphology of the host strain, *Applied Microbiology and Biotechnology*. (2021) 105 5895–5904.

[81] Arnthong, J., Ponjarat, J., Bussadee, P., Deenarn, P., Prommana, P., Phienluphon, A., Charoensri, S., Champreda, V., Zhao, X.Q., Suwannarangsee, S., Enhanced surface display efficiency of *β*-glucosidase in *Saccharomyces cerevisiae* by disruption of cell wall protein-encoding genes YGP1 and CWP2, *Biochemical Engineering Journal*. (2022) 179 108305.

[82] Tang, H., Song, M., He, Y., Wang, J., Wang, S., Shen, Y., Hou, J., Bao, X., Engineering vesicle trafficking improves the extracellular activity and surface

display efficiency of cellulases in *Saccharomyces cerevisiae*, *Biotechnology for Biofuels*. (2017) 10 53.

[83] Arnthong, J., Bussadee, P., Phienluphon, A., Deenarn, P., Tulsook, K., Plupjeen, S.N., Siamphan, C., Tachaapaikoon, C., Champreda, V., Suwannarangsee, S., Overexpression of *LAS21* in cellulase-displaying *Saccharomyces cerevisiae* for high-yield ethanol production from pretreated sugarcane bagasse, *Fermentation*. (2022) 8 652.

[84] Bae, J., Kuroda, K., Ueda, M., Proximity effect among cellulose-degrading enzymes displayed on the *Saccharomyces cerevisiae* cell surface, *Applied and Environmental Microbiology*. (2015) 81 59–66.

[85] Smith, M.R., Gao, H., Prabhu, P., Bugada, L.F., Roth, C., Mutukuri, D., Yee, C.M., Lee, L., Ziff, R.M., Lee, J.K., Wen, F., Elucidating structure–performance relationships in whole-cell cooperative enzyme catalysis, *Nature Catalysis*. (2019) 2 809–819.

[86] Fan, S., Liang, B., Xiao, X., Bai, L., Tang, X., Lojou, E., Cosnier, S., Liu, A., Controllable display of sequential enzymes on yeast surface with enhanced biocatalytic activity toward efficient enzymatic biofuel cells, *Journal of the American Chemical Society*. (2020) 142 3222–3230.

[87] Guo, F., Liu, M., Liu, H., Li, C., Feng, X., Direct yeast surface codisplay of sequential enzymes with complementary anchor motifs: enabling enhanced glycosylation of natural products, *ACS Synthetic Biology*. (2023) 12 460–470.

[88] Liu, Z., Inokuma, K., Ho, S.H., Haan, R. Den, Hasunuma, T., Van Zyl, W.H., Kondo, A., Combined cell-surface display- and secretion-based strategies for production of cellulosic ethanol with *Saccharomyces cerevisiae*, *Biotechnology for Biofuels*. (2015) 8 162.

[89] Underlin, E.N., Frommhagen, M., Dilokpimol, A., van Erven, G., de Vries, R.P., Kabel, M.A., Feruloyl esterases for biorefineries: Subfamily classified specificity for natural substrates, *Frontiers in Bioengineering and Biotechnology*. (2020) 8 332.

[90] Dilokpimol, A., Verkerk, B., Li, X., Bellemare, A., Lavallee, M., Frommhagen, M., Underlin, E.N., Kabel, M.A., Powlowski, J., Tsang, A., de Vries, R.P., Screening of novel fungal carbohydrate esterase family 1 enzymes identifies three novel dual feruloyl/acetyl xylan esterases, *FEBS Letters*. (2022) 596 1932–1943.

Chapter Two

Structural insights into the molecular mechanisms of substrate recognition and hydrolysis by feruloyl esterase from *Aspergillus sydowii*

Reproduced from Ref. *Int. J. Biol. Macromol*., 2023, **253**, 127188 with permission from the Elsevier

2.1 Introduction

Lignocellulosic biomass, derived from plant cell walls, is a renewable resource with significant potential for producing second-generation biofuels and bioproducts without impacting global food security [1]. However, plant cell walls, comprising cellulose, hemicellulose, and lignin, exhibit recalcitrant characteristics [2]. Focusing on polysaccharides, cellulose forms a linear homopolysaccharide consisting of glucose. Hemicelluloses constitute complex, branched heteropolysaccharides comprising hexoses (glucose, mannose, and galactose), pentoses (xylose and arabinose), and various modification groups (ferulic, *p*-coumaric, sinapic, caffeic, and acetic acids) [3]. Efficient and gentle decomposition of lignocellulosic biomass is crucial to acquire valuable key compounds for producing value-added bioproducts.

Due to its intricate structure, the degradation of hemicellulose necessitates a range of enzymes, including glycosyl hydrolases (xylanases, *β*-xylosidases, *α*-arabinofuranosidases, and glucuronidases) and esterases (feruloyl, acetyl xylan, and glucuronoyl esterases) [4]. Feruloyl esterases (FAEs) are enzymes that play a pivotal role in the breakdown of plant biomass and find applications in various industries. FAEs (Enzyme Commission number: EC 3.1.1.73) belong to a subclass of carboxylic ester hydrolases and have the capacity to liberate ferulic acid (FA), sinapic acid (SA), caffeic acid (CA), *p*-coumaric acid (*p*CA), and dimeric FA from natural substrates [5,6]. Within plant cell walls, FAEs serve as auxiliary enzymes, aiding the access of other enzymes to their specific action sites. Recent studies have highlighted the synergistic effect of xylanase and FAE, enhancing the hydrolysis of agricultural biomasses that have

undergone various pretreatment methods [7–9]. For instance, sugarcane bagasse has been subjected to steam explosion and hydrothermal treatment [7], corn cobs to hydrothermal and dilute-acid treatment [8], and oat bran fiber and oat hull to ultrasound-assisted alkali treatment [9]. Consequently, numerous fungal FAEs have been identified and harnessed in diverse biotechnological processes. These applications span industries ranging from pulp and paper to feed, seasonings, alcoholic beverages, and pharmaceuticals [10,11].

Initially, microbial FAEs were classified into four subclasses (A, B, C, and D) based on sequence similarity and their specific activities towards model substrates such as methyl hydroxycinnamates (e.g., methyl esters of FA, SA, CA, and *p*CA) and diferulate [12]. However, with the rapid expansion in the number of sequenced FAEs, particularly from fungal sources, the classification of FAEs has undergone a comprehensive update, resulting in the identification of 13 subfamilies (SFs) [11]. This revised classification suggests that fungal FAEs originate from the highly divergent fungal esterase family, encompassing various subfamilies like tannase (SF1–4, SF9–11), acetyl xylan esterase (AXE; SF5–6), lipase (SF7), and lipase-choline esterase (SF8, SF12–13) [11]. Among the array of fungal FAEs, those belonging to SF6 exhibit promising potential for synthesizing bioactive compounds due to their notably high transesterification activity [13,14].

Previously, fungal FAEs within SF6 have been shown to cleave ester bonds between FA and hemicellulose, releasing FA [6,15–17]. However, the mechanisms through which SF6 FAEs recognize substrates and catalyze hydrolysis remain elusive. Of particular interest is the SF6 FAE derived from *Aspergillus sydowii* (*As*FaeE), which exhibits an enhanced hydrolytic activity towards ester bonds in diverse substrates when compared to SF6 FAEs found in other fungal species [13,18]. *As*FaeE is recognized for its broad substrate specificity, particularly in relation to methyl hydroxycinnamates, and its noteworthy hydrolytic activity towards natural substrates [6,18].

I present the pioneering crystal structures of *As*FaeE, both in its unbound form and in complex with FA or SA **(Fig. 2-1)**. These structures represent the first-ever three-dimensional insights into SF6 fungal FAEs. The structures reveal a conformational change in the loop covering the substrate-binding pocket upon binding, along with the precise binding modes of the MFA and MSA within this pocket. Through mutational

analysis of *As*FaeE, I reveal the roles of specific residues in the substrate-binding pocket and observe intriguing spatial adaptations, potentially guiding the design of tailored ferulate derivatives. These derivatives, tailored for positions 5, 2, or both, offer prospects for optimized hydrolysis or transesterification. The crystal structures of phylogenetically closely related enzymes were determined for fungal FAEs belonging to other subfamilies: *Aspergillus niger* FaeA (*An*FaeA) (SF7) [19,20] and *Anaeromyces mucronatus* Fae1A (*Am*Fae1A) (SF8) [21] pursued by other research groups. My study points towards the possibility of engineering *An*FaeA to expand its substrate range. These structural insights serve as a guide for selecting enzymes for efficient agricultural biomass utilization, ushering advancements in biotechnological applications.

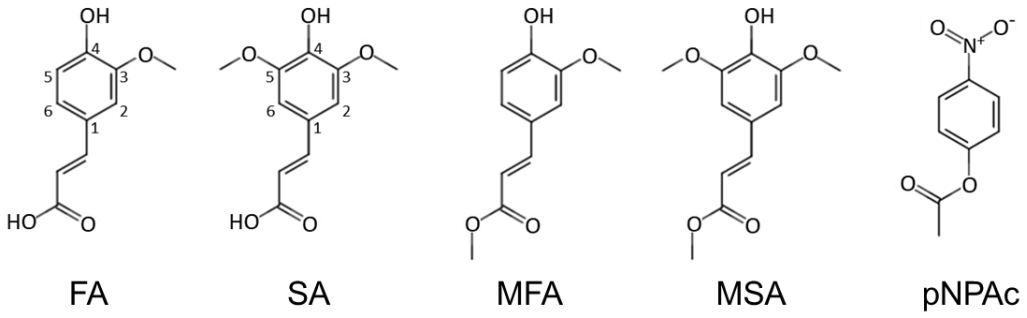

Fig. 2-1. Chemical structures of ferulic acid (FA), sinapic acid (SA), methyl ferulate (MFA), methyl sinapate (MSA), and *p*-nitrophenyl acetate (*p*NPAc).

2.2 Results and discussions

2.2.1 Sequence analysis

*As*FaeE, registered under the UniProtKB/TrEMBL database (primary accession: A0A1L9T9J3) had been previously categorized as a type C or D FAE, based on its substrate specificity [12]. Within the ESTerases and *α*/*β*-Hydrolase Enzymes and Relatives (ESTHER) database, *As*FaeE was designated as a member of the esterase_phb family, situated within the *α*/*β*-hydrolase superfamily [22]. More recently, through a comprehensive phylogenetic analysis (as depicted in **Fig. 2-2** and detailed in **Table 2-1**), *As*FaeE was classified into SF6 subgroup of fungal FAEs [11,18]. To further elucidate its relationship with other proteins, a BLAST search was conducted against the entire repertoire of protein sequences within the Protein Data Bank (PDB) [23]. Among the identified matches, *Al*AXE (PDB: 5X6S), an SF6 member **(Fig. 2-2)** and a constituent of the esterase_phb family in the ESTHER database, demonstrated the highest sequence identity with *As*FaeE displaying 42.8% similarity across 99% of the constituent residues.

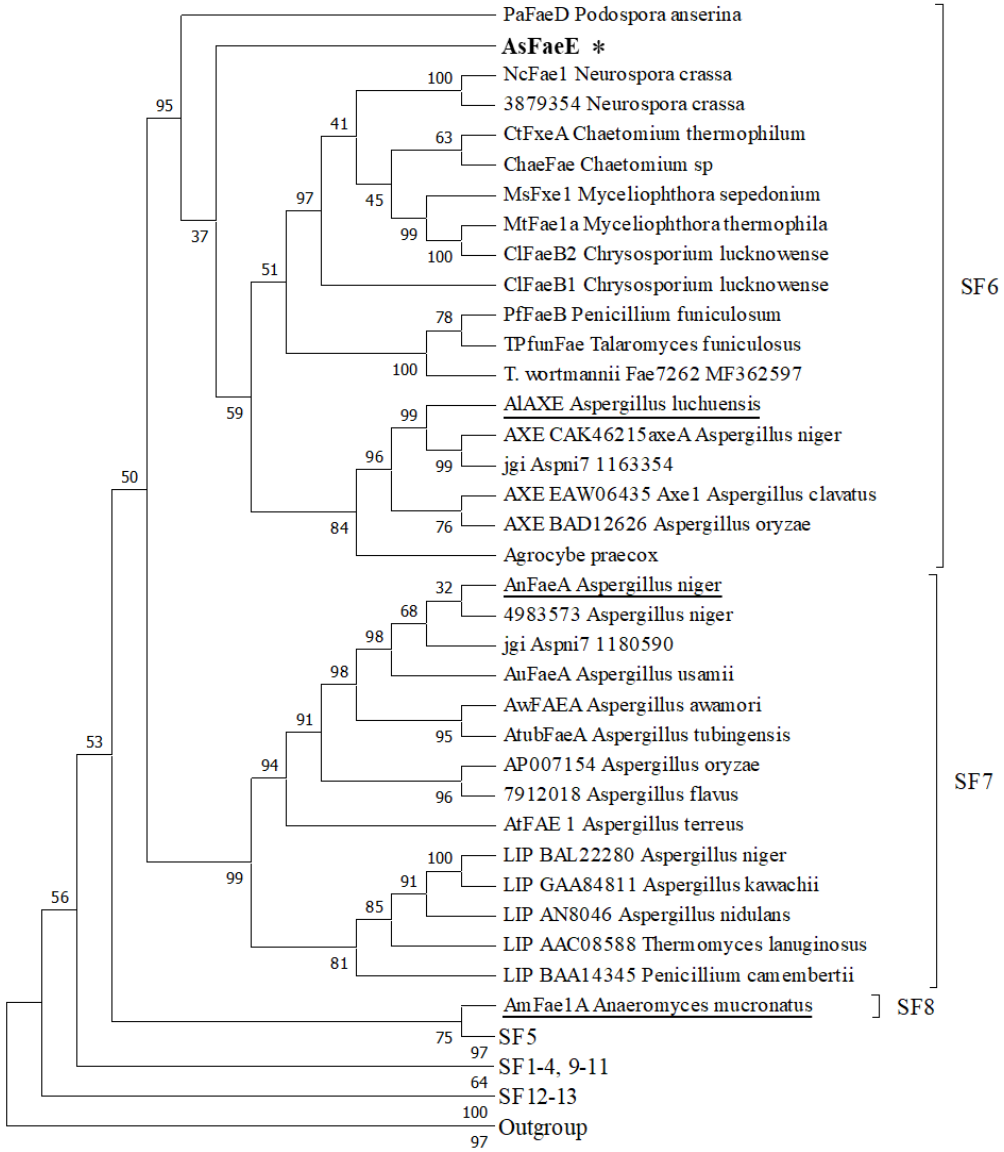

Fig. 2-2. Phylogenetic relationships of feruloyl esterases (FAEs). FAEs were classified into 13 subfamilies (SFs) defined by Benoit *et al*., 2008 [10] and Dilokpimol *et al*., 2016 [11]. Numbers indicate the statistical values (Expected-Likelihood Weight) of internal nodes, given in percentages. *As*FaeE and the FAEs registered in the protein data bank are indicated by an asterisk and underlines, respectively.

Table 2-1. The classification of *As*FaeE and related proteins.

Fungal species	Enzyme	Sequence identity (%)[a]	Classification		PDB ID
			FAE SF[b]	ESTHER family[c]	
Aspergillus sydowii	*As*FaeE	-	6	Esterase_phb	8IY8 (this study)
Aspergillus luchensis	*Al*AXE	42.8	6	Esterase_phb	5X6S
Anaeromyces mucronatus	*Am*Fae1A	24.0	8	Antigen 85c	5CXU
Aspergillus niger	*An*FaeA	21.8	7	Lipase_3	1UZA

[a]Sequence identity with *As*FaeE. [b]Based on [11]. [c]Based on [22].

2.2.2 Preparation of active AsFaeE constructs

Recombinant *As*FaeE (281 residues) comprises an N-terminal 6× histidine tag and the full-length *As*FaeE (275 residues); a signal sequence (18 residues) was omitted in the construct. The theoretical molecular mass (*m*) of the recombinant *As*FaeE was 30,800 Da. *N*- and *O*-glycosylation sites of *As*FaeE were predicted with NetNGlyc 1.0 [24] and NetOGlyc 4.0 [25], respectively; for *N*-glycosylation sites, Asn189 and Asn225, and for an *O*-glycosylation site, Thr160 were obtained. The recombinant *As*FaeE was expressed in *P. pastoris* X-33. SDS-PAGE analysis indicated that *As*FaeE contains various species in a *m* range of 37,000–54,000 Da **(Fig. 2-3)**. *N*-deglycosylation treatment of *As*FaeE with Endo H_f resulted in a single *m* species of ca. 36,500 Da **(Fig. 2-3)**. In contrast, *O*-deglycosylation treatment of *As*FaeE with *O*-glycosylase had no effect on the SDS-PAGE band pattern **(Fig. 2-3)**. Due to the relatively small size of *O*-glycosylation, I was unable to conclusively determine whether *As*FaeE was *O*-glycosylated or not. Activity assaying showed that *As*FaeE exhibits activity towards model substrates for FAE, MFA and MSA, and one for AXE, *p*NPAc **(Fig. 2-1)**. The activities of glycosylated and *N*-deglycosylated *As*FaeE were examined and confirmed to be the same **(Fig. 2-4)**.

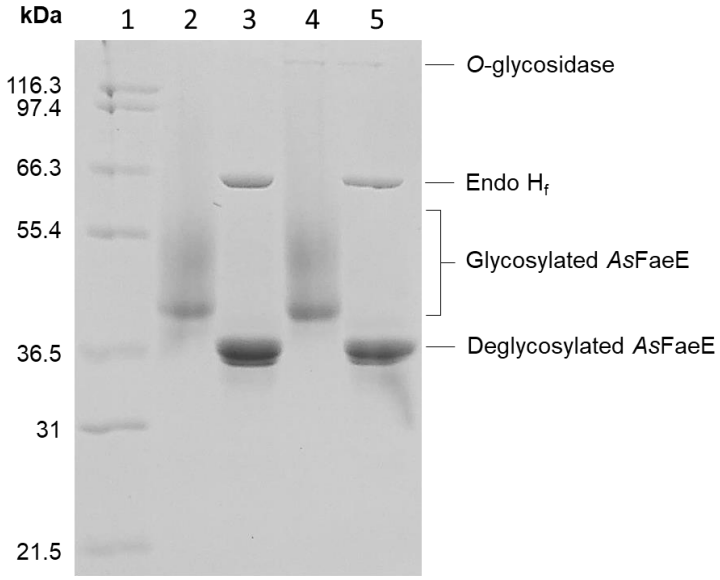

Fig. 2-3. SDS-PAGE analysis of *N*- and *O*-glycosylation of the purified recombinant *As*FaeE, expressed in *Pichia pastoris*. Molecular-mass markers (lane 1), *As*FaeE (lane 2), and *As*FaeE treated with Endo H_f (lane 3), with *O*-glycosidase (lane 4), and with both Endo H_f and *O*-glycosidase (lane 5). The band positions of *As*FaeE before (glycosylated *As*FaeE) and after deglycosylation (deglycosylated *As*FaeE), Endo H_f, and *O*-glycosidase are indicated.

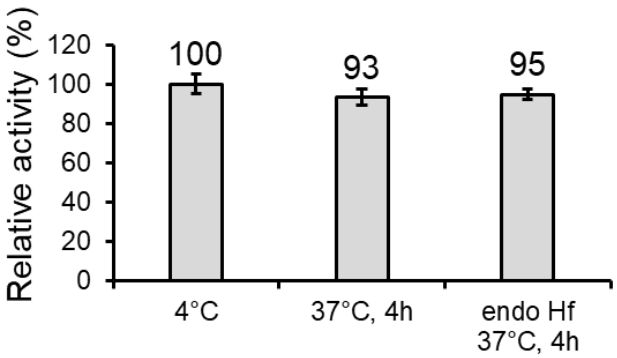

Fig. 2-4. Relative activity of non-deglycosylated and *N*-deglycosylated *As*FaeE determined after incubation for 4 h at temperature 37 °C.

2.2.3 Structure determination

The crystal structure of *As*FaeE was determined using recombinant *As*FaeE treated with Endo H_f. The *As*FaeE crystals with a blade shape were grown within 3 days under the optimized condition: 0.2 M ammonium acetate, 0.1 M sodium acetate (pH 4.6), and 30% (w/v) PEG4000 **(Fig. 2-5 A)**. The crystal structure of *As*FaeE in the free form was determined at 1.50 Å resolution via molecular replacement using *Al*AXE (PDB: 5X6S) as a search model. Iterative model building and refinement were utilized to enhance the quality of the electron density map. The electron density for residues 19–293 of *As*FaeE (residue numbering corresponds to *As*FaeE; UniProtKB/TrEMBL (A0A1L9T9J3)) was unambiguously identified. In chain B, electron density for two *N*-acetylglucosamine (GlcNAc) residues attached to Asn189 and Asn225 was evident, while in chain A, only one GlcNAc residue at Asn189 was found due to ambiguous electron density at GlcNAc near Asn225. The presence of electron density for two GlcNAc residues attached to Asn189 and Asn225 was expected, as Endo H_f cleaves the bond between the first and the second GlcNAc residues proximal to the asparagine. The electron density for the N-terminal 6×His tag was ambiguous. The structure was refined to a final R factor of 11.69% with R free = 16.35% **(Table 2-2)**. The crystal structure contained two *As*FaeE molecules in an asymmetric unit **(Fig. 2-6 A)**. I confirmed the monomeric state of *As*FaeE in solution through calibrated gel filtration **(Fig. 2-7)**.

To determine the crystal structure of *As*FaeE in complex with either FA or SA, I initially attempted the soaking method. However, this resulted in the disruption of the crystals. Consequently, I employed the co-crystallization approach and obtained the prism-shaped crystals of *As*FaeE co-crystallized with FA or SA, as illustrated in **Figs. 2-5 B and C**. The crystal structures of the *As*FaeE/FA and *As*FaeE/SA complexes were determined at 1.50 and 1.55 Å resolution, respectively, using the crystal structure of *As*FaeE determined above as a search model via molecular replacement. Both complex structures contained two molecules of the complex in an asymmetric unit. The electron density for residues 19–293 of *As*FaeE, FA, SA, and two GlcNAc residues attached to Asn189 and Asn225 was identified unambiguously. The structures were refined to a final R factor of 12.15% with R free = 16.88% for the *As*FaeE/FA complex and to a final R factor of 13.39% with R free = 18.57% for the *As*FaeE/SA complex **(Table 2-2)**.

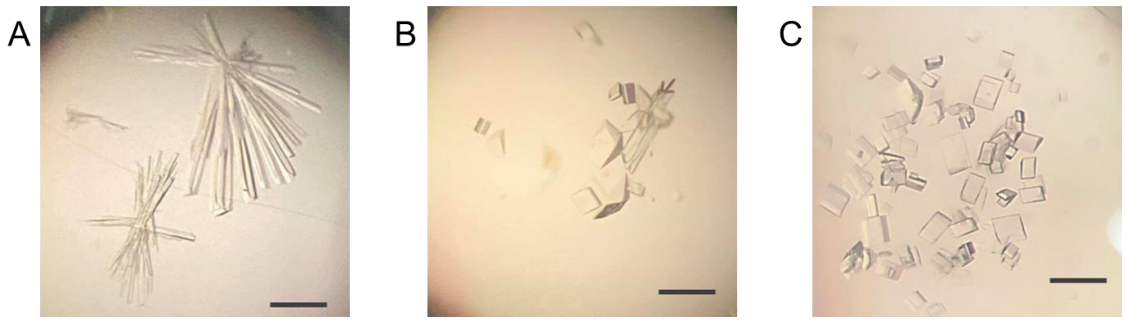

Fig. 2-5. Crystals of *As*FaeE in the absence of any ligands (A), in the presence of ferulic acid (B), and sinapic acid (C). These crystals were obtained under the optimized condition containing 0.2 M ammonium acetate, 0.1 M sodium acetate (pH 4.6), and 30% (w/v) PEG4000. Scale bars represent 0.5 mm.

Table 2-2. Crystallographic data and refinement statistics for *As*FaeE and *As*FaeE in a complex with either ferulic acid (*As*FaeE/FA) or sinapic acid (*As*FaeE/SA).

	***As*FaeE**	***As*FaeE/FA**	***As*FaeE/SA**
PDB ID	**8IY8**	**8IYB**	**8IYC**
A. Diffraction data			
X-ray source	SPring-8/BL26B1	SPring-8/BL26B1	SPring-8/BL26B1
Detector	Dectris Eiger X 4M	Dectris Eiger X 4M	Dectris Eiger X 4M
Wavelength (Å)	1.0	1.0	1.0
Resolution range (Å)*	50.00–1.50 (1.59–1.50)	50.00–1.50 (1.59–1.50)	50.00–1.50 (1.64–1.55)
Space group	$P2_1\,2_1\,2_1$	$P2_1$	$P2_1$
Unit cell (Å)	a = 48.1, b = 75.6, c = 135.6 $\alpha = \beta = \gamma = 90.0°$	a = 37.3, b = 65.5, c = 95. 3, $\alpha = \gamma = 90.0°$, $\beta = 96.7°$	a = 37.4, b = 65.2, c = 95.4 $\alpha = \gamma = 90.0°$, $\beta = 97.0°$
Unique reflections*	79279 (11927)	71862 (10511)	65419 (10051)
Multiplicity*	7.03 (5.09)	4.38 (3.22)	4.51 (4.01)
Completeness (%)*	98.8 (93.3)	98.2 (89.5)	98.2 (93.5)
Mean I/s(I)*	31.45 (4.87)	27.35 (5.89)	26.05 (4.61)
Wilson B-factor (Å^2)	24.8	23.8	27.8
R_{merge} (%)*	3.3 (25.3)	3.0 (15.3)	2.9 (23.8)
R_{meas} (%)*	3.6 (28.1)	3.3 (18.1)	3.3 (27.2)
$CC_{1/2}$ (%)*	100 (95.0)	100 (97.8)	100 (95.5)
B. Structure determination			
Search model in PDB	5X6S	8IY8	8IY8
C. Refinement statistics			
Resolution range*	39.25–1.50 (1.52–1.50)	47.32–1.50 (1.52–1.50)	32.63–1.55 (1.57–1.55)
R_{work} (%)*	11.69 (14.58)	12.15 (13.38)	13.39 (22.18)
R_{free} (%)*	16.35 (21.37)	16.88 (22.42)	18.57 (28.85)
Number of protein residues	278 × 2	275 × 2	275 × 2
ACT/EDO/FER/SIN/NAG/water	6/0/0/0/3/740	5/6/2/0/4/498	8/13/0/2/4/378
RMSD, bond lengths (Å)	0.005	0.005	0.006
RMSD, bond angles (°)	0.740	0.754	0.801
Coordinate error (Å)	0.14	0.13	0.17
Ramachandran favored (%)	97.83	97.99	97.99
Ramachandran outliers (%)	0.00	0.00	0.00
Rotamer outliers (%)	1.43	0.42	0..65
Clash score	2.99	1.85	3.63
Subunit rmsd (Å)	0.58	0.21	0.20

*Values in parentheses are for the highest resolution shell.

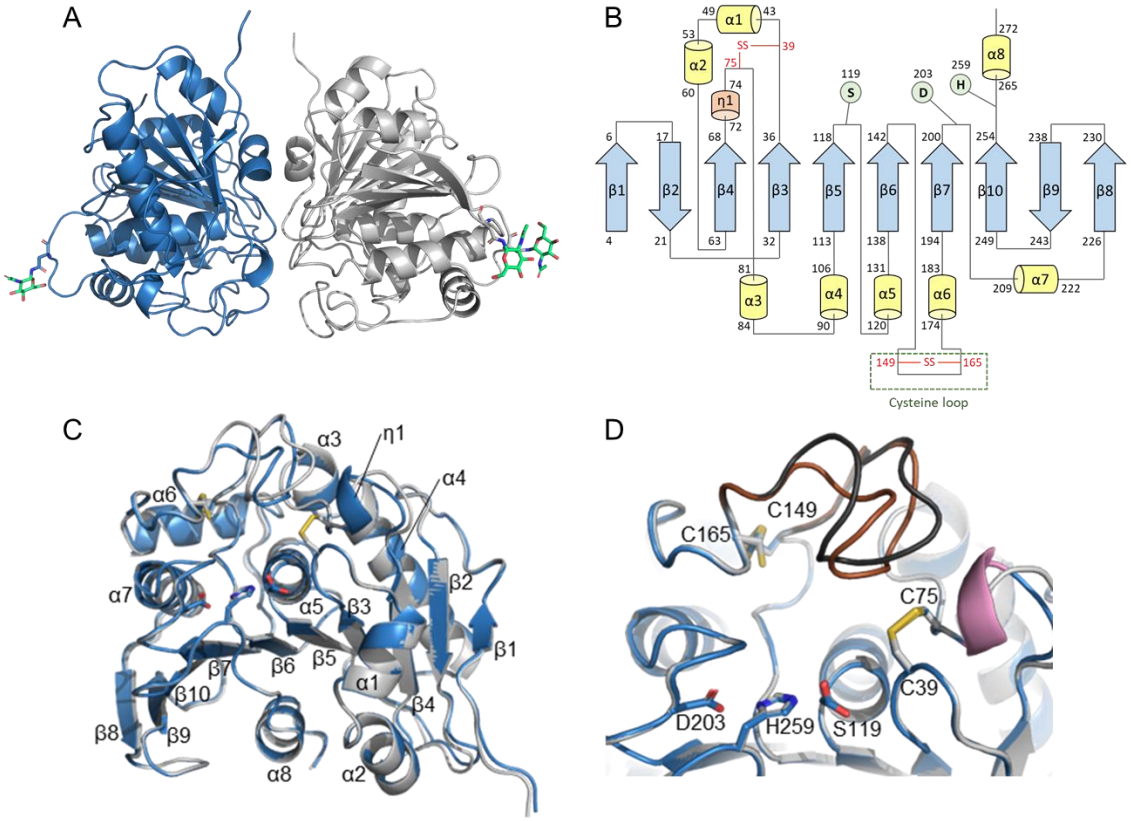

Fig. 2-6. Crystal structures of substrate-free *As*FaeE. (A) Ribbon representation of two *As*FaeE molecules, *As*FaeE chains A (blue) and B (grey), in an asymmetric unit. GlcNAc residues of *N*-glycans are shown as sticks. The carbon, nitrogen, and oxygen atoms are colored green, blue, and red, respectively. (B) Topological diagram of the *As*FaeE structure. The α-helix (α1–8), β-strands (β1–10), and 3_{10} helix (η1) are colored yellow, blue, and orange, respectively. The residues of the S-D-H catalytic triad, disulfide bonds, and the cysteine loop are indicated as green circles, red lines, and a dotted box, respectively. (C) Superimpositioning of *As*FaeE chains A (blue) and B (grey). (D) Close-up view of the S-D-H catalytic triad and cysteine loop formed between Cys149 and Cys165. The cysteine loops of *As*FaeE chains A and B, shown as brown and black ribbon representation, respectively, are poorly superimposed. The 3_{10} helix shown in pink was formed only in chain A. The residues comprising the S-D-H catalytic triad and two disulfide bonds are shown as sticks. The nitrogen, oxygen, and sulfur atoms are colored blue, red and yellow, respectively.

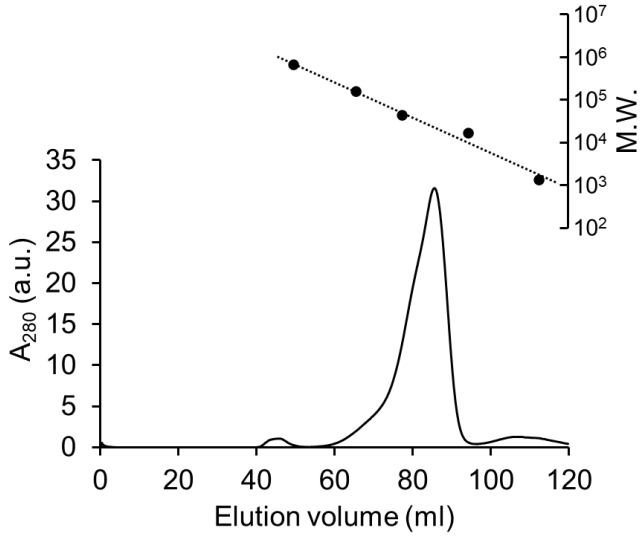

Fig. 2-7. Gel filtration chromatography of the *N*-deglycosylated *As*FaeE using HiLoad 16/600 Superdex 75 column (GE Healthcare). The column was equilibrated with 25 mM potassium phosphate, pH 6.5. The standard curve for the molecular mass was produced using gel filtration standard (BIO-RAD) containing vitamin B12 (1,350 Da), myoglobin (17,000 Da), ovalbumin (44,000 Da), γ-globulin (158,000 Da), and thyroblobulin (670,000 Da).

2.2.4 Overview of the free state

The crystal structure of *As*FaeE is illustrated in **Fig. 2-6**. The structures of the two molecules (chain A and B) in the asymmetric unit superimposed very well **(Fig. 2-6 C)**. When superimposing 275 Cα atoms of all residues between chains A and B using the program ChimeraX [26], the root mean square deviation (RMSD) was 0.58 Å. Particularly, 269 out of the 275 Cα atoms were superimposed very well with an RMSD of 0.33 Å; six Cα atoms were omitted automatically with an RMSD cutoff of 2.0 Å by the program. **Fig. 2-6 B** presents the topological diagram of *As*FaeE, which contains ten β-strands (β1–β10), eight α-helices (α1–α8), and one 3_{10} helix (η1). The β-strands form a continuous β-sheet, consisting of both parallel and anti-parallel portions. Specifically, a six-stranded parallel β-sheet portion (β4-β3-β5-β6-β7-β10) is flanked by two three-stranded anti-parallel β-sheet portions (β1-β2-β4 and β10-β9-β8) **(Fig. 2-6 B)**. The arrangement of α-helices around this β-sheet forms an α-β-α sandwich structure **(Fig. 2-**

6 B), resembling the canonical *α/β*-hydrolase fold [27,28]. Additionally, two disulfide bonds (Cys39−Cys75 and Cys149−Cys165) were identified **(Fig. 2-6 B)**. A classical Ser-Asp-His (S-D-H) catalytic triad is present in *As*FaeE, which comprises Ser119, Asp203, and His259 **(Fig. 2-6 B)**.

Intriguingly, I observed some differences in the regions near the S-D-H catalytic triad of chains A and B **(Figs. 2-6 C and D)**. Firstly, η1 was present only in chain A and absent in chain B. Because this region in chain B was involved in crystal packing, I assume η1 is formed in solution **(Fig. 2-6 D)**. Secondly, the loop between Cys149 and Cys165 did not superimpose well. This loop was previously termed the "cysteine loop" [29] and has been shown to be involved in substate binding.

2.2.5 Overview of the complex states

In both the *As*FaeE/FA and *As*FaeE/SA complexes, I obtained structures of two molecules (chains A and B) in the asymmetric unit. The C*α* superimpositions for all residues in chains A and B for both complexes resulted in RMSDs of 0.21 and 0.20 Å, respectively, indicating that chains A and B are identical in both complexes. The crystal structures of the *As*FaeE/FA (chain A) and *As*FaeE/SA (chain A) complexes are illustrated in **Figs. 2-8 A and D**, respectively. The C*α* superimpositions of all residues in chain A for both complexes resulted in RMSDs of 0.12 Å, indicating that they share the same overall structure. Detailed views of the substrate-binding sites of the *As*FaeE/FA and *As*FaeE/SA complexes, along with the omit map of the FA and SA, are shown in **Figs. 2-8 B and E**, respectively. The structural models of the *As*FaeE/MFA and *As*FaeE/MSA complexes are also depicted in **Figs. 2-8 C and F**, respectively. Importantly, all atoms, excluding the methyl group connected by the ester bond of MFA and MSA, exhibited excellent superimposition. This observation indicates that the structures of the *As*FaeE/FA and *As*FaeE/MFA complexes, as well as those of the *As*FaeE/SA and *As*FaeE/MSA complexes within the ferulic acid and sinapic acid moieties, respectively, are comparable, if not identical. Consequently, all ensuing discussions are based on the structures of the *As*FaeE/FA and *As*FaeE/SA complexes, unless explicitly stated

otherwise. The cysteine loop and the S-D-H catalytic triad are located on the upper and lower sides of the FA or SA, respectively **(Figs. 2-8 B, C, E, and F)**.

Figs. 2-8 H and I, along with **Figs. 2-9 A and B**, display the *As*FaeE residues involved in binding with FA and SA. These residues were chosen based on their sidechain atoms being within a 5 Å proximity to either FA or SA. The bound FA or SA has a planar structure with an extended conjugate π-system and is surrounded by hydrophobic residues (such as Cys39, Leu150, Cys165, and Val205) that engage in hydrophobic interactions. Notably, the sidechain amide group of Gln163 forms a hydrogen bond with a phenolic hydroxyl group of FA or SA. Interestingly, the five-membered ring of Pro158, positioned within the β-turn structure of the cysteine loop and adopting a *cis* conformation, interacts with the aromatic moiety of FA or SA through stacking. The sidechain conformation of Pro158 in FA and SA complexes is Cγ-exo, whereas in the ligand-free *As*FaeE, it is Cγ-endo, indicating that Cγ-exo is favourable for ligand interaction. Furthermore, the aromatic moiety of Phe159 stacks onto the five-membered ring of Pro158, forming three layers; the distance between the rings of FA (or SA) and Pro158, as well as that between the rings of Pro158 and Phe159, measures ~4.3 and ~3.9 Å, respectively **(Figs. 2-8 H and I)**. The three layers are probably formed through CH-π interactions.

Oxyanion holes, commonly found in α/β-hydrolases (e.g., esterases), are pockets in the active sites that stabilize the negative charge on a deprotonated oxygen or alkoxide of the transition state [30,31]. In the case of *As*FaeE, the oxyanion hole is formed by the mainchain amide groups of Cys39 and Ser120 with a carboxylic group of FA or SA occupying the oxyanion hole **(Figs. 2-8 H and I, and 2-9 A and B)**.

The difference in chemical structures between FA and SA lies in the number of methoxy groups in their phenolic rings; FA has one methoxy group, while SA has two. In the *As*FaeE/FA complex structure, Leu150 adopts a χ2 angle of ~115°, allowing the δ1 methyl group to interact with position 5 of FA through van der Waals interactions. On the other hand, in the *As*FaeE/SA complex structure, Leu150 was found to adopt two sidechain conformers, with the χ2 angle being either ~115° or –60°. Because SA has a methoxy group at position 5, a χ2 angle of –60° appears to be more suitable to avoid steric clash. With a χ2 angle of –60°, the methyl groups of Leu150 orient away from the methoxy group at position 5 of SA **(Fig. 2-8 I)**, allowing SA to fit into the substrate-

binding site. Consistent with these observations, the cavity volumes of the substrate-binding pocket for the FA-bound *As*FaeE (the Leu150 χ2 angle of ~115°) and that for the SA-bound *As*FaeE (the Leu150 χ2 angle of –60°) were calculated to be 49.88 and 55.82 $Å^3$, respectively, using CASTp [32]. This suggests that the conformation of the sidechain of Leu150 changes correspondingly with the presence of the methoxy group at position 5 of the phenolic moiety of the substrates.

The residues involved in FA- and SA-binding are also shown for the structure of the substrate-free *As*FaeE in **Fig. 2-8 E**. The Cα superimpositions of all residues between the free form and the *As*FaeE/FA and *As*FaeE/SA complexes resulted in RMSDs of 0.38 and 0.35 Å, respectively, indicating that the overall protein structures are quite similar. However, the orientations of the sidechains of Pro158, Phe159, and Gln163 in the cysteine loop appear to differ between the free and complex forms **(Figs. 2-8 G–I)**. Particularly, it seems that Pro158 and Phe159 are oriented inward towards the substrate-binding pocket in the substrate-free form but are lifted and positioned outward upon binding of FA or SA in the complex form **(Figs. 2-8 G–I)**; this movement of Pro158 and Phe159 appears to be synchronized. Therefore, it is likely that the cysteine loop containing Pro158, Phe159, and Gln163 adopts a closed or semi-open conformation in the substrate-free form, while it takes on an open conformation in the complex forms. The position of the center of Pro158 for the FA- or SA-bound form (open conformation) is shifted by ~1.8 Å from that of the substrate-free form (close or semi-open conformation).

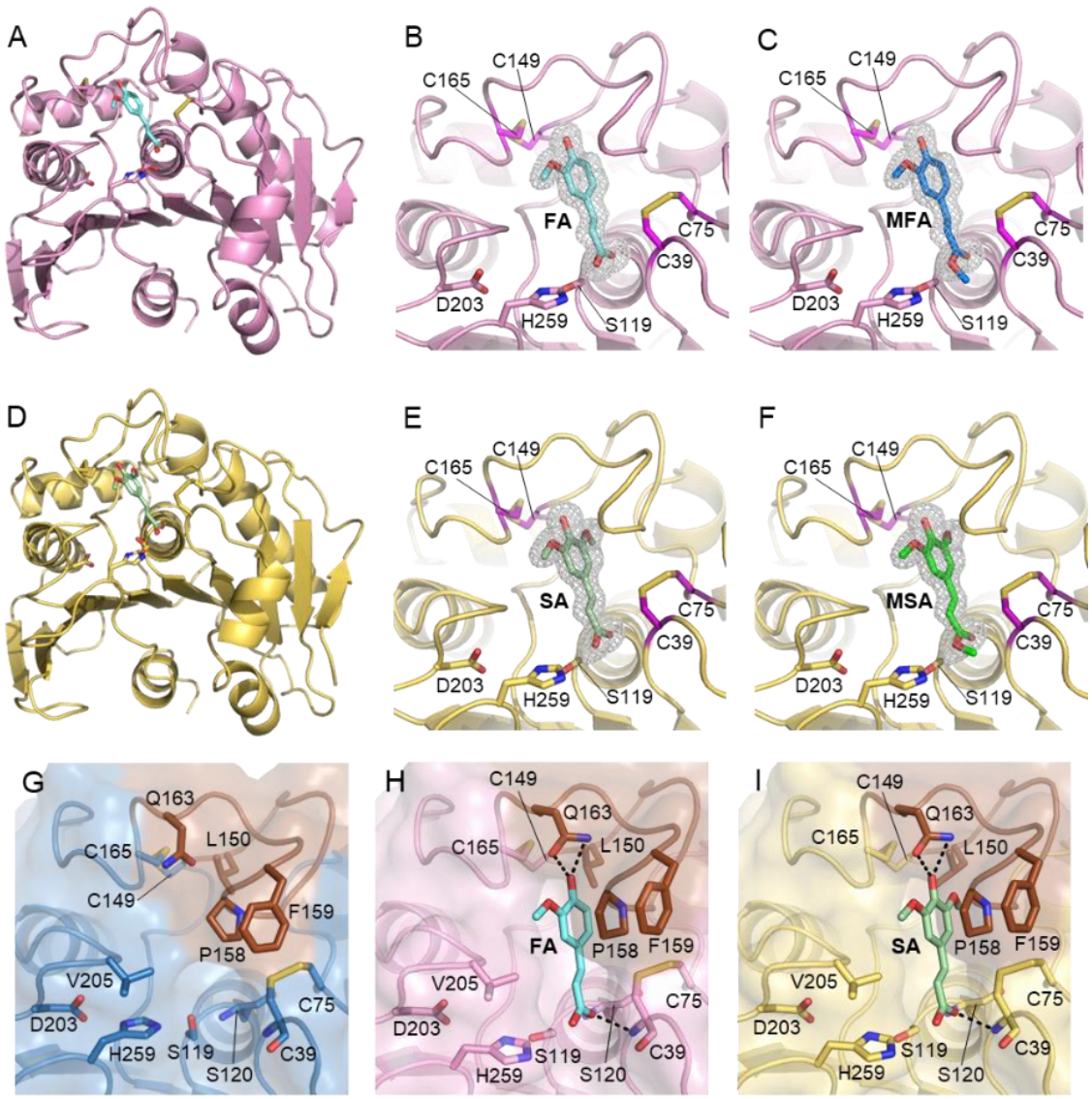
A
B
C165
C149
FA
C75
C39
D203
H259
S119
C
C165
C149
MFA
C75
C39
D203
H259
S119
D
E
C165
C149
SA
C75
C39
D203
H259
S119
F
C165
C149
MSA
C75
C39
D203
H259
S119
G
Q163
L150
C165
C149
F159
P158
V205
C75
D203
H259
S119
S120
C39
H
C149
Q163
C165
L150
FA
P158
F159
V205
C75
D203
S119
C39
H259
S120
I
C149
Q163
C165
L150
SA
P158
F159
V205
C75
D203
S119
C39
H259
S120

Fig. 2-8. Structures of *As*FaeEs bound to ferulic acid (FA) and sinapic acid (SA). Comparison of the *As*FaeE/FA (A, B, and H) and *As*FaeE/SA (D, E, and I) complexes, and substrate-free *As*FaeE (G). (C and F) Model structures of *As*FaeEs bound to methyl ferulate (MFA) (C) and methyl sinapate (MSA) (F). Residues of the S-D-H catalytic triad and cysteine loop (A–I), and those involved in substrate binding (G–I) are shown as sticks. FA (A, B, and H), MFA (C), SA (D, E, and I), and MSA (F) are also shown as sticks. The nitrogen, oxygen, and sulfur atoms are colored blue, red and yellow, respectively. (B, C, E, and F) Close-up views of the S-D-H catalytic triad and cysteine loop (viewed from the same direction as in **Fig. 2-6 C**); cysteine residues forming disulfide bonds are highlighted in purple (B, C, E, and F); the F_o–F_c omit map is contoured at 3σ (represented as a grey mesh). (G–I) Comparison of the S-D-H catalytic triad and cysteine loop (brown) among substrate-free (G), FA-bound (H), and SA-bound (I) *As*FaeE. A proline-aromatic interaction was found between Pro158 and Phe159 in the cysteine loop (G–I). Three layers of the aromatic moiety of either FA or SA, Pro158, and Phe159 were observed in the complex; hydrogen bonds between *As*FaeE and either FA or SA are indicated by black dotted lines, with hydrogen atoms excluded (H, I).

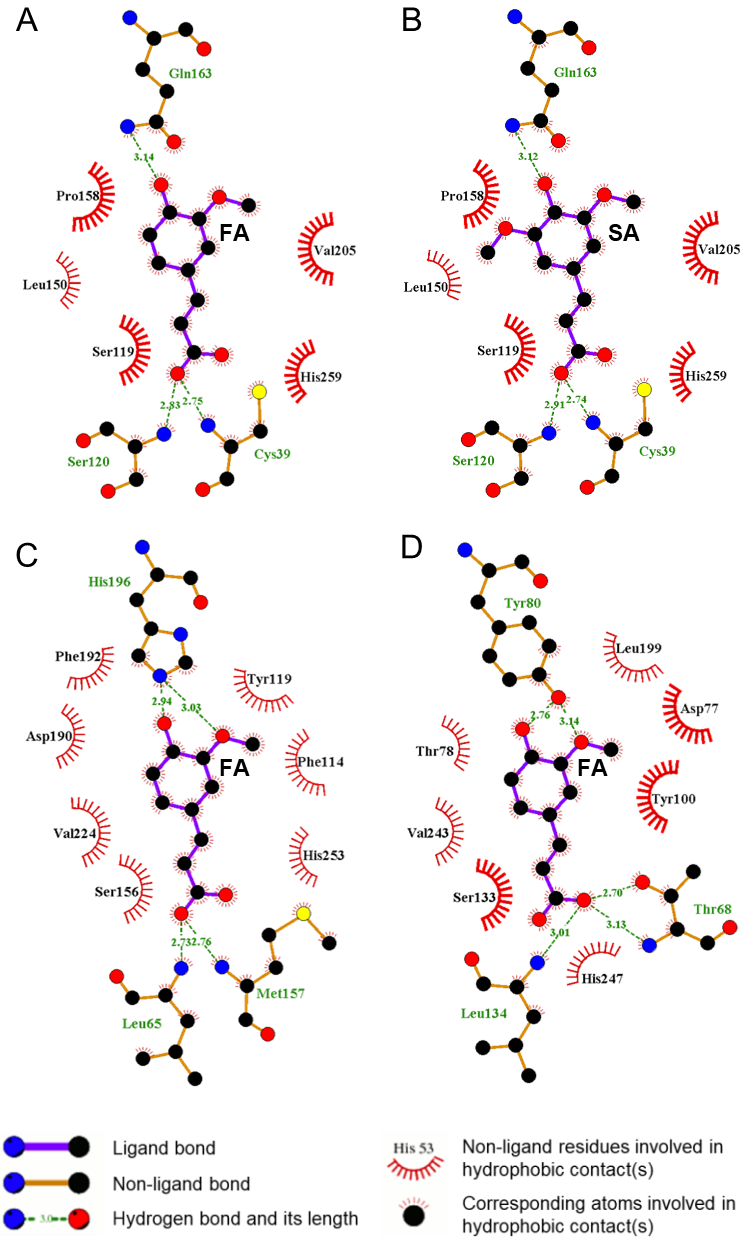

Fig. 2-9. Analysis of Protein-Ligand Interactions using LigPlot. (A) *As*FaeE/FA (PDB: 8IYB), (B) *As*FaeE/SA (PDB: 8IYC), (C) *Am*Fae1A/FA (PDB: 5CXX), and (D) *An*FaeA/FA (PDB: 1UWC). FA and SA are ferulic acid and sinapic acid, respectively. Carbon, oxygen, nitrogen, and sulfur atoms are denoted by black, red, blue, and yellow dots, respectively.

2.2.6 Structure-based functional mutagenesis: the activity toward methyl ferulate (MFA)

To evaluate the roles of the residues surrounding the bound FA or SA in the complex, I conducted structure-based functional mutagenesis on *As*FaeE. I selected the following residues: Cys39, Leu150, Pro158, Phe159, Gln163, and Val205, and substituted them with alanine, valine, serine, or threonine, resulting in mutants C39A, L150A, L150V, P158A, F159A, Q163A, Q163S, V205A, or V205T. The structural integrity of these mutants was confirmed through CD spectroscopy **(Fig. 2-10)**. Since the CD spectral patterns for these mutants closely resembled that of the *As*FaeE WT, it can be inferred that their structures are likely similar, if not identical, to that of the WT.

Subsequently, I determined the kinetic constants of *As*FaeE WT and mutants towards MFA, using Michaelis-Menten analysis **(Table 2-3)**. A lower K_M value signifies a stronger affinity of the enzyme to the substrate, whereas a higher k_{cat} value indicates a greater turnover number of the substrate conversion into product [33]. Collectively, the k_{cat}/K_M value denotes the rate constant for the formation of the enzyme–substrate complex (ES), ultimately influencing product yield [33].

The k_{cat}/K_M value for the WT was 92 s^{-1} mM^{-1}. The k_{cat}/K_M values for the C39A, L150A, and L150V mutants were less than ca. 10% of that of the WT. On the other hand, the values for the P158A, F159A, Q163A, Q163S, and V205A mutants were in the range of 35–65% of that of the WT and that for the V205T mutant was 79% of that of the WT. Notably, the C39A mutant exhibited no activity **(Table 2-3)**, likely due to the disruption of the disulfide bond, which could have compromised the structural integrity of the substrate-binding site.

Regarding the L150A and L150V mutants, their k_{cat}/K_M values were significantly lower than that of the WT, primarily due to an increase in K_M and a decrease in k_{cat}. Based on the K_M values **(Table 2-3)**, the affinity towards MFA followed this order: WT > L150V > L150A. This observation aligns with my previous discussion, suggesting that the space around position 5 of the phenolic moiety of MFA is more occupied in the WT than in the mutants L150V and L150A, resulting in more van der Waals interactions. This rationalizes the observed affinity order. Moreover, for the hydrolysis reaction, it is

essential to position the ester bond moiety of MFA correctly at the S-D-H catalytic triad. The WT seemed to achieve an ideal placement, whereas the mutants' vacant spaces might have led to suboptimal positioning, explaining the correlation between k_{cat} values and affinity order.

Regarding the P158A and F159A mutants, their k_{cat}/K_M values were approximately 40% of that of the WT (39% and 40%, respectively) **(Table 2-3)**. In the WT, Pro158 and Phe159 participated in a three-layer planar ring stacking, aided by CH-π interactions, where Pro158 directly interacted with FA. Substituting Phe159 resulted in a reduced affinity for MFA when compared to the substitution of Pro158, whereas the substitution of Pro158 caused a more reduction in the k_{cat} value than the substitution of Phe159. These findings suggest that the combined presence of the ring structures of phenylalanine and proline is advantageous for establishing CH-π interactions in the context of the three-layer planar ring stacking. I hypothesize that this configuration of a three-layer planar ring stacking could potentially contribute to finely adjusting the positioning of the ester bond of a substrate to the S-D-H catalytic triad.

For the Q163A and Q163S mutants, their k_{cat}/K_M values were 63% and 48% of the WT, respectively **(Table 2-3)**. The intermolecular hydrogen bonding between Gln163 and the phenolic hydroxyl group of FA or SA must be missing in these mutants. Additionally, the sizes of alanine and serine are smaller than that of glutamine. For these two reasons, the ester bond was not ideally placed at the S-D-H catalytic triad in these mutants, resulting in reduced k_{cat}.

Finally, I examined the V205A and V205T mutants, with k_{cat}/K_M values of 35% and 79% of the WT, respectively **(Table 2-3)**. Val205 is located on the opposite side of Leu150 relative to the substrates **(Fig. 2-8 H)**. The sizes of Val205 and Leu150 seem sufficient to accommodate the substrates adequately. As the size of threonine is almost comparable to that of valine, the space available for the substrates at the binding pocket remains nearly the same. This explains the similar K_M for the WT and V205T. The reduction in k_{cat} for V205T is supposed to be moderate for the same reason, ideal placing of the ester bond at the S-D-H catalytic triad being almost achieved. On the other hand, as the size of alanine is rather smaller than that of valine, much larger space is vacant for the substrates for V205A. This can rationalize the reduction in affinity with substrates for

V205A. This can also explain the reduction in k_{cat}, ideal placing of the ester bond at the S-D-H catalytic triad not being achieved.

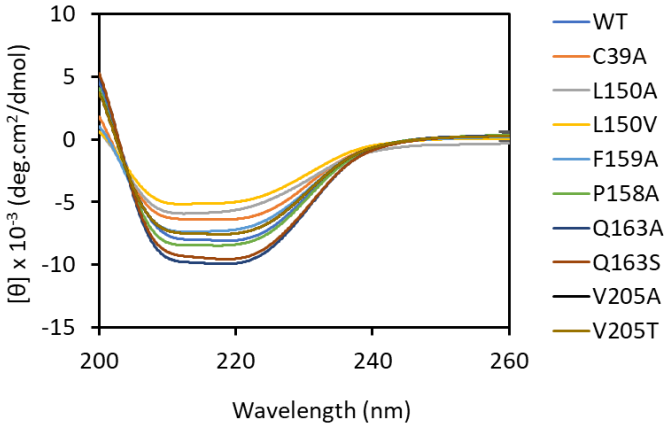

Fig. 2-10. Circular dichroism spectra from 200–260 nm for *As*FaeE wild-type and mutants.

Table 2-3. Kinetic parameters of esterase activity toward MFA of wild-type and mutant *As*FaeE, and other esterases.

Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	Reference
*As*FaeE	0.22 ± 0.01	20.2 ± 0.3	92 ± 5	This study
- C39A	N.D.	N.D.	N.D.	This study
- L150A	1.92 ± 1.20	7.0 ± 3.0	3.6 ± 3	This study
- L150V	0.99 ± 0.30	9.4 ± 1.6	9.5 ± 3	This study
- P158A	0.27 ± 0.02	9.8 ± 0.3	36 ± 3	This study
- F159A	0.38 ± 0.05	14.3 ± 0.7	37 ± 5	This study
- Q163A	0.24 ± 0.02	13.8 ± 0.4	58 ± 5	This study
- Q163S	0.35 ± 0.04	15.6 ± 0.7	44 ± 6	This study
- V205A	0.31 ± 0.07	9.8 ± 0.8	32 ± 7	This study
- V205T	0.21 ± 0.02	15.5 ± 0.5	72 ± 8	This study
*Am*Fae1A	0.0067 ± 0.001	6.05	903	[34]
*An*FaeA	0.78 ± 0.05	70.74 ± 1.44	90.7	[35]

Values are the means ± standard deviation of three independent replicates. N.D., Not detected.

2.2.7 Structure-based functional mutagenesis: change in substrate preference

Next, I determined the specific activities of the mutants towards model substrates for FAE (MFA and MSA) and that for AXE (*p*NPAc) **(Table 2-4)**. Notably, certain mutants exhibited distinct substrate preferences compared to the WT **(Table 2-4)**. The specific activities of the WT towards MFA, MSA, and *p*NPAc were 788.3, 299.1, and 510.2 mU/mg, respectively, and the order of substrate preference was MFA > *p*NPAc > MSA **(Table 2-4)**. The mutants P158A, Q163A, Q163S, V205A, and V205T showed substrate preferences similar to the WT. On the other hand, the substrate preference of L150A and L150V was in the order MSA > MFA ≈ *p*NPAc.

The L150V mutant retained a considerably high specific activity towards MSA (259.5 mU/mg), 87% of the specific activity of the WT. Interestingly, this mutant exhibited a change in substrate preference, with the new order being MSA > MFA ≈ *p*NPAc **(Table 2-4)**. This observed change in substrate preference corresponds well with

the discussion in section **2.2.5**. Notably, in the *As*FaeE/SA complex structure, Leu150 adopted a $\chi 2$ angle of –60°, creating a space for the methoxy group at position 5 of SA to be accommodated **(Fig. 2-8 G)**. In the context of the L150V mutant, the valine substitution appears to generate the necessary space to accommodate the position 5 methoxy group of SA. This adaptation accounts for the reduction in specific activity of L150V towards MFA when compared to the WT. It seems that valine cannot fully occupy the space around the position 5 of FA as effectively as Leu150 does in the WT, thereby affecting the specific activity towards MFA. This reasoning could potentially be extended to the L150A mutant. For the L150A mutant, the unoccupied space around the position 5 for FA is larger than that for SA. This may have caused the reduction in binding affinity and an impaired ideal positioning of the ester bond at the S-D-H catalytic triad for L150A, resulting in lower affinity toward MFA than toward MSA.

Table 2-4. Specific activities of wild-type and mutant *As*FaeE toward methyl ferulate (MFA), methyl sinapate (MSA), and *p*-nitrophenyl acetate (*p*NPAc).

	Specific activity (mU/mg)		
Enzyme	**MFA**	**MSA**	***p*NPAc**
WT	788.3 ± 3.2	299.1 ± 7.0	510.2 ± 5.7
C39A	15.4 ± 3.1	39.8 ± 2.6	N.D.
L150A	44.2 ± 8.8	97.5 ± 5.0	52.5 ± 13.2
L150V	111.8 ± 5.3	259.5 ± 9.1	110 ± 13.7
P158A	263.0 ± 4.8	102.8 ± 2.5	132.3 ± 2.2
F159A	360.2 ± 3.7	112.9 ± 2.2	450.3 ± 22.0
Q163A	410.7 ± 7.3	81.0 ± 1.3	308.5 ± 9.9
Q163S	388.2 ± 2.0	87.2 ± 2.4	253.8 ± 2.6
V205A	251.2 ± 3.8	94.0 ± 5.0	173.9 ± 9.1
V205T	543.9 ± 6.7	167.6 ± 4.4	221.1 ± 3.0

Values are the means ± standard deviation of three independent replicates. N.D., Not Detected.

2.2.8 Structure-based multiple sequence alignment and structural comparison of the members of SF6 FAE

In this section, I conducted a structure-based multiple sequence alignment and structural comparison of the members belonging to SF6, aiming to gain insights into the roles of residues involved in substrate binding in *As*FaeE. The alignment of the nine SF6 members, comprising six FAEs and three AXEs, is presented in **Fig. 2-11**. I observed conservation of the catalytic triad S-D-H and two pairs of disulfide-bond forming cysteine residues. Additionally, the residues participating in secondary structures were either identical or of the same types. Notably, I focused on specific residues (Cys39, Leu150, Pro158, Phe159, Gln163, and Val205) surrounding the bound FA or SA in the complex structures of *As*FaeE. These residues were chosen for structure-based functional mutagenesis, as detailed in the previous section.

Among the SF6 members, I had access to the crystal structure of one of the AXEs, *Al*AXEA, in its free form [PDB: 5X6S]. Consequently, I performed a superimpositioning of the structures of *Al*AXEA and *As*FaeE (free form), shown in **Fig. 2-12**. The Cα superimposition of 271 residues in *As*FaeE (total of 275 residues) and 274 residues in *Al*AXEA resulted in an RMSD of 1.8 Å. Further focusing on 242 residues within the structural regions, the RMSD reduced to 0.8 Å **(Fig. 2-12 B)**. Examining the substrate-binding site, I found a remarkable overlap of the atoms, particularly for residues surrounding the substrates in the *As*FaeE complex structures **(Figs. 2-8 H, I, and 2-12 C)**. However, it was important to note that there were notable differences in the cysteine loop of *As*FaeE compared to the corresponding region in *Al*AXEA.

Cys39 in *As*FaeE acts as a disulfide-bond forming cysteine and also engages in hydrophobic or van der Waals interactions with the substrates. This residue was conserved in all six FAEs and three AXEs.

Within the cysteine loop of *As*FaeE, the sequential residues Pro158-Phe159 play a crucial role in forming three layers with the substrate. These proline-phenylalanine sequential residues were found in three out of the six (3/6) FAEs, while proline-isoleucine (2/6) and proline-glycine (1/6) were present in other FAEs **(Fig. 2-11)**. Considering the discussion in section **2.2.6**, it is reasonable to assume that proline-phenylalanine is

preferred over proline-isoleucine and proline-glycine. Hence, proline-phenylalanine residues are predominant among the six SF6 FAEs.

Gln163 in *As*FaeE plays a crucial role, as its sidechain amide group forms hydrogen bonds with the phenolic hydroxyl group (position 4) of the substrates **(Figs. 2-8 H and I)**. This residue is conserved in all six FAEs **(Fig. 2-11)**, and my functional mutagenesis study **(Table 2-3)** also supports its significance. In contrast, the corresponding residue in all three AXEs is serine. It is intriguing to hypothesize that the hydroxyl group of serine could form a hydrogen bond with the nitro group at position 4 of the model substrate, *p*NPAc.

Val205 is involved in van der Waals interaction with substrates. In the previous section, I suggested that the sizes of Val205 and Leu150 in *As*FaeE may regulate the accommodation of substrates. Among the six FAEs, two had valine (2/6) FAEs, three had leucine (3/6), and one had phenylalanine (1/6) at this position **(Fig. 2-11)**. All these residues can engage in van der Waals interaction with substrates. In contrast, all three AXEs had threonine. The superimposition of the structures of *As*FaeE and *Al*AXEA showed a good overlap of Val205 in *As*FaeE and Thr204 in *Al*AXEA **(Fig. 2-12 C)**, suggesting that Thr204 of *Al*AXEA may also be involved in van der Waals interaction with *p*NPAc, in addition to potential hydrogen bonding interactions.

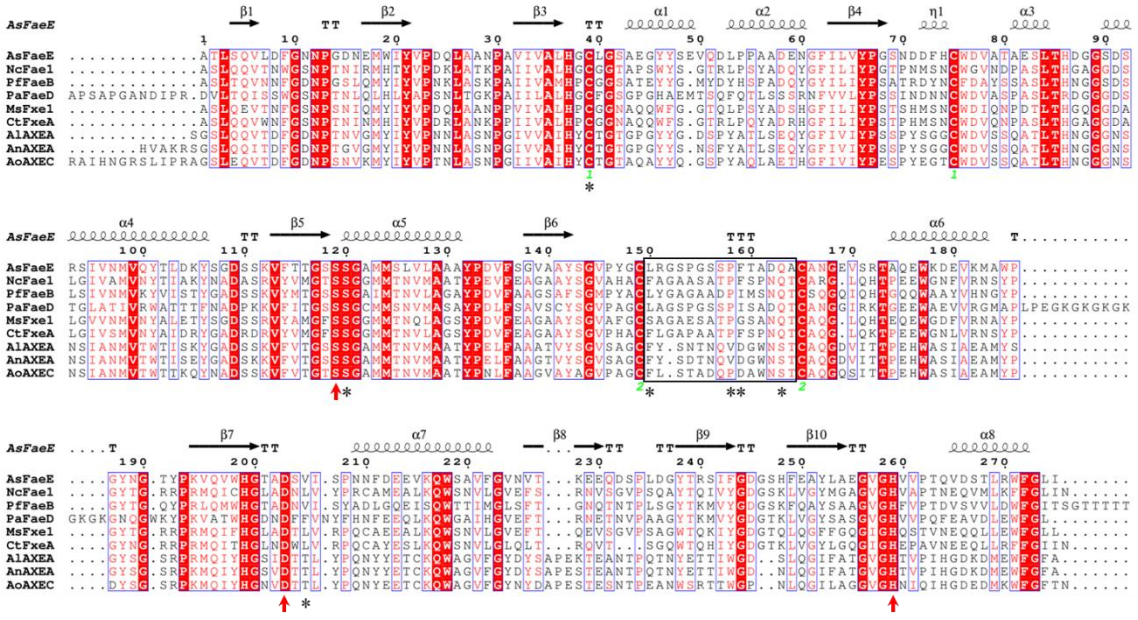

Fig. 2-11. Sequence alignment of *As*FaeE with the esterases in FAE SF6, whose catalytic activities were reported previously. The secondary structures of *As*FaeE are indicated on top (α-helices and a 3_{10} helix (η) by spirals, and β-strands by black arrows). The identical residues are highlighted in red, and similar residues are indicated by red text. The residues of the S-D-H catalytic triad in *As*FaeE are indicated by red arrows. The pairs of cysteine residues that form disulfide bonds in *As*FaeE are indicated by green numbers. The residues of the substrate-binding pocket **(Fig. 2-8)** in *As*FaeE are indicated by asterisks.

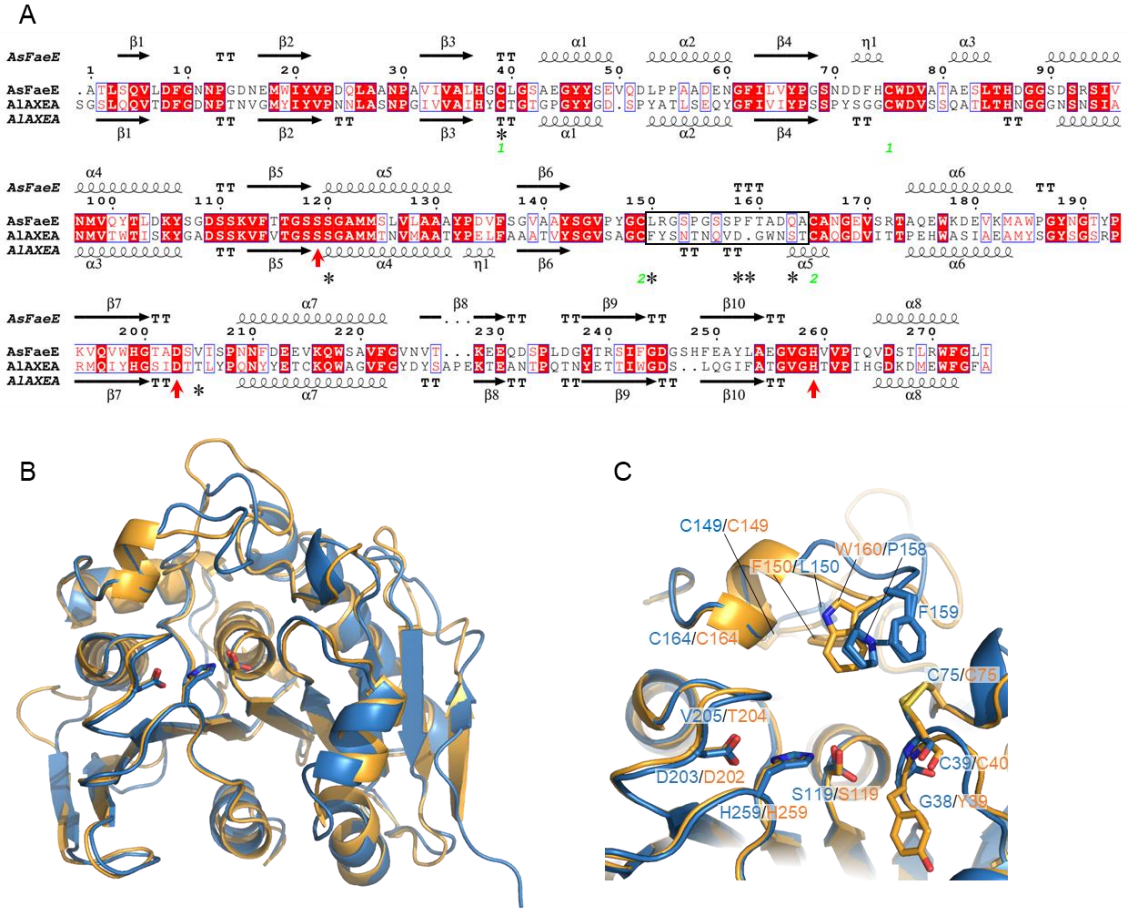

Fig. 2-12. Comparison of *As*FaeE with *Aspergillus luchuensis* AXE (*Al*AXE). (A) Structure-based sequence alignment of *As*FaeE with *Al*AXE. α-helices and a 3_{10} helix (η) are indicated by spirals, and β-strands are indicated by black arrows. The identical residues are highlighted in red, and similar residues are indicated by red text. The residues of the S-D-H catalytic triad are indicated by red arrows. The two pairs of cysteine residues that form disulfide bonds are indicated by green numbers. A black box indicates the region named the cysteine loop. The residues comprising the substrate-binding pocket **(Fig. 2-6)** in *As*FaeE are indicated by asterisks. (B, C) *As*FaeE (blue) and *Al*AXE (PDB: 5X6S) (orange) are superimposed. (C) Close-up view of the substrate-binding pocket of *As*FaeE (blue) and *Al*AXE (orange). The residues comprising the S-D-H catalytic triad (B, C) and cysteine loop (C) are shown as sticks. The nitrogen, oxygen, and sulfur atoms are colored blue, red and yellow, respectively.

2.2.9 Structure-based sequence alignment and structural comparison between the members of SF6, SF7, and SF8

In this section, I conducted a structure-based sequence alignment and structural comparison among the members of SF6, SF7, and SF8 **(Figs. 2-13 A and 2-14)**. My aim was to investigate the differences in the binding mode between these members. The structure-based sequence alignment of *As*FaeE (SF6), *Am*Fae1A (SF8), and *An*FaeA (SF7) indicated notable variations in sequence length and secondary structure distribution **(Fig. 2-14)**. The structural comparisons between the *As*FaeE/FA and *Am*Fae1A/FA complexes **(Fig. 2-13 A)**, and between the *As*FaeE/FA and *An*FaeA/FA complexes **(Fig. 2-13 B)**, provided insights into the similarity and differences in the substrate-binding sites and the mode of accommodation of FA among *As*FaeE, *Am*Fae1A, and *An*FaeA. The RMSD between *As*FaeE/FA and *Am*Fae1A/FA complexes was 1.1 Å among 70 pruned atoms (11.4 Å across all 209 pairs), while that between *As*FaeE/FA and *An*FaeA/FA complexes was 1.1 Å among 36 pruned atoms (12.6 Å across all 216 pairs). The backbone traces of *As*FaeE and *Am*Fae1A exhibited significant overlap, except for the cysteine loop of *As*FaeE and the corresponding region of *Am*Fae1A **(Fig. 2-13 A)**. Additionally, I observed that the phenolic moieties of FA were flipped by 180° around the C1–C4 axis of an aromatic ring with respect to each other **(Figs. 2-13 A, C, and D)**. Conversely, the backbone traces of *As*FaeE and *An*FaeA, as well as the positions and orientations of the bound FA, displayed noticeable distinctions **(Fig. 2-13 B)**.

In **Fig. 2-13 C**, I noticed that the cysteine loop (brown) of *As*FaeE covered the top half of FA, whereas in **Fig. 2-13 D**, the "β-clamp motif" (cyan) and "α3 extension" (yellow) [21] of *Am*Fae1A covered the top half of FA more extensively. These regions of *As*FaeE and *Am*Fae1A, acting as roofs, were observed to cover the substrates. On the other hand, in **Fig. 2-13 E**, *An*FaeA lacked the roof found in *As*FaeE and *Am*Fae1A. Instead, it utilized the so-called "flap region" found in lipase **(Fig. 2-13 E, blue)** and a loop between residues 226–244 **(Fig. 2-13 E, magenta)** to create a deep cleft that enveloped FA [20].

Regarding similarities, all structures displayed the S-D-H catalytic triad and oxyanion hole. S119-D203-H259 in *As*FaeE, S156-D222-H253 in *Am*Fae1A, and S133-

D194-H247 in *An*FaeA formed the S-D-H catalytic triad. The oxyanion holes were formed by the NH groups of Cys39 and Ser120 in *As*FaeE, Leu65 and Met157 in *Am*Fae1A, and Thr68 and Leu134 in *An*FaeA **(Figs. 2-13 C–E, bottom)**. Additionally, in all structures, a valine involved in hydrophobic or van der Waals interactions was identified: Val205 (*As*FaeE), Val224 (*Am*Fae1A), and Val243 (*An*FaeA) **(Fig. 2-9)**.

In *As*FaeE, *Am*Fae1A, and *An*FaeA, FA was surrounded by hydrophobic residues, although the residues are different among these proteins **(Figs. 2-9, and 2-13 C–E, bottom)**. Interestingly, the roof of *As*FaeE played a role in forming the FA-Pro158-Phe159 three layers **(Fig. 2-13 C)**, while in *Am*Fae1A, the roof composed of Phe114 and Phe192 was involved in FA binding **(Fig. 2-13 D)**.

The kinetic parameters for esterase activities against MFA were determined for *As*FaeE, *Am*Fae1A, and *An*FaeA **(Table 2-3)**. The k_{cat}/K_M values followed the order of *Am*Fae1A >> *As*FaeE ≈ *An*FaeA. Similarly, the order of affinity, based on the K_M values, was *Am*Fae1A >> *As*FaeE > *An*FaeA. This order of affinity aligned with the size of the roofs described earlier, which ranked qualitatively as *Am*Fae1A > *As*FaeE > *An*FaeA. Moreover, this order of affinity correlated with the cavity volume of the substrate-binding pockets calculated by CASTp [32], with values of *Am*Fae1A (31.67 $Å^3$) > *As*FaeE (49.26 $Å^3$) > *An*FaeA (52.25 $Å^3$). These findings suggest that a higher number of van der Waals interactions lead to stronger affinity. Notably, *As*FaeE and *Am*Fae1A exhibited esterase activities towards monomeric ferulates but not dimeric ones [6,21]. Conversely, *An*FaeA demonstrated esterase activities towards both monomeric and dimeric substrates [6], likely due to its more open substrate-binding site compared to *As*FaeE and *Am*Fae1A. Thus, the substrate binding modes of SF6, SF7, and SF8 members exhibited significant diversity.

The insights derived from this study hold the potential to inform the engineering of an FAE capable of effectively accommodating larger substrates while enhancing their hydrolytic efficiency. In the case of *As*FaeE, my investigations reveal that the gradual substitution of L150 with valine and alanine leads to a progressive expansion of the accessible spatial region surrounding position 5 of the bound methyl sinapate. This intriguing observation suggests the potential discovery of ferulate derivatives carrying functional groups of appropriate size at position 5. In other words, it is fascinating to

explore derivatives of ferulate with distinct functional groups at position 5 that can be effectively catalyzed by L150V or L150A. Similarly, through the substitution of V205 with threonine and alanine, I observe an enlargement in the available spatial area surrounding position 2 of both methyl ferulate and methyl sinapate substrates. This finding hints at the potential identification of ferulate derivatives with appropriately sized functional groups at position 2. These derivative compounds, tailored for positions 5, 2, or a combination of both, could potentially serve as viable substrates for processes such as hydrolysis or transesterification. Turning to *An*FaeA, the absence of a roof structure combined with its extensively open surface offers an intriguing opportunity. It may be feasible to engineer the shape of this surface to effectively accommodate larger substrates. Such an approach could pave the way for the development of an FAE variant capable of efficiently processing a wider range of substrate sizes.

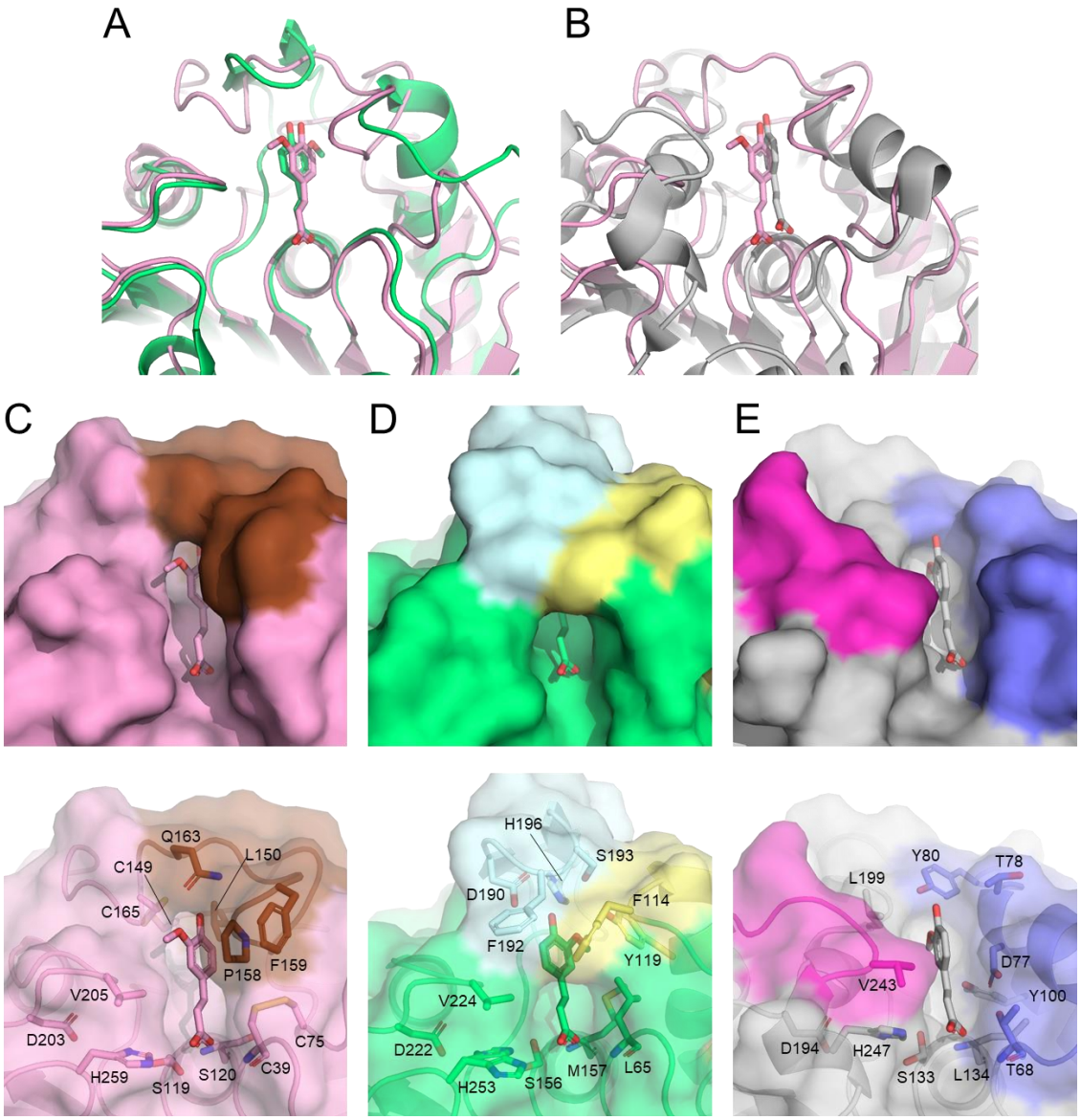
A
B
C
D
E
Q163
L150
C149
C165
P158
F159
V205
D203
C75
H259
S120
C39
S119
H196
S193
D190
F114
F192
Y119
V224
D222
M157
L65
H253
S156
Y80
T78
L199
D77
V243
Y100
D194
H247
S133
L134
T68

Fig. 2-13. Comparison of the FA-binding manner of *As*FaeE (belonging to SF6) with those of *Anaeromyces mucronatus* Fae1A (*Am*Fae1A) (SF8) and *Aspergillus niger* FaeA (*An*FaeA) (SF7). *As*FaeE/FA was superimposed with *Am*Fae1A/FA (PDB: 5CXX) (A) and *An*FaeA/FA (PDB: 1UWC) (B). (C–E, top and bottom) Protein surface representation of *As*FaeE/FA, *Am*Fae1A/FA, and *An*FaeA/FA (C–E). In *As*FaeE/FA (C), an open conformation of the cysteine loop (brown) allows FA to enter and fit into the pocket, resulting in the aromatic (FA)-proline (Pro158)-aromatic (Phe159) interaction. In *Am*Fae1A (D), the β-clamp motif (cyan) and α3 extension (yellow) seem to play a similar role as the cysteine loop of *As*FaeE; FA and Phe114 are interacting but are almost perpendicular to each other. In *An*FaeA (E), the regions equivalent to the cysteine loop of *As*FaeE, and β-clamp motif and α3 extension of *Am*Fae1A seem to be missing. Instead, *An*FaeA has a flap region (blue) and 226–244 loop (magenta), which form a large cleft. (C–E, bottom) Transparent presentation. The residues of the S-D-H catalytic triad and cysteine loop (C), β-clamp motif and α3 extension (D), and flap region and 226–244 loop (E) are shown as sticks. The nitrogen, oxygen, and sulfur atoms are colored blue, red and yellow, respectively.

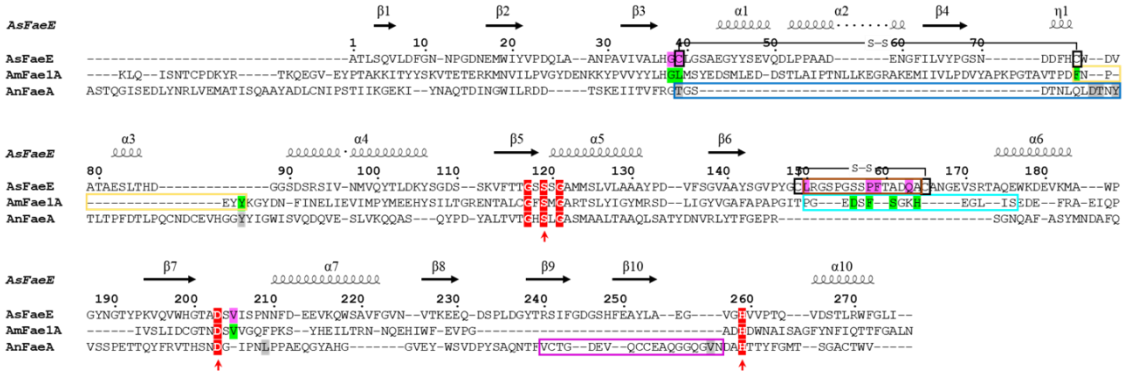

Fig. 2-14. Structure-based sequence alignment of *As*FaeE, *An*FaeA (of SF7), and *Am*Fae1A (of SF8). The secondary structure of *As*FaeE is indicated on top (α-helices and 3_{10} helix (η) by spiral, and β-sheets by black arrows). Red arrows indicate the S-D-H catalytic triad (Ser119-Asp203-His259 of *As*FaeE). The identical residues are highlighted in red. Disulfide bonds formed in AsFaeE are labeled "S-S". The cysteine loop in *As*FaeE is boxed in brown. In *Am*Fae1A, the β-clamp motif and α3 extension [1] are boxed in cyan and yellow, respectively. The flap region and loop 226–244 [2] of *An*FaeA are boxed in blue and magenta, respectively. The colored residues in **Figs. 2-13 C, D, and E** are highlighted with the same color.

2.3 Conclusions

I determined the high-resolution crystal structures of a *As*FaeE, in the free form and in the complex form with either FA or SA. Notably, these structures represent the first three-dimensional structures for the members of SF6 of fungal FAEs. *As*FaeE adopts a canonical α/β-hydrolase fold, featuring a conserved S-D-H catalytic triad. In the crystal structures of the complexes, FA and SA were accommodated in the substrate-binding site through three main types of interactions: (i) hydrophobic or van der Waals interactions; (ii) hydrogen bonding between the sidechain amide group of Gln163 and the phenolic hydroxyl group of the substrates; and (iii) CH-π interaction between Pro158, which adopts the *cis* conformation, and the aromatic moiety of the substrates. An interesting feature is the presence of three layers, consisting of the planar moieties of the substrates, Pro158, and Phe159, which play a unique role in accommodating the substrates. Moreover, my findings indicate that the cysteine loop, containing Pro158, Phe159, and

Gln163, adopts a closed or semi-open form in the absence of substrate and an open form in the presence of substrate. Additionally, I observed that the $\chi 2$ angle of Leu150 is approximately –60° when the phenolic methoxy group is located at position 5 in the substrate. These significant findings are well-supported by my extensive structure-based functional mutagenesis and structure-based multiple sequence alignment analyses.

The gradual replacement of L150 with valine and alanine in *As*FaeE revealed a progressive expansion of the accessible spatial region around position 5 of the bound methyl sinapate. This observation intriguingly hints at the possibility of discovering ferulate derivatives with appropriately sized functional groups at position 5. Similarly, substituting V205 with threonine and alanine resulted in an increased spatial area around position 2 of methyl ferulate and methyl sinapate substrates. This suggests the potential identification of ferulate derivatives with suitable functional groups at position 2. These derivatives, tailored for positions 5, 2, or a combination of both, could serve as viable substrates for processes like hydrolysis or transesterification. Furthermore, the absence of a roof structure in *An*FaeA coupled with its open surface presents an intriguing avenue. Engineering the shape of this surface could lead to the development of an FAE variant capable of effectively handling a broader range of substrate sizes.

The structural information obtained from this study can aid in selecting the appropriate enzyme for efficient agricultural biomass utilization, thereby paving the way for improved biotechnological processes in this field.

2.4 Materials and methods

2.4.1 Sequence analysis

An amino acid sequence of *As*FaeE was retrieved from UniProtKB/TrEMBL [36] (primary accession: A0A1L9T9J3). The signal sequence, predicted with SignalP 5.0, was omitted for further analysis [37]. *As*FaeE was classified by comparing its sequence with either those of the members belonging to 13 subfamilies of fungal FAEs, which were identified by Benoit *et al.* (2008) [10] and Dilokpimol *et al.* (2016) [11], or those of the proteins whose three-dimensional structures were available in the Protein Data Bank (PDB). Information on all sequences used in the alignment and phylogenetic tree analysis

is shown in **Appendix A**. Multiple sequence alignment was performed using the MUSCLE algorithm and construction of the phylogenetic tree was performed using Molecular Evolutionary Genetics Analysis software (MEGA11) [38]. The maximum-likelihood method with the best fit WAG+G substitution model and 95% partial deletion was used for the phylogenetic tree construction. Statistical support for the phylogenetic grouping was obtained by 200 bootstrap re-samplings.

Glycosylation sites of *As*FaeE were predicted with NetNGlyc 1.0 [24] and NetOGlyc 4.0 [25]. The theoretical molecular mass (*m*) and isoelectric point (pI) were calculated using the Expasy-ProtParam tool [39].

2.4.2 Cloning of AsFaeE

The DNA sequence encoding the N-terminal 6× histidine-tag and *As*FaeE, which does not contain a signal sequence, was codon-optimized and synthesized by Thermo Fisher Scientific (Japan) **(Appendix B)**. The DNA fragment was then ligated in frame with an α-factor signal peptide of the pPICZαA expression vector at the EcoRI and NotI restriction sites. The recombinant plasmid was linearized with SacI and introduced into *Pichia pastoris* X-33 (Invitrogen) by electroporation. The *P. pastoris* transformants were selected on the YPD agar medium (1% yeast extract, 2% peptone, 2% glucose, and 1.5% agar) containing 100 μg/mL Zeocin (Invivogen).

2.4.3 Site-directed mutagenesis

Nine mutants of *As*FaeE were constructed by inverse PCR using the following primers: C39A (5′-GCACGGAGCCCTTGGCTCCGCTGAGGGCTA-3′ and 5′-AGC-CAAGGGCTCCGTGCAAAGCCACGATCA-3′), L150A or L150V (5’-TGGTTGT-GYACGTGGCTCTCCAGGCTCTTC-3’ and 5’-AGCCACGTRCACAACCATATG-GAACTCCGG-3’), P158A (5′-CTCTTCCGCCTTCACTGCTGATCAGGCATG-3′ and 5′-CAGTGAAGGCGGAAGAGCCTGGAGAGCCAC-3′), F159A (5’-TTCCCCCGCC-ACTGCTGATCAGGCATGTGC-3’ and 5’-CAGCAGTGGCGGGGGAAGAGCCT-GGAGAG-3’), Q163A or Q163S (5′-TGCTGATKCCGCATGTGCCAATGGAGAG-

GT-3′ and 5′-CACATGCGGMATCAGCAGTGAAGGGGGAAG-3′), and V205A or V205T (5′-CGATTCCRCCATTTCACCAAATAATTTTGACG-3′ and 5′-GTGAAA-TGGYGGAATCGGCGGTTCCATGCC-3′) (mutated sites underlined). R is an A or G nucleotide, Y a C or T nucleotide, K a G or T nucleotide, and M an A or C nucleotide.

2.4.4 Protein expression and purification

To induce protein expression, the *P. pastoris* transformants were initially grown overnight in YPD medium (1% yeast extract, 2% peptone, and 2% glucose) at 30 °C with 250 rpm shaking. An overnight culture was transferred to 1 L of BMGY (buffered glycerol complex medium: 1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate, pH 6.0, 1.34% (w/v) yeast nitrogen base (YNB), and 4×10^{-5}% (w/v) biotin) supplemented with 1% (v/v) glycerol and grown at 30 °C with 250 rpm for 3 days. The cells were harvested by centrifugation at 8,000 × *g* for 5 min at room temperature and inoculated into 500 mL of BMMY containing 0.5% methanol. To continue the induction of protein expression, 0.5% methanol was added to the yeast culture once daily for 3 days. After centrifugation at 8,000 × *g* for 5 min at 4 °C, the supernatant was collected and the pH of the supernatant was adjusted to 8.0.

Protein purification was performed at 4 °C. For the enzyme activity assay, *As*FaeE and its mutants were purified using a Ni^{2+}-NTA Sepharose column (Cytiva); protein elution was performed with 50 mM potassium phosphate, pH 6.5, and 200 mM imidazole. The obtained protein solution was concentrated and the buffer was changed with an Amicon Ultra-0.5 (3 kDa molecular mass cut-off (MMCO), Millipore) to 25 mM potassium phosphate, pH 6.5. The purified protein was then stored at 4 °C until use. For protein crystallization, I applied further purification using HiTrap Q HP (GE Healthcare), equilibrated with 25 mM potassium phosphate, pH 6.5; protein elution was performed with a linear gradient (60 mL) of 0–1 M NaCl in the same solution. The obtained protein solution was concentrated with an Amicon Ultra-4 (10 kDa MMCO, Millipore). Subsequently, the protein solution was treated with Endoglycosidase H (Endo H) fused with maltose binding protein (Endo H_f) (New England BioLabs, USA), and then loaded onto HiLoad 16/600 Superdex 75 column (GE Healthcare) equilibrated with 25 mM

potassium phosphate, pH 6.5. The purity of the protein fractions obtained at each purification step was analyzed by 12% SDS-PAGE. The obtained protein was stored at 4 °C for further use.

I determined the protein concentration by measuring UV absorbance of the samples at 280 nm. The extinction coefficient was theoretically determined based on the amino acid composition of *As*FaeE and its mutants [40].

2.4.5 Circular dichroism

*As*FaeE wild-type (WT) and mutants (5 μM) were individually dissolved in the 25 mM potassium phosphate, pH 6.5. Circular dichroism (CD) spectra in the far-UV region (200–260 nm) were recorded using a Jasco J720A spectrometer (Japan Spectroscopic Co., Japan) with an optical path cell of 0.1 cm. All spectra underwent background correction, smoothing, and conversion to molar ellipticity [θ] ($deg \cdot cm^2 \cdot dmol^{-1}$) [41].

2.4.6 Enzyme activity assay

Enzyme activity assays were conducted for *As*FaeE and its mutants, without undergoing Endo H_f and *O*-glycosylase treatment for deglycosylation. Three model substrates were employed: two methyl hydroxycinnamates, specifically methyl ferulate (MFA; LKT Laboratories, Inc., USA) and methyl sinapate (MSA; Biosynth, Switzerland), alongside *p*-nitrophenyl acetate (*p*NPAc; TCI Chemicals, Japan). I followed a modified version of the protocols reported by Dilokpimol *et al.* (2017) [42] and Xu *et al.* (2021) [43]. In summary, 100 ng of *As*FaeE was added to a solution consisting of 80 mM potassium phosphate, pH 6.5, and either 0.12 mM methyl hydroxycinnamate (MFA or MSA) or 0.6 mM *p*NPAc. The reaction proceeded at 37 °C for 30 min.

I employed a microplate reader (Tecan Infinite 200 Pro, Switzerland) to monitor the reaction through UV spectra, measuring UV absorbance at 340 nm for methyl hydroxycinnamates and 410 nm for *p*NPAc, with readings taken every 2 min. This allowed us to determine the initial reaction rates of hydrolysis. Extinction coefficients of

both model substrates and phenolic products, FA (MP Biomedicals, USA), SA (Cayman Chemical Company, USA), and *p*-nitrophenol (TCI Chemicals, Japan), were experimentally determined. For clarity, I defined one unit (U) of enzyme activity as the amount that catalyzes the production of 1 μmol of *p*-nitrophenol from *p*NPAc or 1 μmol of hydroxycinnamic acids from methyl hydroxycinnamates per min. All assays were performed in triplicate.

The initial reaction rates of MFA hydrolysis at varying concentrations (0.02–1.2 mM) were determined for *As*FaeE and its mutants in a solution comprising 100 mM potassium phosphate, pH 6.5. Curve fitting was conducted using Kaleidagraph (Synergy, Reading, PA, USA) [44]. Data were fitted to the Michaelis-Menten equation, from which kinetic parameters (K_M and k_{cat}) were derived. Nonlinear regression analyses yielded the fitting errors.

2.4.7 X-ray crystallography

The crystallization of *As*FaeE was performed by the sitting-drop vapor-diffusion method at 293 K. An *As*FaeE crystal was obtained by mixing 1 μL of 10 mg/mL enzyme solution dissolved in 25 mM potassium phosphate, pH 6.5, with an equal volume of a reservoir solution consisting of 0.2 M ammonium acetate, 0.1 M sodium acetate (pH 4.6), and 30% (w/v) PEG4000. The enzyme-ligand complex crystals were obtained by mixing 1 μL of 10 mg/mL enzyme solution with an equal volume of a reservoir solution containing 10 mM FA or SA. A droplet was equilibrated against 100 μL reservoir solution. Diffraction data were collected at 100 K at a wavelength of 1.0 Å using an Eiger X 4M detector (Dectris) on beamline BL26B1 at SPring-8 (Hyogo, Japan) after flash-cooling in the cold nitrogen gas stream at 100 K in the presence (FA and SA complex) or absence (apo) of 25% (v/v) ethylene glycol.

The collected data were processed using the XDS program package [45]. The structure of *As*FaeE was solved by the molecular replacement method with the MOLREP program in the CCP4 package [46] using the structure of AXE from *A. luchuensis* (PDB: 5X6S [29]) as a starting model. Model rebuilding and refinement were achieved using COOT [47] and PHENIX [48]. The secondary structure of *As*FaeE was assigned with

DSSP & Stride plugin [49] for the PyMOL. The figures for crystal structures were prepared with PyMOL software [50]. The 2D representation of ligand-protein interactions was visualized using LigPlot^{+} v.2.2 (EMBL-EBI, Cambridgeshire, UK) [51]. Structure-based sequence alignment of *As*FaeE with either *Al*AXE, FA-bound *Am*Fae1A (*Am*Fae1A/FA; PDB: 5CXX [21]) or FA-bound *An*FaeA (*An*FaeA/FA; PDB: 1UWC [19]) was performed using PROMALS3D [52]. The cavity volumes of the substrate-binding pocket were calculated using CASTp analysis [32]. The secondary structure alignments, examined in the previous step, were visualized using an ESPript online tool [53]. The structure of *As*FaeE in complex with either the model substrates, MFA or MSA, was modelled using the COOT. MFA and MSA were superimposed onto the electron density map of FA and SA within the structures of *As*FaeE/FA and *As*FaeE/SA.

2.5 References

[1] Kurowska, K., Marks-Bielska, R., Bielski, S., Kryszk, H., Jasinskas, A., Food security in the context of liquid biofuels production, *Energies*. (2020) 13 6247.

[2] Zoghlami, A., Paës, G., Lignocellulosic biomass: understanding recalcitrance and predicting hydrolysis, *Frontiers in Chemistry*. (2019) 7 874.

[3] Scheller, H.V., Ulvskov, P., Hemicelluloses, *Annual Review of Plant Biology*. (2010) 61 263–289.

[4] Østby, H., Várnai, A., Hemicellulolytic enzymes in lignocellulose processing, *Essays in Biochemistry*. (2023) 67 533–550.

[5] Szwajgier, D., Waśko, A., Targoński, Z., Niedźwiadek, M., Bancarzewska, M., The use of a novel ferulic acid esterase from *Lactobacillus acidophilus* K1 for the release of phenolic acids from brewer's spent grain, *Journal of the Institute of Brewing*. (2010) 116 293–303.

[6] Underlin, E.N., Frommhagen, M., Dilokpimol, A., van Erven, G., de Vries, R.P., Kabel, M.A., Feruloyl esterases for biorefineries: subfamily classified specificity for natural substrates, *Frontiers in Bioengineering and Biotechnology*. (2020) 8 332.

[7] Braga, C.M.P., Delabona, P. da S., Lima, D.J. da S., Paixão, D.A.A., Pradella, J.G. da C., Farinas, C.S., Addition of feruloyl esterase and xylanase produced on-site improves sugarcane bagasse hydrolysis, *Bioresource Technology*. (2014) 170 316–324.

[8] Mkabayi, L., Malgas, S., Wilhelmi, B.S., Pletschke, B.I., Evaluating feruloyl esterase-xylanase synergism for hydroxycinnamic acid and xylo-oligosaccharide production from untreated, hydrothermally pre-treated and dilute-acid pre-treated corn cobs, *Agronomy*. (2020) 10 688.

[9] Schmitz, E., Leontakianakou, S., Norlander, S., Nordberg Karlsson, E., Adlercreutz, P., Lignocellulose degradation for the bioeconomy: the potential of enzyme synergies between xylanases, ferulic acid esterase and laccase for the production of arabinoxylo-oligosaccharides, *Bioresource Technology*. (2022) 343 126114.

[10] Benoit, I., Danchin, E.G.J., Bleichrodt, R.J., De Vries, R.P., Biotechnological applications and potential of fungal feruloyl esterases based on prevalence, classification and biochemical diversity, *Biotechnology Letters*. (2008) 30 387–396.

[11] Dilokpimol, A., Mäkelä, M.R., Aguilar-Pontes, M.V., Benoit-Gelber, I., Hildén, K.S., De Vries, R.P., Diversity of fungal feruloyl esterases: updated phylogenetic classification, properties, and industrial applications, *Biotechnology for Biofuels*. (2016) 9 231.

[12] Crepin, V.F., Faulds, C.B., Connerton, I.F., Functional classification of the microbial feruloyl esterases, *Applied Microbiology and Biotechnology*. (2004) 63 647–652.

[13] Antonopoulou, I., Dilokpimol, A., Iancu, L., Mäkelä, M.R., Varriale, S., Cerullo, G., Hüttner, S., Uthoff, S., Jütten, P., Piechot, A., Steinbüchel, A., Olsson, L., Faraco, V., Hildén, K.S., de Vries, R.P., Rova, U., Christakopoulos, P., The synthetic potential of fungal feruloyl esterases: a correlation with current classification systems and predicted structural properties, *Catalysts*. (2018) 8 242.

[14] Antonopoulou, I., Papadopoulou, A., Iancu, L., Cerullo, G., Ralli, M., Jütten, P., Piechot, A., Faraco, V., Kletsas, D., Rova, U., Christakopoulos, P., Optimization of enzymatic synthesis of L-arabinose ferulate catalyzed by feruloyl esterases from *Myceliophthora thermophila* in detergentless microemulsions and assessment of its antioxidant and cytotoxicity activities, *Process Biochemistry*. (2018) 65 100–108.

[15] Kroon, P.A., Williamson, G., Fish, N.M., Archer, D.B., Belshaw, N.J., A modular esterase from *Penicillium funiculosum* which releases ferulic acid from plant cell walls and binds crystalline cellulose contains a carbohydrate binding module, *Journal of Biochemistry* (2000) 267 6740–6752.

[16] Crepin, V.F., Faulds, C.B., Connerton, I.F., A non-modular type B feruloyl esterase from *Neurospora crassa* exhibits concentration-dependent substrate inhibition, *Biochemical Journal* (2003) 370 417.

[17] Kühnel, S., Pouvreau, L., Appeldoorn, M.M., Hinz, S.W.A., Schols, H.A., Gruppen, H., The ferulic acid esterases of *Chrysosporium lucknowense* C1: purification,

characterization and their potential application in biorefinery, *Enzyme and Microbial Technology*. (2012) 50 77–85.

[18] Dilokpimol, A., Mäkelä, M.R., Varriale, S., Zhou, M., Cerullo, G., Gidijala, L., Hinkka, H., Brás, J.L.A., Jütten, P., Piechot, A., Verhaert, R., Hildén, K.S., Faraco, V., de Vries, R.P., Fungal feruloyl esterases: functional validation of genome mining based enzyme discovery including uncharacterized subfamilies, *New Biotechnology*. (2018) 41 9–14.

[19] McAuley, K.E., Svendsen, A., Patkar, S.A., Wilson, K.S., Structure of a feruloyl esterase from *Aspergillus niger*, *Acta Crystallographica Section D: Biological Crystallography*. (2004) 60 878–887.

[20] Hermoso, J.A., Sanz-Aparicio, J., Molina, R., Juge, N., González, R., Faulds, C.B., The crystal structure of feruloyl esterase A from *Aspergillus niger* suggests evolutive functional convergence in feruloyl esterase family, *Journal of Molecular Biology*. (2004) 338 495–506.

[21] Gruninger, R.J., Cote, C., McAllister, T.A., Abbott, D.W., Contributions of a unique *β*-clamp to substrate recognition illuminates the molecular basis of exolysis in ferulic acid esterases, *Biochemical Journal*. (2016) 473 839–849.

[22] 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 Research*. (2013) 41 D423–D429.

[23] Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Research*. (1997) 25 3389–3402.

[24] Gupta, R., Brunak, S., Prediction of glycosylation across the human proteome and the correlation to protein function, *Pacific Symposium on Biocomputing*. (2002) 7 310–322.

[25] Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T.B.G., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., Gupta, R., Paul Bennett, E., Mandel, U., Brunak, S., Wandall, H.H., Levery, S.B., Clausen, H., Precision mapping of the human *O*-GalNAc glycoproteome through SimpleCell technology, *EMBO Journal*. (2013) 32 1478–1488.

[26] Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., UCSF ChimeraX: structure visualization for researchers, educators, and developers, *Protein Science*. (2021) 30 70–82.

[27] Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G.,

Goldman, A., The α/β hydrolase fold, *Protein Engineering, Design and Selection*. (1992) 5 197–211.

[28] Denesyuk, A., Dimitriou, P.S., Johnson, M.S., Nakayama, T., Denessiouk, K., The acid-base-nucleophile catalytic triad in ABH-fold enzymes is coordinated by a set of structural elements, *PLoS ONE*. (2020) 15 e0229376.

[29] 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*, *Applied and Environmental Microbiology*. (2017) 83 e01251-17.

[30] Simón, L., Goodman, J.M., Enzyme catalysis by hydrogen bonds: the balance between transition state binding and substrate binding in oxyanion holes, *Journal of Organic Chemistry*. (2010) 75 1831–1840.

[31] Yang, H., Wong, M.W., Oxyanion hole stabilization by C-H·O Interaction in a transition state—A three-point interaction model for cinchona alkaloid-catalyzed asymmetric methanolysis of meso-cyclic anhydrides, *Journal of the American Chemical Society*. (2013) 135 5808–5818.

[32] Tian, W., Chen, C., Lei, X., Zhao, J., Liang, J., CASTp 3.0: computed atlas of surface topography of proteins, *Nucleic Acids Research*. (2018) 46 W363–W367.

[33] Park, C., Visual interpretation of the meaning of k_{cat}/K_M in enzyme kinetics, *Journal of Chemical Education*. (2022) 99 2556–2562.

[34] Qi, M., Wang, P., Selinger, L.B., Yanke, L.J., Forster, R.J., Mcallister, T.A., Isolation and characterization of a ferulic acid esterase (Fae1A) from the rumen fungus *Anaeromyces mucronatus*, *Journal of Applied Microbiology*. (2011) 110 1341–1350.

[35] Faulds, C.B., Molina, R., Gonzalez, R., Husband, F., Juge, N., Sanz-Aparicio, J., Hermoso, J.A., Probing the determinants of substrate specificity of a feruloyl esterase, *An*FaeA, from *Aspergillus niger*, *FEBS Journal*. (2005) 272 4362–4371.

[36] The UniProt Consortium, UniProt: the Universal protein knowledgebase in 2023, *Nucleic Acids Research*. (2023) 51 D523–D531.

[37] Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., Nielsen, H., SignalP 5.0 improves signal peptide predictions using deep neural networks, *Nature Biotechnology*. (2019) 37 420–423.

[38] Tamura, K., Stecher, G., Kumar, S., MEGA11: molecular evolutionary genetics analysis version 11, *Molecular Biology and Evolution*. (2021) 38 3022–3027.

[39] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A., Protein identification and analysis tools on the ExPASy server, in: John

M. Walker (Ed.), The Proteomics Protocols Handbook, Humana Press, 2005: pp. 571–607.

[40] Simonian, M.H., Spectrophotometric determination of protein concentration, Current Protocols in Cell Biology. (2002) 15 A.3B.1-A.3B.7.

[41] Kelly, S.M., Jess, T.J., Price, N.C., How to study proteins by circular dichroism, *Biochimica et Biophysica Acta - Proteins and Proteomics.* (2005) 1751 119–139.

[42] Dilokpimol, A., Mäkelä, M.R., Mansouri, S., Belova, O., Waterstraat, M., Bunzel, M., de Vries, R.P., Hildén, K.S., Expanding the feruloyl esterase gene family of *Aspergillus niger* by characterization of a feruloyl esterase, FaeC, *New Biotechnology*. (2017) 37 200–209.

[43] Xu, J., Zhao, X., Yao, Q., Zong, W., Dai, S., Deng, Z., Liu, S., Yun, J., Yang, X., Li, H., Cloning, characterization of a novel acetyl xylan esterase, and its potential application on wheat straw utilization, *All Life.* (2021) 14 622–635.

[44] Kirsch, P.D., Ekerdt, J.G., KaleidaGraph: graphing and data analysis. version 3.5 for windows synergy software, 2457 Perkiomen Ave., Reading, PA 19606-2049. www.Synergy.com. $155.00, *Journal of the American Chemical Society*. (2000) 122 11755.

[45] Kabsch, W., XDS, *Acta Crystallographica Section D Biological Crystallography*. (2010) 66 125–132.

[46] Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., Overview of the CCP4 suite and current developments, *Acta Crystallographica Section D: Biological Crystallography*. (2011) 67 235–242.

[47] Emsley, P., Cowtan, K., Coot: model-building tools for molecular graphics, *Acta Crystallographica Section D: Biological Crystallography*. (2004) 60 2126–2132.

[48] Adams, P.D., Afonine, P. v., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D: Biological Crystallography*. (2010) 66 213–221.

[49] Kabsch, W., Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers*. (1983) 22 2577–2637.

[50] Delano, W.L., The PyMOL molecular graphics system, in: Delano Scientific, San Carlos, 2002.

[51] Laskowski, R.A., Swindells, M.B., LigPlot+: multiple ligand-protein interaction diagrams for drug discovery, *Journal of Chemical Information and Modeling*. (2011) 51 2778–2786.

[52] Pei, J., Kim, B.H., Grishin, N. v., PROMALS3D: a tool for multiple protein sequence and structure alignments, *Nucleic Acids Research*. (2008) 36 2295–2300.

[53] Robert, X., Gouet, P., Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Research*. (2014) 42 W320–W324.

Chapter Three

Structure-based characterization and improvement of an enzymatic activity of *Acremonium alcalophilum* feruloyl esterase

Reproduced from Ref. *ACS Sustain. Chem. Eng.*, 2024 (In press) with permission from the American Chemical Society

3.1 Introduction

Lignocellulosic biomass, such as sugarcane, grass, and straw, is an abundant and renewable resource that exhibits potential for biofuel and biochemical production, which could help alleviate food insecurity [1]. However, its recalcitrant structure poses a significant obstacle to enzymatic conversion [2]. Composed of cellulose, hemicellulose, and lignin, lignocellulosic biomass presents particular challenges in hemicellulose breakdown due to its complex, branched structure. Hemicellulose contains a variety of sugars (e.g., xylose, arabinose, and glucose) and carboxylic acids (e.g., acetic acid and ferulic acid (FA)) [3]. FA and its dimeric forms (diferulates) play important roles in biomass recalcitrance, forming ester bonds between hemicellulose and lignin, as well as between two hemicelluloses [3,4]. These ester bonds present major obstacles for the breakdown of hemicellulose by a single enzyme. Therefore, a combination of hemicellulases, including feruloyl esterase (FAE) and acetyl xylan esterase (AXE), is required to improve hemicellulose degradability [5–7].

The general function of FAE (EC 3.1.1.73) is to catalyze the hydrolysis of the feruloyl ester linkage between FA and polysaccharides. FAE is a member of carbohydrate esterase family 1 (CE1), a family defined in the carbohydrate-active enzymes database (CAZy). Fungal esterases can be classified phylogenetically into five subfamilies (SF1–5) **(Fig. 3-1)** [8,9]. Some members of SF1, SF2, and SF5 have been previously studied and characterized, such as AXEs belonging to SF1 [8]. It should be noted that different classification systems were proposed previously. In the ESTerases and *α*/*β*-Hydrolase Enzymes and Relatives (ESTHER) database, fungal CE1 members are classified into the

Esterase_phb, Antigen85c, and FaeC families [10]. The Esterase_phb family comprises members of SF1 and SF2, which are closely related **(Fig. 3-1)**. The Antigen85c family comprises members of SF3, while the FaeC family has recently been shown to include members of SF5. There is currently no corresponding ESTHER family for members of SF4.

Studying the members of SF5 is important as they have the ability to catalyze hydrolysis of a wide range of natural substrates, releasing various chemicals such as acetic acid, hydroxycinnamic acids, and diferulate [11–13]. Moreover, certain SF5 members were shown to exhibit transesterification activity towards feruloyl derivatives, which is attractive for chemical synthesis [14,15]. Understanding the molecular mechanism by which SF5 members catalyze these reactions is crucial when it comes to applying these enzymes in industrial processes. However, despite structural and structure-function relationship studies of members of SF1 [8], SF2 (Esterase_phb) [16], and SF3 (Antigen85c) [17], no such studies have been conducted on the members of SF5 (FaeC family in the ESTHER database).

This study is significant because it is the first to determine the structure of an SF5 FAE. Through biochemical analyses, I characterized *Aa*FaeD from *Acremonium alcalophilum*. I obtained the crystal structures of both the free form and a complex form with FA, the product from common substrate methyl ferulate, revealing the substrate-binding mode for SF5 FAE. I also conducted structure-based functional mutagenesis to determine the roles of amino acid residues located around the catalytic core in binding affinity and catalytic rate. Structural comparison between SF5 and other families in CE1 highlights the unique features of SF5, providing insight into how it can bind to substrates of different sizes. Finally, the release of bulkier 5,5'-dehydrodiferulate (5,5'-diFA) from destarched wheat bran (DSWB) by SF5 FAE was also investigated by chromatographic separation and mass spectrometry. These findings offer valuable structural insights into SF5 FAE, paving the way for innovative biotechnology solutions for sustainable biomass degradation.

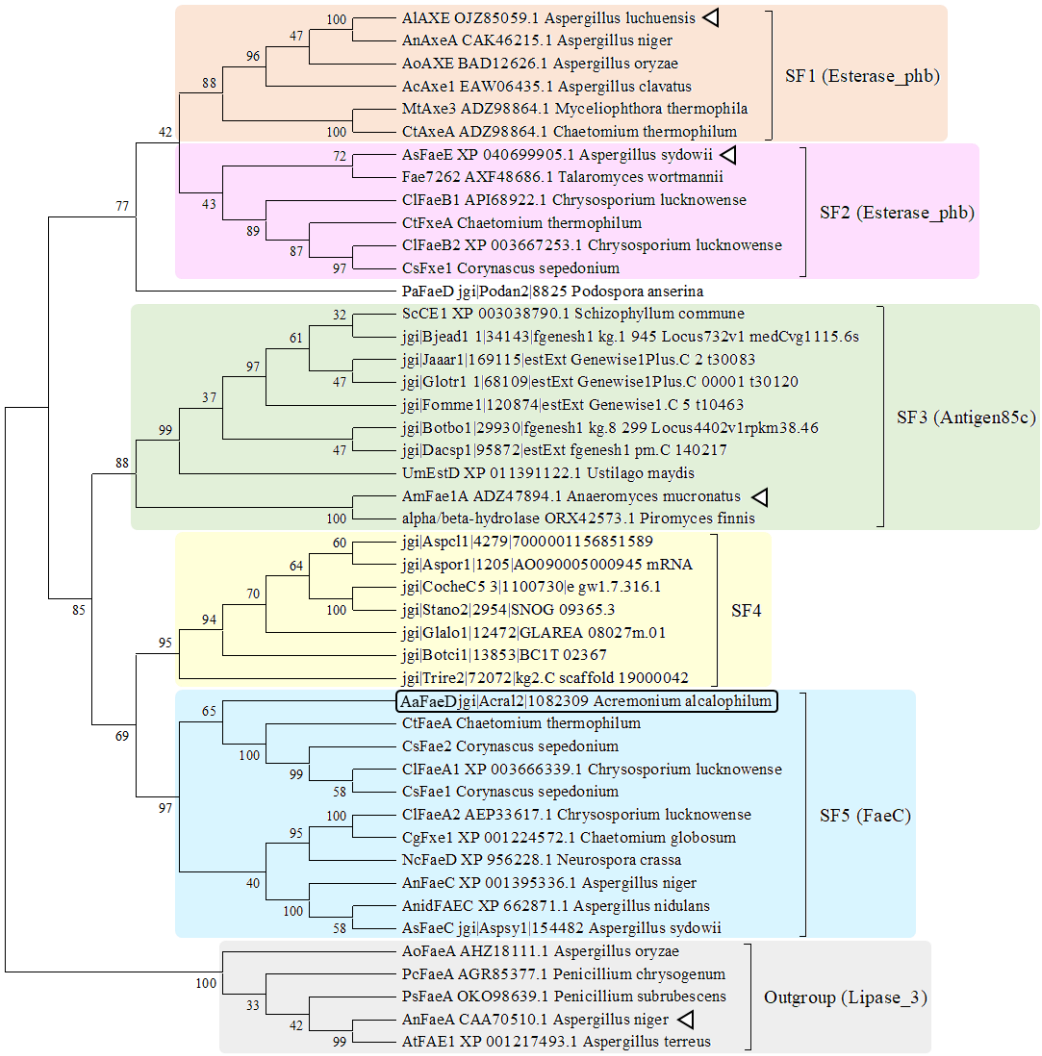

Fig. 3-1. Phylogenetic relationships among fungal CE1 members. The phylogenetic analysis was conducted with a WAG+G+I substitution model and the maximum-likelihood method with 100 bootstrap resampling. Numbers indicate the statistical values (Expected-Likelihood Weight) of internal nodes, given in percentages. The selected FAEs were classified into five subfamilies (SFs) based on the previous reports (Komiya, *et al.*, 2017 [8] and Dilokpimol, *et al.*, 2022 [12]). Five FAEs from SF7 (Dilokpimol, *et al.* 2016) [18] were used as an outgroup. The names of the esterase classes according to the ESTerases and *α*/*β*-Hydrolase Enzymes and Relatives (ESTHER) classifications system are shown in parentheses. The FAEs whose structures were solved previously are indicated by arrowheads. *Aa*FaeD is boxed.

3.2 Results and discussions

3.2.1 Identification of A. alcalophilum FaeD

The gene of my interest from *A. alcalophilum*, belonging to the ascomycetes, encodes a putative FAE (*Aa*FaeD) (JGI protein ID: Acral2|1082309), which is newly classified as a member of SF5 of CE1 in the CAZy database **(Fig. 3-1)** [19]. In the ESTHER database, *Aa*FaeD belongs to the FaeC family [10]. Amino acid sequence alignment for the members of SF5 showed moderate sequence identity of approximately 40–55% between *Aa*FaeD and ten previously studied FAEs of ascomycetes species, e.g., *Neurospora crassa* [20], *Aspergillus niger* [21], *Aspergillus sydowii*, *Aspergillus nidulans* [22], *Chrysosporium lucknowense* [23], *Chaetomium globulosum*, *Chaetomium thermophilum*, and *Corynascus sepedonium* [12] (Table S1).

3.2.2 Preparation of active AaFaeD

Sequence analysis of *Aa*FaeD indicated that it comprises a catalytic domain (CD), a linker region, and carbohydrate-binding module 1 (CBM1) **(Fig. 3-2 A)**. I made two expression constructs for N-terminal 6× histidine-tagged *Aa*FaeD, the full-length *Aa*FaeD (*Aa*FaeD), and that containing only the CD (*Aa*FaeD-CD). They were individually cloned and expressed in *P. pastoris* X-33. The calculated molecular mass (*m*) of the recombinant *Aa*FaeD and *Aa*FaeD-CD were 34,900 and 27,400 Da, respectively. *N*- and *O*-glycosylation sites of *Aa*FaeD were predicted with NetNGlyc 1.0 and NetOGlyc 4.0. For an *N*-glycosylation site, Asn109 (within CD), and for *O*-glycosylation sites, Ser250, Thr254, Ser255, Thr259, Thr264, Thr269, and Thr274 (all in the linker region), were obtained. As indicated by SDS-PAGE analysis, *Aa*FaeD and *Aa*FaeD-CD contained various species with different *m* ranges, i.e., 35,000–40,000 Da and 28,000–32,000 Da, respectively **(Fig. 3-2 B)**. *N*-deglycosylation treatment with Endo H_f resulted in a single *m* species of ca. 28,000 Da for *Aa*FaeD-CD, while *Aa*FaeD still contained different *m* species **(Fig. 3-2 B)**.

The activity assay for *Aa*FaeD and *Aa*FaeD-CD showed that they both have activity towards model substrates, methyl hydroxycinnamates (methyl *p*-coumarate

(M*p*CA), methyl caffeate (MCA), methyl ferulate (MFA), and methyl sinapate (MSA)), and *p*-nitrophenyl acetate (*p*NPAc) **(Fig. 3-2 C)**.

I conducted an AlphaFold2 [24] structural prediction of the full-length *Aa*FaeD, as illustrated in **Fig. 3-3**. The prediction indicated that the *Aa*FaeD-CD and CBM1 regions are structured, while the proline-rich linker is unstructured **(Fig. 3-2 A and 3-3)**. Consequently, *Aa*FaeD-CD and CBM1 regions appear to function independently. This implies that CBM1 does not regulate substrate binding to the substrate binding site of *Aa*FaeD-CD, consistent with my observation of comparable activities between *Aa*FaeD and *Aa*FaeD-CD. Therefore, in this study, I focus on functional and structural study of *Aa*FaeD-CD.

It was also suggested that deglycosylation has no effect on the enzymatic activity of both *Aa*FaeD and *Aa*FaeD-CD **(Fig. 3-4)**. Therefore, biochemical and structural analyses were carried out for *Aa*FaeD-CD, which was not deglycosylated.

For crystal structure determination, *Aa*FaeD-CD obtained after size-exclusion chromatography was used **(Fig. 3-5)**. Specifically, the fractions 40–53 comprising bands II and III were pooled and used for the crystallization of *Aa*FaeD-CD **(Fig. 3-5)**. It is worth noting that *Aa*FaeD-CD consists of three different sizes due to differences in glycosylation **(Fig. 3-2 B)**.

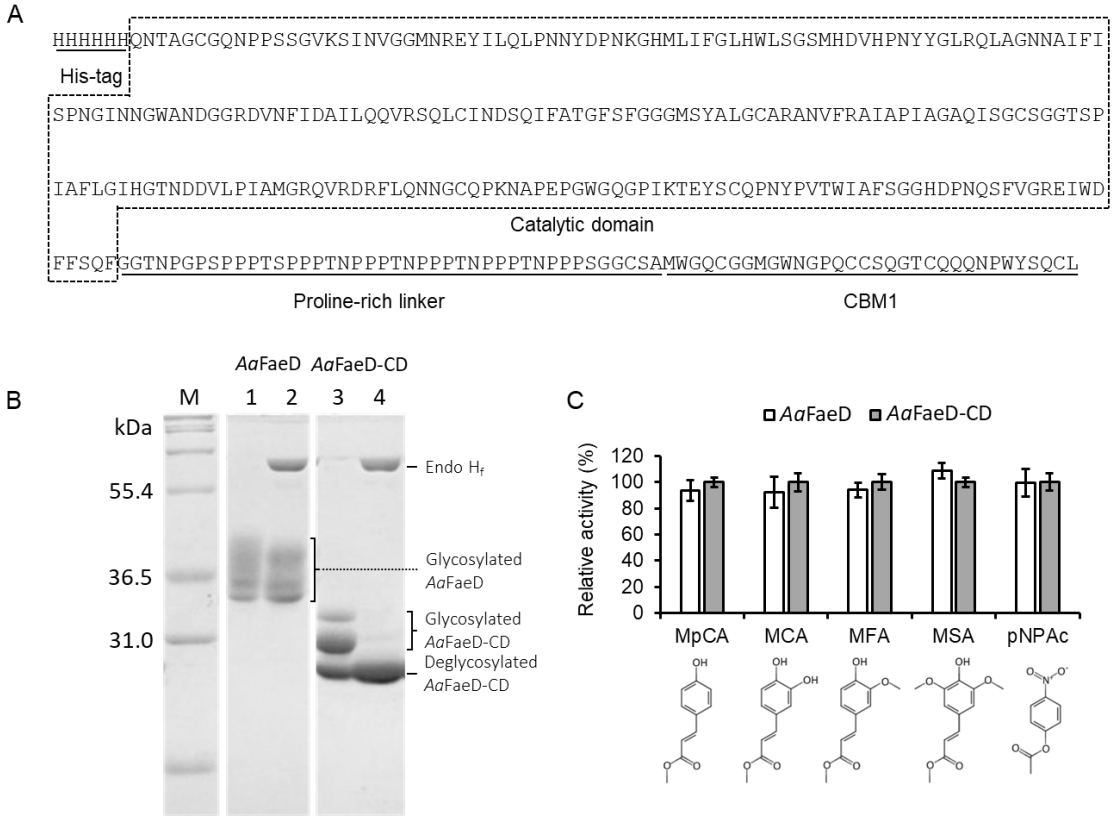

Fig. 3-2. Amino acid sequence, purification, and enzymatic activity of the catalytic domain of *Aa*FaeD. (A) Amino acid sequence of full-length *Aa*FaeD. The N-terminal histidine tag, catalytic domain (CD), proline-rich linker, and carbohydrate-binding module1 (CBM1) are either boxed or underlined with a label. (B) SDS-PAGE analysis of the purified full-length *Aa*FaeD (*Aa*FaeD) and CD of *Aa*FaeD (*Aa*FaeD-CD). The *Aa*FaeD and *Aa*FaeD-CD not treated (lane 1 and 3, respectively) or treated (lane 2 and 4, respectively) with Endo H_f; molecular-mass marker (lane M). The bands of *N*-glycosylated and deglycosylated proteins, and Endo H_f are indicated. (C) Activity of *Aa*FaeD and *Aa*FaeD-CD against model substrates. *Aa*FaeD-CD activity was set as 100%. The chemical structures of the model substrates are displayed below. Methyl *p*-coumarate (M*p*CA), methyl caffeate (MCA), methyl ferulate (MFA), and methyl sinapate (MSA) are model substrate for FAEs, while *p*-nitrophenyl acetate (*p*NPAc) is a model substrate for esterases.

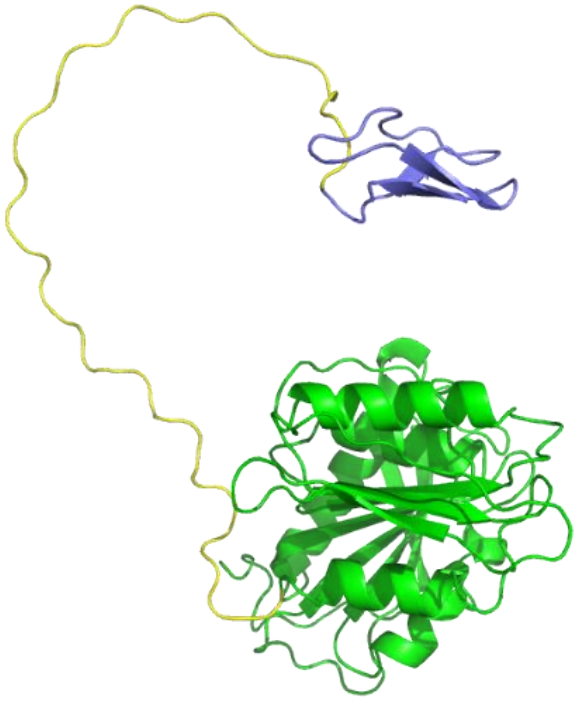

Fig. 3-3. Structure of *Aa*FaeD full-length predicted by AlphaFold2 [24]. The catalytic domain, proline-rich linker, and carbohydrate-binding domain 1 (CBM1) were colored green, yellow, and blue, respectively.

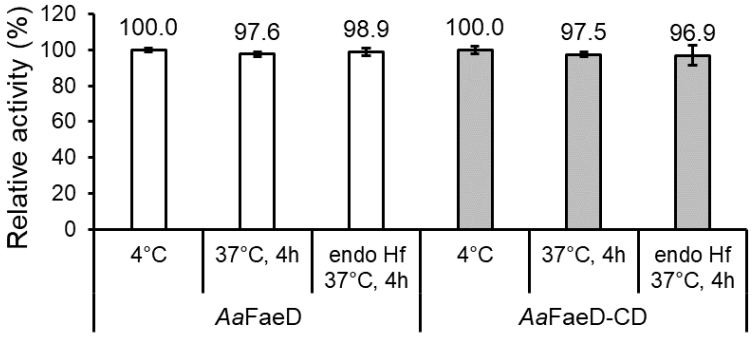

Fig. 3-4. Relative activity of non-deglycosylated and *N*-deglycosylated *Aa*FaeD and *Aa*FaeD-CD determined after incubation for 4 h at temperature 37 °C.

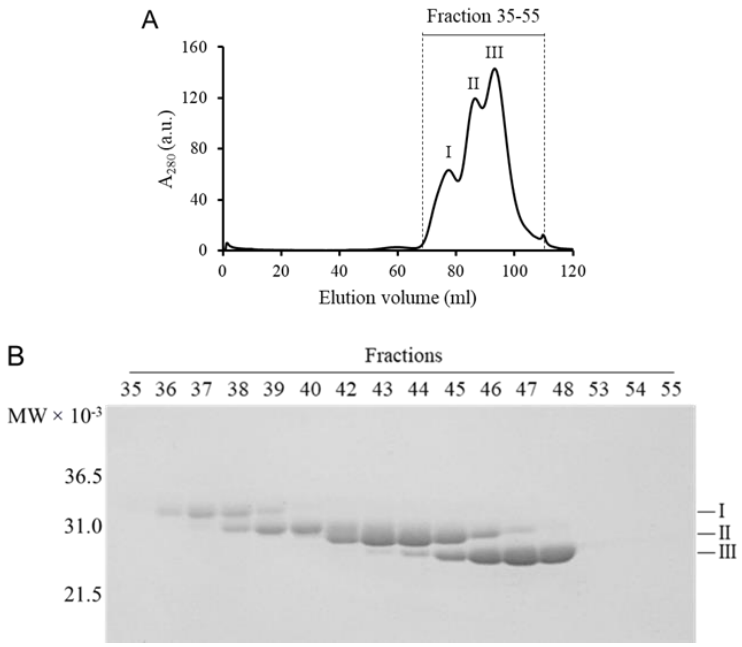

Fig. 3-5. Size-exclusion chromatography (SEC), an additional purification step for protein crystallization. (A) The size-exclusion chromatogram of *Aa*FaeD-CD; the protein solution applied to SEC was the one obtained after Ni^{2+} affinity chromatography (Lane 3, **Fig. 3-2 B**). A HiLoad 16/600 Superdex 75 column (GE Healthcare) equilibrated with 25 mM potassium phosphate, pH 6.5, was used. The flow rate was 1 mL/min. (B) SDS-PAGE analysis of the SEC fractions. The fraction numbers (35–55) are indicated on top. *Aa*FaeD-CD with high *m*-glycosylation, low *m*-glycosylation, and no glycosylation is indicated as I, II, and III, respectively. Crystallization of *Aa*FaeD-CD was performed using a solution comprising fractions 40–53.

3.2.3 Biochemical properties of AaFaeD-CD

Here the effects of temperature and pH on activity of *Aa*FaeD-CD were determined. The activity assay of *Aa*FaeD-CD at pH 7.5 and different temperatures (15–65 °C) indicated that *Aa*FaeD-CD exhibits the highest activity at 40 °C (the optimum temperature) **(Fig. 3-6 A)**. At temperatures between 20 and 55 °C, the activity was higher than 60% of the activity at the optimum temperature. The activity decreased to below 20% when the temperature was above 65 °C. Next, to determine thermal stability, *Aa*FaeD-CD was incubated for 2 h at pH 7.5 and different temperatures (15–65 °C); after

incubation, the activity was measured at the optimum temperature and pH 7.5. *Aa*FaeD-CD incubated between 15 and 45 °C retained the highest level of activity but the activity significantly decreased when *Aa*FaeD-CD was incubated at 55 and 65 °C **(Fig. 3-6 B)**.

The enzymatic activity of *Aa*FaeD-CD measured at different pHs (pH 6.5 and 9.0) at 40 °C indicated that *Aa*FaeD-CD exhibits the highest activity at pH 7.5 (the optimum pH). At pHs between 6.5 and 9.0, the activity was higher than 60% of the activity at the optimum pH **(Fig. 3-6 C)**. To determine pH stability, *Aa*FaeD-CD was incubated for 24h at different pHs (6.5–9.0) and the optimum temperature, 40 °C; after incubation, the activity was measured at the optimum pH, 7.5, and at the optimum temperature. *Aa*FaeD-CD incubated between pH 5.0–9.0 retained more than 90% of the activity at the optimum pH and temperature **(Fig. 3-6 D)**.

The optimum pH (7.5) and temperature (40 °C) of *Aa*FaeD-CD turned out to be similar to those of other cellulolytic enzymes from *A. alcalophilum* [25] and those of other FAEs in SF5 [9,12,21,23]. Here, I have revealed that *Aa*FaeD-CD exerts its enzymatic activity at the highest level at pHs around 7.0–8.5 and temperatures around 40–45 °C.

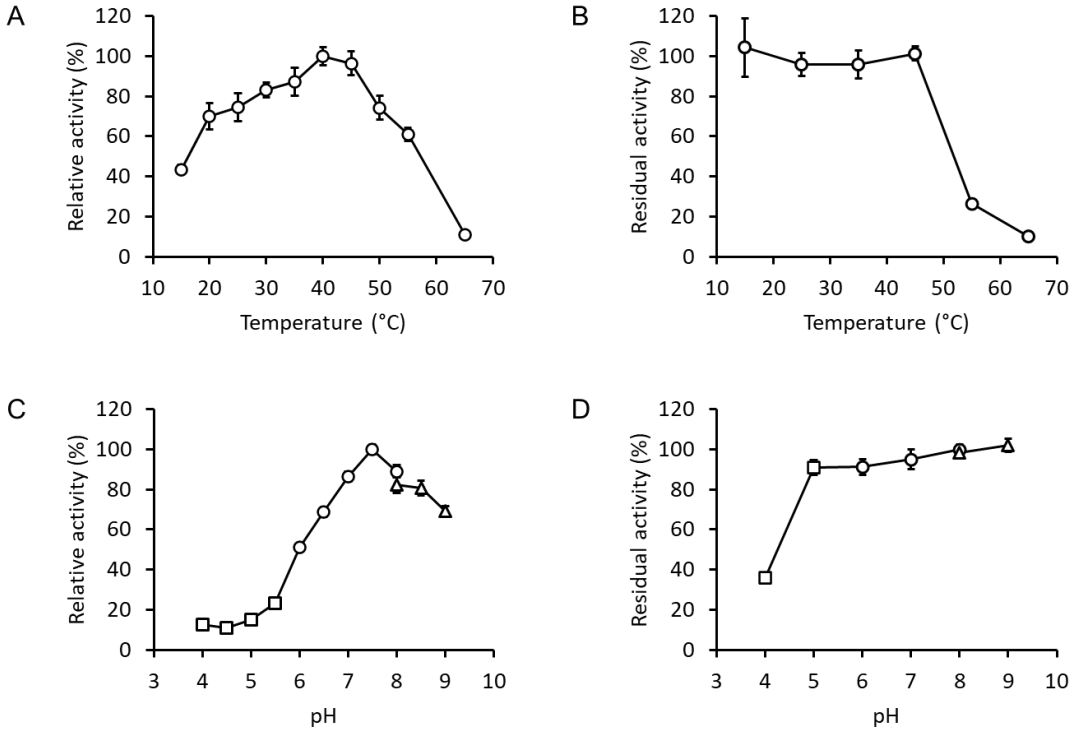

Fig. 3-6. Temperature and pH profiles of the enzymatic activity of *Aa*FaeD-CD against methyl ferulate (MFA). (A) Optimal temperature. The relative enzymatic activity of *Aa*FaeD-CD measured at the temperature indicated on the x-axis is plotted. (B) Thermostability. *Aa*FaeD-CD was incubated at various temperatures for 2 h and then the activity was measured at 40 °C. The residual activity is plotted against the temperature at which *Aa*FaeD-CD was incubated. (C) Optimal pH. The relative enzymatic activity of *Aa*FaeD-CD measured at the pH indicated on the x-axis is plotted. (D) pH stability. *Aa*FaeD-CD was incubated at various pHs for 24 h and then the activity was measured at pH 7.5. The residual activity is plotted against the pH at which *Aa*FaeD-CD was incubated: squares, pH 4.0–5.5 (80 mM sodium acetate); circles, pH 6.0–8.0 (80 mM potassium phosphate); triangles, pH 8.0–9.0 (80 mM Tris-HCl).

3.2.4 Structure determination of AaFaeD-CD and the AaFaeD-CD/FA complex

The *Aa*FaeD-CD crystal with a prism shape were grown under the optimized condition: 0.2 M $MgCl_2.6H_2O$, 0.1 M Tris-HCl (pH 8.5), and 30% (w/v) PEG4000 **(Fig.**

3-7). The crystal structure of *Aa*FaeD-CD in a free form was refined at 1.55 Å resolution after the structure was determined by molecular replacement using a predicted structure obtained by AlphaFold2 [24] as a search model. Iterative model building and refinement decreased R-factor and enhanced the quality of the electron density map. Unambiguous identification of the electron density for residues 20–261 of *Aa*FaeD was achieved. Additionally, the electron density for two *N*-acetylglucosamine (NAG) residues attached to Asn109 was also determined. The electron density of the N-terminal 6×His tag was ambiguous. The final model also contained 235 water molecules, one diethylene glycol molecule, and two magnesium ions. Crystallographic refinement converged with a final R_{work} of 14.87% with R_{free} of 22.39% **(Table 3-1)**.

To determine the crystal structure of the *Aa*FaeD-CD/FA complex, the capillary method after soaking was used to avoid the effect of cryo-solvent. The crystal structure of the *Aa*FaeD-CD/FA complex was determined at 2.00 Å resolution via molecular replacement using the crystal structure of *Aa*FaeD-CD determined above as a search model. Unambiguous identification of the electron density for residues 20–261, FA, and two NAG residues attached to Asn109 was achieved. The final model contained 133 water molecules, one ferulic acid molecule, and one magnesium ion. Crystallographic refinement converged with a final R_{work} of 13.20% with R_{free} of 18.23% **(Table 3-1)**. In this structure, the occupancy of FA was calculated to be 0.78, suggesting somehow weak FA binding.

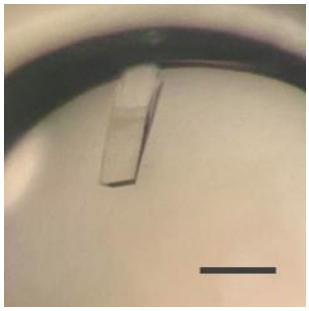

Fig. 3-7. The *Aa*FaeD-CD crystal grown under the optimized condition containing 0.2 M $MgCl_2.6H_2O$, 0.1 M Tris-HCl (pH 8.5), and 30% (w/v) PEG4000. Scale bars represent 0.5 mm.

Table 3-1. Crystallographic data and refinement statistics for *Aa*FaeD-CD and *Aa*FaeD-CD in a complex with ferulic acid (*Aa*FaeD-CD/FA).

	***Aa*FaeD-CD**	***Aa*FaeD-CD/FA**
PDB ID	**8JH8**	**8JH9**
A. Diffraction data		
X-ray source	SPring-8/BL26B1	SPring-8/BL26B1
Detector	Dectris Eiger X 4M	Dectris Eiger X 4M
Wavelength (Å)	1.0	1.0
Measurement temperature (K)	100 (cryo)	297 (capillary)
Resolution range (Å)*	50.00–1.55 (1.64–1.55)	50–2.00 (2.12–2.00)
Space group	C 1 2 1	C 1 2 1
Unit cell (Å)	a = 99.5, b = 37.7, c = 77.7 β = 125.9°	a = 101.3, b = 38.6, c = 77.9 β = 125.6°
Unique reflections*	33124 (4350)	16580 (2438)
Multiplicity*	5.23 (3.46)	4.21 (3.18)
Completeness (%)*	95.9 (78.3)	97.7 (89.8)
Mean I/σ (I)*	18.61 (2.84)	30.64 (10.18)
Wilson *B*-factor (Å^2)	29.8	31.8
R_{merge} (%)*	4.7 (29.3)	3.3 (13.3)
R_{meas} (%)*	5.1 (34.5)	3.8 (15.7)
$CC_{1/2}$ (%)*	99.9 (87.5)	99.9 (98.4)
B. Structure determination	Alphafold2 model	8JH8
C. Refinement statistics		
Resolution range*	37.78–1.55 (1.59–1.55)	41.19–2.00 (2.12–2.00)
R_{work} (%)*	14.87 (22.34)	13.20 (15.06)
R_{free} (%)*	22.39 (29.03)	18.23 (19.40)
Number of protein residues	242	242
PEG/MG/FER/NAG/water	1/2/0/2/235	0/1/1/2/133
R.m.s.d., bond lengths (Å)	0.005	0.006
R.m.s.d., bond angles (°)	0.734	0.769
Coordinate error (Å)	0.2	0.17
Ramachandran favored (%)	98.33	98.33
Ramachandran allowed (%)	1.67	1.67
Ramachandran outliers (%)	0	0
Rotamer outliers (%)	2.01	0.93
Clash score	2.41	4.09

* Values in parentheses are for the highest resolution shell.

3.2.5 Overview

The crystal structure of *Aa*FaeD-CD is illustrated in **Fig. 3-8 A**. *Aa*FaeD-CD forms an α-β-α sandwich composed of nine β-strands (β1–β9), five α-helices (α1–α5), and three 3_{10} helices (η1–η3) **(Fig. 3-8)**, which closely resembles a canonical α/β-hydrolase fold [26,27]. A characteristic feature of *Aa*FaeD-CD is the presence of a central β-sheet displaying a left-handed superhelical twist, which leads to an almost perpendicular alignment of β1 and β8. A classical Ser-His-Asp (S-H-D) catalytic triad, conserved among the members of CE, is also found in *Aa*FaeD, which comprises Ser119, Asp167, and His224. The well conserved disulfide bonds were also found in *Aa*FaeD, i.e., Cys6–Cys107, Cys130–Cys150, and Cys188–Cys209 **(Fig. 3-8 B)**.

The crystal structures of the *Aa*FaeD-CD/FA complex and free *Aa*FaeD-CD are superimposed in **Fig. 3-9 A**. In the crystal structure of FA-free *Aa*FaeD-CD, a diethylene glycol was observed to be bound in the hydrophobic cleft instead of FA **(Fig. 3-9 B)**. This appears to be the reason why FA could not bind to the active site during the cryo data collection after soaking. The average root-mean square deviation (RMSD) of the Cα atoms (residues 7–248) of the *Aa*FaeD-CD/FA complex and free *Aa*FaeD-CD was 0.22 Å, indicating that the overall structures are the same.

The crystal structure of the *Aa*FaeD-CD/FA complex, as depicted in **Fig. 3-9 C**, reveals that FA binds within the hydrophobic cleft, which is formed by a right wall consisting of Leu50 and Phe120, and a left wall composed of Val169, Leu170, Met174, and His224. Additionally, Ile147 was identified at the bottom of the hydrophobic cleft. The catalytic triad S-H-D, consisting of Ser119, Asp167, and His224, is positioned at the forefront facing the reader **(Fig. 3-9 B and C)**.

The binding of FA in the hydrophobic cleft is achieved through van der Waals interaction, including CH-π interaction between the methyl group of Ile147 and the aromatic ring of FA, and π-π T-shaped interaction between the aromatic rings of Phe120 and FA. Ser119 and Phe120 are contained in the GXSXG pentapeptide, which is conserved among esterases and lipases [28,29], and serine hydrolases [15]. It is known that GXSXG pentapeptide forms a “catalytic elbow” that is important for activity [30,31]. In addition, the carboxyl group of FA was found to be located inside the oxyanion hole

[32,33] formed by the backbone NH groups of Leu50 and Phe120, nitrogen atoms being colored blue in **(Fig. 3-9 C)**.

I examined the hydrogen bond network surrounding the carboxyl group of FA in the *Aa*FaeD-CD/FA complex, as illustrated in **Fig. 3-10 B**. The distances between the carboxyl oxygen of FA (proton acceptor) and the backbone nitrogen of Leu50 or Phe120 (proton donor) were approximately 3.0 Å, indicating the hydrogen bonding. These hydrogen bonds are the key interactions within the oxyanion hole.

Next, I analyzed the hydrogen bond network between the S-H-D catalytic triad residues in both free *Aa*FaeD-CD and the *Aa*FaeD-CD/FA complex (**Fig. 3-10 A and B**, respectively). In both cases, the distance between the $O^{\delta1}/O^{\delta2}$ of Asp167 and the $N^{\delta1}$ of His224, as well as the distance between the $N^{\varepsilon2}$ of His224 and the O^{γ} of Ser119, were approximately 2.5-2.6 and 3.0 Å, respectively, confirming the presence of hydrogen bonds. This demonstrates that free *Aa*FaeD-CD and the *Aa*FaeD-CD/FA complex share the same geometric arrangement for the S-H-D catalytic triad residues.

Typically, when the S-H-D catalytic triad residues catalyze the reaction, the serine residue serves as the nucleophile, attacking the ester bond of a substrate. The histidine residue acts as a base, deprotonating the serine residue and enhancing its nucleophilicity. The aspartate residue functions as an acid, stabilizing the protonated histidine residue. Therefore, the hydrogen bond network that I observed in both free *Aa*FaeD-CD and the *Aa*FaeD-CD/FA complex can reflect the initial and final stages, respectively. In other words, the *Aa*FaeD-CD/FA complex represents the final stage of the reaction, with the product FA poised to dissociate and the S-H-D catalytic triad residues forming the hydrogen bond network characteristic of the initial state, ready to accept the next substrate.

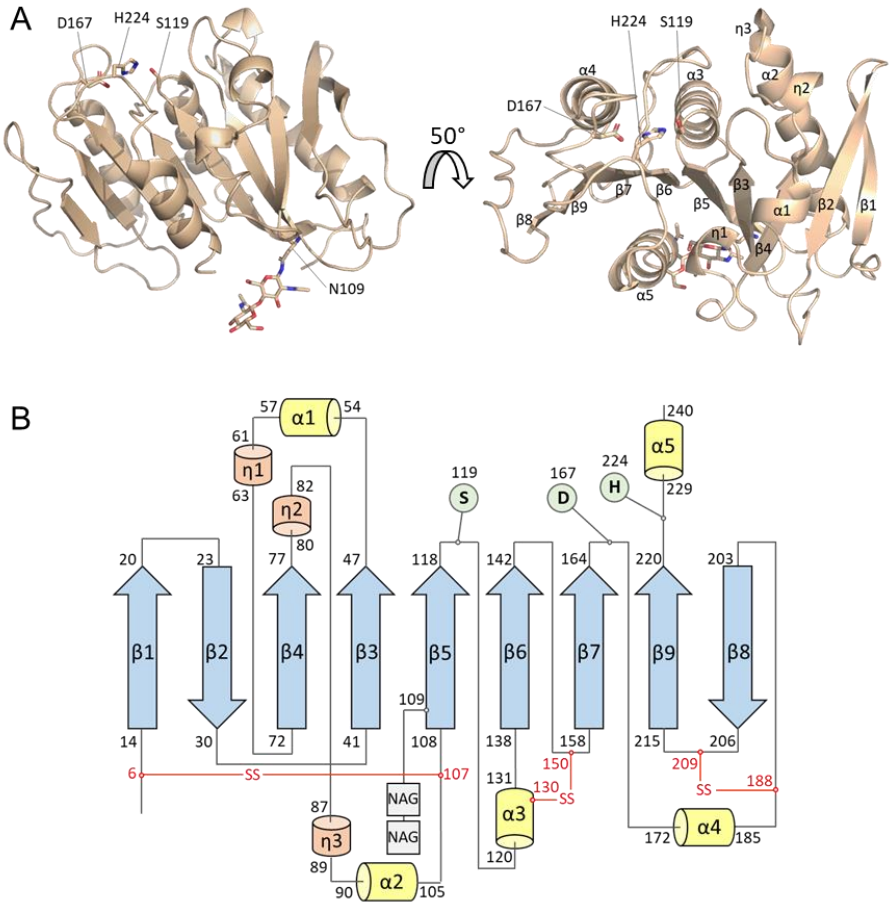

Fig. 3-8. Crystal structure of the catalytic domain of *Aa*FaeD, *Aa*FaeD-CD. (A) Ribbon model of *Aa*FaeD-CD. The S-H-D catalytic triad (Ser119, Asp167, and His224) and two *N*-acetyl glucosamine residues linked to Asn109 are shown as a stick model. The carbon, nitrogen, and oxygen atoms are colored beige, blue, and red, respectively. (B) Topological diagram of *Aa*FaeD-CD. β-strands (β1–β9), α-helices (α1–α5), and 3_{10} helices (η1–η3) are shown in blue, yellow, and orange, respectively. The residues forming the S-H-D catalytic triad and the disulfide bonds are indicated by green circles and red lines, respectively. *N*-acetylglucosamines (NAGs) are indicated by grey boxes.

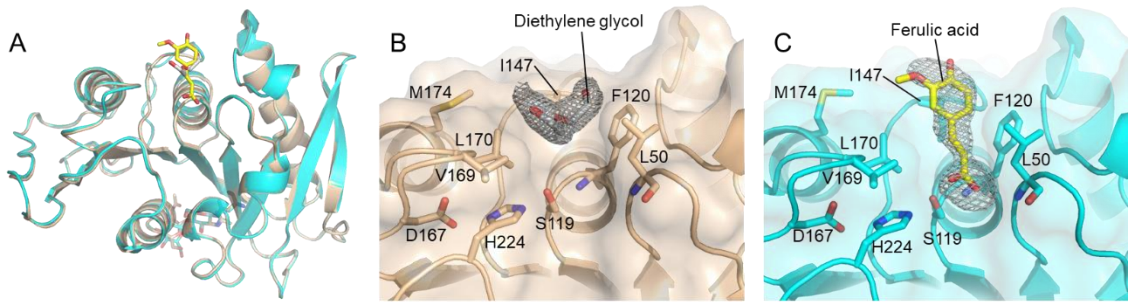

Fig. 3-9. Interaction between *Aa*FaeD and substrates revealed by the crystal structure of the complex between *Aa*FaeD-CD and ferulic acid (FA). (A) Superimposition of ribbon representations of the structures of *Aa*FaeD-CD in substrate-free (beige) and FA-bound forms (cyan). (B, C) Close-up views of the substrate-binding site of *Aa*FaeD-CD in substrate-free (beige) and FA-bound forms (cyan), respectively. The residues forming the S-H-D catalytic triad and those involved in substrate binding are shown as a stick model. A diethylene glycol (white stick) and FA (yellow stick), each bound at the substrate-binding site, are overlaid with the F_o-F_c omit map contoured at 3.0 σ in (B) and (C), respectively. The nitrogen, oxygen, and sulfur atoms are colored blue, red, and yellow, respectively.

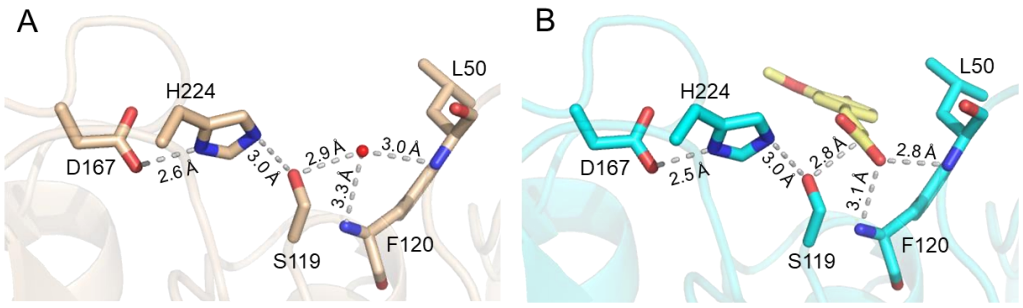

Fig. 3-10. Hydrogen bond network involving the S-H-D catalytic triad residues in free *Aa*FaeD-CD and the *Aa*FaeD-CD/FA complex. Close-up views of the S-H-D catalytic triad residues are depicted for free *Aa*FaeD-CD (A) and the *Aa*FaeD-CD/FA complex (B). The sticks represent the S-H-D catalytic triad residues, the residues involved in the oxyanion hole, a water oxygen atom (depicted as a red sphere in A), and ferulic acid (yellow in B). Nitrogen and oxygen atoms are colored blue and red, respectively. The dashed lines illustrate the distances between hydrogen bond donor and acceptor atoms. The distance between the O^{γ} of Ser in the S-H-D and the carbonyl carbon of FA is also presented. In (A) and (B), the structure is rotated by an angle of 80° away from the reader around the x-axis compared to the orientation shown in **Fig. 3-9**.

3.2.6 Structural sequence alignment

Since structure of *Aa*FaeD-CD was the first to be solved for the members of SF5, I performed structure-based sequence alignment to investigate conservation of the binding-site among SF5 members **(Fig. 3-11)**. It was shown that the S-H-D catalytic triad is conserved among the members. For the aforementioned GXSXG pentapeptide, in which Ser119 and Phe120 of *Aa*FaeD-CD correspond to the third and fourth positions, it was found that X is Phe, Tyr, or Trp for SF5 FAEs **(Fig. 3-11)**. Among the positions of the residues forming the hydrophobic cleft in *Aa*FaeD-CD, the positions of Val169 and Leu170 (left wall) were conserved completely, while those of Phe120 (right wall) and Ile147 (bottom) contained similar amino acid in some cases. Notably, compared to other SF5 FAEs, Phe120 is typically replaced by Tyr or Trp, whereas Ile147 is often substituted with Leu or Val **(Fig. 3-11)**. The consistent property of position 120 being an aromatic residue and position 147 being a bulky aliphatic residue, as observed in other SF5 FAEs, suggests their significant roles in shaping the hydrophobic cleft **(Fig. 3-9)**. On the other hand, the positions for Leu50 (right wall) and Met174 (left wall) showed considerable variability.

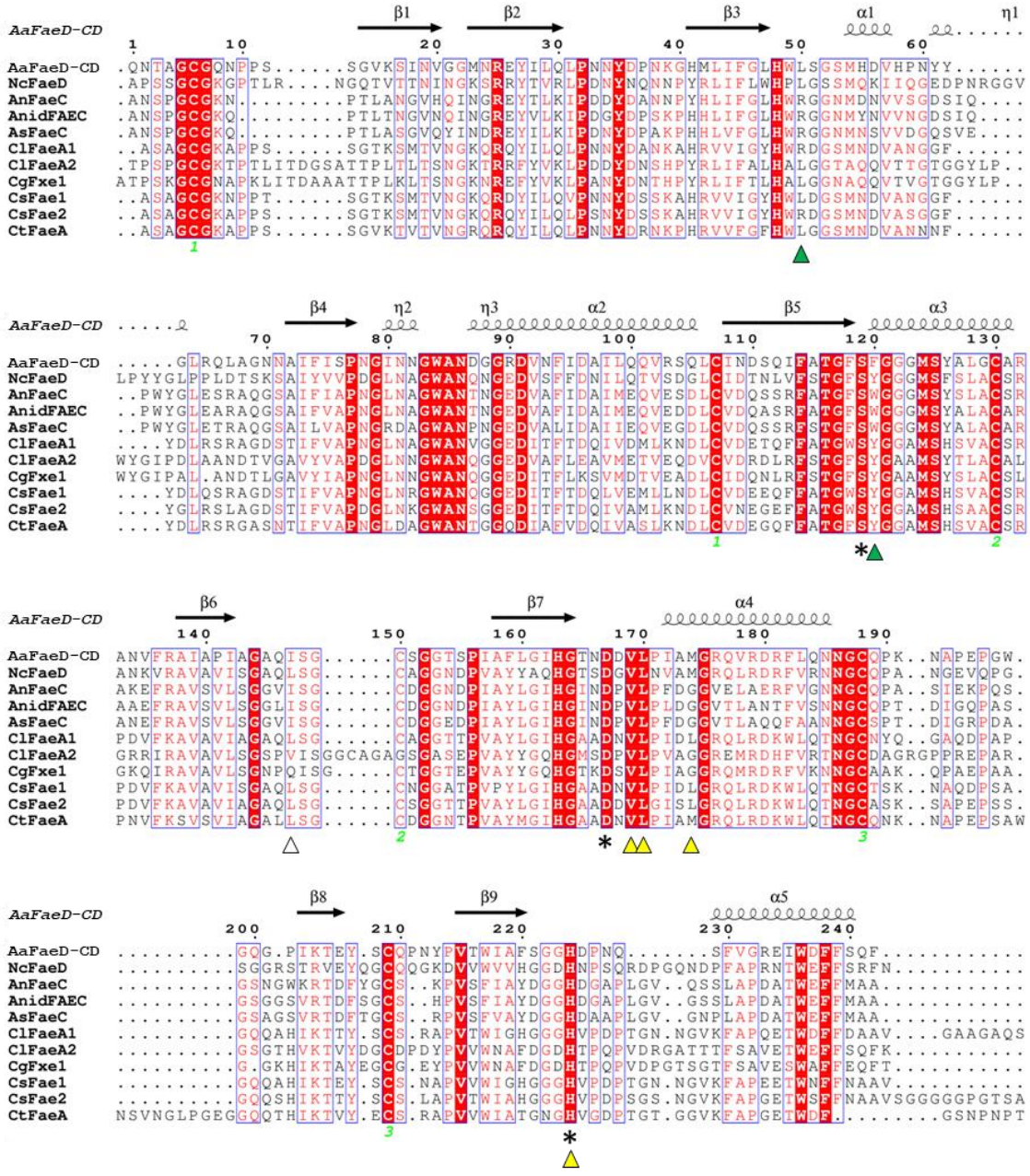

Fig. 3-11. Sequence alignment of CDs of FAEs belonging to SF5 of CE1. The secondary structures, β-strand, α-helix, and 3_{10} helix (η), of *Aa*FaeD are indicated on top. The identical residues are highlighted in red with white characters and similar residues in red characters. The residues forming the S-H-D catalytic triad are indicated by asterisks. The three pairs of cysteine residues that form disulfide bonds are indicated by green numbers 1, 2, and 3. The residues suggested to form a substrate-binding site in *Aa*FaeD are indicated by triangles. Green and yellow triangles specifically indicate amino acids constituting the right and left walls, respectively, corresponding to **Fig. 3-9**.

3.2.7 Structure-based functional mutagenesis

To evaluate the roles of the hydrophobic residues contained in the hydrophobic cleft, I selected Leu50, Phe120, Ile147, Val169, and Leu170 **(Fig. 3-9)**, and performed structure-based functional mutagenesis. Leu50, Ile147, Val169, and Leu170 were replaced with alanine, giving L50A, I147A, V169A, L170A single mutants, and a V169A/L170A double mutant. On the other hand, Phe120 was replaced with tyrosine, giving a F120Y single mutant. Then I determined the specific activities of these mutants against model substrates for FAE (M*p*CA, MCA, MFA, and MSA) and that for AXE (*p*NPAc) **(Table 3-2)**.

The specific activities of all alanine mutants, L50A, I147A, V169A, L170A, and V169A/L170A, decreased against all substrates, M*p*CA, MCA, MFA, MSA, and *p*NPAc, as compared with that of the wild type *Aa*FaeD-CD; the specific activities of the alanine mutants were ca. 0.06- to 0.81-fold of that of the wild type **(Table 3-2)**.

Interestingly, the Phe120 to tyrosine mutant exhibited higher specific activities against MpCA, MCA, MFA, and MSA than the wild type **(Table 3-2)**; the activities of F120Y were ca. 1.3- to 1.5-fold of that of the wild type **(Table 3-3)**. The specific activity of F120Y against pNPAc, on the other hand, was just slightly higher than that of the wild type (ca. 1.1 fold) **(Table 3-2)**.

Kinetic parameters of the activity against MFA were determined for the wild type and all mutants **(Table 3-3)**. For the alanine mutants, the k_{cat}/K_M values were less than ca. 0.5-fold of that of the wild type, except for V169A, whose k_{cat}/K_M value was the same as that of the wild type. In the case of V169A, although the k_{cat}/K_M value was the same as that of the wild type, both k_{cat} and K_M were lower than those of the wild type. Interestingly, in the case of L50A and L170A, k_{cat} was higher than that of the wild type, but K_M was also higher than that of the wild type. These results indicate that the mutations of Leu50 and Leu170 to alanine have a negative effect on affinity of the substrate-binding but a positive effect on turnover number. Substitution of these leucine residues to alanine is supposed to widen the hydrophobic cleavage, and thereby the fitting of MFA to the hydrophobic cleavage may have changed (loosened).

In the case of F120Y, the k_{cat}/K_M value was higher than that of the wild type (ca. 1.2-fold), which turned out to be due to a lower K_M, namely higher affinity of the MFA towards the hydrophobic cleft. This increased affinity of F120Y may result from the formation of water-mediated hydrogen bonds between the hydroxyl groups of tyrosine and MFA, a phenomenon observed in various structures [34]. Similarly, Tyr80 and Tyr119 within the hydrophobic cleft of *A. niger* FaeA and *Anaeromyces mucronatus* Fae1A, respectively, also interact with FA through hydrogen bonding interaction [17,35,36].

Table 3-2. Specific activities of the wild-type and mutant *Aa*FaeD-CDs.

	Specific activity (mU/mg)				
Enzyme	**M*p*CA**	**MCA**	**MFA**	**MSA**	***p*NPAc**
*Aa*FaeD-CD (WT)	104 ± 5	90 ± 5	175 ± 5	126 ± 6	1,060 ± 62
-L50A	37 ± 18	41 ± 6	62 ± 18	50 ± 9	638 ± 31
-F120Y	146 ± 3	116 ± 1	262 ± 4	164 ± 2	1,108 ± 30
-I147A	39 ± 6	31 ± 2	53 ± 8	55 ± 10	80 ± 30
-V169A	72 ± 5	74 ± 14	92 ± 6	83 ± 2	783 ± 145
-L170A	40 ± 2	29 ± 3	39 ± 2	42 ± 13	261 ± 28
-V169A/L170A	43 ± 5	34 ± 5	43 ± 12	56 ± 1	136 ± 35

Values are the means ± standard deviation of three independent replicates.

Table 3-3. Kinetic parameters of esterase activity against MFA of the wild-type and mutant *Aa*FaeD-CDs.

Enzyme	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
WT	1.2 ± 0.1	18.7 ± 1.1	16.0 ± 1.9
L50A	2.8 ± 1.1	20.4 ± 6.2	7.2 ± 3.6
F120Y	0.8 ± 0.1	18.1 ± 1.1	21.9 ± 2.9
I147A	1.3 ± 0.4	9.7 ± 2.0	7.5 ± 3.1
V169A	0.8 ± 0.1	13.0 ± 1.2	16.3 ± 3.3
L170A	4.9 ± 2.8	30.3 ± 14.5	6.2 ± 4.6
V169A/L170A	2.0 ± 0.6	13.2 ± 3.0	6.7 ± 2.7

Values are the means ± standard deviation of three independent replicates.

3.2.8 Structural comparison with other esterases in the CE1 family

To investigate the structural and functional features of *Aa*FaeD-CD, the crystal structures and enzymatic activities of members belonging to different subfamilies of CE1 were compared. Firstly, I compared the crystal structures of *Aspergillus luchuensis* AXE (*Al*AXE; belonging to SF1) [PDB ID: 5X6S], *Aspergillus sydowii* FaeE (*As*FaeE; SF2) [8IYB], *A. mucronatus* Fae1A (*Am*Fae1A; SF3) [5CXX], and *Aa*FaeD-CD (SF5, [8JH9], this study) **(Fig. 3-12)**. These proteins shared the α/β-hydrolase fold; the central regions were superimposed very well. The substrate-binding sites are positioned on top in **Fig. 3-12**. It can be seen clearly that *Al*AXE, *As*FaeE, and *Am*Fae1A have extra motifs containing secondary structures and loops, which *Aa*FaeD does not have. These extra motifs comprise a region connecting β6 and β7. The corresponding region in *Aa*FaeD is much shorter than those in the others. These extra motifs have different configurations and are reportedly involved in substrate-binding [17,37]. For further detailed comparison of the structures around the substrate-binding sites including the extra motifs, I chose FAEs, namely *As*FaeE, *Am*Fae1A, and *Aa*FaeD, for which substrate-bound structures are available.

Close-up views of the substrate-binding sites for *As*FaeE, *Am*Fae1A, and *Aa*FaeD with bound FA are presented in **Fig. 3-13 A, B, and C**, respectively (same orientation as in **Fig. 3-9** and **3-11**. It is noted that the substrate binding sites of *As*FaeE and *Am*Fae1A are small pockets **(Fig. 3-13 B and C)**, but that of *Aa*FaeD has more space **(Fig. 3-13 A)**. In the case of *Aa*FaeD, the bound FA is wedged between two hydrophobic walls (hydrophobic cleft) and the upper region to the bound FA is open **(Fig. 3-13 A)**. In the case of *As*FaeE and *Am*Fae1A, the bound FA is surrounded by three walls **(Fig. 3-13 B and C)**; two walls on the sides of the bound FA are hydrophobic but the upper regions, which correspond to the extra motifs discussed above, contain both hydrophobic and hydrophilic residues. The extra motifs of *As*FaeE and *Am*Fae1A, which were previously named the cysteine loop [16] and β-clamp motif [17] (Figures S7B and S7C), respectively, are reportedly involved in substrate-binding. The hydroxyl group of FA forms hydrogen bonds with a sidechain amide group of Gln163 in *As*FaeE, and with a sidechain imidazole group of His196 in *Am*Fae1A. Additionally, the extra motifs cover the aromatic ring of FA like a lid. Thus, *As*FaeE and *Am*Fae1A each has a small substrate-binding pocket, while the hydrophobic cleft of *Aa*FaeD has more space for substrate binding. There are more interactions with FA for *As*FaeE and *Am*Fae1A than for *Aa*FaeD to accommodate FA. This explains why the affinity against FA is much higher for *As*FaeE and *Am*Fae1A than *Aa*FaeD **(Table 3-4)**. It should be added that the differences in k_{cat} turned out to be relatively moderate among them **(Table 3-4)**.

It was shown that *Aa*FaeD, a member of SF5, has more space for substrate binding than *As*FaeE and *Am*Fae1A, members of SF2 and SF3, respectively. The latter two, having small substrate-binding pockets, exhibit higher activity due to the higher affinity towards small substrates such as MFA. This, on the other hand, suggests that the members of SF5 might be able to exert higher activity than the members of SF2 and SF3 toward larger substrates, due to higher affinity to them. In fact, the specific activity of SF5 had been shown to be higher towards larger or bulkier ferulic acid derivative compounds than those of SF2 and SF3 [13].

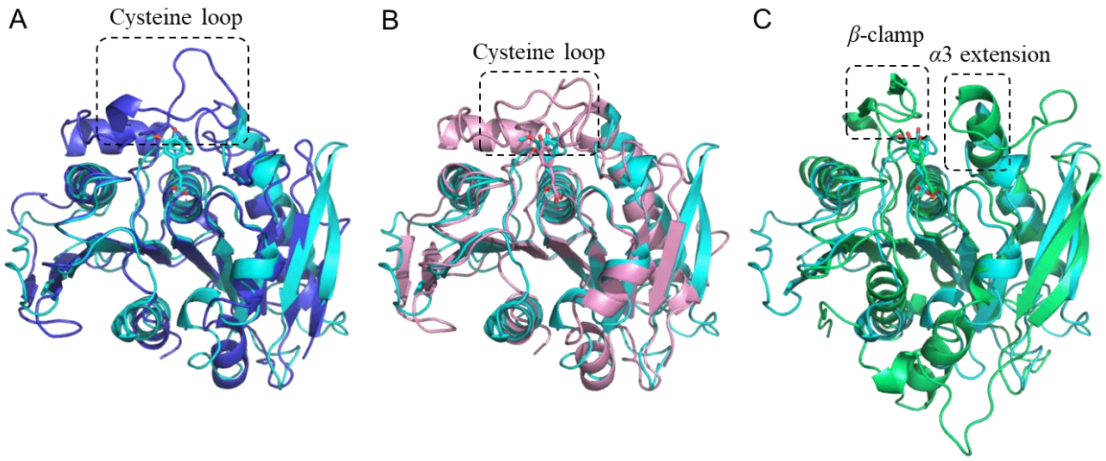

Fig. 3-12. Structural comparison of the crystal structure of *Aa*FaeD-CD, ribbon representation, with *Aspergillus luchuensis* AXE (*Al*AXE), *Aspergillus sydowii* FaeE (*As*FaeE), and *Anaeromyces mucronatus* Fae1A (*Am*Fae1A). Superimposition of *Aa*FaeD-CD/FA crystal structures (cyan), ribbon representation, with *Al*AXE (blue; PDB ID: 5X6S) belonging to SF1 of CE1 (A), *As*FaeE/FA (pink; PDB ID: 8IYB) belonging to SF2 of CE1 (B), and *Am*Fae1A/FA (green; PDB ID: 5CXX) belonging to SF3 of CE1 (C), respectively. Ferulic acids are indicated as stick models. The cysteine loops, β-clamp, and α3 extension are indicated by dotted boxes.

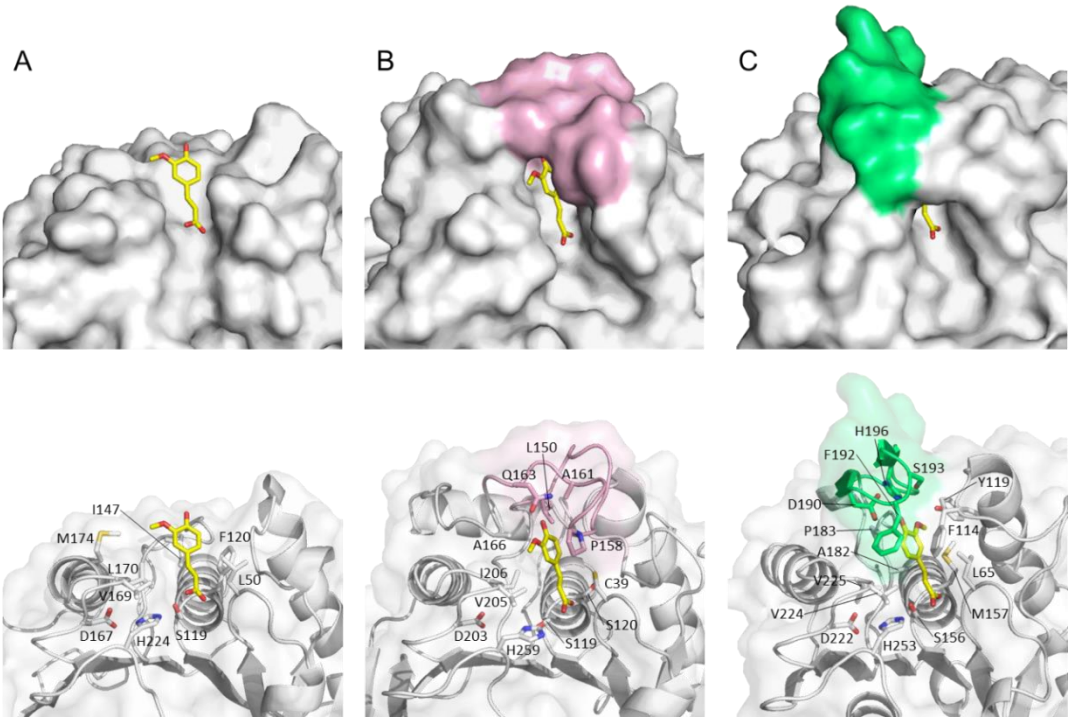

Fig. 3-13. Structural comparison of the substrate-binding sites of *Aa*FaeD, *As*FaeE, and *Am*Fae1A. (A-C) Surface representations, in either non-transparent (top) or transparent (bottom), of the substrate-binding sites of the crystal structures of *Aa*FaeD (PDB ID: 8JH9) (A), *As*FaeE (PDB ID: 8IYB) belonging to SF2 of CE1 (B), and *Am*Fae1A (PDB ID: 5CXX) belonging to SF3 of CE1 (C), all in a complex with ferulic acid (FA; yellow sticks). FA, the S-H-D catalytic triad, and the amino acid residues that are within 5 Å from FA are indicated by stick models. The cysteine loop of *As*FaeE [16] and *β*-clamp of *Am*Fae1A [17] are colored pink and green, respectively. The nitrogen, oxygen, and sulfur atoms are colored blue, red, and yellow, respectively. The backbone is shown as a ribbon for at bottom.

Table 3-4. Kinetic parameters of esterase activity against MFA of *Aa*FaeD-CD, *As*FaeE, and *Am*Fae1A belonging to different SFs of CE1.

Sample	CE1-SF	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	Reference
*Aa*FaeD-CD	5	1.2 ± 0.1	18.7 ± 1.1	16.0 ± 1.9	This study
*As*FaeE	2	0.22 ± 0.01	20.2 ± 0.3	91.5 ± 5.1	[16]
*Am*Fae1A	3	0.0067 ± 0.001	6.05	903	[38]

Values are the means ± standard deviation of three independent replicates.

3.2.9 Enzyme activity of AaFaeD-CD towards the natural substrate for releasing bulkier 5,5'-dehydrodiferulate

To examine whether the open substrate-binding site of *Aa*FaeD-CD could react with the bulkier 5,5'-diFA, FAE activity on natural substrate, DSWB, was evaluated. When DSWB was incubated with *Aa*FaeD-CD alone, FA was detected as a major product with an elution time of approximately 18.8 min **(Fig. 3-14)**. In the co-presence of Celluclast and *Aa*FaeD-CD, 5,5'-diFA was clearly detected with an elution time of ca. 21.0 min, indicating the release of 5,5'-diFA from DSWB **(Fig. 3-14)**. It turned out that a modest amount of 5,5'-diFA was also released when *Aa*FaeD-CD alone was present **(Fig. 3-14)**. This observation suggests that Celluclast contributes to the relaxation of the intricate structure of DSWB, enabling FAE to effectively target bulkier phenolic substrates, including 5,5'-diFA. Additionally, the identification of 5,5'-diFA as a target implies that *Aa*FaeD is a D-type FAE, aligning with the classification proposed by Crepin *et al.* (2004) [39].

Parallel experiments were performed with *As*FaeE, characterized by a "closed" substrate binding site [16], and it was observed that *As*FaeE did not release 5,5'-diFA from DSWB **(Fig. 3-14)**, consistent with the previous finding of Underlin *et al.* (2020) [13]. This outcome could be attributed to steric hindrance induced by the cysteine loop in *As*FaeE, potentially impeding the binding of *As*FaeE to a bulkier substrate such as 5,5'-diFA ester [16].

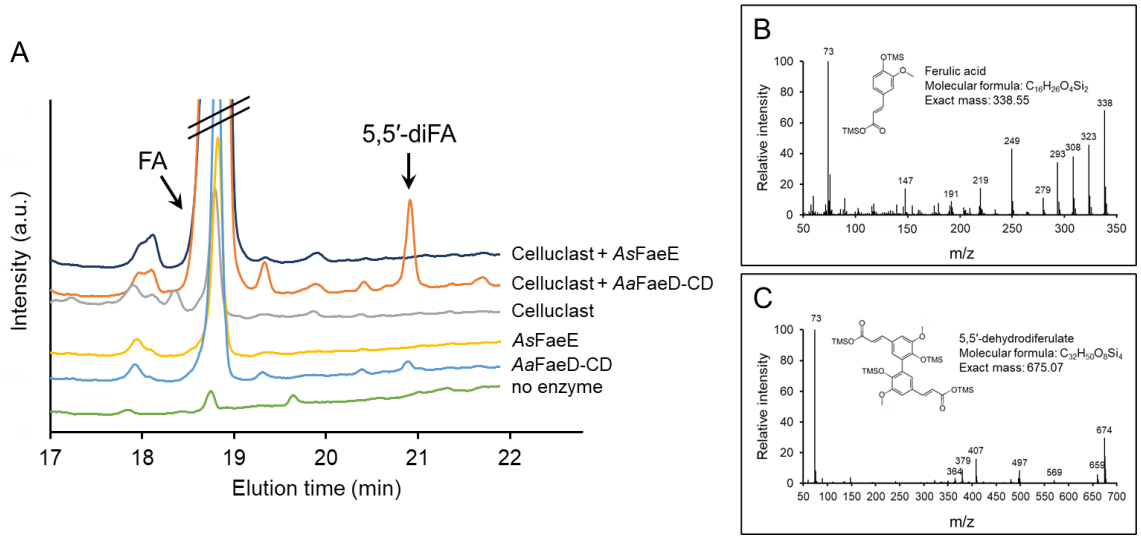

Fig. 3-14. (A) Reverse-phase HPLC elution profiles for destarched wheat bran (DSWB) incubated with the following various kinds of combination of enzymes; no enzyme, *Aa*FaeD-CD, *As*FaeE, Celluclast, Celluclast+*Aa*FaeD-CD, or Celluclast+*As*FaeE. The phenolic acids released from DSWB were monitored by a UV detector at 310 nm. EI mass spectra of ferulic acid (FA) (B) and 5,5′-dehydrodiferulate (5,5′-diFA) (C) obtained from RP-HPLC of 'Celluclast+*Aa*FaeD-CD'. Corresponding compounds, trimethylsilylated FA (B) and trimethylsilylated 5,5′-diFA (C), were identified based on the NIST mass spectral library [40] and mass spectra reported by Bunzel, *et al*. (2001) [41], respectively.

3.3 Conclusions

This study provides valuable insights into the structure and function of SF5 FAEs that can be utilized in biotechnology for sustainable biomass degradation. By determining the crystal structure of the catalytic domain (CD) of a fungal feruloyl esterase (FAE) from *Acremonium alcalophilum*, *Aa*FaeD, the first SF5 FAE to be structurally characterized, I have identified notable features of this enzyme family, including wide substrate specificity and an open substrate-binding site. Moreover, through structure-based functional mutagenesis, I uncovered the key roles of hydrophobic residues in the hydrophobic cleft, influencing binding affinity and catalytic rate. It was remarkable that the F120Y mutant exhibited higher enzymatic activity towards various model substrates than the wild type. The structural comparison of *Aa*FaeD-CD, belonging to SF5, with other FAEs in SF2 and SF3, indicates that SF5 FAEs may have a wide spectrum of

substrate specificity due to their more open substrate-binding site. This idea is supported by the reported catalysis of SF5 FAEs on not only small compounds like ferulate derivatives but also larger compounds, including 5,5'-diFA derivatives. These findings pave the way for the development of innovative biotechnology solutions for biomass degradation and utilization, with implications for biofuels, biorefining, and bioproducts. Overall, my study represents a major advance in understanding SF5 FAEs and opens exciting new avenues for future research and applications in biotechnology.

3.4 Materials and methods

3.4.1 Substrates

Methyl *p*-coumarate (M*p*CA), methyl caffeate (MCA), *p*-nitrophenyl acetate (*p*NPAc), caffeic acid, and *p*-nitrophenol were obtained from TCI Chemicals (Tokyo, Japan). Methyl ferulate (MFA), methyl sinapate (MSA), ferulic acid (FA), sinapic acid (SA), and *p*-coumaric acid (*p*CA) were purchased from LKT Laboratories (St. Paul, Minnesota, USA), Biosynth (Staad, Switzerland), MP Biomedicals (Santa Ana, California, USA), Cayman Chemical Company (Ann Arbor, Michigan, USA), and Nacalai Tesque (Kyoto, Japan), respectively.

Wheat bran was obtained from Alishan (Saitama, Japan). The bran was destarched using the modified protocol of Aguedo, *et al.* (2013) [42] was initially soaked in 50 mM potassium phosphate pH 7.0 in a ratio of 1:10 (w/v) and 25 μL/g bran of thermostable α-amylase (Sigma-Aldrich, St. Louis, Missouri, USA) was added to remove starch from the bran. The suspension was then incubated at 50 °C for 24 h. The enzyme was denatured at 100 °C for 10 min. The medium was cooled down, washed several times with distilled water, and freeze-dried. The destarched wheat bran (DSWB) was cut by cutting mill, sieved to obtain particles in the size range of 0.5–1 mm, and stored at room temperature (RT) until further use.

3.4.2 Bioinformatics

An amino acid sequence of an SF5 FAE from *A. alcalophilum* (*Aa*FaeD) was retrieved from the Joint Genome Institute database (JGI) (protein ID: Acral2|1082309). The signal sequence was predicted using SignalP 5.0 and was omitted for further analyses [43]. *Aa*FaeD was classified by comparing its sequence with those of the functionally validated FAEs and representative FAEs belonging to the five subfamilies of fungal CE1 that were proposed in the studies of Komiya, *et al.* (2017) [8] and Dilokpimol, *et al.* (2022) [12]. The sequences of AXE and FAE used in the alignment and phylogenetic analysis are shown in **Appendix C**. Multiple sequence alignment was carried out using the MUSCLE algorithm and construction of the phylogenetic tree was carried out using Molecular Evolutionary Genetics Analysis (MEGA11) [44]. The tree was constructed using the maximum-likelihood method with the best fit WAG+G+I substitution model and 95% partial deletion. Statistical support for the phylogenetic grouping was obtained by 100 bootstrap re-samplings.

The glycosylation sites of *Aa*FaeD were predicted with NetNGlyc 1.0 [45] and NetOGlyc 4.0 [46]. The theoretical molecular mass (*m*) and isoelectric point (pI) were deduced using the Expasy-ProtParam tool [47].

3.4.3 Structural prediction of the full-length AaFaeD

Structural prediction of the full-length *Aa*FaeD was performed by AlphaFold2 [24] **(Fig. 3-3)**.

3.4.4 Cloning

The DNA sequence encoding the N-terminal 6x histidine-tag and the full-length *Aa*FaeD, which does not contain a signal sequence, was codon-optimized and synthesized by Thermo Fisher Scientific (Japan) **(Appendix D)**. The DNA sequences for the recombinant proteins encoding the N-terminal 6x histidine-tag with the full-length *Aa*FaeD (*Aa*FaeD) and that with catalytic domain of *Aa*FaeD (*Aa*FaeD-CD) **(Fig. 3-2 A)** were amplified by polymerase chain reaction (PCR) using a forward primer (5'-

AGCTGAATTCCACCATCACCATCACCAT-3') and reverse primers (5'-TGAGCGG-CCGCTCACAGGCA-3' and 5'-GTGAGCGGCCGCTCAGAACTGACTAAAAAA-GTCCCAGATC-3', for *Aa*FaeD and *Aa*FaeD-CD, respectively), EcoRI and NotI restriction sites being underlined. The DNA amplicons were ligated in frame with the *α*-factor signal peptide of the pPICZ*α*A expression vector at the EcoRI and NotI restriction sites. The obtained plasmids were linearized with SacI and introduced into *Pichia pastoris* X-33 (Invitrogen) by electroporation.

3.4.5 Site-directed mutagenesis

Six mutants of *Aa*FaeD-CD were constructed by inverse PCR using the following primers: L50A (5′-CTACACTGGGCATCCGGCAGTATGCACGACG-3′ and 5′-ACTGCCGGATGCCCAGTGTAGGCCGAAGATT-3′), F120Y (5'-GGATTC-TCTTACGGAGGTGGAATGTCCTATG-3' and 5'-TCCACCTCCGTAAGAGAAT-CCGGTAGCAAAA-3'), I147A (5′-GGTGCACAAGCAAGTGGATGCAGTGGCG-GC-3′ and 5′-GCATCCACTTGCTTGTGCACCGGCGATTGG-3′), V169A (5'-AAC-GACGACGCTCTGCCAATAGCAATGGGA-3' and 5'-TATTGGCAGAGCGTCGT-CGTTAGTTCCGTG-3'), and L170A or V169A/L170A (5′-AACGACGACGYTGCT-CCAATAGCAATGGGAAGA-3′ and 5′-TGCTATTRGAGCAACGTCGTCGTTAG-TTCCGTG-3′) (mutated sites are underlined). R is an A or G nucleotide, and Y is a C or T nucleotide.

3.4.6 Protein preparation

Protein expression was induced in the *P. pastoris* transformants for following the procedure in section **2.4.4**. The recombinant *Aa*FaeD or *Aa*FaeD-CD expressed and secreted in the medium was purified at first with a Ni^{2+}-NTA sepharose column (Cytiva), as described in section **2.4.4**. The purified proteins were used for biochemical analyses. For protein crystallization, one more purification step was added. I used a HiLoad 16/600 Superdex 75 column (GE Healthcare) equilibrated with 25 mM potassium phosphate, pH 6.5. The purity of the protein fractions obtained in each purification step was analyzed by 12% SDS-PAGE. The purified proteins were stored at 4 °C for further use.

The mutant proteins were produced in the same way as the wild type. The mutant proteins were purified using a Ni^{2+}-NTA Sepharose column (Cytiva). The purified protein solution was concentrated, and the buffer changed with an Amicon Ultra-0.5 (3 kDa, Millipore) to 25 mM potassium phosphate, pH 6.5.

I determined the protein concentration by measuring UV absorbance of the samples at 280 nm. The extinction coefficient was theoretically determined based on the amino acid composition of *Aa*FaeD-CD and its mutants [48].

3.4.7 Investigation of enzyme activities and biochemical properties

Activity assays were carried out using model substrates, methyl hydroxycinnamates (M*p*CA, MCA, MFA, and MSA) and *p*NPAc **(Fig. 3-2 C)**, following modified versions of the protocols reported by Dilokpimol, *et al.* (2017) [21] and Xu, *et al.* (2021) [49]. Briefly, 100 ng of either *Aa*FaeD or *Aa*FaeD-CD was added to a solution comprising 80 mM potassium phosphate, pH 7.0, and 0.12 mM methyl hydroxycinnamates (or 0.6 mM *p*NPAc), followed by incubation at 37 °C for 30 min. The reaction was monitored with a microplate reader (Tecan Infinite 200 Pro, Switzerland) at 340 nm for methyl hydroxycinnamates and 410 nm for *p*NPAc with 2 min intervals, and the initial reaction rates of hydrolysis were obtained. The extinction coefficients of both model substrates and product phenolic compounds (FA, *p*CA, CA, SA, and *p*-nitrophenol) were determined experimentally. One unit (U) is defined as the activity that produces 1 μmol of hydroxycinnamic acid from methyl hydroxycinnamates or 1 μmol of *p*-nitrophenol from *p*NPAc per min. All assays were performed in triplicate.

The optimum temperature and thermostability of *Aa*FaeD-CD were inspected by measuring the activity towards MFA. The optimum temperature for enzyme activity was investigated between 15 and 65 °C in 80 mM potassium phosphate, pH 7.5. For thermostability, *Aa*FaeD-CD dissolved in 80 mM potassium phosphate, pH 7.5, was incubated for 120 min at temperatures ranging from 15 to 65 °C, after which the activity was measured at 40 °C as described above. For determination of the optimal pH for enzyme activity, *Aa*FaeD-CD was dissolved in various solutions with different pHs, 80 mM sodium acetate (pH 4.0–5.5), 80 mM potassium phosphate (pH 6.0–8.0), or 80 mM

Tris-HCl (pH 8.0–9.0), and then the activity towards MFA was measured at 40 °C. To determine the pH stability of *Aa*FaeD-CD, the enzyme was dissolved in the same solutions as those used to determine the optimal pH. The solutions were then incubated for 24 h, after which the activity of the enzyme was measured at 40 °C in 80 mM potassium phosphate buffer at pH 7.5.

The initial reaction rates of the hydrolysis of MFA at different concentrations (0.02–1.2 mM) were determined for *Aa*FaeD-CD and its mutants in a solution comprising 100 mM potassium phosphate, pH 7.0. Curve fitting was conducted for a plot of each saturation curve using a Kaleidagraph (Synergy, Reading, PA, USA) [50]. The data were fitted to the Michaelis-Menten equation, by which kinetic parameters (K_M and k_{cat}) were obtained. The fitting errors were obtained by nonlinear regression analyses.

3.4.8 Enzyme activity towards destarched wheat bran

The activity of *Aa*FaeD-CD on bulkier substrates, releasing phenolic acids, was assessed using DSWB as the substrate. In comparison, *As*FaeE, known for its specificity to monomeric ferulate, was obtained from my previous study [16]. The reaction solutions comprised 10 mg/mL destarched bran, with the following combinations: 10 μg of purified FAE alone, 10 U/mg substrate of Celluclast 1.5L (Sigma-Aldrich, St. Louis, Missouri, USA) alone, or a mixture of 10 μg FAE and 10 U/mg substrate of Celluclast 1.5L. These solutions were prepared in 100 mM potassium phosphate buffer at pH 7.0 and were mixed at 40 °C for 24 h. The sample was then analyzed by HPLC analysis.

3.4.9 X-ray crystallography

The crystallization of *Aa*FaeD-CD was conducted by the sitting-drop vapor-diffusion method at 293 K. A crystal was obtained by mixing 1.2 μL of 10 mg/mL *Aa*FaeD-CD dissolved in 25 mM potassium phosphate (pH 6.5) with 0.8 μL of a reservoir solution consisting of 0.2 M $MgCl_2.6H_2O$, 0.1 M Tris-HCl (pH 8.5), and 30% (w/v) PEG4000. The droplet was equilibrated against 100 μL reservoir solution. Diffraction data for *Aa*FaeD-CD free were collected at 100 K by flash-cooling in a cold nitrogen gas

stream after a brief soaking in a bottom solution containing 25% diethylene glycol as a cryo-protectant. The complex with FA was obtained by soaking the crystal with 10 mM FA in the reservoir solution for 20 min at RT. Since I were unable to obtain the FA-complex data using the cryo-method due to incorrect binding of the cryo-protectant, I employed the capillary method for data collection at 298 K. The crystal soaked with 10 mM FA was placed into a glass capillary (Mark-tube, f=0.8 mm, Hilgenberg), and both ends were sealed with dental wax after adding the soaked solution in one end. Both the cryo and capillary crystal data were collected at a wavelength of 1.0 Å using an MX225HE detector (Rayonix) on beamline BL26B1 at SPring-8 (Hyogo, Japan). The collected data were processed using the XDS program package [51]. Then, the *Aa*FaeD-CD crystal structure was solved by the molecular replacement method with the MOLREP program in the CCP4 package [52] using the *Aa*FaeD-CD structure predicted by Alphafold2 [24]. Model rebuilding and refinement were achieved using COOT [53] and PHENIX [54]. The secondary structure of *Aa*FaeD-CD was assigned by DSSP & Stride plugin [55] for PyMOL. Figures of crystal structures were prepared with PyMOL software [56]. The secondary structure alignments, examined in the previous step, were visualized by using an ESPript online tool [57].

3.4.10 HPLC analysis

To analyze the phenolic compounds in the hydrolyzed reaction, 200 µL of the sample was mixed with 600 µL of 100% acetonitrile. Phenolic acids were analyzed using RP-HPLC (Shimadzu, Japan) equipped with a reversed-phase column (4.6 mm × 15 cm, TSKgel ODS-80 TM; Tosoh, Tokyo, Japan) and a UV detector (SPD-20A UV/VIS detector; Shimadzu, Japan). The binary mobile phase consisted of (A) Milli-Q water + 0.1% trifluoroacetic acid (TFA) and (B) acetonitrile + 0.1% TFA. The elution profile was as follows: isocratic at 5% B for 5 min, B linearly increased from 5–70% for 20 min, isocratic at 70% B for 10 min, and isocratic at 5% B for 10 min. The flow rate was 1.0 mL/min. UV spectra of the released FA and phenolic compounds were monitored at 310 nm. Commercial FA was utilized as a standard for estimating the retention time during analysis. Data were processed using LabSolution software (Shimadzu, Japan). The

presence of FA and 5,5'-dehydrodiferulate in the enzymatic reaction was validated by gas chromatography-mass spectrometry (GC-MS).

3.4.11 GC-MS analysis

To obtain 5,5'-dehydrodiferulate, the fractions at the retention time of 21 min **(Fig. 3-14)** were pooled and extracted into ethyl acetate (three times). The extracts were combined and dried using a centrifugal evaporator (EYSLA, China). For GC-MS analysis, the dried extract was silylated by adding 10 µL of pyridine and 40 µL of N,O-Bis(trimethylsilyl)trifluoroacetamide with 1 vol% chlorotrimethylsilane. The trimethylsilylated sample was dried under nitrogen gas and dissolved in ethyl acetate before GC-MS analysis.

For GC-MS analysis, the trimethylsilylated samples were separated by GCMS-QP2010SE system (Shimadzu, Japan) equipped with a DB-5MS column (30 m × 0.25 mm id, 0.25 µm film thickness; Agilent Technologies, USA). Helium was used as a carrier gas with a flow rate of 1.0 mL/min. The initial column temperature was held at 80 °C for 3 min, then ramped at a rate of 5 °C/min to 300 °C and held for 20 min. The mass spectrometer was operated in electron impact ionization mode with an ionizing energy of 70 eV, an ion source temperature of 250 °C, and an interface temperature of 300 °C. The mass scanning range was set to m/z 40–800.

Compounds corresponding to mass spectra were identified based on the NIST mass spectral library [40] and mass spectra reported by Bunzel, *et al*. (2001) [41].

3.5 References

[1] Kurowska, K., Marks-Bielska, R., Bielski, S., Kryszk, H., Jasinskas, A., Food security in the context of liquid biofuels production, *Energies*. (2020) 13 6247.

[2] Zoghlami, A., Paës, G., Lignocellulosic biomass: Understanding recalcitrance and predicting hydrolysis, *Frontiers in Chemistry*. (2019) 7 874.

[3] Bunzel, M., Ralph, J., Funk, C., Steinhart, H., Structural elucidation of new ferulic acid-containing phenolic dimers and trimers isolated from maize bran, *Tetrahedron Letters*. (2005) 46 5845–5850.

[4] Manga-Robles, A., Santiago, R., Malvar, R.A., Moreno-González, V., Fornalé, S., López, I., Centeno, M.L., Acebes, J.L., Álvarez, J.M., Caparros-Ruiz, D., Encina, A., García-Angulo, P., Elucidating compositional factors of maize cell walls contributing to stalk strength and lodging resistance, *Plant Science*. (2021) 307.

[5] Mkabayi, L., Malgas, S., Wilhelmi, B.S., Pletschke, B.I., Evaluating feruloyl esterase-xylanase synergism for hydroxycinnamic acid and xylo-oligosaccharide production from untreated, hydrothermally pre-treated and dilute-acid pre-treated corn cobs, *Agronomy*. (2020) 10 688.

[6] Mafa, M.S., Malgas, S., Pletschke, B.I., Feruloyl esterase (FAE-1) sourced from a termite hindgut and GH10 xylanases synergy improves degradation of arabinoxylan, *AMB Express*. (2021) 11 21.

[7] Zhang, Y., Yang, H., Yu, X., Kong, H., Chen, J., Luo, H., Bai, Y., Yao, B., Synergistic effect of acetyl xylan esterase from Talaromyces leycettanus JCM12802 and xylanase from Neocallimastix patriciarum achieved by introducing carbohydrate-binding module-1, *AMB Express*. (2019) 9.

[8] 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*, *Applied and Environmental Microbiology*. (2017) 83 e01251-17.

[9] Mäkelä, M.R., Dilokpimol, A., Koskela, S.M., Kuuskeri, J., de Vries, R.P., 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, *Microbial Biotechnology*. (2018) 11 869–880.

[10] 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 Research*. (2013) 41 D423–D429.

[11] Li, X., Griffin, K., Langeveld, S., Frommhagen, M., Underlin, E.N., Kabel, M.A., de Vries, R.P., Dilokpimol, A., Functional validation of two fungal subfamilies in carbohydrate esterase family 1 by biochemical characterization of esterases from uncharacterized branches, *Frontiers in Bioengineering and Biotechnology*. (2020) 8.

[12] Dilokpimol, A., Verkerk, B., Li, X., Bellemare, A., Lavallee, M., Frommhagen, M., Underlin, E.N., Kabel, M.A., Powlowski, J., Tsang, A., de Vries, R.P., Screening of novel fungal carbohydrate esterase family 1 enzymes identifies three novel dual feruloyl/acetyl xylan esterases, *FEBS Letters*. (2022) 596 1932–1943.

[13] Underlin, E.N., Frommhagen, M., Dilokpimol, A., van Erven, G., de Vries, R.P., Kabel, M.A., Feruloyl esterases for biorefineries: Subfamily classified specificity for natural substrates, *Frontiers in Bioengineering and Biotechnology*. (2020) 8 332.

[14] Antonopoulou, I., Leonov, L., Jütten, P., Cerullo, G., Faraco, V., Papadopoulou, A., Kletsas, D., Ralli, M., Rova, U., Christakopoulos, P., Optimized synthesis of novel prenyl ferulate performed by feruloyl esterases from *Myceliophthora thermophila* in microemulsions, *Applied Microbiology and Biotechnology*. (2017) 101 3213–3226.

[15] Antonopoulou, I., Dilokpimol, A., Iancu, L., Mäkelä, M.R., Varriale, S., Cerullo, G., Hüttner, S., Uthoff, S., Jütten, P., Piechot, A., Steinbüchel, A., Olsson, L., Faraco, V., Hildén, K.S., de Vries, R.P., Rova, U., Christakopoulos, P., The synthetic potential of fungal feruloyl esterases: A correlation with current classification systems and predicted structural properties, *Catalysts*. (2018) 8 242.

[16] Phienluphon, A., Kondo, K., Mikami, B., Nagata, T., Katahira, M., Structural insights into the molecular mechanisms of substrate recognition and hydrolysis by feruloyl esterase from *Aspergillus sydowii*, *International Journal of Biological Macromolecules*. (2023) 253 127188.

[17] Gruninger, R.J., Cote, C., McAllister, T.A., Abbott, D.W., Contributions of a unique *β*-clamp to substrate recognition illuminates the molecular basis of exolysis in ferulic acid esterases, *Biochemical Journal*. (2016) 473 839–849.

[18] Dilokpimol, A., Mäkelä, M.R., Aguilar-Pontes, M.V., Benoit-Gelber, I., Hildén, K.S., De Vries, R.P., Diversity of fungal feruloyl esterases: Updated phylogenetic classification, properties, and industrial applications, *Biotechnology for Biofuels*. (2016) 9 231.

[19] Tulsani, N.J., Jakhesara, S.J., Hinsu, A.T., Jyotsana, B., Dafale, N.A., Patil, N. v., Purohit, H.J., Joshi, C.G., Genome analysis and CAZy repertoire of a novel fungus *Aspergillus sydowii* C6d with lignocellulolytic ability isolated from camel rumen, *Electronic Journal of Biotechnology*. (2022) 59 36–45.

[20] Crepin, V.F., Connerton, I.F., Faulds, C.B., Identification of a type-D feruloyl esterase from *Neurospora crassa*, *Applied Microbiology and Biotechnology*. (2004) 63 567–570.

[21] Dilokpimol, A., Mäkelä, M.R., Mansouri, S., Belova, O., Waterstraat, M., Bunzel, M., de Vries, R.P., Hildén, K.S., Expanding the feruloyl esterase gene family of *Aspergillus niger* by characterization of a feruloyl esterase, FaeC, *New Biotechnology*. (2017) 37 200–209.

[22] Dilokpimol, A., Mäkelä, M.R., Varriale, S., Zhou, M., Cerullo, G., Gidijala, L., Hinkka, H., Brás, J.L.A., Jütten, P., Piechot, A., Verhaert, R., Hildén, K.S., Faraco, V., de Vries, R.P., Fungal feruloyl esterases: Functional validation of genome

mining based enzyme discovery including uncharacterized subfamilies, *New Biotechnology*. (2018) 41 9–14.

[23] Kühnel, S., Pouvreau, L., Appeldoorn, M.M., Hinz, S.W.A., Schols, H.A., Gruppen, H., The ferulic acid esterases of Chrysosporium lucknowense C1: Purification, characterization and their potential application in biorefinery, *Enzyme and Microbial Technology*. (2012) 50 77–85.

[24] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., Highly accurate protein structure prediction with AlphaFold, *Nature*. (2021) 596 583–589.

[25] Hayashi K, Nimura Y, Ohara N, Uchimura T, Suzuki H, Komacata K, Kozaki M., Low-temperature-active cellulase produced by *Acremonium alcalophilum* JCM 7366, *Seibutsu-Kogaku Kaishi*. (1996) 74 7–10.

[26] Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G., Goldman, A., The *α/β* hydrolase fold, *Protein Engineering, Design and Selection*. (1992) 5 197–211.

[27] Denesyuk, A., Dimitriou, P.S., Johnson, M.S., Nakayama, T., Denessiouk, K., The acid-base-nucleophile catalytic triad in ABH-fold enzymes is coordinated by a set of structural elements, *PLoS ONE*. (2020) 15 e0229376.

[28] Brenner, S., The molecular evolution of genes and proteins: a tale of two serines, *Nature*. (1988) 334 528–530.

[29] Akoh, C.C., Lee, G.C., Liaw, Y.C., Huang, T.H., Shaw, J.F., GDSL family of serine esterases/lipases, *Progress in Lipid Research*. (2004) 43 534–552.

[30] Liu, X., Zhou, M., Xing, S., Wu, T., He, H., Bielicki, J.K., Chen, J., Identification and biochemical characterization of a novel hormone-sensitive lipase family esterase est19 from the antarctic bacterium *Pseudomonas* sp. E2-15, *Biomolecules*. (2021) 11.

[31] Bai, Q., Cai, Y., Yang, L., Wu, Q., Su, M., Tang, L., Substitution of L133 with methionine in GXSXG domain significantly changed the activity of Pe*nicillium expansum* lipase, *Catalysis Letters*. (2022) 152 2047–2055.

[32] Simón, L., Goodman, J.M., Enzyme catalysis by hydrogen bonds: The balance between transition state binding and substrate binding in oxyanion holes, *Journal of Organic Chemistry*. (2010) 75 1831–1840.

[33] Yang, H., Wong, M.W., Oxyanion hole stabilization by C-H·O Interaction in a transition state -a three-point interaction model for cinchona alkaloid-catalyzed asymmetric methanolysis of meso-cyclic anhydrides, *Journal of the American Chemical Society*. (2013) 135 5808–5818.

[34] Lu, Y., Wang, R., Yang, C.Y., Wang, S., Analysis of ligand-bound water molecules in high-resolution crystal structures of protein-ligand complexes, *Journal of Chemical Information and Modeling*. (2007) 47 668–675.

[35] McAuley, K.E., Svendsen, A., Patkar, S.A., Wilson, K.S., Structure of a feruloyl esterase from *Aspergillus niger*, *Acta Crystallographica Section D: Biological Crystallography*. (2004) 60 878–887.

[36] Faulds, C.B., Molina, R., Gonzalez, R., Husband, F., Juge, N., Sanz-Aparicio, J., Hermoso, J.A., Probing the determinants of substrate specificity of a feruloyl esterase, *An*FaeA, from *Aspergillus niger*, *FEBS Journal*. (2005) 272 4362–4371.

[37] Lai, K.K., Stogios, P.J., Vu, C., Xu, X., Cui, H., Molloy, S., Savchenko, A., Yakunin, A., Gonzalez, C.F., An inserted *α/β* subdomain shapes the catalytic pocket of *lactobacillus johnsonii* cinnamoyl esterase, *PLoS ONE*. (2011) 6 e23269.

[38] Qi, M., Wang, P., Selinger, L.B., Yanke, L.J., Forster, R.J., Mcallister, T.A., Isolation and characterization of a ferulic acid esterase (Fae1A) from the rumen fungus *Anaeromyces mucronatus*, *Journal of Applied Microbiology*. (2011) 110 1341–1350.

[39] Crepin, V.F., Faulds, C.B., Connerton, I.F., Functional classification of the microbial feruloyl esterases, *Applied Microbiology and Biotechnology*. (2004) 63 647–652.

[40] NIST/NIH/EPA Mass Spectral Library, Standard Reference Database 1, NIST 11. Standard Reference Data Program,National Institute of Standards and Technology: Gaithersburg, MD, USA, 2011., n.d.

[41] Bunzel, M., Ralph, J., Marita, J.M., Hatfield, R.D., Steinhart, H., Diferulates as structural components in soluble and insoluble cereal dietary fibre, *Journal of the Science of Food and Agriculture*. (2001) 81 653–660.

[42] Aguedo, M., Vanderghem, C., Goffin, D., Richel, A., Paquot, M., Fast and high yield recovery of arabinose from destarched wheat bran, *Industrial Crops and Products*. (2013) 43 318–325.

[43] Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., Nielsen, H., SignalP 5.0 improves signal peptide predictions using deep neural networks, *Nature Biotechnology*. (2019) 37 420–423.

[44] Tamura, K., Stecher, G., Kumar, S., MEGA11: Molecular evolutionary genetics analysis version 11, *Molecular Biology and Evolution*. (2021) 38 3022–3027.

[45] Gupta, R., Brunak, S., Prediction of glycosylation across the human proteome and the correlation to protein function, *Pacific Symposium on Biocomputing*. (2002) 7 310–322.

[46] Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager, K.T.B.G., Lavrsen, K., Dabelsteen, S., Pedersen, N.B., Marcos-Silva, L., Gupta, R., Paul Bennett, E., Mandel, U., Brunak, S., Wandall, H.H., Levery, S.B., Clausen, H., Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology, *EMBO Journal*. (2013) 32 1478–1488.

[47] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A., Protein identification and analysis tools on the ExPASy server, in: John M. Walker (Ed.), The Proteomics Protocols Handbook, Humana Press, 2005: pp. 571–607. http://www.expasy.org/tools/.

[48] Simonian, M.H., Spectrophotometric determination of protein concentration, *Current Protocols in Cell Biology*. (2002) 15 A.3B.1-A.3B.7.

[49] Xu, J., Zhao, X., Yao, Q., Zong, W., Dai, S., Deng, Z., Liu, S., Yun, J., Yang, X., Li, H., Cloning, characterization of a novel acetyl xylan esterase, and its potential application on wheat straw utilization, *All Life*. (2021) 14 622–635.

[50] Kirsch, P.D., Ekerdt, J.G., KaleidaGraph: graphing and data analysis. version 3.5 for windows synergy software, 2457 Perkiomen Ave., Reading, PA 19606-2049. www.Synergy.com. $155.00, *Journal of the American Chemical Society*. (2000) 122 11755.

[51] Kabsch, W., XDS, *Acta Crystallographica Section D Biological Crystallography*. (2010) 66 125–132.

[52] Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., Overview of the CCP4 suite and current developments, *Acta Crystallographica Section D: Biological Crystallography*. (2011) 67 235–242.

[53] Emsley, P., Cowtan, K., Coot: Model-building tools for molecular graphics, *Acta Crystallographica Section D: Biological Crystallography*. (2004) 60 2126–2132.

[54] Adams, P.D., Afonine, P. v., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., PHENIX: A comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D: Biological Crystallography*. (2010) 66 213–221.

[55] Kabsch, W., Sander, C., Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers*. (1983) 22 2577–2637.

[56] Delano, W.L., The PyMOL molecular graphics system, in: Delano Scientific, San Carlos, 2002.

[57] Robert, X., Gouet, P., Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Research*. (2014) 42 W320–W324.

Chapter Four

Boosting sugarcane trash decomposition: Synergistic action of xylanase and feruloyl esterase co-displayed on the cell surface of *Pichia pastoris*

Reproduced from Ref. *Sustain. Energy Fuels*, 2024, DOI: 10.1039/D3SE01482G with permission from the Royal Society of Chemistry.

4.1 Introduction

Lignocellulosic biomass, derived from plant materials such as agricultural waste, wood, and grasses, represents a vast and renewable resource with significant potential for producing biofuels and valuable chemicals [1]. However, its intricate structure, which stabilizes the plant cell wall, presents challenges for effective breakdown into constituent sugars and aromatic compounds [2]. Hemicellulose plays a crucial role in bridging cellulose fibers/microfibrils through hydrogen bonding, while lignin fills the spaces between cellulose and hemicellulose in the cell wall [3]. Hemicellulose is a branched heteropolysaccharide made up from monomers (glucose, mannose, galactose, xylose, and arabinose) and various modification groups (ferulic, *p*-coumaric, and acetic acids) [3]. The complex composition and cross-linked nature of hemicellulose pose challenges for its enzymatic degradation.

Xylanase (XYN) and feruloyl esterase (FAE) are essential enzymes with distinct catalytic functions that play a critical role in efficient hemicellulose degradation [4,5]. XYN is an enzyme that specifically hydrolyzes the β-1,4-glycosidic linkages within the xylan backbone, yielding shorter xylooligosaccharides and xylose molecules **(Fig. 4-1)** [4]. This enzymatic activity significantly enhances the accessibility of hemicellulose for further degradation and conversion. In parallel, FAE specializes in cleaving the ester linkages between hemicellulose and phenolic moieties, liberating ferulic acid (FA) and disrupting the lignin-carbohydrate complex **(Fig. 4-1)** [5]. By hydrolyzing the ester bonds within the lignocellulosic matrix, FAE not only improves the accessibility of other enzymes to the substrate for biomass degradation [6] but also generates a valuable

phenolic compound, FA, which can be used in the food, pharmaceutical, and nutraceutical industries [7–9].

Previous studies have shown that XYN and FAE work synergistically in biomass degradation, underscoring their potential to enhance the efficiency of lignocellulosic biomass conversion for sustainable bioprocessing [6,10–12]. For instance, the addition of a crude extract from *Aspergillus oryzae*, containing FAE, to Celluclast 1.5L, which includes xylanase, cellulase, and pectinase, resulted in observable synergistic effects [10]. Additionally, synergistic effects were demonstrated through the incorporation of FAEs from termite metagenomes with XYN11 from *Thermomyces lanuginosus* [11]. Likewise, the combined utilization of FAE-1 and various XYNs also showed synergistic effects [12].

Yeast surface display (YSD) is a technique employed for the presentation of enzymes on the yeast cell surface through fusion with an anchoring domain [13]. This method has been employed to create whole-cell biocatalysts capable of producing valuable chemicals and biofuels from lignocellulosic biomass [14]. The potential of YSD has been further augmented through genetic engineering strategies, including the optimization of promoters [15], signal peptides [16], anchoring proteins [17,18], yeast cell wall modification [18,19], and modification of the secretory pathway [20,21].

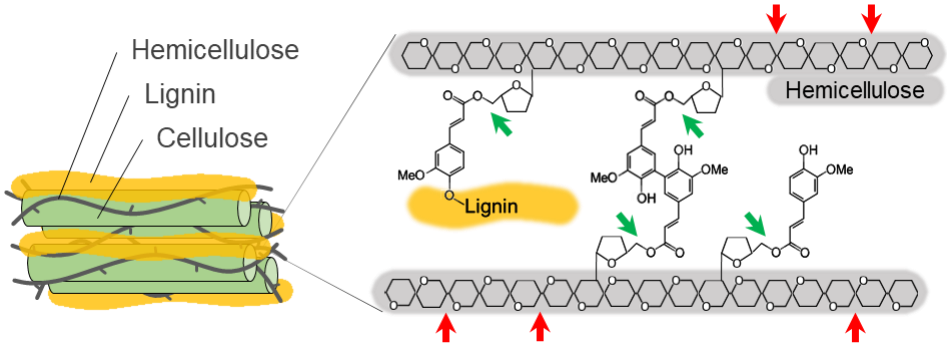

Fig. 4-1. Illustration of the feruloylated arabinoxylan from the lignocellulosic biomass and its hydrolyses. The hydrolyses by xylanase (XYN) and feruloyl esterase (FAE) are represented by red and green arrows, respectively.

YSD offers several potential advantages over enzyme secretion for applications in lignocellulosic biomass degradation [22]. Immobilization on the cell surface can enhance enzyme stability and prevent degradation compared to free enzymes in solution [23,24]. Concentration of enzymes on the cell surface can result in higher local enzyme concentrations, thereby increasing catalytic efficiency [25,26]. Yeast cells with surface-displayed enzymes can be easily separated from the reaction mixture and reused, simplifying enzyme recycling [27–29].

Furthermore, YSD facilitates the co-display of multiple different enzymes on the same cell surface, introducing the concept of a proximity effect among these enzymes [30,31]. This concept has gained prominence due to its ability to accelerate the reaction rates of enzyme cascades [28,31–34]. Notably, the close proximity of three different kinds of cellulase co-displayed on the yeast cell surface has been reported to increase ethanol yield through the simultaneous saccharification and fermentation of phosphoric acid-swollen cellulose [28,30,35], microcrystalline cellulose [36] and alkali-pretreated sugarcane bagasse [21]. Additionally, scaling up of the simultaneous saccharification and fermentation of pretreated bagasse in a 1-L bioreactor using cellulase-displaying yeast together with commercial cellulase achieved fermentability of up to 86.5% of the theoretical yield [21]. It is worth noting that the depolymerization of lignocellulosic biomass can be enhanced by sequential actions using peroxidases and cellulases [37].

Previously, YSD of XYNs was shown to be effective in biomass degradation. Xylanase Orf6-un$_m$, when displayed on the yeast cell surface, exhibited enhanced xylanolytic activity compared to the enzyme purified from *Escherichia coli*, leading to improved corn stover digestion by rumen cultures [38]. In a separate study, the surface display of a XYN from *Lentinula edodes* on yeast cells demonstrated its capacity to efficiently hydrolyze wheat residue [39].

However, it is worth noting that, to date, there have been no reported instances of YSD for FAEs. YSD of both XYN and FAE on the same yeast surface has not been previously investigated, either. Since the coexistence of XYN and FAE leads to their synergistic action, I hypothesized that employing YSD to immobilize XYN and FAE on the same yeast surface could introduce not only synergistic effects but also proximity effects.

Recent studies have highlighted the utilization of the engineered *Pichia pastoris* as a versatile chassis biocatalyst to produce biofuels, including ethanol [40], isobutanol [41], and biodiesel [42].

In this study, I individually fused XYN from *Thermomyces lanuginosus* and FAE from *Acremonium alcalophilum* with the *Saccharomyces cerevisiae* SED1 anchoring domain [43]. This anchoring domain is known for its ability to effectively localize fusion proteins on the *P. pastoris* cell surface [29,39,44–46]. After successfully generating strains that display XYN alone (X-Pichia), FAE alone (F-Pichia), and both enzymes together (X/F-Pichia) **(Fig. 4-2)**, I examined the localization, stability, and relative activity of the displayed enzymes under various conditions, including temperature, pH level, and the presence of different additives.

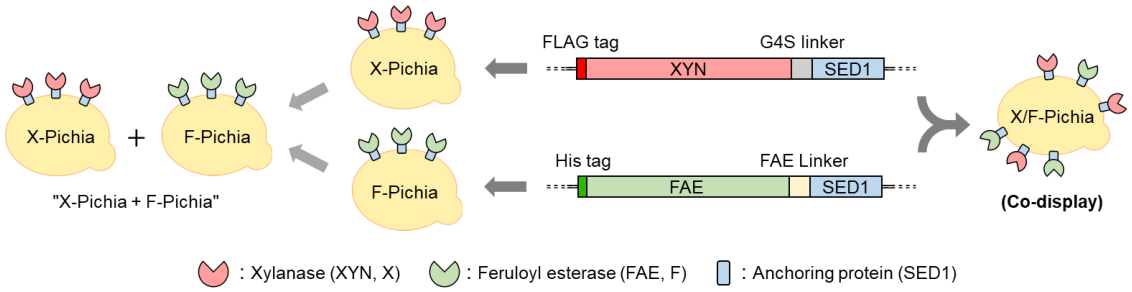

Fig. 4-2. Schematic presentation of *P. pastoris* strains. The X-Pichia strain expresses xylanase (XYN), the F-Pichia strain expresses feruloyl esterase (FAE), and the X/F-Pichia strain expresses both XYN and FAE, all on the cell surface. The term "X-Pichia + F-Pichia" refers to a mixture of X-Pichia and F-Pichia. SED1 corresponds to the anchoring protein.

This study represents the first instance of displaying FAE on the yeast cell surface and exploring the synergistic and proximity effects between XYN and FAE in biomass degradation. Specifically, sugarcane trash, defined as the upper shoots and leaf sheath with the leaves remaining on the ground and constituting approximately 15% of the total aboveground biomass [47], was used as the substrate. When I used a mixture of X-Pichia and F-Pichia ("X-Pichia + F-Pichia"), I observed synergistic enhancement of

the hydrolysis of acid-pretreated sugarcane trash, leading to increased production of reducing sugars compared to X-Pichia alone. Moreover, X/F-Pichia, benefiting from the proximity effect between XYN and FAE, further amplified the synergistic effect observed in "X-Pichia + F-Pichia", resulting in higher reducing sugar production. Additionally, X/F-Pichia produced more ferulic acid than F-Pichia alone or "X-Pichia + F-Pichia". These findings strongly suggest that the strategy of co-displaying the combination of FAE and XYN enables the utilization of their synergistic and proximity effects for efficiently utilizing lignocellulosic biomass and lays the foundation for sustainable bioprocessing with substantial potential.

4.2 Results and discussions

4.2.1 Construction of P. pastoris strains

In this study, I employed wild-type *P. pastoris* X-33 as the starting strain for displaying fungal XYN and FAE enzymes. Plasmid pPICZαA served as the foundational DNA for constructing pPICZαA-XYN-SED and pPICZαA-FAE-SED **(Table 4-1)**. The XYN and FAE genes, along with α-factor secretion signal sequences from the plasmid backbone, were expressed under the control of the *AOX*1 promoter. To anchor the enzymes onto the yeast cell surface, I fused the *Sc*SED1 anchoring protein to the C-terminus of FLAG-tagged XYN and 6× His-tagged FAE **(Fig. 4-2)**. The enzyme-displaying yeast strains, namely X-Pichia, F-Pichia, and X/F-Pichia, were generated by introducing the gene cassettes of pPICZαA-XYN-SED and/or pPICZαA-FAE-SED into the *P. pastoris* X-33 genome, as depicted in **Fig. 4-2**.

Table 4-1. Microbial strains, plasmids, and primers utilized in this study.

Strain, plasmid, or primer	Relevant feature or sequence
Strains	
P. pastoris strains	
X-33 (Invitrogen)	Wild type (WT)
X-Pichia	Display of *Thermomyces lanuginosus* XYN
F-Pichia	Display of *Acremonium alcalophilum* FAE
X/F-Pichia	Display of XYN and FAE
S. cerevisiae strain	
Cen-PK2-1C (Euroscarf)	
Bacterial strain	
E. coli DH5*α* (Invitrogen)	
Plasmids	
pPICZ*α*A (Invitrogen)	*Zeo*^{r} expression vector, *AOX1* promoter
pPICZ*α*A-XYN-SED	Cell surface display of *T. lanuginosus* XYN with FLAG tag
pPICZ*α*A-FAE-SED	Cell surface display of *A. alcalophilum* FAE with His tag
Primers	
AaFAE-EcoRI-F	AGCTGAATTCCACCATCACCATCACCAT
AaFAE-SpeI-R	TAAACTAGTTCCTCCACTGGGTGGGGGA
TlXYN-EcoRI_F	GAAGCTGAATTCGACTACAAAGATG
TlXYN-G4S-SpeI_R	AATTGACTAGTAGAGCCTCCGC
SpeI-SED1_F	ATCACTAGTCAATTTTCCAACAGTACATCTGCTTCTTC
NotI-SED1_R	CTGGCGGCCGCTTATAAGAATAACATAGCAACACCAGCC

4.2.2 Localization of fusion proteins in P. pastoris

I visualized XYN and FAE displayed on the yeast cell surface via immunofluorescence analysis. I employed a FLAG-tag antibody labeled with Alexa Fluor 594 for XYN and His-probe antibody labeled with Alexa Fluor 488 for FAE. The results showed strong red fluorescence on the surface of X-Pichia and X/F-Pichia, while F-Pichia and X/F-Pichia displayed intense green fluorescence **(Fig. 4-3)**. These findings confirm the successful anchoring of XYN and FAE on the cell surface of recombinant *P. pastoris* X-33.

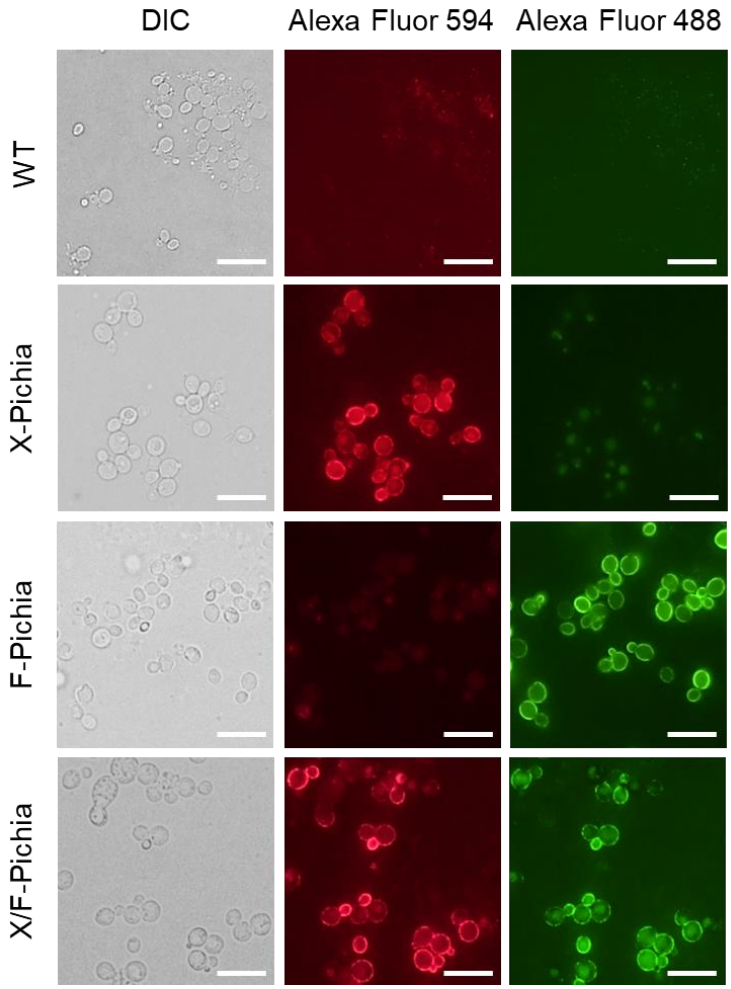

Fig. 4-3. Visualization of enzyme localization by immunofluorescence. Immunofluorescent labeling of cell surface-displayed XYN and FAE was performed using Alexa Fluor 594-labeled FLAG-tag antibody and Alexa Fluor 488-labeled His-probe antibody. DIC: differential interference contrast. Scale bars represent 10 μm.

4.2.3 Enzyme activity of XYN and FAE displayed on yeast cell surface

After induction for protein expression, I collected both the *P. pastoris* cell fraction and the media fraction. Activity assays were conducted on the cell fraction using beechwood xylan and MFA as substrates to assess the functionality of the enzymes displayed on the *P. pastoris* cell surface. The results revealed XYN activity levels of 6.0 mU/OD_{590} unit for X-Pichia and 5.0 mU/OD_{590} unit for X/F-Pichia, while FAE activity levels were 2.0 mU/OD_{590} unit for F-Pichia and 1.8 mU/OD_{590} unit for X/F-Pichia **(Table**

4-2, left). One OD_{590} unit represents the amount of enzyme included in 1 mL of the solution whose optical density at 590 nm is 1.0. These findings confirm the functional activity of the enzymes displayed on the cell surface of X-Pichia, F-Pichia and X/F-Pichia.

Further investigation of enzyme activity in the media fraction indicated that the fusion proteins in my study are secreted into the culture media **(Fig. 4-4)**. This suggests that the overexpression of proteins conjugated with the SED1 anchoring protein led to a minor fraction of fusion proteins being released into the growth medium [48].

Table 4-2. Activities of XYN and FAE displayed on the *P. pastoris* cell surface (left) and experimental setup of *P. pastoris* strain in optical density (OD_{590} unit) used for hydrolysis of acid-pretreated sugarcane trash (right).

***P. pastoris* strain**	**Enzyme activity (mU/OD_{590} unit)**[a]		**Sample name / *P. pastoris* strain**	**Amounts of *P. pastoris* (OD_{590} units) used for a 1 mL hydrolysis of acid-pretreated sugarcane trash**				
	XYN[b]	**FAE**[c]		**X-33**[d]	**X-Pichia**[e]	**F-Pichia**[f]	**"X-Pichia + F-Pichia"**	**X/F-Pichia**
X-33	ND	ND	X-33	17.7	9.3	8.4	-	7.7
X-Pichia	6.0 ± 0.1	ND	X-Pichia	-	8.4	-	8.4	-
F-Pichia	ND	2.0 ± 0.0	F-Pichia	-	-	9.3	9.3	-
X/F-Pichia	5.0 ± 0.4	1.8 ± 0.1	X/F-Pichia	-	-	-	-	10

[a] Values are the means ± standard deviation of three independent replicates. ND: not detected.

[b] XYN activity assay using beechwood xylan as substrate.

[c] FAE activity assay using MFA as substrate.

[d] Reference reaction was carried out using 17.7 OD_{590} units of *P. pastoris* X-33.

[e] 8.4 OD_{590} units of X-Pichia exerts 50 mU XYN activity, which is equivalent to the XYN activity of 10 OD_{590} units of X/F-Pichia.

[f] 9.3 OD_{590} units of F-Pichia exerts 18 mU FAE activity, which is equivalent to the FAE activity of 10 OD_{590} units of X/F-Pichia.

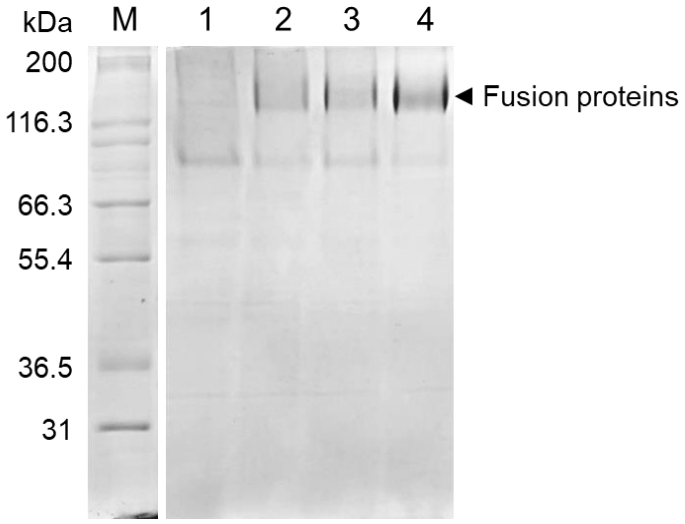

Fig. 4-4. SDS-PAGE analysis of the crude proteins secreted by *P. pastoris* X-33 after two days induction. Lanes: M, molecular mass markers; 1, X-33 (WT); 2, X-Pichia; 3, F-Pichia; 4, X/F-Pichia.

4.2.4 Temperature and pH stability of XYN and FAE on the P. pastoris cell surface

To assess the robustness of immobilized enzymes on the *P. pastoris* cell surface, I investigated the temperature and pH stability of XYN and FAE on X-Pichia and F-Pichia, respectively. X-Pichia and F-Pichia were subjected to various temperatures and pH conditions for 24 h, after which I examined the residual XYN and FAE activities.

The results indicate that surface-displayed XYN remains stable within a temperature range of 20 to 70 °C, retaining over 90% of its activity **(Fig. 4-5 A)**. However, its activity significantly decreases on incubation at 80 °C for 24 h **(Fig. 4-5 A)**. Additionally, XYN retains high activity over a broad pH range of 4.0 to 9.0, with residual activity exceeding 85% **(Fig. 4-5 B)**. Notably, the XYN displayed on the cell surface in this study exhibits superior thermal stability compared to surface-displayed *Lentinula edodes* XYN, which remains stable only within a temperature range of 20 to 30 °C [39]. This difference in stability may be attributed to the adaptation of *T. lanuginosus* to high-temperature environments, such as hot springs [49].

Surface-displayed FAE, on the other hand, demonstrates stability within a narrower temperature range, specifically between 20 to 40 °C, and exhibits almost no activity after exposure to temperatures exceeding 50 °C **(Fig. 4-5 A)**. Additionally, it was observed that surface-displayed FAE maintains robust stability at pH levels ranging from 6.0 to 9.0, consistently retaining over 90% of its activity. However, it displays lower stability under mildly acidic conditions, with activity falling below 50% **(Fig. 4-5 B)**. Overall, the temperature and pH stability of surface-displayed XYN and FAE closely align with those of the crude *Tl*XYN [50] and secreted *Aa*FaeD in **Chapter 3**, respectively.

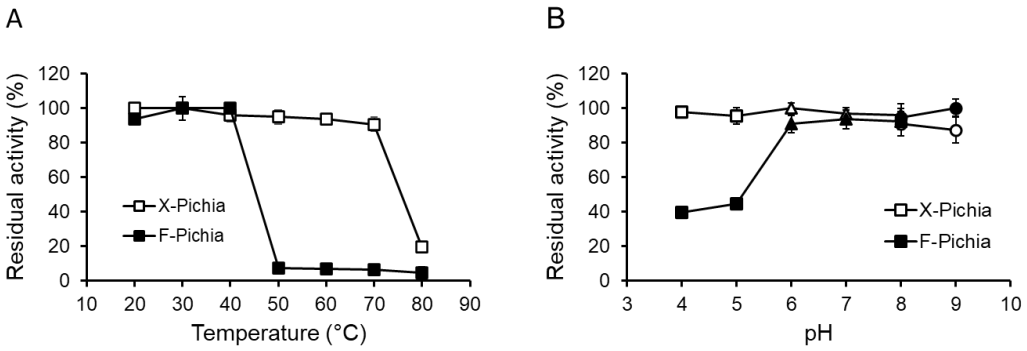

Fig. 4-5. Temperature and pH stability of enzymes displayed on the cell surface of X-Pichia and F-Pichia. The activity of enzymes displayed on the cell surface of X-Pichia and F-Pichia was assessed concerning temperature and pH, using beechwood xylan and methyl ferulate (MFA), respectively. (A) Thermostability: X-Pichia and F-Pichia were incubated at various temperatures for 24 h, and the activity was measured at 30 °C. The residual activity is plotted against the incubation temperature. (B) pH stability: X-Pichia and F-Pichia were incubated at various pH values for 24 h, and the activity was measured at pH 7.0. The residual activity is plotted against the incubation pH: squares represent the residual activities for pH 4.0–5.0 (sodium acetate), triangles for pH 6.0–8.0 (potassium phosphate), and circles for pH 8.0–9.0 (Tris-HCl).

4.2.5 Effect of additives on the enzyme activities of XYN and FAE on the P. pastoris cell surface

In previous studies, it has been reported that certain metal ions and organic reagents can enhance the activities of XYN [51,52] and FAE [53]. Therefore, in this study, I investigated the impact of additives on the activities of XYN and FAE on the P. pastoris cell surface.

The addition of $CoCl_2$, $NiCl_2$, and $CuCl_2$ to X-Pichia led to an increase in XYN activity. Specifically, at a concentration of 5 mM, these salts enhanced the activity by 88%, 23%, and 81%, respectively **(Table 4-3)**. The increase in XYN activity could be due to the structural change in the active site, resulting in the enlargement of substrate-binding pocket, induced by binding with the metal ion [52]. In contrast, the introduction of ethylenediaminetetraacetate (EDTA) had an adverse effect on XYN activity, resulting in a 35% reduction at the same 5 mM concentration. Additionally, the inclusion of β-mercaptoethanol (2-ME) and dithiothreitol (DTT) showed a slight improvement in XYN activity, with increases of 12% and 16% at the 5 mM concentration, respectively (Table 3). On the other hand, X-Pichia exhibited a 10% decrease in XYN activity when sodium dodecyl sulfate (SDS) and Triton X-100 were added, while the presence of Tween 20 and Tween 80 had no discernible impact on XYN activity.

The addition of $MgCl_2$ and $CaCl_2$ to F-Pichia resulted in an increase in FAE activity. At a concentration of 5 mM, these salts increased the activity by 20% and 34%, respectively **(Table 4-3)**. Conversely, the addition of $CoCl_2$, $NiCl_2$, $CuCl_2$, and $ZnCl_2$, had an adverse effect on FAE activity, resulting in 34%, 63%, 100%, and 84% reduction at the same 5 mM concentration, respectively. The addition of EDTA led to a slight improvement in FAE activity, with an increase of 15%. On the other hand, F-Pichia exhibited reductions of 87%, 44%, 66%, and 41% in FAE activity when SDS, Triton X-100, Tween 20, and Tween 80 were added, respectively.

Table 4-3. Effects of additives on the activities of XYN and FAE displayed on the *P. pastoris* cell surface.

	Relative activity (%)[a]			
	XYN activity (X-Pichia)[b]		FAE activity (F-Pichia)[c]	
Additive	**1mM**	**5mM**	**1mM**	**5mM**
Metal ions				
LiCl	100 ± 3	104 ± 3	94 ± 9	97 ± 4
NaCl	100 ± 2	100 ± 4	84 ± 10	100 ± 5
KCl	96 ± 2	100 ± 3	93 ± 4	89 ± 6
$MgCl_2$	97 ± 2	104 ± 3	93 ± 8	120 ± 8
$CaCl_2$	98 ± 2	105 ± 5	113 ± 9	134 ± 12
$MnCl_2$	99 ± 4	108 ± 4	95 ± 5	103 ± 5
$CoCl_2$	141 ± 5	188 ± 7	58 ± 6	66 ± 16
$NiCl_2$	104 ± 7	123 ± 8	46 ± 6	37 ± 6
$CuCl_2$	119 ± 6	181 ± 5	24 ± 4	N.D.
$ZnCl_2$	96 ± 2	109 ± 4	18 ± 10	16 ± 12
Inhibitors				
EDTA	95 ± 8	65 ± 4	102 ± 5	115 ± 9
2-ME	111 ± 1	112 ± 1	112 ± 5	101 ± 3
DTT	106 ± 3	116 ± 3	102 ± 5	106 ± 5
Furfural	102 ± 1	105 ± 1	115 ± 4	110 ± 7
Detergents				
SDS	96 ± 9	90 ± 7	80 ± 5	13 ± 2
Triton X-100	96 ± 7	93 ± 6	84 ± 8	56 ± 9
Tween 20	100 ± 8	100 ± 6	34 ± 5	34 ± 4
Tween 80	106 ± 7	100 ± 7	88 ± 3	59 ± 4

[a] Reference reactions were carried out for each additive in the absence of X-Pichia or F-Pichia. The values obtained from the reference reactions were subtracted to derive the values presented in the Table. The activity of X-Pichia and F-Pichia in the absence of additives was set as 100%. Values are the means ± standard deviation of three independent replicates. ND: Not Detected.

[b] XYN activity assay was conducted using beechwood xylan as substrate.

[c] FAE activity assay was conducted using MFA as substrate.

4.2.6 Sample preparation for investigating synergistic hydrolysis of XYN and FAE displayed on the P. pastoris cell surface

This section outlines the sample preparation process for the subsequent investigation into the hydrolysis of acid-pretreated sugarcane trash by different *P. pastoris* strains: a single-displaying strain alone (X-Pichia or F-Pichia), a mixture of single-displaying strains ("X-Pichia + F-Pichia"), and the co-displaying strain (X/F-Pichia). Specifically, I examine the hydrolysis of beechwood xylan by cell-surface displayed XYN (referred to as XYN activity) and the hydrolysis of MFA by cell-surface displayed FAE (referred to as FAE activity). In this section, I calculate the relative XYN and FAE activities of each *P. pastoris* strain to determine the proportions in which they need to be mixed for hydrolysis of natural biomass, sugarcane trash.

As established in the previous section, the XYN activities (activities of hydrolysis of beechwood xylan) of X-Pichia and X/F-Pichia were approximately 6.0 and 5.0 mU/OD_{590} units, respectively **(Table 4-2, left)**. Therefore, when one OD_{590} unit of X/F-Pichia is used, 0.84 OD_{590} units of X-Pichia are required to equalize their XYN activities. Similarly, the FAE activities (MFA hydrolysis) of F-Pichia and X/F-Pichia were approximately 2.0 and 1.8 mU/OD_{590} units, respectively **(Table 4-2, left)**. In this case, when one OD_{590} unit of X/F-Pichia is used, 0.93 OD_{590} units of F-Pichia are needed to equalize their FAE activities.

Firstly, I set the amount of X/F-Pichia to 10 OD_{590} units (**Table 4-2, right**, X/F-Pichia column). Consequently, to achieve equivalent XYN and FAE activities, 8.4 OD_{590} units of X-Pichia and 9.3 OD_{590} units of F-Pichia are used. Secondly, the total amount of *P. pastoris* for "X-Pichia + F-Pichia" becomes 17.7 OD_{590} units (**Table 4-2, right**, "X-Pichia + F-Pichia" column). In order to adjust the total *P. pastoris* amount to 17.7 OD_{590} units, original *P. pastoris* X-33, which contains no XYN or FAE, is added to X-Pichia, F-Pichia, and X/F-Pichia in quantities of 9.3, 8.4, and 7.7 OD_{590} units, respectively (**Table 4-2, right**, X-Pichia, F-Pichia, and X/F-Pichia columns). Finally, a sample of 17.7 OD_{590} units of X-33 was prepared as a control.

4.2.7 Synergistic hydrolysis of XYN and FAE on P. pastoris cell surface

To explore the synergistic action of XYN and FAE displayed on the *P. pastoris* cell surface during biomass degradation, I conducted activity assays of a single-displaying strain alone (X-Pichia or F-Pichia), a mixture of single-displaying strains ("X-Pichia + F-Pichia"), and the co-displaying strain (X/F-Pichia), using acid-pretreated sugarcane trash as the substrate. Measurements were taken at both 24 and 48 h, and the product amounts after 48 h were used for assessment.

After 48 h, the concentration of reducing sugar was ca. 0.04 mg/mL for X-Pichia, ca. 0.01 mg/mL for F-Pichia, ca. 0.05 mg/mL for "X-Pichia + F-Pichia", and ca. 0.06 mg/mL for X/F-Pichia **(Fig. 4-6 A)**. Notably, "X-Pichia + F-Pichia" released 1.22 times more reducing sugar than X-Pichia alone, a clear synergistic action between XYN and FAE in the breakdown of hemicellulose being observed **(Fig. 4-6 A)**. Furthermore, X/F-Pichia released 1.21 times more reducing sugar than "X-Pichia + F-Pichia" and 1.47 times more than X-Pichia **(Fig. 4-6 A)**. In X/F-Pichia, XYN and FAE are co-displayed on the *P. pastoris* cell surface, bringing these enzymes into close proximity. This physical co-localization enhances the synergistic effect between XYN and FAE. Within hemicellulose, there are numerous sites linked with FA. When FAs at these sites are hydrolyzed and released by FAE, the hemicellulose at these sites becomes exposed, and the local structure of hemicellulose relaxes **(Fig. 4-1)**. These products have improved accessibility and become better substrates for XYN [6,10–12]. Thus, synergistic and proximity effects allow products generated by one enzyme to immediately become substrates for the other, resulting in an observed increase in reducing sugar production.

After 48 h, the concentration of FA was ca. 0.07 μg/mL for X-Pichia, ca. 0.66 μg/mL for F-Pichia, ca. 0.66 μg/mL for "X-Pichia + F-Pichia", and ca. 0.73 μg/mL for X/F-Pichia **(Fig. 4-6 B)**. The amount of FA produced by X-Pichia (ca. 0.07 μg/mL) closely matched the initial content in the acid-pretreated sugarcane trash. Interestingly, "X-Pichia + F-Pichia" released the same amount of FA as F-Pichia alone, while X/F-Pichia released 1.10 times more FA than F-Pichia alone **(Fig. 4-6 B)**. This finding underscores the importance of the proximity effect in facilitating the synergistic action of XYN and FAE in producing FA.

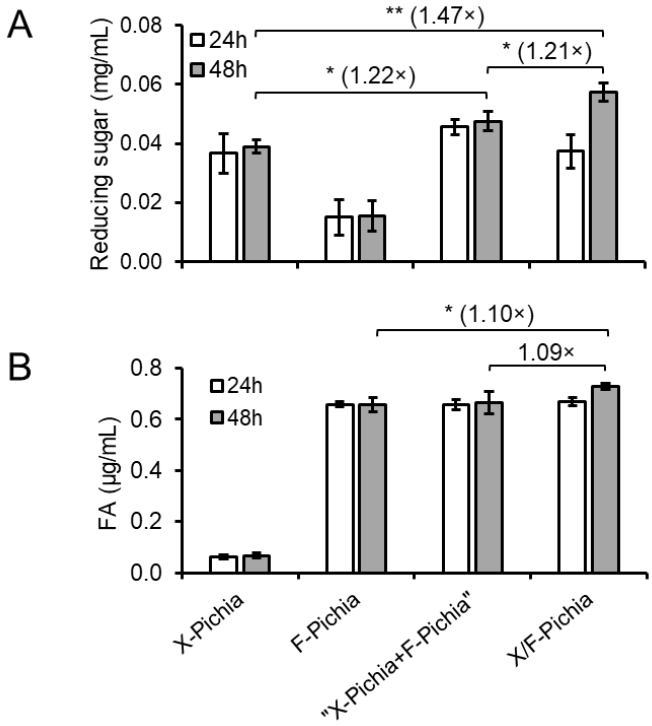

Fig. 4-6. Synergistic and proximity effects on degradation of acid-pretreated sugarcane trash by X/F-Pichia. Measurement of reducing sugar release (A) and ferulic acid release (B) was conducted following 24-h and 48-h hydrolysis of 1% acid-pretreated sugarcane trash using different *P. pastoris* strain combinations: X-Pichia, F-Pichia, "X-Pichia + F-Pichia," and X/F-Pichia. The values presented in the figure are the mean ± standard deviation (n = 3). Significance is indicated by * for $p < 0.05$ and ** for $p < 0.01$.

4.3 Conclusions

In this pioneering study, I successfully harnessed YSD technology to co-display fungal XYN and FAE enzymes, thereby exploring the intriguing interplay of synergistic and proximity effects in biomass degradation. My findings shed light on the potential of this innovative strategy for efficient lignocellulosic biomass utilization.

I successfully constructed *P. pastoris* strains, namely X-Pichia, F-Pichia, and X/F-Pichia, expressing FLAG-tagged XYN and 6x His-tagged FAE, anchored to the yeast cell surface using *Sc*SED1. My investigation demonstrated robust enzyme stability across varying temperatures and pH levels, expanding the versatility of this co-display

system. Furthermore, the impact of specific salts and chemicals on XYN and FAE activities provided valuable insights for optimizing their catalytic efficiency in biomass degradation.

The most significant findings emerged from my analysis of synergistic hydrolysis. Activity assays revealed that "X-Pichia + F-Pichia" released more reducing sugar than X-Pichia alone, indicating a clear synergistic effect between XYN and FAE in hemicellulose breakdown. Impressively, X/F-Pichia, displaying both XYN and FAE on the same cell surface, demonstrated even greater enhancements in releasing reducing sugar and FA than "X-Pichia + F-Pichia". This underscores the amplified synergistic effect achieved through the proximity effect, where the physical co-localization of XYN and FAE facilitated effective substrate turnover, resulting in increased product yields.

In summary, my study has revealed a potent approach to maximize the utilization of lignocellulosic biomass by harnessing the synergistic and proximity effects of co-displayed XYN and FAE on the yeast cell surface. This strategy holds significant promise for sustainable bioprocessing, as it has the potential to increase the production of reducing sugars and valuable phenolic compounds such as ferulic acid [7–9]. As I progress, this innovative approach offers the potential to advance biotechnology applications and contribute to a more environmentally friendly and resource-efficient future.

4.4 Materials and methods

4.4.1 Strains and culture conditions

Yeast and bacterial strains used in this study are summarized in **Table 4-1**. *Pichia pastoris* strains were cultured in Yeast extract Peptone Dextrose (YPD) agar (20 g/L peptone, 10 g/L yeast extract, 10 g/L D-glucose, and 15 g/L bacteriological agar) at 30 °C. *P. pastoris* transformants were screened on YPD supplemented with 100 μg/mL Zeocin (Invivogen, USA). For protein expression, *P. pastoris* transformants were pre-cultivated overnight in YPD at 30 °C with 250 rpm shaking. The overnight cultures were transferred to 100 mL of buffered glycerol complex medium (BMGY) (1% (w/v) yeast extract, 2% (w/v) peptone, 100 mM potassium phosphate, pH 6.0, 1.34% (w/v) yeast nitrogen base (YNB), and 4×10^{-5}% (w/v) biotin) supplemented with 1% (v/v) glycerol

and grown at 30 °C with 250 rpm for 3 days. The cells were harvested by centrifugation at 8,000 × *g* for 5 min at room temperature (RT) and inoculated into 50 mL of BMMY containing 0.5% methanol. To maintain the induction of protein expression, methanol was added to the yeast culture once daily for 2 days to make its concentration 0.5%.

Escherichia coli DH5*α* (Invitrogen) was used for plasmid propagation. Bacterial cells were grown in Luria-Bertani broth (10 g/L peptone, 5 g/L yeast extract, and 5 g/L NaCl) containing 25 μg/mL Zeocin for transformant selection.

4.4.2 Plasmid construction

PCR amplification was carried out using KOD-plus DNA polymerase (Toyobo, Osaka, Japan) with the primers listed in **Table 4-1**. The anchoring domain *Sc*SED1 was amplified from *S. cerevisiae* strain Cen-PK2-1C **(Appendix E)** using SpeI-SED1_F and NotI-SED1_R primers **(Table 4-1)**. DNAs encoding *Thermomyces lanuginosus* xylanase (*Tl*XYN; UniProt ID: O43097) and *Acremonium alcalophilum* FaeD (*Aa*FaeD; JGI ID: Acral2|1082309) were codon-optimized and synthesized by Thermo Fisher Scientific (Japan). *Tl*XYN, fused with an N-terminal FLAG tag and C-terminal $(G4S)_3$ linker, and *Aa*FaeD, fused with an N-terminal 6× histidine tag and its native C-terminal linker, were amplified using TlXYN-EcoRI_F and TlXYN-G4S-SpeI_R primers, and AaFAE-EcoRI-F and AaFAE-SpeI-R primers, respectively **(Appendix E, Table 4-1)**. The PCR amplicons were then ligated with the *Sc*SED1 anchoring domain (Scheme 1) and cloned into pPICZ*α*A at the EcoRI and NotI restriction sites, resulting in pPICZ*α*A-XYN-SED and pPICZ*α*A-FAE-SED, respectively **(Table 4-1)**. To construct the single-displayed strains, X-Pichia and F-Pichia, pPICZ*α*A-XYN-SED and pPICZ*α*A-FAE-SED were linearized using SacI restriction enzyme and integrated into the *P. pastoris* X-33 genome. The double-displayed strain, X/F-Pichia, was created by co-integrating the linearized pPICZ*α*A-XYN-SED and pPICZ*α*A-FAE-SED into *P. pastoris* X-33.

4.4.3 Immunofluorescence

To visualize the protein localization of XYN and FAE on the *P. pastoris* cell surface, induced *P. pastoris* transformants were collected by centrifugation at 8,000 × *g* for 5 min and washed twice with 1×PBS (0.8% NaCl, 0.2% KCl, 0.144% Na_2HPO_4, and 0.024% KH_2PO_4). Subsequently, cells were incubated with a 100-fold dilution of His-probe antibody (AD1.1.10) conjugated with Alexa Fluor 488 (Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA) and a 100-fold dilution of OctA-Probe antibody (H-5), an IgG1 κ mouse monoclonal anti-FLAG tag antibody, conjugated with Alexa Fluor 594 (Santa Cruz Biotechnology Inc.) in 1×PBS for 1h at room temperature. Following incubation, the cells were washed twice with 1×PBS and mounted on a glass slide using 50% glycerol in 1×PBS. Fluorescence detection was performed using a fluorescence microscope (FluoView™ FV1000, Olympus, Tokyo, Japan) with excitation wavelengths of 488-nm (argon laser line) and 594-nm (helium/neon laser line). The fluorescent emission channels were set using filters with wavelengths of 510-nm and 610-nm, respectively. Cells were imaged using a 100× oil immersion objective lens, and the images were acquired and processed using the cellSens Standard software (Olympus, Tokyo, Japan).

4.4.4 Determination of activities of enzymes displayed on the P. pastoris cell surface towards beechwood xylan and methyl ferulate (MFA)

To assess enzyme activities on the yeast cell surface, induced yeast cells were washed three times with 25 mM potassium phosphate at pH 7.0. Following this, the washed cells were adequately diluted (based on OD_{590} unit) in the washing solution before commencing the assay.

The XYN activity assay was performed using beechwood xylan (Sigma, USA) as the substrate. One mL of reaction mixture contained 1% (w/v) beechwood xylan and 10 OD_{590} units of *P. pastoris* cells in 100 mM potassium phosphate, pH 7.0. The reaction was carried out at 30 °C for 30 min, and enzymatic activity was stopped by heating the mixture at 100 °C for 10 min. The concentration of reducing sugars was measured using the 3,5-dinitrosalicylic acid (DNS; Wako, Japan) method [54]. Briefly, a mixture

containing 200 µL of the sample and 600 µL of DNS was boiled for 10 min and then cooled on ice. The mixture was centrifuged at 10,000 × *g* for 1 min to pellet the residual cells and insoluble substrate. The reducing sugar content in the supernatant was then measured as the absorbance at 540 nm against a xylose standard curve (ranging from 0.15 to 1.5 mg/mL). One unit (U) of enzyme activity was defined as the amount of enzyme that released 1 µmole of reducing sugars per min.

FAE activity was assayed using methyl ferulate (MFA; LKT Laboratories, Inc., USA) following a modified version of my previous protocol [55]. The 250 µL reaction mixture, comprising 0.05 OD_{590} units of *P. pastoris* cells in 80 mM potassium phosphate, pH 7.0, and 0.12 mM MFA, was incubated at 30 °C for 30 min. The reaction was monitored using a microplate reader (Tecan Infinite 200 Pro, Switzerland) at 340 nm with a 2-min interval to obtain the initial reaction rates of hydrolysis. The extinction coefficients of MFA and FA were determined experimentally. One unit of activity was defined as the production of 1 µmole of FA from MFA per min.

To assess the thermostability and pH stability of X-Pichia and F-Pichia, I measured their activity towards beechwood xylan and MFA, respectively. For thermostability assessment, X-Pichia and F-Pichia were dissolved in 50 mM potassium phosphate, pH 6.5, and incubated for 24 h at temperatures ranging from 20 to 80 °C. After incubation, the activity was measured at 30 °C using the method described above. To determine the pH stability, *P. pastoris* cells were dissolved in various solutions with different pH values: 100 mM sodium acetate (pH 4.0–5.0), 100 mM potassium phosphate (pH 6.0–8.0), or 100 mM Tris-HCl (pH 8.0–9.0). The cell suspensions were then incubated at 30 °C for 24 h, and the activities were measured at 30 °C in 100 mM potassium phosphate, pH 7.0.

To assess the tolerance of the displayed enzymes to some metal ions, inhibitors, and detergents, I measured the activity of X-Pichia and F-Pichia at 30 °C in the presence of these additives under 100 mM Tris-HCl and pH 7.0 conditions. The enzyme activity in the presence of an additive was determined by subtracting the rate of the blank reaction, which contained only the additive, from the individual reaction rate. The activity obtained without the presence of any additives was used as a control and defined as 100%.

4.4.5 Preparation of sugarcane trash

Sugarcane trash was obtained from Eastern Sugar and Industries Ltd. (Sa Kaeo, Thailand). Initially, it was ground to a range of approximately 1–8 mm in length and 0.5–2 mm in thickness. Subsequently, the sugarcane trash was meticulously cut and sieved to obtain particles in the size range of 0.5–1 mm. The sugarcane trash was pretreated using the diluted acid pretreatment protocol of Mkabayi et al. (2020) [11]. Briefly, the cut sugarcane trash was suspended in a solution of 0.5% (w/w) sulfuric acid at a solid-to-liquid ratio of 1:10. The suspension was then autoclaved at 121 °C for 20 min. Following acid pretreatment, the sugarcane trash was thoroughly washed with Milli-Q water until reaching neutral pH. The washed sugarcane trash was freeze-dried and stored at –20 °C until further used.

4.4.6 Degradation of natural biomass, sugarcane trash

The degradation of biomass was conducted using *P. pastoris* cells displaying enzymes on their surface. Acid-pretreated sugarcane trash was used as the substrate. To compare the hydrolysis activity of X/F-Pichia with that of X-Pichia, F-Pichia, and the mixture of X-Pichia and F-Pichia, "X-Pichia + F-Pichia", I adjusted the amounts of X-Pichia and F-Pichia to match the XYN and/or FAE activities of X/F-Pichia (refer to **Table 4-2**). The XYN and/or FAE activities of X-Pichia, F-Pichia, and X/F-Pichia were measured in the previous section. A 1 mL reaction mixture was prepared, comprising 1% (w/v) acid-pretreated sugarcane trash and *P. pastoris* cells in 100 mM potassium phosphate buffer with a pH of 7.0. As a control, a 1 mL reaction mixture comprising 1% (w/v) acid-pretreated sugarcane trash and *P. pastoris* X-33 cells was also prepared. The reactions were conducted at 30 °C for 24 and 48 h, respectively.

4.4.7 Determination of reducing sugar and ferulic acid contents obtained on the hydrolysis of natural biomass, sugarcane trash

The total reducing sugar content resulting from enzymatic hydrolysis of acid-pretreated ST was determined using the DNS method as mentioned above.

For the analysis of FA and *p*CA contents, 200 µL of the sample was mixed with 600 µL of 100% acetonitrile. FA and *p*CA contents were then quantified using RP-HPLC (Shimadzu, Japan) equipped with a reversed phase column (4.6 mm × 15 cm, TSKgel ODS-80 TM; Tosoh, Tokyo, Japan) and a UV detector (310 nm, SPD-20A UV/VIS detector; Shimadzu, Japan). The binary mobile phase consisted of (A) Milli-Q water + 0.1% trifluoroacetic acid (TFA) and (B) acetonitrile + 0.1% TFA. The elution profile was as follows: isocratic at 5% B for 0.0–5 min, B linearly increased from 5–70% over 5–30 min, isocratic at 70% B for 30–35 min, and isocratic at 5% B for 35–45 min. The flow rate was 1 mL/min. The FA (MP Biomedicals, USA) and *p*CA (Nacalai Tesque, Inc., Japan) contents were estimated using standard curves. Data were processed using LabSolution software (Shimadzu, Japan).

The reducing sugar and FA contents that were obtained for the control (a 1 mL reaction mixture comprising 1% (w/v) acid-pretreated sugarcane trash and *P. pastoris* X-33 cells) were subtracted from the corresponding contents obtained for reaction mixtures with various combinations of *P. pastoris* described above.

4.4.8 Statistical analysis

All experiments were performed in triplicate, and the values are reported as mean ± standard deviation. Statistical analyses were conducted using Excel Software, employing the *t*-test for significance testing.

4.5 References

[1] Mujtaba, M., Fernandes Fraceto, L., Fazeli, M., Mukherjee, S., Savassa, S.M., Araujo de Medeiros, G., do Espírito Santo Pereira, A., Mancini, S.D., Lipponen, J., Vilaplana, F., Lignocellulosic biomass from agricultural waste to the circular economy: A review with focus on biofuels, biocomposites and bioplastics, *Journal of Cleaner Production*. (2023) 402 136815.

[2] Zoghlami, A., Paës, G., Lignocellulosic biomass: Understanding recalcitrance and predicting hydrolysis, *Frontiers in Chemistry*. (2019) 7 874.

[3] Rao, J., Lv, Z., Chen, G., Peng, F., Hemicellulose: Structure, chemical modification, and application, *Progress in Polymer Science*. (2023) 140 101675.

[4] Chaudhary, R., Kuthiala, T., Singh, G., Rarotra, S., Kaur, A., Arya, S.K., Kumar, P., Current status of xylanase for biofuel production: a review on classification and characterization, *Biomass Conversion and Biorefinery*. (2021) 13 8773–8791.

[5] Dilokpimol, A., Mäkelä, M.R., Aguilar-Pontes, M.V., Benoit-Gelber, I., Hildén, K.S., De Vries, R.P., Diversity of fungal feruloyl esterases: Updated phylogenetic classification, properties, and industrial applications, *Biotechnology for Biofuels*. (2016) 9 231.

[6] Schmitz, E., Leontakianakou, S., Norlander, S., Nordberg Karlsson, E., Adlercreutz, P., Lignocellulose degradation for the bioeconomy: The potential of enzyme synergies between xylanases, ferulic acid esterase and laccase for the production of arabinoxylo-oligosaccharides, *Bioresource Technology*. (2022) 343 126114.

[7] Mathew, S., Abraham, T.E., Ferulic acid: An antioxidant found naturally in plant cell walls and feruloyl esterases involved in its release and their applications, *Critical Reviews in Biotechnology*. (2004) 24 59–83.

[8] Ou, S., Kwok, K.C., Ferulic acid: Pharmaceutical functions, preparation and applications in foods, *Journal of the Science of Food and Agriculture*. (2004) 84 1261–1269.

[9] Antonopoulou, I., Sapountzaki, E., Rova, U., Christakopoulos, P., Ferulic acid from plant biomass: A phytochemical with promising antiviral properties, *Frontiers in Nutrition*. (2022) 8 777576.

[10] Braga, C.M.P., Delabona, P. da S., Lima, D.J. da S., Paixão, D.A.A., Pradella, J.G. da C., Farinas, C.S., Addition of feruloyl esterase and xylanase produced on-site improves sugarcane bagasse hydrolysis, *Bioresource Technology*. (2014) 170 316–324.

[11] Mkabayi, L., Malgas, S., Wilhelmi, B.S., Pletschke, B.I., Evaluating feruloyl esterase-xylanase synergism for hydroxycinnamic acid and xylo-oligosaccharide

production from untreated, hydrothermally pre-treated and dilute-acid pre-treated corn cobs, *Agronomy*. (2020) 10 688.

[12] Mafa, M.S., Malgas, S., Pletschke, B.I., Feruloyl esterase (FAE-1) sourced from a termite hindgut and GH10 xylanases synergy improves degradation of arabinoxylan, *AMB Express*. (2021) 11 21.

[13] Fujita, Y., Takahashi, S., Ueda, M., Tanaka, A., Okada, H., Morikawa, Y., Kawaguchi, T., Arai, M., Fukuda, H., Kondo, A., Direct and efficient production of ethanol from cellulosic material with a yeast strain displaying cellulolytic enzymes, *Applied and Environmental Microbiology*. (2002) 68 5136–5141.

[14] Teymennet-Ramírez, K. V., Martínez-Morales, F., Trejo-Hernández, M.R., Yeast surface display system: Strategies for improvement and biotechnological applications, *Frontiers in Bioengineering and Biotechnology*. (2022) 9 794742.

[15] Inokuma, K., Hasunuma, T., Kondo, A., Efficient yeast cell-surface display of exo- and endo-cellulase using the SED1 anchoring region and its original promoter, *Biotechnology for Biofuels*. (2014) 7 8.

[16] Kajiwara, K., Aoki, W., Ueda, M., Evaluation of the yeast surface display system for screening of functional nanobodies, *AMB Express*. (2020) 10 51.

[17] Phienluphon, A., Mhuantong, W., Boonyapakron, K., Deenarn, P., Champreda, V., Wichadakul, D., Suwannarangsee, S., Identification and evaluation of novel anchoring proteins for cell surface display on *Saccharomyces cerevisiae*, *Applied Microbiology and Biotechnology*. (2019) 103 3085–3097.

[18] Inokuma, K., Kitada, Y., Bamba, T., Kobayashi, Y., Yukawa, T., Den Haan, R., Van Zyl, W.H., Kondo, A., Hasunuma, T., Improving the functionality of surface-engineered yeast cells by altering the cell wall morphology of the host strain, *Applied Microbiology and Biotechnology*. (2021) 105 5895–5904.

[19] Arnthong, J., Ponjarat, J., Bussadee, P., Deenarn, P., Prommana, P., Phienluphon, A., Charoensri, S., Champreda, V., Zhao, X.Q., Suwannarangsee, S., Enhanced surface display efficiency of *β*-glucosidase in *Saccharomyces cerevisiae* by

disruption of cell wall protein-encoding genes *YGP1* and *CWP2*, *Biochemical Engineering Journal*. (2022) 179 108305.

[20] Tang, H., Song, M., He, Y., Wang, J., Wang, S., Shen, Y., Hou, J., Bao, X., Engineering vesicle trafficking improves the extracellular activity and surface display efficiency of cellulases in *Saccharomyces cerevisiae*, *Biotechnology for Biofuels*. (2017) 10 53.

[21] Arnthong, J., Bussadee, P., Phienluphon, A., Deenarn, P., Tulsook, K., Plupjeen, S.N., Siamphan, C., Tachaapaikoon, C., Champreda, V., Suwannarangsee, S., Overexpression of *LAS21* in cellulase-displaying *Saccharomyces cerevisiae* for high-yield ethanol production from pretreated sugarcane bagasse, *Fermentation*. (2022) 8 652.

[22] Zhang, C., Chen, H., Zhu, Y., Zhang, Y., Li, X., Wang, F., *Saccharomyces cerevisiae* cell surface display technology: Strategies for improvement and applications, *Frontiers in Bioengineering and Biotechnology*. (2022) 10 1056804.

[23] Shiraga, S., Kawakami, M., Ishiguro, M., Ueda, M., Enhanced reactivity of *Rhizopus oryzae* lipase displayed on yeast cell surfaces in organic solvents: Potential as a whole-cell biocatalyst in organic solvents, *Applied and Environmental Microbiology*. (2005) 71 4335–4338.

[24] Li, X., Jin, X., Lu, X., Chu, F., Shen, J., Ma, Y., Liu, M., Zhu, J., Construction and characterization of a thermostable whole-cell chitinolytic enzyme using yeast surface display, *World Journal of Microbiology and Biotechnology*. (2014) 30 2577–2585.

[25] Fan, L.H., Zhang, Z.J., Yu, X.Y., Xue, Y.X., Tan, T.W., Self-surface assembly of cellulosomes with two miniscaffoldins on *Saccharomyces cerevisiae* for cellulosic ethanol production, *Proceedings of the National Academy of Sciences of the United States of America*. (2012) 109 13260–13265.

[26] Liu, Z., Ho, S.H., Sasaki, K., Den Haan, R., Inokuma, K., Ogino, C., Van Zyl, W.H., Hasunuma, T., Kondo, A., Engineering of a novel cellulose-adherent cellulolytic *Saccharomyces cerevisiae* for cellulosic biofuel production, *Scientific Reports*. (2016) 6 24550.

[27] Yuzbasheva, E.Y., Yuzbashev, T. V., Laptev, I.A., Konstantinova, T.K., Sineoky, S.P., Efficient cell surface display of Lip2 lipase using C-domains of glycosylphosphatidylinositol-anchored cell wall proteins of *Yarrowia lipolytica*, *Applied Microbiology and Biotechnology*. (2011) 91 645–654.

[28] Liu, Z., Inokuma, K., Ho, S.H., Haan, R. Den, Hasunuma, T., Van Zyl, W.H., Kondo, A., Combined cell-surface display- and secretion-based strategies for production of cellulosic ethanol with *Saccharomyces cerevisiae*, *Biotechnology for Biofuels*. (2015) 8 162.

[29] Yang, S., Lv, X., Wang, X., Wang, J., Wang, R., Wang, T., Cell-surface displayed expression of trehalose synthase from *Pseudomonas putida* ATCC 47054 in *Pichia pastoris* using Pir1P as an anchor protein, *Frontiers in Microbiology*. (2017) 8 2583.

[30] Matano, Y., Hasunuma, T., Kondo, A., Display of cellulases on the cell surface of *Saccharomyces cerevisiae* for high yield ethanol production from high-solid lignocellulosic biomass, *Bioresource Technology*. (2012) 108 128–133.

[31] Bae, J., Kuroda, K., Ueda, M., Proximity effect among cellulose-degrading enzymes displayed on the *Saccharomyces cerevisiae* cell surface, *Applied and Environmental Microbiology*. (2015) 81 59–66.

[32] Smith, M.R., Gao, H., Prabhu, P., Bugada, L.F., Roth, C., Mutukuri, D., Yee, C.M., Lee, L., Ziff, R.M., Lee, J.K., Wen, F., Elucidating structure–performance relationships in whole-cell cooperative enzyme catalysis, *Nature Catalysis*. (2019) 2 809–819.

[33] Fan, S., Liang, B., Xiao, X., Bai, L., Tang, X., Lojou, E., Cosnier, S., Liu, A., Controllable display of sequential enzymes on yeast surface with enhanced biocatalytic activity toward efficient enzymatic biofuel cells, *Journal of the American Chemical Society*. (2020) 142 3222–3230.

[34] Guo, F., Liu, M., Liu, H., Li, C., Feng, X., Direct yeast surface codisplay of sequential enzymes with complementary anchor motifs: enabling enhanced glycosylation of natural products, *ACS Synthetic Biology*. (2023) 12 460–470.

[35] Fujita, Y., Ito, J., Ueda, M., Fukuda, H., Kondo, A., Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain codisplaying three types of cellulolytic enzyme, *Applied and Environmental Microbiology*. (2004) 70 1207–1212.

[36] Liu, Z., Inokuma, K., Ho, S.-H., Den Haan, R., Van Zyl, W.H., Hasunuma, T., Kondo, A., Improvement of ethanol production from crystalline cellulose via optimizing cellulase ratios in cellulolytic *Saccharomyces cerevisiae*, *Biotechnol. Bioeng*. (2017) 114 1201–1207.

[37] Teo, K.S.K., Kondo, K., Saito, K., Iseki, Y., Watanabe, T., Nagata, T., Katahira, M., Enhanced depolymerization of beech wood lignin and its removal with peroxidases through continuous separation of lignin fragments, *Green Chemistry*. (2023) 25 7682–7695.

[38] Wang, J.K., He, B., Du, W., Luo, Y., Yu, Z., Liu, J.X., Yeast with surface displayed xylanase as a new dual purpose delivery vehicle of xylanase and yeast, *Animal Feed Science and Technology*. (2015) 208 44–52.

[39] Liu, C., Zhang, W., Li, Y., Pan, K., OuYang, K., Song, X., Xiong, X., Zang, Y., Wang, L., Qu, M., Zhao, X., Characterization of yeast cell surface displayed *Lentinula edodes* xylanase and its effects on the hydrolysis of wheat, *International Journal of Biological Macromolecules*. (2022) 199 341–347.

[40] Dong, C., Wang, X., Sun, W., Chen, L., Li, S., Wu, K., Wu, K., Ma, L., Ma, L., Liu, Y., Liu, Y., Engineering *Pichia pastoris* with surface-display minicellulosomes for carboxymethyl cellulose hydrolysis and ethanol production, *Biotechnology for Biofuels*. (2020) 13 108.

[41] Siripong, W., Wolf, P., Kusumoputri, T.P., Downes, J.J., Kocharin, K., Tanapongpipat, S., Runguphan, W., Metabolic engineering of *Pichia pastoris* for production of isobutanol and isobutyl acetate, *Biotechnology for Biofuels*. (2018) 11 1.

[42] Liu, Y., Huang, L., Zheng, D., Fu, Y., Shan, M., Xu, Z., Ma, J., Lu, F., Development of a *Pichia pastoris* whole-cell biocatalyst with overexpression of mutant lipase I

PCL^{G47I} from *Penicillium cyclopium* for biodiesel production, *RSC Advances*. (2018) 8 26161–26168.

[43] Shimoi, H., Kitagaki, H., Ohmori, H., Iimura, Y., Sed1p Is a major cell wall protein of *Saccharomyces cerevisiae* in the stationary phase and is involved in lytic enzyme resistance, *Journal of Bacteriology*. (1998) 180 3381–3387.

[44] Su, G., Zhang, X., Lin, Y., Surface display of active lipase in *Pichia pastoris* using Sed1 as an anchor protein, *Biotechnology Letters*. (2010) 32 1131–1136.

[45] Li, W., Shi, H., Ding, H., Wang, L., Zhang, Y., Li, X., Wang, F., Cell surface display and characterization of *Rhizopus oryzae* lipase in *Pichia pastoris* using Sed1p as an anchor protein, *Current Microbiology*. (2015) 71 150–155.

[46] Yang, J., Huang, K., Xu, X., Miao, Y., Lin, Y., Han, S., Cell surface display of *Thermomyces lanuginosus* lipase in *Pichia pastoris*, *Frontiers in Bioengineering and Biotechnology*. (2020) 8 544058.

[47] Khattab, S.M.R., Okano, H., Kimura, C., Fujita, T., Watanabe, T., Efficient integrated production of bioethanol and antiviral glycerolysis lignin from sugarcane trash, *Biotechnology for Biofuels and Bioproducts*. (2023) 16 82.

[48] Yang, S., Shen, J., Deng, J., Li, H., Zhao, J., Tang, H., Bao, X., Engineering cell polarization improves protein production in *Saccharomyces cerevisiae*, *Microorganisms*. (2022) 10 2005.

[49] Singh, S., Madlala, A.M., Prior, B.A., *Thermomyces lanuginosus*: Properties of strains and their hemicellulases, *FEMS Microbiology Reviews*. (2003) 27 3–16.

[50] Adam, M., Gomes, I., Mohiuddint, G., Hoq, M.M., Production and characterization of thermostable xylanases by *Thermomyces lanuginosus* and *Thermoascus aurantiacus* grown on lignocelluloses, *Enzyme and Microbial Technology*. (1994) 16 298–302.

[51] Gaffney, M., Carberry, S., Doyle, S., Murphy, R., Purification and characterisation of a xylanase from *Thermomyces lanuginosus* and its functional expression by *Pichia pastoris*, *Enzyme and Microbial Technology*. (2009) 45 348–354.

[52] Hou, M., Liang, C., Fei, Y., Yang, D., Zhang, N., Lu, Y., Wang, L., Xing, Z., Zhao, Z., Analysis of the effect of metal ions on the ability of xylanase to hydrolyze wheat bran by molecular dynamics simulations, *Frontiers in Bioengineering and Biotechnology*. (2023) 11 1142873.

[53] Zhang, Y., Feng, Z., Xiang, H., Zhang, X., Yang, L., Characterization of feruloyl esterase from *Klebsiella oxytoca* Z28 and its application in the release of ferulic acid from de-starching wheat bran, *Microorganisms*. (2023) 11 989.

[54] Miller, G.L., Use of dinitrosalicylic acid reagent for determination of reducing Sugar, *Analytical Chemistry*. (1959) 31 426–428.

[55] Phienluphon, A., Kondo, K., Mikami, B., Nagata, T., Katahira, M., Structural insights into the molecular mechanisms of substrate recognition and hydrolysis by feruloyl esterase from *Aspergillus sydowii*, *International Journal of Biological Macromolecules*. (2023) 253 127188.

Chapter Five

General conclusions

In conclusion, the collective findings presented in these research studies significantly advance my understanding of the molecular mechanisms and applications of FAEs in lignocellulosic biomass degradation. FAEs play a crucial role in depolymerizing lignin-polysaccharide ester bonds, a key step in biomass utilization for various industrial applications, including biofuel and biochemical production. These studies explore FAEs from different subfamilies and organisms, shedding light on their structural characteristics, substrate recognition mechanisms, catalytic activities, and potential engineering strategies.

In **chapter two**, I focus on *Aspergillus sydowii* FaeE (*As*FaeE) , a member of FAE SF6, providing structural insights into its substrate recognition and hydrolysis mechanisms. Through X-ray crystallography and kinetic analysis, I identify a catalytic triad comprising Ser119, Asp203, and His259 in *As*FaeE, a characteristic feature of *α/β*-hydrolase enzymes. Notably, the study reveals structural insights by examining the complex structures of *As*FaeE associated with FA and SA. These structural analyses uncover a significant conformational change occurring in the loop that covers the substrate-binding pocket. Particularly, an interesting feature of the three layers consisting of planar moieties of the substrates, Pro158 and Phe159, is identified to play a unique role in accommodating the substrates. An intriguing finding is the impact of functional mutagenesis, particularly the substitution of Leu150 with a smaller amino acid, which leads to an altered substrate preference from MFA to MSA. This study emphasizes the distinct substrate-binding mechanisms employed by *As*FaeE in comparison to other FAEs, thus providing valuable insights for the potential engineering of FAE variants capable of more effectively processing a broader range of substrates.

In **chapter three**, I characterize the structure of *Acremonium alcalophilum* FaeD (*Aa*FaeD) , a member of FAE SF5 known for its ability to target larger substrates. Determining the crystal structure of *Aa*FaeD catalytic domain in complex with FA reveals a hydrophobic cleft site formed by two facing hydrophobic walls that accommodate the

substrate. Structure-based mutagenesis experiments highlight the functional roles of key residues in catalytic activity. Surprisingly, the study identifies an F120Y mutant with enhanced catalytic activity, and comparisons with other subfamilies (SF2 and SF3) reveal differences in substrate-binding site accessibility, explaining ability of SF5 FAE to hydrolyze both monomeric and dimeric phenolic substrates.

In **chapter four**, I apply yeast surface display technology to engineer a whole-cell biocatalyst for lignocellulosic biomass degradation. Through genetic engineering, I construct a *Pichia pastoris* strain capable of co-displaying XYN and FAE on the cell surface. This co-display generates a synergistic effect, significantly enhancing the breakdown of acid-pretreated sugarcane trash and notably increasing the reducing sugar yields. Closer proximity between XYN and FAE on the cell surface produces the highest results, suggesting promising implications for sustainable and efficient bioconversion processes for lignocellulosic biomass.

In conclusion, these studies collectively provide comprehensive insights into the structural, functional, and application aspects of FAEs in lignocellulosic biomass degradation. The findings not only enhance my understanding of the molecular mechanisms governing FAEs but also offer valuable strategies for designing more efficient biocatalysts for biomass conversion, with potential implications for sustainable biofuel and biochemical production. These insights represent a significant contribution to the field and pave the way for further research and development in lignocellulosic biomass utilization.

Appendix A

Amino acid sequences of fungal FAEs and related proteins used for phylogenetic analysis in Chapter 2

>4322286_*Aspergillus_terreus*

RQDAFEAKCHSFANKIHLPNVHVNFASYVPGGTNLTLADNPSSCGATSQSVSADVCRV
AMAVATSNSSEITLEAWFPRNYTGRFLSTGNGGLSGCIQYYDMAYTTGFGFATVGANN
GHNGTSGEPFYHHPEVLEDFAYRSIHTGVVIGKKLTKMFYEEGFNKSYYLGCSTGGRQ
GFKSVQKYPNDFDGVVAGAPAFNFANLISWSAHFYPITGPPGSDTYLSPAMWKVAHDE
IIRQCDQIDGAKDGIIEDPSLCNPIMETIICKPGASSDNCLSAAQAKTVREVLYPLYGVNG
TLLYPRMQPGSEVLAAPIMYNGQPFAYSTDWYRYVVYNDPNWNGTTFDVQDAAAAL
AQNPYNIQTWDADLTPFRKSGGKVLTYHGLQDQLISSENSKLYYARVAETMGMPPEEL
DEFYRFFQISGMGHCGGGDGAYGIGNGLATYSGKDPENNVLMAMVQWVEKGIAPETV
RGAKFANGPGSTVEYSRKHCRYPRRNVFKGPGNYTDENAWECVV

>AAHF01000006_*Neosartorya_fumigata*

KQDAFQAKCASFGRKIKLPNVHVNFVEYVPGGTNLTLPDNDVTCGASSQVVSADMCR
VAMAVDTSKSSQITLEAWFPREYTGRFLSTGNGGLSGCIQYYDLAYTAGLGFATVGAN
NGHNGTSGKPFYQHPEVIEDFAYRSIHTGVVVGKQLTKMFYKEGFDKSYYLGCSTGGR
QGFKSIQKYPNDFDGVVAGAPAFNFVNLISWSIHFYSITGSNTSDTYLSPESWKVVHDEI
VRQCDEIDGAKDGIIEDTDLCQPVIETIICKPGASDKTNCITGAQAKTVRNVLSPFYGVN
GNLLYPRMQPGSELFASSVVYNGQPFRYSTDWYRYVVYNNPDWDATKWTVEDAAVA
LAQNPYNIQTWDADISSFQKAGGKVLTYHGMQDQLISSDNSKLYYARVAEEMGLGPE
ELDDFYRFFPVSGMAHCTGGDGAYGIGNGLRTYNGAEPENNVLMAMVQWVEKGIAP
EFIRGAKFSNGVGSSVEYTRKHCRYPRRNVYKGPGNYSDENAWECV

>AAHF01000018_*Neosartorya_fumigata*

ASSASFESRCQHFHKEIHLQNVHVLSTTYVPIGSNIPMVYNPPICGGTASSSISTIQFCQV
ALNVTTSDKSQFFMEAWLPSNYTGRFLSTGNGGLNGCVSYADMVYATQYGFATIGTN
NGHFGDTGQYFLNNPEVIEDFAYRALHTGTVVGKALTKLFYPQGYKNSYYLGCSTGG
RQGWKSIQRFPDDFDGVVAGAPAINFVNLCSWGSRFLKITGPPGSETFVTSAQWSAVH
NEILRQCDALDGAVDGIIEDTDLCQPVFETLLCNSTAVDKTSCLTGVQANTVNEVFSAM
YGLDGKWLYPRMQPGSELAASFIYYSGNGFKYSDDWYKYVVYNDSNWGHSTWTLAD
AAAAAAQDPFQISSFDGNISGFQKAGGKVLHYHGLEDAIITSDSSKAYYKHVADTMGL
SPSELDHFYRLFPISGMGHCSPGTGAASIGQGSSTYAGDDPQDNVLMAIVQWVEKGIAP
EYVRGSKMSRDGTIDYRRKHCKYPKRNRYVGPGKYTDENAWKCV

>AN1772_*Aspergillus_nidulans*

AVLVQDQFQTRCENFAGKIDLPNVKVNFASYIPGSTNLTLDNVPTCDQSQVVSSDICRV
AMAVTTSNASEITLEAWFPRDYTGRFLSTGNGGLGGCIQYSDLDYASRLGFATVGANN
GHNGTSGEPFYKAPEVLEDFVYRSVHTGIVVGKQLTKLFYDEGFDTSYYLGCSTGGRQ
GFKLAQDFPGEVDGIIAGAPAINFVGLLSWSAHFYPITGPVGSATYLSLDDWDLVHEEIL
RQCDGLDGAEDGIIEDPDLCHPNATTLLCSPGATSGSCLTATQVNTVHEVYAPLLSSNS

TLIYPRMQPGGEQFAAPAMYNGQPFQYSKDWWRYVVYSDPTWNATKWTIRDAEAAL
RQNPYNIQTWNADLSPLRDSGSKLLTYHGLQDQLISSDDSKLYYHRLMKTMGVTSNQL
DEFYRFFQISGMAHCQDGDGAYGIGNRAETEFSTEPEDNVLMAMVRWVEEGIAPETVR
GAKFSDGVGSEVEYYRKHCRYPRRNVYKGPGDYTDETAWECV

>*An*FaeB_*Aspergillus_niger*

ATDPFQSRCNEFQNKIDIANVTVRSVAYVAAGQNISQAEVASVCKASVQASVDLCRVT
MNISTSDRSHLWAEAWLPRNYTGRFVSTGNGGLAGCVQETDLNFAANFGFATVGTNG
GHDGDTAKYFLNNSEVLADFAYRSVHEGTVVGKQLTQLFYDEGYNYSYYLGCSTGGR
QGYQQVQRFPDDYDGVIAGSAAMNFINLISWGAFLWKATGLADDPDFISANLWSVIHQ
EIVRQCDLVDGALDGIIEDPDFCAPVIERLICDGTTNGTSCITGAQAAKVNRALSDFYGP
DGTVYYPRLNYGGEADSASLYFTGSMYSRTEEWYKYVVYNDTNWNSSQWTLESAKL
ALEQNPFNIQAFDPNITAFRDRGGKLLSYHGTQDPIISSTDSKLYYRRVANALNAAPSEL
DEFYRFFQISGMGHCGDGTGASYIGQGYGTYTSKAPQVNLLRTMVDWVENGKAPEYM
PGNKLNANGSIEYMRKHCRYPKHNIHTGPGNYTDPNSWTCV

>*Ao*FaeB_*Aspergillus_oryzae*

AAIDSTSSSNGSDHHGSSFQAECESFKAKINVTNANVHSVTYVPAGVNISMADNPSICG
GDEDPITSTFAFCRIALNVTTSSKSQIFMEAWLPSNYSGRFLSTGNGGLGGCVKYDDMA
YAAGYGFATVGTNNGHFGNNGVSFYQNTEVVEDFAYRALHTGVVVGKELTKNFYPQ
GYNKSYYLGCSTGGRQGWKSVQTFPDDFDGVVAGAPAFNFINLTSWGARFLTLTGDSS
AETFVTETQWTAVHNEIIRQCDSLDGAKDGIIEDPDLCQPIIEALLCNATQSSTSGTCLTG
AQVKTVNGVFSATYGLNGSFLYPRMQPGSELAAYSSYYSGTPFAYAEDWYRYVVFNN
TNWDVATWTVQDAAIANAQDPYQISTWNGDLSPFQKKGGKVLHYHGMEDAIISSESS
KVYYKHVADTMNLSPSELDSFYRFFPISGMAHCANADGPSAIGQGTGTFAGNNPQDNV
LLAMVQWVEEGVAPDFVRGAKLNGSTVEYRRKHCKYPKRNRYVGPGSYTDENAWEC
V

>*Ao*FaeC_Aspergillus_oryzae

SQDTFQGKCTGFADKINLPNVRVNFVNYVPGGTNLSLPDNPTSCGTTSQVVSEDVCRIA
MAVATSNSSEITLEAWLPQNYTGRFLSTGNGGLSGCIQYYDLAYTSGLGFATVGANSG
HNGTSGEPFYHHPEVLEDFVHRSVHTGVVVGKQLTKLFYEEGFKKSYYLGCSTGGRQG
FKSVQKYPNDFDGVVAGAPAFNMINLMSWSAHFYSITGPVGSDTYLSPDLWNITHKEIL
RQCDGIDGAEDGIIEDPSLCSPVLEAIICKPGQNTTECLTGKQAHTVREIFSPLYGVNGTL
LYPRMQPGSEVMASSIMYNGQPFQYSADWYRYVVYENPNWDATKFSVRDAAVALKQ
NPFNLQTWDADISSFRKAGGKVLTYHGLMDQLISSENSKLYYARVAETMNVPPEELDE
FYRFFQISGMAHCSGGDGAYGIGNQLVTYNDANPENNVLMAMVQWVEKGIAPETIRG
AKFTNGTGSAVEYTRKHCRYPRRNVYKGPGNYTDENAWQCV

>CH476596_*Aspergillus_terreus*

SQSFEERCTDFRDSINSLPNVQATIVEYVAGSQNVSLPDNDPSCNQSSQFVTADICRAAM
VVKTSNSSQIVMEAWFPRNYTGRFLATGNGGFGGCIRYPELDYTTRLGFAAVATNNGH
NGTSAEAFLNSPEVLRDFADRSIHTAAKVGKELTKRFYAEGFRKSYYLGCSTGGRQGF
KSVQQYPHDFDGVVAGAPAVHEVNLISWAGHIYEITGNKSEETYLPPALWNIVHSEVM

RQCDGLDGAQDNLIEDPDLCHPTFENIMCPSDNKSNNGSLSCITEAQANTVIQLMSPYY
NTDGSMLFPGMQPGSETVSSALLYTGVPTPYAKEWFRYVVYNDTNWDPTTFNIKDAQ
AALKQNPFNIQTWEGDLSRFQNAGGKIITYHGMQDFLVSSFNSREYYKHVHETMGLAP
DQLDEFYRYFRISGMAHCYYGDGASYIGGSAPSAYSDDPEDNVLMAMVEWVEKGIAP
EFIRGTKLDQDGHPQYTRKHCRYPRRNVYRGPGSYLDENAWECVL

>DS027060_*Aspergillus_clavatus*

KKDTFQTKCAALQHKVKLPNVHVNFVEYVPGGTNLDLPDNAPSCGASSQAVSTDMCR
IAMAVDTSDSSQITLEAWFPRDYTGRFLSTGNGGLSGCIQYYDLAYAAGLGFATVGAN
NGHNGTSGEPFYQHPEVVEDFAHRSVHTGVVVGKQLTKLFYDKGFKKSYYLGCSTGG
RQGFKSVQKYPKDFDGIVAGAPAFNFVNLISWSAYFYSLTGSNTSESYLSPAMWKIAH
DEIVRQCDELDGAKDGIIEDTDLCHPRLETIICKPGAKDTANCLTGAQAKTVRDVLSPM
YGVNGTLLYPRMQPGSEVYAAGIMYNGEPFQYSTDWYRYVVYNNPDWDDTKWSVE
DAAAALAQNPYDIQTFDADISSFRGAGGKVLTYHGLQDQMISSDNSKLYYARVAETM
KLPPSELDEFYRFFPVSGMTHCAGGDGAYGIGNGLGSYNGVDPENNVLMAMVQWVE
KGIAPEFIRGAKFAEGPGSAVEYTRKHCRYPRRNVYKGPGNYTDENAWECV

>DS027698_*Neosartorya_fischeri*_1

SFESRCQHFHKEIHLQNVHVHSTTYVPIGSNISMAYNPPICGGTSSSSISTIEFCQVALNVT
TSDKSQFFMEAWLPSNYTGRFLSTGNGGLNGCVGYGDMIYASQYGFATIGTNNGHFG
DTGQYFLNNPEVIEDFAYRALHTGTVVGKALTKLFYPQGYKNSYYLGCSTGGRQGWK
SIQRFPDDFDGVVAGAPAFNFVNLCNWGSRFLKITGPPDSDTFVTSAQWSIIHNEIIRQC
DALDGAVDGTIEDTDLCQPIFETLICNSTAVNKTSCLTGVQANTVNEVFSAMYGLDGK
WLYPRMQPGSELAASFIYYSGNGFKYSDDWFKYVVYNDSNWDHSTWTLADAAAADA
QDPFQISTFDGDISGFQKAGGKVLHYHGLEDAIITSDSSKAYYKHVADTMGLSPSDLDQ
FYRFFPISGMGHCSPGTGAASIGQGSSTYAGDDPQDNVLMAMVQWVEKGIAPEYVRGS
KKSIDGQTEYRRKHCKYPKRNRYVGPGKYTDENAWKCV

>DS027698_*Neosartorya_fischeri*_2

KQDAFQAKCASFGHRIKLPNVHVNFVEYVPGGTNLTLPDNHVTCGASSQIVSADMCRV
AMAVDTSKSSQITLEAWFPRNYTGRFLSTGNGGLSGCIQYYDLAYTAGLGFATVGANN
GHNGTSGKPFYQHPEVIEDFAYRSIHTGVVVGKQLIKMFYSEGFDKSYYLGCSTGGRQ
GFKSIQKYPNDFDGVVAGAPAFNFVNLISWSIHFYSITGSNTSDTYLSPASWKVVHDEIV
RQCDGIDGAKDGIIEDTDLCHPILETIICKPGASSTTNCITGTQAKTVRNVLSPFYGVNGT
LLYPRMQPGSELFASSIMYNGQPFSYSTDWYRYVVYNNPNWDATKWTVEDAAVALA
QNPYNIQTWDADISSFQKAGGKVLTYHGIQDQLISSDNSKLYYARVAETMGLGPEELD
DFYRFFPVSGMAHCSGGDGAYGIGNGLRTYNGAEPENNVLMAMVQWVEKGVAPEFIR
GAKFSNGVGSPVEYTRKHCKYPRRNVYKGPGNYSDENAWECV

>DS499599_*Neosartorya_fumigata*

KQDAFQAKCASFGRKIKLPNVHVNFVEYVPGGTNLTLPDNDVTCGASSQVVSADMCR
VAMAVDTSKSSQITLEAWFPREYTGRFLSTGNGGLSGCIQYYDLAYTAGLGFATVGAN
NGHNGTSGKPFYQHPEVIEDFAYRSIHTGVVVGKQLTKMFYKEGFDKSYYLGCSTGGR
QGFKSIQKYPNDFDGVVAGAPAFNFVNLISWSIHFYSITGSNTSDTYLSPESWKVVHDEI

VRQCDEIDGAKDGIIEDTDLCQPVIETIICKPGASDKTNCITGAQAKTVRNVLSPFYGVN
GTLLYPRMQPGSELFASSVVYNGQPFRYSTDWYRYVVYNNPDWDATKWTVEDAAVA
LAQNPYNIQTWDADISSFQKAGGKVLTYHGMQDQLISSDNSKLYYARVAEEMGLGPE
ELDDFYRFFPVSGMAHCTGGDGAYGIGNGLRTYNGAEPENNVLMAMVQWVEKGIAP
EFIRGAKFSNGVGSSVEYTRKHCRYPRRNVYKGPGNYSDENAWECV

>DS499602_*Neosartorya_fumigata*

ASSASFESRCQHFHKEIHLQNVHVLSTTYVPIGSNIPMVYNPPICGGTASSSISTIQFCQV
ALNVTTSDKSQFFMEAWLPSNYTGRFLSTGNGGLNGCVSYADMVYATQYGFATIGTN
NGHFGDTGQYFLNNPEVIEDFAYRALHTGTVVGKALTKLFYPQGYKNSYYLGCSTGG
RQGWKSIQRFPDDFDGVVAGAPAINFVNLCSWGSRFLKITGPPGSETFVTSAQWSAVH
NEILRQCDALDGAVDGIIEDTDLCQPVFETLLCNSTAVDKTSCLTGVQANTVNEVFSAM
YGLDGKWLYPRMQPGSELAASFIYYSGNGFKYSDDWYKYVVYNDSNWDHSTWTLAD
AAAAAAQDPFQISSFDGNISGFQKAGGKVLHYHGLEDAIITSDSSKAYYKHVADTMGL
SPSELDHFYRLFPISGMGHCSPGTGAASIGQGSSTYAGDDPQDNVLMAIVQWVEKGIAP
EYVRGSKMSRDGTIDYRRKHCKYPKRNRYVGPGKYTDENAWKCV

>EQ963481_*Aspergillus_flavus*

SQDTFQGKCTGFADKINLPDVRVNFVNYVPGGTNLSLPDNPTSCGTTSQVVSEDVCRIA
MAVATSNSSEITLEAWLPQNYTGRFLSTGNGGLSGCIQYYDLAYTSGLGFATVGANSG
HNGTSGEPFYHHPEVLEDFVHRSVHTGVVVGKQLTKLFYEEGFKKSYYLGCSTGGRQG
FKSVQKYPNDFDGVVAGAPAFNMINLMSWSAHFYSITGPVGSDTYLSPDLWNITHKEIL
RQCDGIDGAEDGIIEDPSLCSPVLEAIICKPGQNTTECLTGKQAHTVREIFSPLYGVNGTL
LYPRMQPGSEVMASSIMYNGQPFQYSADWYRYVVYENPNWDATKFSVRDAAVALKQ
NPFNLQTWDADISSFRKAGGKVLTYHGLMDQLISSENSKLYYARVAETMNVPPEELDE
FYRFFQISGMAHCSGGDGAYGIGNQLVTYNDANPENNVLMAMVQWVEKGIAPETIRG
AKFTNGTGSAVEYTRKHCRYPRRNVYKGPGNYTDENAWQCV

>jgi_Aspni7_1085850_e_gw1.303.126.1

ATDPFQSRCNEFQNKIDIANVTVRSVAYVAAGQNISQAEVASVCKASVQASVDLCRVT
MNISTSDRSHLWAEAWLPRNYTGRFVSTGNGGLAGCVQETDLNFAANFGFATVGTNG
GHDGDTAKYFLNNSEVLADFAYRSVHEGTVVGKQLTQLFYDEGYNYSYYLGCSTGGR
QGYQQVQRFPDDYDGVIAGSAAMNFINLISWGAFLWKATGLADDPDFISANLWSVIHQ
EIVRQCDLVDGALDGIIEDPDFCAPVIERLICDGTTNGTSCITGAQAAKVNRALSDFYGP
DGTVYYPRLNYGGEADSASLYFTGSMYSRTEEWYKYVVYNDTNWNSSQWTLESAKL
ALEQNPFNIQAFDPNITAFRDRGGKLLSYHGTQDPIISSTDSKLYYRRVANALNAAPSEL
DEFYRFFQISGMGHCGDGTGASYIGQGYGTYTSKAPQVNLLRTMVDWVENGKAPEYM
PGNKLNANGSIEYMRKHCRYPKHNIHTGPGNYTDPNSWTCV

>*Pc*FAE_I_*Penicillium_chrysogenum*

SQDAFQSKCVEFGAQIDIPNVKVNFAEFVQGGTNLSLADNPPSCGRSNQAVPVDLCRV
AMAVSTSNSSEITLEAWFPREYKGRFLSTGNGGISGCIQYYDLAYTAQLGFATVGANN
GHNGTSGKPFYRQPEVIEDYAYRSVHTGVVIGKELTKQFYEEGFDKSYYLGCSTGGRQ
GWKSVQKYPNDFDGVVAGAPAIGLVNLFSWSAHFYPITGSPTSDTFLSPAEWKIVHEEII

RQCDAIDGAEDGIIEDTDLCHPVLETLTCDPSASSKTSCLTSAQVNTAQQVLSPFYGING
TLLYPRMQPGSEILAAPIMYSGAPFQYSQDWYRYVVYNNPAWSGANFTVKDAAVALR
QNPYNIQTWDGDISSFKNGGGKILHYHGLQDQLISSDNSKLYYSHVSKTMKLPPNKLDE
FYRFFPISGMAHCGDGDGAYGIGQGASTYAGTDPEDNVLMAMVRWVEEGKAPETVR
GTKFSDGPGSEVEYKRRHCRWPRRNVFKGPGNYTDENAWQCV

>*Ts*FaeC_*Talaromyces_stipitatus*

DDSSRENFSNRCDQLAKEIHIPNVTVNFVEYVANGTNVTLADNPPSCGQSNQVVLADL
CRVAMEVTTSNQSQITLEAWFPENYTGRFLSTGNGGLAGCIQYVDMAYASSMGFATV
GANGGHNGTSGESFYHNPDIVEDLSWRSVHTGVVVGKELTKKFYHEGFHKSYYLGCS
TGGRQGFKAVQEFVHDFDGVVAGCPAFNFVNLNSWSGHFYPITGNSSADTFLTTAQW
TLVQQSVMEQCDSLDGAVDGVIEAIDQCHPVFEQLICRPGQNASECLTGKQVNTAQLV
LSPIYGTKGEFLYPRMQPGVENVDMYITYNGDPFAYSTDWYKYVVFSDPNWDPATLN
AQDYEIALAQNPSNIQTFEGDLSAFRDAGAKVLTYHGTADPIITGETSKVYYRHVAETM
NAAPEELDEFYRYFRIGGMSHCGGGTGATAIGNVLSAQWSNDPDANVLMAMVRWVE
EGVAPEYIRGASLGSGPGAKVEYTRRHCKYPTRNVYVGPGNWTDENAWKCIL

>EAT81026_*Phaeosphaeria_nodorum*

MTACDQGLEQPGLNVFTILFKMKFEQSLWTSLLGATNVFAAPAEPTYADCPDTSAGGD
FNSKCAAIASTLAIEHGTVHFSEFVAAGTNLSLSDNDPTCESSVVVLADICRIALSVATS
NSSGLNMEAWLPSNWTGRFLSGGNGGLNGCVDYASLAYTSSLGFSSVGTNNGHNGTS
GLPFFNKPEVLEDFAYRAMHTGVVTGKQITEDFYGKPHDKSYYLGCSTGGRQGFKSAQ
DFPDDFDGIISGAPALAFNNLTSWSGHFYTITGPANSSTFVTMPQWVMVHEDVLKQCD
ELDGYADGIIEDPRLCEYDPAGLTCAEGSTNSSSCLTPTQVETVRAVFAPLLNAQGDLV
YPRFQPGPEVIASAILFNGQPFPYTTDWFRYAIYNDPNWDPSTLNTTDYDNAARINPYNI
QTWKGDLSGIQNRGSKVIHWHGAADMIISYDNSPRYYEHVSSTMGLSPSELDDFYRFFT
VSGTGHCSGGSGAHAIGQASDELKSLDPKENILMALVDWVENGNAPESLVGTKFVNDT
VSLGVEFQRAHCKYPKRNQYKGTGDVNAIESWECVDL

>EDJ93992_*Magnaporthe_oryzae*

QDAPAFPESCLKLADQLNGKLPNTTVWFSEAVLAGTNITLKDTDPTCGGGNQRVDVDL
CRVAMFVTTSPTSNVSVEAWLPVDWKGDRFLSAGNGGLSGCIAYDDLAYTTARGFAS
VSGNAGHNGTSGVAFGYSDEVVEDFAWRSVHTGVVLGKQIVEEYYGRPHQKSYYLGC
STGGRQGFKSAQDFPEDFDGIVAGAPAFAFNNLSSWSGHFYTVTGKPGSPTFLTPDQW
MAVQADVLKQCDGLDGHEDGILEDPDMCAYRPEDLICGSGNNSTTCITGTQAETVRKV
FGAIYGVDGSFTYPGMQYGVDNSFIFVYLTGNPFVYSYEWFQYAIYHNPQWDPATLGP
KDYAASASADFGGISSFKGDLSGVRDRGAKILHYHGLQDQIITSTNSARYYNHVSRTM
NLPSGDLDSFYRYFRVSGMGHCNGGSGAAAIGNRGGDSYVSEDAADNVLTAIVDWVE
KGVAPETITGTRYQNGTRSSGVIEYKRRHCKYPLRNVFKGGDASNPDSWQCV

>7921857_*Aspergillus_flavus*

VGTDFPASCQAFSPDTRAANAHREFTEYVPAGTNLSLPYNDATCARPNQVVTVDLCRV
ALYVETSNRSSVTTEIWLPRNWTGRFLGTGNGGIDGCIKYEDLAYGAANGFAVVGSNN
GHNGTTAASFYQNSDVLADFAWRALHLSTVIGKEITQAFYGEPHRKSYYLGCSLGGRQ

GINSAVEFPDDFDGIIAGSPAVDFNSLVSWRASFFPITGSANSTDFISVSTWKDLIHAEVL
TQCDTLDCVNDGIIEDPSLCNFCPEALKCTDDRINNCLSPAQVEIVRKVFSPMYGEDGQ
LIFPAMQPGSELEAADQLYTGKPFRYSKEWFQYVVYNPSWDPAEFDIHDAKVADDLNP
QNIRTWPNDLSNYEKRGGKIITFHGQQDGKITSFNTERFYNHLATAMNMSSSELDNFFR
FFRISGMSHCSSGPGAWAFGQGGSPAPAMTPFNGNENILAALVAWVEHGVAPETITGT
KYVDDNPELGISIRRSHCRFVIQLNHGRHELCY

>BAE61129_*Aspergillus_oryzae*

VGTDFPASCQAFSPDTRAANAHREFTEYVPAGTNLSLPYNDATCARPNQVVTVDLCRV
ALYVETSNRSSVTTEIWLPRNWTGRFLGTGNGGIDGCIKYEDLAYGAANGFAVVGSNN
GHNGTTAASFYQNSDVLADFAWRALHLSTVIGKEITQAFYGEPHRKSYYLGCSLGGRQ
GINSAVEFPDDFDGIIAGSPAVDFNSLVSWRASFFPITGSANSTDFISVSTWKDLIHAEVL
TQCDTLDCVNDGIIEDPSLCNFCPEALKCTDDRINNCLSPAQVEIVRKVFSPMYGEDGQ
LIFPAMQPGSELEAADQLYTGKPFRYSKEWFQYVVYNPSWDPAEFDIHDAKVADDLNP
QNIRTWPNDLSNYEKRGGKIITFHGQQDGKITSFNTERFYNHLATAMNMSSSELDNFFR
FFRISGMSHCSSGPGAWAFGQGGSPAPAMTPFNGNENILAALVAWVEHGVAPETITGT
KYVDDNPELGISFRRSHCRYPLRNTYIGGNHWECTLP

>EAT91631_*Phaeosphaeria_nodorum*

MFLDLCRAVLNVNTSSSSTDRIEAWLPDNWNGRFLATGNGGTGGCIDYGNLQTGASY
GFATFGTNGGHDGQAGYDFFLNKPENINDFGYRAIHAEAETGKKVVEHYYGRPGDYN
YYMGCSTGGRQGYEEAILYPEDFDGILLGSMATDWLHIVASKAILAHRLGWPDLSSPS
YVSAAQWAAIVAAQVALLDPLDGVKDGVIDNPTLHNFDPSILACGTGVLNSSVCLTAA
QVNTVREIYQPLANSSGHIVYPSFELGTPTGVFSTNTKVDNGTTIPSLNYGLVDDFWRG
AVYNDTSFSSLNFTTKHMDFATNLNPGGIQIAGDKSHNVSAFHARGGKLLAYHGRADP
TVMSGVASRTFQKTAEALNLTLNEMHEFFRLFFVPGMGHCAGGNGGQWAIGQPGVERP
TAAGFNDTQHNILLAMVDWVENDRAPISLVGTKFLNDTILDGVVESQRTHCVYPNASR
WDGVGDSKKAESWTCQLEGVL

>4353506_*Aspergillus_terreus*

QSSGCGKQPTLTNGVQNINGREYILNIPEGYDSSKQYKLIFGLHWLGGSMNDVVSGNSI
EPWYGLESRAEGSAIFVAPNGLNAGWANNGGEDTALMDAIIEAVEADLCIDQSSRFAT
GFSFGGGMSYALACARASKFRAVSVLSGGVISGCDGGNDPIAYLGIHGINDPVLPIDGGI
EMANKFVQNNGCQSADVGRPNSGSGQSVRTDFQGCSRPVSFIAYDGGHEGAPLGVGN
PLAPDATWEFFTGA

>4706641_*Aspergillus_clavatus*

EARSPGCGKLPNLSNGVHKINGREYTLKVPSNYDKNKAYHLVFGLHWRGGNMGAVV
KGEGVQPWYGLEKRAHGSAIFVAPNGKSAGWANTGGQDIAFLDAIIKEVESNYCVDQS
SRFATGFSWGGGMSYALACSRAKQFKAVSVLSGGLISGCSGGNDPIAYLAIHGIKDNVL
PFNGGISLANRFVKNNRCHQSEIHTPKSGSRSSVRTDFRGCSKPVSFIAYDGGHDAAPLG
VGSSLAPDATWKFFMAA

>5076993_*Neosartorya_fumigata*

QTRSSGCGKQPSLANGVHNINGREYILKVPDNYDKNKAHHLVFGLHWRGGNMWNIV
DGQSIQPWYGLETRAQGSAIFVAPNGKNAGWANYGGEDIAFIDAIIKQVESDLCVDQSS
RFATGFSWGGGMSYSLACSRAKQFKAVSVLSGGVISGCDGGNDPIAYLGIHGINDGVL
PFQGGVNLAQKFVRNNGCQQSNVGTPQPGSRGSVRTDFKGCSKPVSFIAYDGGHDAAP
LGVGSSLAPDATWRFFMAA

>*Ac*FAE_*Aspergillus_clavatus*

ETRSSGCGKHPSLANGVIHLNGREYILKLPDRYDNNHAYHLVFGLHWRGGNMQNVAN
GESIQPWYGLETRAQGSTIFIAPNGKNAGWANNGGEDVAFIDAIIKQVEADLCVDQSSR
FATGFSWGGGMSYSLACSRAKQFKAVSVLSGGVISGCDGGHDPIAYLGIHGINDGVLPF
NGGVGLAQKFVQNNGCQQANIGAPPSGSKSSVRTDFKGCSKPVSFIAYDGGHDSAPLG
VGSSLAPDATWKFFMAA

>*Anid*FAE_*Aspergillus_nidulans*

ANSPGCGKQPTLTNGVNQINGREYVLKIPDGYDPSKPHHLIFGLHWRGGNMYNVVNG
DSIQPWYGLEARAQGSAIFVAPNGLNAGWANTNGEDVAFIDAIMEQVEDDLCVDQAS
RFATGFSWGGGMSYALACARAAEFRAVSVLSGGLISGCDGGNDPIAYLGIHGINDPVLP
LDGGVTLANTFVSNNGCQPTDIGQPASGSGGSVRTDFSGCSHPVSFIAYDGGHDGAPLG
VGSSLAPDATWEFFMAA

>AP007157_*Aspergillus_oryzae*

ANSSGCGKQPTLVNGVHKINDREYILKVPDNYNANKPHHLIFGLHWRGGNMNSVVNG
ESVEPWYGLETRAQGSAILVAPNGRNAGWANINGEDVALIDAIIKQVEDDLCIDQSSRF
ATGFSWGGGMSYALACARAKEFRAVSVLSGGVISGCEGGHDPIAYLGIHGISDPVLPFD
GGVTLANKFAANNGCQQTYVGKPGLGSHSSVQTDFKGCSRPVSFIAYDGGHDAAPLG
VGNPLAPDATWKFFMAA

>*Cl*FaeA1_*Chrysosporium_lucknowense*

ASAGCGKAPPSSGTKSMTVNGKQRQYILQLPNNYDANKAHRVVIGYHWRDGSMNDV
ANGGFYDLRSRAGDSTIFVAPNGLNAGWANVGGEDITFTDQIVDMLKNDLCVDETQFF
ATGWSYGGAMSHSVACSRPDVFKAVAVIAGAQLSGCAGGTTPVAYLGIHGAADNVLP
IDLGRQLRDKWLQTNGCNYQGAQDPAPGQQAHIKTTYSCSRAPVTWIGHGGGHVPDP
TGNNGVKFAPQETWDFFDAAVGAAGAQSPMT

>*Cl*FaeA2_*Chrysosporium_lucknowense*

ATPSPGCGKTPTLITDGSATTPLTLTSNGKTRRFYVKLPDDYDNSHPYRLIFALHALGGT
AQQVTTGTGGYLPWYGIPDLAANDTVGAVYVAPDGLNNGWANQGGEDVAFLEAVM
ETVEQDVCVDRDLRFSTGFSYGAAMSYTLACALGRRIRAVAVLSGSPVISGGCAGAGS
GASEPVAYYGQHGMSDPVLPVAGGREMRDHFVRTNGCDAGRGPPREPARGSGTHVK
TVYDGCDPDYPVVWNAFDGDHTPQPVDRGATTTFSAVETWEFFSQFK

>DS027696_*Neosartorya_fischeri*

QTRSSGCGKQPSLTNGVHNINGREYILKVPDNYDKNKAHHLVFGLHWRGGNMWNIVD
GQSVQPWYGLETRAQGSTIFVAPNGKNAGWANNGGEDIAFIDAIIKQVEGDLCVDQSS

RFATGFSWGGGMSYSLACSRAKQFRAVSVLSGGVISGCDGGNDPIAYLGIHGINDGVLP
FQGGVNLAQKFVKNNRCQQANVGTPQPGSHGSVRTDFKGCSKPVSFIAYDGGHDAAP
LGVGSSLAPDATWKFFMAA

>EQ963475_*Aspergillus_flavus*

ANSSGCGKQPTLVNGVHKINDREYILKVPDNYNANKPHHLIFGLHWRGGNMNSVVNG
ESVEPWYGLETRAQGSAILVAPNGRNAGWANTNGEDVALIDAIIKQVEDDLCIDQSSRF
ATGFSWGGGMSYALACARAKEFRAVSVLSGGVISGCEGGHDPIAYLGIHGISDPVLPFD
GGVTLANKFAANNGCQQAYVGKPGLGSHSSVQTDFKGCSRPVSFIAYDGGHDAAPLG
VGNPLAPDATWKFFMAA

>*Fo*FaeC_*Fusarium_oxysporum*

ASAGCGKQPPSSGVKTMQVNGKNREYTLQLPNNYQNNKPHRLVFGYHWLSGNMGNV
VQGGYYGLRNLAGDSTIFIAPNGLNAGWANQGGEDITFTDQMLAFAKQNLCIDEKQVF
ATGFSYGGAMSHSVACSRPNDFAAVAVISGALLSGCNGGNTPVSYLHIHGSADNVLSIQ
QGRQLRDKWIGTNGCQQKQVNDPAPGAQNYVKTSYTCSRKPVTWIGHGGGHVADPT
ANGQKFAPGETWSFFNAAAGTSAKLRC

>jgi_Aspni7_1182746_estExt_Genemark1.C_chr_3010193

ANSPGCGKNPTLANGVHQINGREYTLKIPDDYDANNPYHLIFGLHWRGGNMDNVVSG
DSIQPWYGLESRAQGSAIFIAPNGLNAGWANTNGEDVAFIDAIMEQVESDLCVDQSSRF
ATGFSWGGGMSYSLACSRAKEFRAVSVLSGGVISGCDGGNDPIAYLGIHGINDPVLPFD
GGVELAERFVGNNGCQPASIEKPQSGSNGWKRTDFYGCSKPVSFIAYDGGHDGAPLGV
QSSLAPDATWEFFMAA

>*Nc*FaeD_*Neurospora_crassa*

APSSGCGKGPTLRNGQTVTTNINGKSRRYTVRLPDNYNQNNPYRLIFLWHPLGSSMQKI
IQGEDPNRGGVLPYYGLPPLDTSKSAIYVVPDGLNAGWANQNGEDVSFFDNILQTVSD
GLCIDTNLVFSTGFSYGGGMSFSLACSRANKVRAVAVISGAQLSGCAGGNDPVAYYAQ
HGTSDGVLNVAMGRQLRDRFVRNNGCQPANGEVQPGSGGRSTRVEYQGCQQGKDVV
WVVHGGDHNPSQRDPGQNDPFAPRNTWEFFSRFN

>*Pf*FaeA_*Penicillium_funiculosum*

QQSLWGQCGGTGWTGPTTCVSGACCQEQNPYYSQCIQGNCSPASSTSSTSSTKTTTTSA
SSTTTSTSASGTSLSSGCGKALSLKSGTYTTTVAGQQRQYTLTLPSNYNPNKAYQLIFGY
HWLGGTMGNVVSGSYYGIQPLAGDNAIFVAPQGLNNGWGNTNGDDIIFTDQMLSTLE
NALCIDETQIYSMGWSYGGAMSYALACARPDVFRAVAVMSGANLSGCSPGTQPVAYY
GQHGVSDTVLPFSLGEGIRDTFVKDDHCTPTNPPAPAAGSGTHIKTEYSGCDSEHPVW
WIAFDGPHEPLATDAGASSSWTPGQIWSFFSLFH

>A0A1L9T9J3_*As*FaeE

ATLSQVLDFGNNPGDNEMWIYVPDQLAANPAVIVALHGCLGSAEGYYSEVQDLPPAA
DENGFILVYPGSNDDFHCWDVATAESLTHDGGSDSRSIVNMVQYTLDKYSGDSSKVFT
TGSSSGAMMSLVLAAAYPDVFSGVAAYSGVPYGCLRGSPGSSPFTADQACANGEVSRT

AQEWKDEVKMAWPGYNGTYPKVQVWHGTADSVISPNNFDEEVKQWSAVFGVNVTK
EEQDSPLDGYTRSIFGDGSHFEAYLAEGVGHVVPTQVDSTLRWFGLI

>3879354_*Neurospora_crassa*

ASLQQVTNWGSNPTNIRMYTYVPDKLATKPAIIVALHGCGGTAPSWYSGTRLPSYADQ
YGFILIYPGTPNMSNCWGVNDPASLTHGAGGDSLGIVAMVNYTIAKYNADASRVYVM
GTSSGGMMTNVMAATYPEVFEAGAAYSGVAHACFAGAASATPFSPNQTCARGLQHTP
EEWGNFVRNSYPGYTGRRPRMQIYHGLADNLVYPRCAMEALKQWSNVLGVEFSRNV
SGVPSQAYTQIVYGDGSKLVGYMGAGVGHVAPTNEQVMLKFFGLIN

>*Agrocybe_praecox*

LTGSVLSVDADFGPNPTNVQMLQYRPAQLITPTPLIIALHYCTGRGIAYFSTTSLARLAD
QYGYIVVYPTAPSSDGCWDVSSNATLTHNGGGDSLGIASMVRYAIANWGVDASRVFA
TGTSSGAMMTNVLMGAYPDLFAAGSLYSGVPYGCFASPSAYAWNAQCATGQLIKTAQ
QWGDQVRSGFPGFTGTRPKVQFWHGAIDITLYPQNFWEEIKQWTNIFGVSETPTSNLTD
NPQSGYSRASFGPNVQAILAQGVGQPVPENETDTLDWFGLTNLTVPAPTTTIAPAPTGV
TSPQWAQCGGLGWTGPTVCAAPYTCIVENAYFSHCY

>*Al*AXE_*Aspergillus_luchuensis*

SGSLQQVTDFGDNPTNVGMYIYVPNNLASNPGIVVAIHYCTGTGPGYYGDSPYATLSE
QYGFIVIYPSSPYSGGCWDVSSQATLTHNGGGNSNSIANMVTWTISKYGADSSKVFVTG
SSSGAMMTNVMAATYPELFAAATVYSGVSAGCFYSNTNQVDGWNSTCAQGDVITTPE
HWASIAEAMYSGYSGSRPRMQIYHGSIDTTLYPQNYYETCKQWAGVFGYDYSAPEKT
EANTPQTNYETTIWGDSLQGIFATGVGHTVPIHGDKDMEWFGFA

>AXE_BAD12626_*Aspergillus_oryzae*

RAIHNGRSLIPRAGSLEQVTDFGDNPSNVKMYIYVPTNLASNPGIIVAIHYCTGTAQAYY
QGSPYAQLAETHGFIVIYPESPYEGTCWDVSSQATLTHNGGGNSNSIANMVTWTTKQY
NADSSKVFVTGTSSGAMMTNVMAATYPNLFAAGVAYAGVPAGCFLSTADQPDAWNS
TCAQGQSITTPEHWASIAEAMYPDYSGSRPKMQIYHGNVDTTLYPQNYEETCKQWAG
VFGYNYDAPESTESNTPEANWSRTTWGPNLQGILAGGVGHNIQIHGDEDMKWFGFTN

>AXE_CAK46215axeA_*Aspergillus_niger*

HVAKRSGSLQQITDFGDNPTGVGMYIYVPNNLASNPGIVVAIHYCTGTGPGYYSNSFYA
TLSEQYGFIVIYPSSPYSGGCWDVSSQATLTHNGGGNSNSIANMVTWTISEYGADSKKV
FVTGSSSGAMMTNVMAATYPELFAAGTVYSGVSAGCFYSDTNQVDGWNSTCAQGDV
ITTPEHWASIAEAMYPGYSGSRPKMQIYHGSVDTTLYPQNYYETCKQWAGVFGYDYS
APESTEANTPQTNYETTIWGDNLQGIFATGVGHTVPIHGDKDMEWFGFA

>AXE_EAW06435_Axe1_*Aspergillus_clavatus*

ASVLESRSSALLPRAGSLQQVTNFGDNPTNVGMYIYVPNNLASNPGIIVAIHYCTGTAE
AYYNGSPYAKLAEKHGFIVIYPESPYQGKCWDVSSRASLTHNGGGNSNSIANMVKWTI
KKYKTNTSKVFVTGSSSGAMMTNVMAATYPDMFAAGVVYSGVAAGCFMSNTNQQA
AWNSTCAHGKSIATPEAWAHVAKAMYPGYDGPRPRMQIYHGSADTTLYPQNYQETC

KEWAGVFGYDYNAPRSVENNKPQANYKTTTWGKELQGIYATGVGHTVPINGDRDMA
WFGFAK

>*Chae*Fae_*Chaetomium*_sp

ASLQQVTNFGNNPTNIQMHIYVPDRVASNPAIIVALHPCGGNAQQWFGGTRLPSYADQ
HGFILIYPSTPHMSNCWDVHNPASLTHGQGGDALGIVSMVNYALNQYNGDRNRVYAM
GFSSGGMMTNVLAGSYPDVFEAGAAYSGVPHACFLGAPAATPFSPNQTCAQGLQKTE
QEWGDLVRNSYPGYTGRRPRMQITHGLADFLVRPQCAYEALKQWSNVLGVQLTREVR
GVPSPQFTQLIYGDGTQLQGFLGDGVGHEPSVNEEQMLRFFGLIN

>*Cl*FaeB1 *Chrysosporium lucknowense*

MRLLRSLTKMAAMAGLAAGASLQPVTNFGDNPTGLQMYVYVPDKVAVSPAIIVALHP
CGGSAQGWYSQTRLPSYADQLGFILIYAGTTKMSNCWDVQNPASLTHNGGGDAGGIV
SMVKYALKQYNGDASRVYVMGGSSGAMMTNVLAGSYPDVFEAGAAFSGVAHACFL
GADSATPFSPNQTCAQGRIQRSAREWGDLVRNSFPAYDGRRPRMQIFHGNADFLVHPE
CAHQALAQWADVLGLQLTQTNKGVPSAEYTQEVYGDGTQLQGFFGDGVGHIAPVNEP
VMLRFFGLMN

>*Cl*FaeB2_*Chrysosporium_lucknowense*

ASLQEVTEFGDNPTNIQMYIYVPDQLDTNPPVIVALHPCGGSAQQWFSGTQLPSYADD
NGFILIYPSTPHMSNCWDIQNPDTLTHGQGGDALGIVSMVNYTLDKHSGDSSRVYAMG
FSSGGMMTNQLAGSYPDVFEAGAVYSGVAFGCAAGAESATPFSPNQTCAQGLQKTAQ
EWGDFVRNAYAGYTGRRPRMQIFHGLEDTLVRPQCAEEALKQWSNVLGVELTQEVSG
VPSPGWTQKIYGDGTQLQGFFGQGIGHQSTVNEQQLLQWFGLI

>*Ct*FxeA_*Chaetomium_thermophilum*

ASLQQVWNFGSNPTNINMHIYVPDRLANKPPIIVALHPCGGNAQQWFSGTRLPQYADR
HGFILIYPSTPHMSNCWDVQNPASLTHGAGGDALGIVSMVNYAIDRYGADRDRVYVM
GFSSGGMMTNVLAGSYPDVFEAGAAYSGVPHACFLGAPAATPFSPNQTCAQGLQKTPE
EWGNLVRNSYPGYNGRRPRMQITHGLNDWLVRPQCAYESLKQWSNVLGLQLTRQVT
SGQWTQHIYGDGTKLVGYLGQGIGHEPAVNEEQLLRFFGIIN

>jgi_Aspni7_1163354_estExt_fgenesh1_pg.C_chr_302_t10192

HVAKRSGSLQQITDFGDNPTGVGMYIYVPNNLASNPGIVVAIHYCTGTGPGYYSNSPYA
TLSEQYGFIVIYPSSPYSGGCWDVSSQATLTHNGGGNSNSIANMVTWTISEYGADSKKV
FVTGSSSGAMMTNVMAATYPELFAAGTVYSGVSAGCFYSDTNQVDGWNSTCAQGDV
ITTPEHWASIAEAMYPGYSGSRPKMQIYHGSVDTTLYPQNYYETCKQWAGVFGYDYS
APESTEANTPQTNYETTIWGDNLQGIFATGVGHTVPIHGDKDMEWFGFA

>*Ms*Fxe1_*Myceliophthora_sepedonium*

ASLQEVTNFGSNPSNIQMYIYVPDQLAANPPVIVALHPCGGNAQQWFGGTQLPSYADS
HGFILIYPSTSHMSNCWDIQNPDTLTHGQGGDALGVVNMVNYALETYSGDSSRVYAM
GFSSGGMMTNQLAGSYPDVFEAGAAYSGVAFGCSAGAESATPGSANQTCAQGLQHTE

QEWGDFVRNAYPGYTGRRPRMQIFHGLADTLVRPQCAEEALKQWSNVLGVEFTQEVS
GVPSAGWTQKIYGDGTQLQGFFGQGIGHQSTVNEQQLLEWFGLL

>*Mt*Fae1a_*Myceliophthora_thermophila*

ASLQEVTEFGDNPTNIQMYIYVPDQLDTNPPVIVALHPCGGSAQQWFSGTQLPSYADD
NGFILIYPSTPHMSNCWDIQNPDTLTHGQGGDALGIVSMVNYTLDKHSGDSSRVYAMG
FSSGGMMTNQLAGSYPDVFEAGAVYSGVAFGCAAGAESATPFSPNQTCAQGLQKTAQ
EWGDFVRNAYAGYTGRRPRMQIFHGLEDTLVRPQCAEEALKQWSNVLGVELTQEVSG
VPSPGWTQKIYGDGTQLQGFFGQGIGHQSTVNEQQLLQWFGLI

>*Nc*Fae1_*Neurospora_crassa*

ASLQQVTNWGSNPTNIRMHTYVPDKLATKPAIIVALHGCGGTAPSWYSGTRLPSYADQ
YGFILIYPGTPNMSNCWGVNDPASLTHGAGGDSLGIVAMVNYTIAKYNADASRVYVM
GTSSGGMMTNVMAATYPEVFEAGAAYSGVAHACFAGAASATPFSPNQTCARGLQHTP
EEWGNFVRNSYPGYTGRRPRMQICHGLADNLVYPRCAMEALKQWSNVLGVEFSRNVS
GVPSQAYTQIVYGDGSKLVGYMGAGVGHVAPTNEQVMLKFFGLIN

>*Pa*FaeD_*Podospora_anserina*

APSAPGANDIPRDVLTQISSWGSNPTNLQLHLYAPSNLTGKPAIILALHGCFGSGPGHAE
MTSQFQTLSSSRNFVVLYPSSINDNNCWDVASPASLTRDGGGDSTGLATIVRWATTTFN
ADPKKVFITGSSSGCMMSNVMASAYPDLFSAVSCYSGVPAGCLAGSPGSSPISADQTCA
NGGIRKTGEEWAEVVRGMAPLPEGKGKGKGKGKGKGNQGWKYPKVATWHGDNDFF
VNYFHNFEEQLKQWGAIHGVEFTRNETNVPAAGYTKMVYGDGTKLVGYSASGVGHV
VPQFEAVDLEWFGL

>*Pf*FaeB_*Penicillium_funiculosum*

ASLTQVNNFGDNPGSLQMYIYVPNKLASKPAIIVAMHPCGGSATEYYGMYDYHSPAD
QYGYILIYPSATRDYNCFDAYSSASLTHNGGSDSLSIVNMVKYVISTYGADSSKVYMTG
SSSGAIMTNVLAGAYPDVFAAGSAFSGMPYACLYGAGAADPIMSNQTCSQGQIQHTGQ
QWAAYVHNGYPGYTGQYPRLQMWHGTADNVISYADLGQEISQWTTIMGLSFTGNQT
NTPLSGYTKMVYGDGSKFQAYSAAGVGHFVPTDVSVVLDWFGITSGTTTTTTPTTTPT
TSTSPSSTGGCTAAHWAQCGGIGYSGCTACASPYTCQKANDYYSQCL

>*T. wortmannii* Fae7262 MF362597

MAIPLVLVLAWLLPAVLAASLTQVNNFGDNPGSLQMYIYVPNTLASKPAVIVAMHPCG
GSATEYYGMYDYHSPADQYGYILIYPSATRDYNCFDAYSSSSLTHNGGSDSLSIVNMV
KYVISTYGADSSKVYMTGSSSGAIMTNVLAGAYPDVFAAGSAFSGMPYACLYGAGAA
DPIMSNQTCSQGQIQHTGQQWAAYVHNGYPGYTGRYPRLQMWHGTADNVISYADLG
QEISQWTTVMGLSFTGNQTNTPLSGYTKMVYGDGSQFQAYSAAGVGHFVPTDVSVVL
DWFGITSGTTTTTTSKTTSATTSTTSSAPSSTGGCTAAHWAQCGGIGYTGCTACVSPYT
CQKSNDYYSQCL

>TPfunFae_*Talaromyces_funiculosus*

ASLTQVNNFGDNPGSLQMYIYVPNKLASKPAIIVAMHPCGGSATEYYGMYDYHSPAD
QYGYILIYPSATRDYNCFDAYSSASLTHNGGSDSLSIVNMVKYVISTYGADSSKVYMTG
SSSGAIMTNVLAGAYPDVFAAGSAFSGMPYACLYGAGAADPIMSNQTCSQGQIQHTGQ
QWAAYVHNGYPGYTGQYPRLQMWHGTADNVISYADLGQEISQWTTIMGLSFTGNQT
NTPLSGYTKMVYGDGSKFQAYSAAGVGHFVPTDVSVVLDWFGITSGTTTTTTPTTTPT
TSTSPSSTGGCTAAHWAQCGGIGYSGCTACASPYTCQKANDYYSQCL

>4983573_*Aspergillus_niger*

ASTQGISEDLYNRLVEMATISQAAYADLCNIPSTIIKGEKIYNAQTDINGWILRDDTSKEI
ITVFRGTGSDTNLQLDTNYTLTPFDTLPQCNDCEVHGGYYIGWISVQDQVESLVKQQAS
QYPDYALTVTGHSLGASMAALTAAQLSATYDNVRLYTFGEPRSGNQAFASYMNDAFQ
VSSPETTQYFRVTHSNDGIPNLPPAEQGYAHGGVEYWSVDPYSAQNTFVCTGDEVQCC
EAQGGQGVNDAHTTYFGMTSGACTW

>7912018_*Aspergillus_flavus*

AITQGISEGTYSRIVEMATISQAAYANLCNIPSTITSAGKIYNAETDINGWVLRDDSRQEII
TVFRGTGSDTNLQLDTNYTQAPFDTLPQCSGCAVHGGYYVGWISVKDQVEGLVQQQA
SQYPDYSLVITGHSLGASMAAITAAQLSATYNNITVYTFGEPRTGNQAYASYVDETFQA
TNPDATKFYRVTHTNDGIPNLPPTSQGYVHHGTEYWSVEPHGPQNMYLCLGDEVQCC
EAQGGQGVNDAHVTYFGMASGACTW

>*An*FaeA_*Aspergillus_niger*

ASTQGISEDLYNRLVEMATISQAAYADLCNIPSTIIKGEKIYNAQTDINGWILRDDTSKEI
ITVFRGTGSDTNLQLDTNYTLTPFDTLPQCNDCEVHGGYYIGWISVQDQVESLVKQQAS
QYPDYALTVTGHSLGASMAALTAAQLSATYDNVRLYTFGEPRSGNQAFASYMNDAFQ
VSSPETTQYFRVTHSNDGIPNLPPADEGYAHGGVEYWSVDPYSAQNTFVCTGDEVQCC
EAQGGQGVNDAHTTYFGMTSGACTW

>AP007154_*Aspergillus_oryzae*

AITQGISEGTYSRIVEMATISQAAYANLCNIPPAITSAGKIYNAETDINGWVLRDDSRQEI
ITVFRGTGSDTNLQLDTNYTQAPFDTLPQCSGCAVHGGYYVGWVSVKDQVEGLIHQQ
ASQYPDYSLVVTGHSLGASMAAITAAQLSATYNNITVYTFGEPRTGNQAYASYVDETF
QATNPDATKFYRVTHTNDGIPNLPPTSQGYVHHGTEYWSVEPHGPQNMYLCLGDEVQ
CCEAQGGQGVNDAHVTYFGMASGACTW

>*At*FAE_1_*Aspergillus_terreus*

ASTQGISEDLYNRLVEMATISQAAYANMCNIPSTITVGEKIYNAQTDINGWVLRDDSTK
EIITVFRGTGSDTNLQLDTNYTLTPFSTFSECSGCEVHGGYFIGWSSVQDQVMSLVKEQ
ADQYPDYTLTVTGHSLGASMATLAAAQLSGTYDNITLYTFGEPRSGNEAFASYMNDKF
TATSADTTKYFRVTHSNDGIPNLPPAEQGYVHGGVEYWSVDPYSAQNTYVCTGDEVQ
CCEAQGGQGVNDAHTTYFGMTSGACTW

>*Atub*FaeA_*Aspergillus_tubingensis*

ASTQGISEDLYSRLVEMATISQAAYADLCNIPSTIIKGEKIYNSQTDINGWILRDDSSKEII
TVFRGTGSDTNLQLDTNYTLTPFDTLPQCNSCEVHGGYYIGWISVQDQVESLVQQQVS
QFPDYALTVTGHSLGASLAALTAAQLSATYDNIRLYTFGEPRSNQAFASYMNDAFQAS
SPDTTQYFRVTHANDGIPNLPPADEGYAHGVVEYWSVDPYSAQNTFVCTGDEVQCCE
AQGGQGVNNAHTTYFGMTSGHCTW

>*Au*FaeA_*Aspergillus_usamii*

ASTQGISEDLYNRLVEMATISQAAYADLCNIPSTIIKGEKIYNAQTDINGWILRDDTSKEI
ITVFRGTGSDTNLQLDTNYTLTPFDTLPQCNDCEVHGGYYIGWISVQDQVESLVKQQAS
QYPDYALTVTGHSLGASMAALTAAQLSATYDNVRLYTFGEPRSGNQAFASYMNDAFQ
VSSPETTQYFRVTHSNDGIPNLPPAEQGYAHGGVEYWSADPYSAQNTFVCTGDEVQCC
EAQGGQGVNDAHTTYFGMTSGACTW

>*Aw*FAEA_*Aspergillus_awamori*

ASTQGISEDLYTRLVEMATISQAAYADLCNIPSTIIKGEKIYNSQTDINGWILRDDSSKEII
TVFRGTGSDTNLQLDTNYTLTPFDTLPQCNGCEVHGGYYIGWVSVQDQVESLVKQQV
SQYPDYALTVTGHSLGASLAALTAAQLSATYDNIRLYTFGEPRSGNQAFASYMNDAFQ
ASSPDTTQYFRVTHANDGIPNLPPVEQGYAHGGVEYWSVDPYSAQNTFVCTGDEVQC
CEAQGGQGVNNAHTTYFGMTSGACTW

>jgi_Aspni7_1180590_estExt_Genemark1.C_chr_101_t10028

ASTQGISEDLYNRLVEMATISQAAYADLCNIPSTIIKGEKIYNAQTDINGWILRDDTSKEI
ITVFRGTGSDTNLQLDTNYTLTPFDTLPQCNDCEVHGGYYIGWISVQDQVESLVKQQAS
QYPDYALTVTGHSLGASMAALTAAQLSATYDNVRLYTFGEPRSGNQAFASYMNDAFQ
VSSPETTQYFRVTHSNDGIPNLPPAEQGYAHGGVEYWSVDPYSAQNTFVCTGDEVQCC
EAQGGQGVNDAHTTYFGMTSGACTW

>LIP_AAC08588_*Thermomyces_lanuginosus*

SPIRREVSQDLFNQFNLFAQYSAAAYCGKNNDAPAGTNITCTGNACPEVEKADATFLYS
FEDSGVGDVTGFLALDNTNKLIVLSFRGSRSIENWIGNLNFDLKEINDICSGCRGHDGFT
SSWRSVADTLRQKVEDAVREHPDYRVVFTGHSLGGALATVAGADLRGNGYDIDVFSY
GAPRVGNRAFAEFLTVQTGGTLYRITHTNDIVPRLPPREFGYSHSSPEYWIKSGTLVPVT
RNDIVKIEGIDATGGNNQPNIPDIPAHLWYFGLIGTCL

>LIP_AN8046_*Aspergillus_nidulans*

APAPLNRRDVSTEALNQLTLFAEYSAASYCTPNIGSVGDKLTCASGNCPTVEAADTTTL
AEFYQENEYGDVAGFLAADTTNELLVLSFRGSRTIDTWIANLDFGLESVEEICSGCKAH
GGFWKAWQVVADSLTSAIESATATYPGYAIVFTGHSFGGALATLGAAQLRKAGYAIEL
YPYGSPRVGNEALAQYITDQGANYRVTHTNDIVPRLPPMLLGFSHLSPEYWITSDNEVT
PTTTDIQVIEGVGSRDGNAGEAAQSVEAHSWYLIDITACQ

>LIP_BAA14345_*Penicillium_camembertii*

LPGKLQSRDVSTSELDQFEFWVQYAAASYYEADYTAQVGDKLSCSKGNCPEVEATGA
TVSYDFSDSTITDTAGYIAVDHTNSAVVLAFRGSYSVRNWVADATFVHTNPGLCDGCL

AELGFWSSWKLVRDDIIKELKEVVAQNPNYELVVVGHSLGAAVATLAATDLRGKGYP
SAKLYAYASPRVGNAALAKYITAQGNNFRFTHTNDPVPKLPLLSMGYVHVSPEYWITS
PNNATVSTSDIKVIDGDVSFDGNTGTGLPLLTDFEAHIWYFVQVDAGKGPGLPFKRV

>LIP_BAL22280_*Aspergillus_niger*

APAPLDVRSVSTSTLDELQLFAQWSAAAYCSNNIDSDDSNLTCTANACPSVEEASTTML
LEFDLTNDFGGTAGFLAADNTNKRLVVAFRGSSTIENWVANLDFILEDNDDLCTGCKV
HTGFWKAWESAADDLTSKIKSAMSTYSGYTLYFTGHSLGGALATLGATVLRNDGYSV
ELYTYGCPRIGNYALAEHITSQGSGANFRVTHLNDIVPRVPPMDFGFSQPSPEYWITSGT
GASVTASDIEVIEGINSTAGNAGEATVSVLAHLWYFFAISECLL

>LIP_GAA84811_*Aspergillus_kawachii*

APAPLAVRSVSTSTLDELQLFAQWSAAAYCSNNIDSKDSNLTCTANACPSVEEASTTML
LEFDLTNDFGGTAGFLAADNTNKRLVVAFRGSSTIENWIANLDFILEDNDDLCTGCKVH
TGFWKAWESAADELTSKIKSAMSTYSGYTLYFTGHSLGGALATLGATVLRNDGYSVE
LYTYGCPRIGNYALAEHITSQGSGANFRVTHLNDIVPRVPPMDFGFSQPSPEYWITSGNG
ASVTASDIEVIEGINSTAGNAGEATLSHPRVATTIASGASVVLLYYSIRRPTFHPHPAPPQ
NLPESVNSLEPPPASTSPTPDRLRSSNRAAMFPAPSR

>*Am*Fae1A_*Anaeromyces_mucronatus*

MSKLQISNTCPDKYRTKQEGVEYPTAKKITYYSKVTETERKMNVILPVGYDENKKYPV
VYYLHGLMSYEDSMLEDDSTLAIPTNLLKEGRAKEMIIVLPDVYAPKPGTAVTPDFNPE
YYKGYDNFINELIEVIMPYMEEHYSILTGRENTALCGFSMGARTSLYIGYMRSDLIGYV
GAFAPAPGITPGEDSFSGKHEGLISEDEFRAEIQPIVSLIDCGTNDSVVGQFPKSYHEILTR
NNQEHIWFEVPGADHDWNAISAGFYNFIQTTFGALNN

>BAE66413_*Aspergillus_oryzae*

CEPSFFAPYLTSNVSIVYARTFTSTETFEGPMGNYTGLPEYCALYVNVSSSTTSAYEFGL
WLPTRTWNKRYMAYGNGGFTGQVAFADMAPGLNYGFAVVSTNTGHNSSVQEAGDA
GWALNNPETRTDWGWRALHGSVALGRVLTEAFYDNNIEYSYYAGCSTGGRQGLKEV
QMFPDDFDGVLAGACAWWTSHQQNWDLKVALDNLPNNASHHVTAELMDVLATEVL
RQCDVQDGVKDNIIMDGYACQFRLEALLCSSPRTNTSTCFSADQLDTIHRGSEAQVSLM
IITDDDSVDEPSGIAYARDFVYNDASWPAEHLDYATIQLSEYLDPGNATADTYDLRPFY
QRGGKLIHYHGFADGEIPTGSSIYYYKQVEKTMIPLGYDLDDFYRFFLIPGMQHCSGSV
HGAPWYIGSANQPSALTGSNIWGVPGFRDTKHDAIMALMAWVENGTAPAELIATKYID
DTPALGVQAQRRTCPYPQRAAYVGGNWNQTSSFKCE

>jgi_Aspni7_1082974_e_gw1.202.1306.1

GSRTVHAVLDCDWVAVNAILPQNVSLSFVTAIPENGTFTVPREDTGWPTDPINLPSLCAI
GAMVLGETANSSFGFGLFLPTKWNGRTLTVGNGGLAGGVNWVDMGTGVKYGHAVFS
TDTGHNSTATDASWSYQNIESQVNWGWRALHETVIYSKLITEAFYGIKPSFNYYQGCST
GGRQGFKEAEAFPSDFDGIIAGAPAWWTSHQQIWQFWVGYVNYISNLTMIPTSKFDVIG
KEVLRQCDGQDGLVDTIISDPLGCNFNPEMLLCSSDVTDTTNAECLTAEQLQTFDILTR
DFLETNSTLVFPAWLRGSEHFWNLNIDGGAPNIIGLGYIQYMLGLGPDWKWESFDADI

VQLSERLNPGQADASYYNLREFYEGGRKLIHYHGLSDGGIATGASYYLYDQIDRALTP
QGIDVNDFYRFFPIPGMGHCVDTANYVNAPYYIAGISMTNNGQHSVPHYQDAKHDVLF
ALMDWVENGTAPDAIIGTAYANFSTMDAITRQRPICAYPKQAKYRGKGDPNIPENWRC
QLLY

>jgi_Aspni7_1119031_estExt_Genewise1Plus.C_chr_302_t10316

QFAPSGRGPTSGGAFDCLVDDIKSLLPDGSKVLYASHYDAGDRFTPPLEHNSGAFRISSY
TLPRSACVFQANLTLPGETEHSVGVVLPDDWNGRFLTAGNGGFSGSVSWGSIIDASWY
GFAAVSTNTGHESSGAEWAYHNQEALANWGYRAMHDAVVNGKSVTEQYYGNKIAY
SYYRGCSAGGKQGLKEVEMFPDDFDGVIAGAPAWWTTHLQLWNMIVGIWNSPADSP
HHLLNEHIDLLANEVIRQCDPQDGVEDGIVINPDACIFRPEALLCARDVRDLSTCLNTEQ
IKTVYKLHNDWMEANDTFIFPHLTLSSEWKWSGLVSSPASLGTDYVKWMLQLGEDWS
WEDWDPELIKLSDELNPGNATTDDYDLSPFYNKGGKLLHYHGWSDWSIAAGSSIYFYD
QVVNALRPKGINPGDFYRFYLIPGMGHCGGTSDLMNAPWVIAGDGQSSSLGRDGNIHS
VPGHEDAKHDVVLAMMNWVENGTSPDELIATKFVNDKDPEEGVYRQRPLCVYPAEA
QYKGSGDINAAESWECREL

>jgi_Aspni7_1134144_fgenesh1_pg.chr_304_405

RNNCSIETFQAFLDANGTAARVVYARHYAENSTFVNPNISSVTNPTQLPAACAIQVNAT
TDVNTHFSFGVFLPDTWNERFFHAAQEGEDINWVDMAVGLRYGFASVATDTGHTGSD
MDGDWWKNPESINDWAWRANHLTSVLGKFLVERFYGQKPKYNYFAGCSTGGRQAM
KETQKFPRDYDGVIAACPAWWTTHQQLYNLKQTTYQAPQGSDHTIPMEMFAVIGAEAI
RQCDPQDGLVDNIISDPLGCHFDYLPLLCTETKNTSCLTGPQIETLHKIYNDWVETNQTF
VYTHPHYGSEAMWNESGVIGNGSVGNEDSQLWYPQHILGLANFTVQDFDYSVVQYAD
QTDPGNSSANSYDLAEFYRHGGKLIHYHGHSDAIVPPGIGLYYRDHMAATLAPQGIDID
DFYRFFAIPGMEHCYHTPSTMGAPWYIAGPTQASTLSSSLYSVPGYQDAKHDIFLALIN
WVENGTAPDYIIGSRFNSSVEYTSRTIDKQRPICAYPKLAQYVGSGDIDDAANFECKAL
Y

>jgi_Aspni7_1153767_fgenesh1_pm.chr_602_124

QAIATSRCTPETFTPHLSSNASLVYAVPYGPNDTFVGPLGNYSGLPEYCAVYVNVSSSA
TSSYEFGLWLPINTWNKRYMAYGNGGLTGMIAFADMAPGLFYGFATVSTNTGHNSSV
TEAGDAGWALNDPETRTDWGWRALHGSVALGKELTEAFYSEDISYSYYAGCSTGGRQ
GLKSVQMFPDDFDGVLAGACAWWTTHQQNWDLKLATYNLPNNASHHITETQVDALA
TEVLRQCDEVDGVKDNIIMNGYACDFRPETLLCDGPSANKSTCFTADQIDTVYKIYSPW
VETNDSFVFPSFAYGAETQFAQMIMTDDDEQDSPSGIAYPRDFVYNDANWPWQNLDY
STIELSDRLDPGNATAGDFDLRPFKNKGGKLIHYHGFADGEIPTGSSIYYYKQVQKMLT
PLGIDLDDFYRFFLIPGMEHCLDSVHDAPWYIASANQPGGLTGTNVWGVPGYRDRKHD
AILALMAWVENGIAPDELIATKFVDDTPAEGVQAQRPLCVYPKEAKYVGGNWNKTSA
FECR

>jgi_Aspni7_1180848_estExt_Genemark1.C_chr_101_t10424

KSTTPLLAAAGVAAAKNCSVDTIAKYLPKNATVFYANHYEKGYNFTPPIEYNYGLTSD
PMGISAYELPLAGCVAQANISLPNNTQHSVGLVLPDEWNGRFMAVGNGEFAGSVGWS

SIINTMWYGFASVSTDTGHEGNNGSFGYHNEAALTNWGYRALHDAVVNGKKVTEGY
YGKDISYSYYRGCSAGGKQGFKEVEMFPDDFDGVVAGAPAWWTSHQQLWNVLTAIW
NLPETADYHVSDAQMTAVQDEILKQCDPQDGLKDNILQNPFGCVFDPVPVMCNATSSN
DTCVTPAQLKTVNKLFNPWYEANDTLIFPGYTLGTEVGAPSLDDDFVTYIQYMLQIGG
DWTWKDWNPDLVALSDKINPGNATADDFDISPFYKKGGKLLHYHGYSDPSIATGSSVY
LYNHIQEALRPQDIPIDDFYRFFLIPGMEHCTGTPSDQDAPYYMNGDSQAASLSGTVFG
VPGFNDPKHDLVLAIMNWVENGTAPDYLIPTKFKNDDVADGVDKQRPICPHPQLAKY
KGSGDVDKAENWYCGTLY

>jgi_Aspni7_1210062_MIX9935_4_17

QFKCSLPAGDICAYENATTYAAVNAPNCTTSGLQAFLNNLPGLEGTTLSVPYAFYIDSN
ATLDEFTDYVGATTFTTDIPSVCAFRVNGTNEQGATWGLGALLPTHFNGSIMMAAPDS
GLSWGVASAGLRYAFATFASNSGHDEPDYIPGWETPNGLLDWSHRALHLSTVTAKAIV
KAWYDVEPKYSYITGCSDGGRQVMKSIQTYPEDYDGAMAGAPTWWMTHQAFYNYK
QTTIGGSARSNSSIPPSLYPVITDEVLRQCDPQDNLTDSIISNPRGCFFNPEALACSSNKTS
NCLTAPQLDTLHKIYNDWTTFDQDLIYPGVWFGSEPTWNTTFGSPHSKSWYIENVLGF
KNFTAQDLTYEIIAEADARDPAHANADDFDLSPFFNRGGKFFHWHGMSDSVVSPGSSV
YYHHKATYAALEKGIDIDQHYQLYLIPGLEHCTGTPSCMAALWYLAGPYQAGNFFNST
PANVVDNYVHDSFLTLISWVEGGHAPSYMVTTNFEDNADPTTLLHNRKICPYPKQAN
WTKTGVSIHEDYWDCIYV

>jgi_Aspni7_1090214_e_gw1.501.133.1

SNVSICSPAAIQTPVVYGAEILSLSAAWVTNYTDYIPSSFNYNGGTVDLENAKFCNITVK
YTHPGYEDNITVETWLPEAANWNGRLQATGGGGWAAGRFVLSEFFMGGAIGEGFATT
TTDAGLGKDTTGPREWALTSPGNVDWVAVENFGSRAFSDQAIIGKSLINSFYDRAPEYS
YWSGCSQGGRQGMMIAERYPTAYDGIAASAPAQSFTKFTSSLYYPLLMRIWHNVKPLV
CELDFLTSEAIKYCDPLDGVVDGLISNMTACDYNPYTAVNKTFTCDSLNRTIALSQGAA
LIAEAAWSGAHTTDGHQLWFGYNPGSDIGSTFGAQPETNISSLATTKDEWFNLFVAKNI
SFNTMGLSHEEYQEFFNLITSEYGSAWNADDANLQAFKKNGGKLLTYHGMADPSIPTK
GTEYLYNKARALFPDVQDFWRFFESPGLGHCAGGLGGQPTTVMKALQRWVENGTAP
DTLPVEYPSLGTALHRNLCPYPSQIEYIGGNISLAESFRCT

>jgi_Aspni7_1112728_estExt_Genewise1.C_chr_801_t10084

WQPKQQPCGSLTLPHIPGVDIISMTRQEVHNYSVATSPPYNLADVHNLDFCNVSVTLTH
PGEDDTVSVSVWLPLNNWNGRFQATGGGGLAAGIIGLAPAHPVSQGYATAATDGGLT
LNGTSNPQTGAWILRPDGSINTALLLNFAHRSIYDMTIIGKTLAERFYGSPPRYSYWSGC
STGGRQGYFAAAKYPNLFDGVLAGAPALNFPRLIGYMFWPPVHMFHSAAPPQCVFDTF
WKAIIDECDPLDGATDGLISDYNPQSCPFNPETLVGHTVTCPEMGSDSPVTITAQHATLV
KQILQGPDLQDHPDLWTGLPPGASFRGTANTQVINGSIVRPVPFFPIIGWIKNFVYRDPD
YNVFDMTFDDFNTAYRLTLDGYNGILGSDDLNLSEFRRAGGKLLTWHGLADELIPASW
TNAFWEGIDKTDGADVDEFYRVFLAPGLGHCSAGHGPKPVDLLGALVRWVEEGIAPD
RLSAAAVKADGKEVTRELCRYPATLVVYIFFAVLTRKSKRPPLPPGPRRKPIVGNLWDL
PDPSQQDWQHWLKHKDRYGPISSLSIMGQTIIVLNDARLAVELLESRSSIHSSRPQQHFA
EMAGWNNVLGAVKQSQRFRATRKNLHREIGSNVSVARFNEIQTAEVGRFLLRVLDAP
DKLMKHIRKEAGAIILKVGYGYTIEPHDQDPLVDLADKAMMDFSMAMLPATWAVDFF

PPLKYLPSWFPGTEFMKIAQRYRKNVTAFSDIPYAFVKEQMRTGRFVPSFLSNLLESSDL
EPGSEEENTVKWSAGSLYAGGADTTVSSIASFFLAMALFPEVQRKAQQELDTVIGTDRL
PQYADREQLPYINALVKETFRWHPVVPMSLTHTSTADDVCEGYFIPKGSSVLANIWAFT
HDPAAYHDPMSFKPERFLGPKPERDPHFLVFGFGRRVCPGRTLADVNVYLTVAQALAV
FEISKPVENGKVKDVQPEFLPGVISHPAPFDVSIRPRSAKHLELLRSLEQKYPWEKSNAE
DLKNI

>jgi_Aspni7_1149971_fgenesh1_pm.chr_202_535

YEYLAGYILAQYEPIIPRLEAVKRGYVKAKPLLPKNSSRMQHLRKGEHLASEMIVAAN
NVVAARSGFDHRYPTAYDDLATIQALALLATILSTFYAHFANGYHIPKPNIPGLKILSITA
FPLRDYTITGVQVNPIQSHTVNISFWNVTVTYTHPGWDDLIHVHVWVPLSGWNNRLQA
VGGGGWAGLLDYGTLALPVYQGFAAAGSDMGHDRNPWSAESWALEDASGHVNFAQ
LIDFFSTCLNELSLLAKYVIQQIYGLPPAYSYWNGCSTGGRQGLEIAQRWPDAFDGILAG
APTINWAQFVPANYWPTFVMNQLGVYPSPCVMEKITDAAVASCDGQDGVIDGIIAVPS
LCNFNPLSLIGMEVWCECIPSIITAEVALLAESYWAGLTAEDGTPLWYGFNQGTPWNYT
ICTSTDDIGNHRCHSLPDYLSVDWLQLFVAQDPELDLVDMDQADLESLFHQSVAGYKA
ITDANNPDLYDFKVRGGKIIHWHGLADWQVPPNGSVDYYRRVQLRDPSVQDFYLYFE
APGVEHCGHGEGLAPTGLMESLMAWVEGGVAPEYLRAVSKNGRMSRILCPYPAIAQY
VEGDPTVASSFVCQ

>jgi_Aspni7_1150671_fgenesh1_pm.chr_302_196

SRCLPAFIKKPDVSGASVINIEAHEAHNFSAVSLAPGSNEGGLYTISFCNVTVTHTHPGW
NDTVHTQVWLPLEGWNGRFQALGGGGYSLDLGTTYITYAVAMGFASASTDGGLPAG
NGKDSSIPTDLSWALSDENNVNWQLLENYASKATNDMAMIGQQIIKSYYNKSASYSYF
AGCSGGGRQGLLMAQKFPDVFDGILAVSPAINMPAFIPAGQWASLVMREIGFFPSPCEI
KAFTQEAVKACDHLDGVEDGIISYPDSCHVKAADFVGKNYTCDGVQKAFTASSAQVIQ
AAWSGSRSVSERYGWYGVNKDASIGASYVSTECSANITCHSSGSDLFGSWLKYLVAKD
STFSSSNMTRNEFFEALHTSIADYTSMLATNDPDLSKFKANGGKMITWHGLADEVIPPN
GTVAYYDEVLKNDPNAHDFYRFFEAPGVGHCYGGLGPCPNGAMSQLIEWVEKDHAPS
VLHATKGNNNTERALCPYPLQQKFIGGDPRNATSFTCAKRY

>jgi_Aspni7_1187426_estExt_Genemark1.C_chr_701_t10120

SSPCTEGTFSSLNLSNIDILSLNVTAARNYATDASLSSGTTSAPVGQMPSTVDICIVSLQY
THPGQNDIVNTYIGLPLNDSNWNSRFLMDGGGGWVAGGLDEIIAPVASGYSSSSTDGG
HNSTASTADWGFVSEGNTNWPALWDFSSVALGEAAVLGKLATEIYFGSPPKYSYWNG
CSTGGRQGHMMAQRFPTYFDGIVGGSPAINWDKFQLAEFWPAFLAQLLDTQPPACVL
DAFTDAAIDACDLLDGVKDDIISLPGQCHFEASSIVGQTVNCSDPDGQIAITNKMAELV
QGMWDGPRSLEDQFEWYGLGYDADLAALLTTTCTSVDNCTVTPFSISADWARIFLARN
SSFSIQGLTRQDFDDLYRESVDQYASVIGTRNLDLTGFKLAGGKMISWHGMQDELIPTN
GTVDYYSRVMELDPSVADYYRFFLAPGVEHCEGGNGFDPNDYVFETLRAWVENSTVP
DTLEATAVAVAGSNSSSTRTAYLCPYPKIFTYVGGDPNDASSFSCV

>jgi_Aspni7_1082298_e_gw1.202.1754.1

AGLSATAAHAASLADVCTTSNVKAALPSSVYGLTMIPSSVTASPVYNASTSGQVFFPDA
TYDYCGVTFNYTHNGRGDTVKLQYWLPAPESFENRFLATGGMAYMITGGSSYLPGGV
MYGAVAGTTDGGFGGELDTAFLLANGTINYEALYSMGYHGIGELTIVGKEFTKNFYSM
GDEKLYTYYQGCSEGGREGWSQVQKYPGVYDGVIPGAPAIRYGQQQANHLYSNVVE
KTLGYYPPPCELEKIVNATIDACDSLDGKVDGVVARTDLCQLHFNINSTIGLPYHCAAS
SSSSIGLNYGKRSTDPAINGTVSAQGVAVAAEILKGLHDSKGRRAYISYQPTASFDDAET
SYNSETGEYELSIASSGGEWVAKFLELRDADNLSTLDNVTYDTLRDWMELGWKRYED
VMQTTWPDLTEFEKAGGKIITFHGESDQSIPTGSSVHFYNSVRSIMYPDMSYNASVAAM
GDWYRLFFVPGAAHCATNDLQANGPFPQTNLAVMINWVEKGVVPHTLNATHLAGEW
EGQNAQLCSWPLRPLWTENGTKFNCVYDQASVDTWDYNFDSYKVPLY

>jgi_Aspni7_1098047_e_gw1.802.689.1

KSLADVCTVSHVQSSLPLDAFTGLEIKTSSVTASPVYNVSVSDQTFYPDGTFDYCNVTF
AYSHAGRDDSILLGIWLPAPDAYQNRWLSTGGGGFAINLEGDLMPAGLLYGAAAGITD
GGFGAFTTDLDAVFPSANGTINWDAIYNFGYQSHHELSVIGKAFTKNFFNTTASSSKLY
SYYMGCSEGGREGWSQVQRFPEEWDGAIIGAPAIRYGQQQVNHLFPNVVEQTLGYYPP
PCELEKIMNLTIAACDELDGRKDGVVSRTDLCQLHFNLNSTIGKPYYCAAETTSENLKK
RSAHAKRLTVSNITPAQNGTVSAKGAEAAATMLKGLHTLDGKRAYIWFQITSAFSDAE
TQYNSDTGSYELDIDALSGEWVARFLELVEADNLSSLDNVTYDTLRDWMELGWRRYG
DSLQTTYPDLSAFKDAGGKIIHYHGESDPSIPTGSSVHYHESVRKIMYPDLTYNQSTDAL
ASFYRLYLVPGASHCATNAYEPGPWPQDNMAVMIDWVEKGVVPVTLNATYLSGEEK
GSNAQICAWPLRPLWKNNGTVMECVYDQASIDTWQYDFNAYSLPLY

>jgi_Aspni7_1114588_estExt_Genewise1Plus.C_chr_101_t10144

RQYSRAAAAAVALAASARAANSSSLTSLCTVSHVQSALPANGTLLGIDLIPSSVTAKTT
TEDSTEYCAVSVSYVHTGTTNEIVLNYGFPSPADFKNRFYVAGGFGYSLSTSSTGGLEY
GAVGGATDAGYGALNGTTVDEVNLSGNGTINWDPIFMFAYQALGEMTTVGKPLTRAF
YGLDDDAKVYTYYEGCSDGGRQGMSQIQRYGDEYDGAIIGAPAFRYGQQQVNHLFSS
VVEQTLDYYPPTCELEKIVNATIAACDPLDGRTDGVISRTDLCQLDFNLTSIIGEPYHCA
AEVSTSLGFNFNKRSDGTSTSYTPAQNGNVTAEGVAVAQTIYDGLFNTAGERAYLSYRI
GSEFSDGETAYDNSTGTWGVDIPSTGGEFVTRFIELVELSNLETLSSAMVRYMDTLQTT
LPDLTTFKSTGGKVIHYHGENDPSVPTASSVHYWQSVMSIMYPKLGSSKAIEQMKEWY
QLYLVPGAAHCGRNTLQPNGPFPEDNMATMIKWVEQGVQPTGLNATVDAGKYDGEV
QSLCQFPKRPVWKSNKWECESDDEGADTFYYTFPAFKVPIY

>jgi_Aspni7_1123092_estExt_Genewise1Plus.C_chr_4030028

VTTSTTLDDVCTPAYVRSKLPVDNVYPGITISHTNLSANPVYNVSTSGSAVYPATTIDYC
NVTLAYSHNGRDDTVQMRFWLPSPDKFANRYLSTGGADYYVNQGTQQLPGGVMYGA
VSGSTDGGFGGFDVGVDEILLLANGTLDWQNLYMFGYEAHHELATIGKRFTRNFFNM
TVNEEDQKLYSYYQSCSEGGREGWSQVQRYGDQFDGAVIGAPAIRYSFQMTMHLWA
NVVEKTLGYYPSQCEIEAVVNETITACDGMDGRMDGVVSRSDLCKMQFNLTETIGKG
YYCAADGDVPVQNSTISAQAVEVFEKILDGMTDDDGKQVYLSYQPGALFWDAQATY
NTTSGNWGLDINSMGGEFVTKFLQLLDLDNLPTLENVTYETLKEWVQHGWQMYEDSL
HTTWPDLTPFHEAGGKVLHYHGEQDGSIPTASSVHYYDSVRRTMYPTLGYNASHEAL

GQWYRLFLVPGASHCMNNDLQPNGPFPQTSLQTLIEWVEQGVVPVRLNGTVLSGTFEG
ETQEICAWPLRPMWYDNGTSMECEYDQKSLKTWEWDLDAFDLPVY

>jgi_Aspni7_1123274_estExt_Genewise1Plus.C_chr_501_t10189

TSLSDLCTVSNVQSALPSNGTLLGINLIPSAVTANTVTDASSGMGSSGSYDYCNVTVTY
THTGKGDKVVVKYALPAPSDFKNRFYVAGGGGFSLSSDATGGLEYGAASGATDAGYD
AFSYSYDEVVLYGNGSINWDATYMFGYQALGEMTKIAKPLTRGFYGLSSDKKIYTYYE
GCSDGGREGMSQVQRWGDEYDGVIAGAPAFRFAQQQVHHVFPATIEHTMDYYPPPCE
LDKIVNATIEACDPLDGRTDGVVSRTDLCMLNFNLTSIIGEPYYCAAENYTSLGFGFSKR
AEGSTTSYQPAQNGSVTAEGVALAQAIYDGLHDSNGKRAYLSWQIAAELSDGDTEYDS
TTDSWTLSIPSTGGEYVTKFVQLLNIDNLENLDNVTYDTLVEWMNIGMIRYIDSLQTTIP
DLTTFQKSGGKMIHYHGESDPSIPTASSVHYWQAVRQAMYPNTTYTQSLKEMSDWYQ
LYLVPGAAHCGTNDLQPGPYPEDNMEIMIDWVENGNKPSRLNATVSSGYYAGETQML
CQWPSRPLWTSNSSFSCVHDSKSLATWDYTFDAFKMPVF

>jgi_Aspni7_1125571_estExt_Genewise1Plus.C_chr_601_t20364

ATLEDVCTTAHAKAAVPVSGLVQGVVTNPSSITAAPVYNESTAGSDYYPAGTYDYCNI
TMTYSHAGKDDSVLLQFWMPTPDDFQNRWVSTGGFGFAINVASNVPAGLPYGAVAG
RTDGGFGSFDTQFSSVYPLVNGTANYDALYMFGYQAHHEMSSIGKAFTKNFYSMSDK
LYAYYQGCSEGGREGWSQVQRFPEEWDGAVIGAPAIRYGQQQVNHLVPQVVEQTLDY
YPENCEFSEMVTLIIDACDELDGRKDGVISRTDLCQLNFDLKSTIGKPYACNASTSMGV
TTPAQNGTITAEGVRVAQTILDGLKDSEGRQAYFWYRIGSEFTDAETSYDSTTGTYGIDI
SELGGEWVTHGLQLLDLDNLSNLDNVTYDTLRDWMLEGLQRYEDVLQTTWPDLSRF
QAAGGKVIHYHGESDNSIPPASSVRYFESVRDTMFPNKTYNASVEALNEFYRLYLVPG
ASHCNTNSLQPNGPYPQNSLGDLMNWVEKGIEPVTLNGTVLSGDYEGEEQKICSWPLR
PLWKNNGTVMDCVYDQKSIDTWLYDLDAYDIPIY

>jgi_Aspni7_1138908_fgenesh1_pg.chr_701_295

TTLQDVCTVSFARNSLPADSVYPGVSIDRSSVTAVAAYNVSSTDQVYWPDATFDYCNV
TFAYSHLGRDDRVVVSYWMPAPSKFANRYLSTGGSGYDITLGSTSLPGGVMYGAVAG
TTDGGFGGFENGSDDVWLLANGTLNYETVYMFGYQAIHELTVIGKELTKRFFNTSSIYS
YYQGCSEGGRDGFSQIQRFAEDFDGAVTGAPGFRFTFQQMQHLWSPVVEQTLDYYPPP
CELEKIVNETIKACDALDGKEDGVVSRTDLCKLHFDANSTIGEPYSCGASSLTTYKRRK
RSETYELPAQNGTVSAKGVEVAKEILRGMHDTKGNQVYFSYQPSAAFSDAETHYDYD
TGSWTLDITESNSEWVEKYINLLNATEMPTLRDVTYDNLKEWVRIGWQKYDDTLQAT
YPDLSPYSEAGGKVLHYHGESDYTVPTASSVHYYESVRRIMYPDMSTNASYAALGDW
YRLFLVPGAAHCGPDDSQPNGPWPQKTLETMIRWVEDGVFPETLNATVLQGEHEGEE
QKLCAWPLRPLWTNNGTSMECVNDPEGMKTWVYDLNAFKMPVY

>jgi_Aspni7_1138949_fgenesh1_pg.chr_701_336

LESSLASVCTTSYIQSTLPADGTLMGITIKPSSVTAGAVYNNSIASNAFFPATTVDYCGV
TFNYTHKGRDDNVQLTYWLPSPANFENRFLATGGEAYNINEDSMILAGGVMYGAAAG
LTDGGWGSTSFDEVFLLANGTINWTDTYMFGYQAIHELTLIGREFTRNFYGRSKDTKIY
TYYQGCSEGGREGWSQAQRFGEDYDGIITGAPAFHYAQQQINHLYSNVVEQTMNYYP

PPCELEKIVNATIEACDPLDGKADGVIARSDLCKLHFDLSKLVGTPYACAPSATGGIGLN
YGKRQSGSSSMNPAQNGTISAQGVAVAKKILDGLKDSQGRRGYLSYQPGADFDDAQT
AYDDATENWTLKIASSGGEFVQKFVHLVDANNLANLTDVTYDTLIEWMQLAWNRYQ
DSLQTGNPDLSTIQKAGSRILHFHGESDPSIPTGSSVHYYEAVRQIMFPHLGVEEGASQL
NDFYRLFLVPGGAHCAANPEQPNGPMPNTNMAVLIDWVEKGVTPITLNATILQGEREG
ENQQLCAWPKRPLWKTGTMECVFDRPSYETWIYDFDAYKVPLY

>jgi_Aspni7_1149420_fgenesh1_pm.chr_201_162

RNLVPLAAAIGAAHGVSLSDICNDAYIKAALPADGTFQGISLISSSVSTTIHSNVSISSSNF
YPAATISYCDVAFNYTHIGRNDRVALTYWLPAPADFRNRFLTTGGEAYEINEGSGTSGS
LPGGLMYGAVAGLTDGGFNGQSFDDVFLLANGTVDWQNTFMFGYQGIHEMMEIGRQ
LTRNVYNTGDDKVYTYYQSCSEGGREGWSQAQRYGYDYDGIIVGAPAFRFAHQQINH
LVPNVVEQTMNYYPPPCELELIVNATIAACDPLDGRADGVVARSDLCKLHFNYTSLIGA
PYYCAASSGGSSLGLGFSKRQSMSSTPAQNGTVTKEAVELVQTLLAGLKDTQGRQAYL
YFQPGADFDDAETTYDSETGTWGLSIDGNGGEFVAKFVNLQDVDSLTTLDNVTYDTIR
DWMQMGWMMYEDTLMAHNPDLSVLQQSGGKILHFHGEQDPSVPTASSAHYYEAVR
KIMFPNESLNTSAAALGEFYRLYLVPGMAHCGTNSEQPNGPFPQTNMAVMIDWVENG
VVPVTLNATVLQGENEGEHQQICGWPLRPLWSNNGTTMDCVFDQASYDTWIYDFDAY
KLPLY

>jgi_Aspni7_1153838_fgenesh1_pm.chr_602_195

ATPSTLAELCTDSIVKAALPPSEFIQGITIDSDSVTTEVVTNSSVSSEFYPSATINYCNVTF
AYSHDGIDGDQVLLEIWLPAPTDFQNRWLSTGGGGYAINSGDQSLPGGVMYGAASGM
TDGGFGGFSNNADTAMLLANGTLDYETLYMFAYKAHRELSLIGKALTRNVYGMSDSD
KLYAYYQGCSEGGREGWSQVQRFGDEWDGAIIGAPAFRWSFQQTQHLYSNVVEKTLD
YYPPPCELDKIVNETIAACDAMDGKVDWVVARTDLCLLDFDISTIEGKPYSCAASRGTP
AQNGTVSAKGIEVAKTIINGLHDSQGRRVYFSYQPTAAFDDAETQYNSTTGQWGLDVD
QLGGEYIALLVDKNGTTLDSLDGVTYDTLKDWMISGLQEYYSTLQTTWPDLTPFHEAG
GKVIHFHGDADFSIPTAASIRYWESVRSIMYPNQDYNSSAEALNEWYRLYTVPGAGHC
ATNDAMPNGPFPQTNMAVMIDWVENGVVPTTLNATVLQGENEGQNQQLCAWPLRPL
WTNNGTTMECVYNQRSIDSWHYDLDAVPMPVY

>jgi_Aspni7_1185597_estExt_Genemark1.C_chr_5010172

TSLADVCTSSYARSVAPSSPLMGVTVDTTSITANPVYNYSSASTAFWPASTFDFCNVTL
AYTHDGLDDQVLLQFWLPAPADFKNRWLSTGGFGYAINGDSADLPGGVMYGAASGQ
TDGGFGSFSTEFDAVFLKANGTIDRQALYMFGYQAHYELSAVGKAFTKNFFNLGSSKL
YSYYQGCSEGGREGWSQVQRFGEEWDGAVIGAPAIRYGQQQPNHLYPGLVEYTMGY
YPPPCELEKINNLTIAACDALDGKKDGVVARTDLCKLHFNVNSTIGAPYYCPAETTTTV
LKKRLSTSNTTPAQNGTVTKEGAAVAAKILDGLKTLDGKRAYIWYQPSATFDDAETEY
NSDTDSWELDITSLGGEWVAKFLELRDVDNLSTLDNVTYDVMTNWMSLGWQRYEDV
LQTTWPDLTPFHSNGGKIIHVHGESDPSIPTGSSVHYHESVRKTMYSSLSFNESSDALND
WNRLYLVPGAAHCSYSTDQPNGGWPQSTLPTLFDWVENGQKPDRLNATVQEGENYG
QVQQLCAWPLRPYYVNNGTELTCVFDEESYQTWVYDFDAYELPLY

>jgi_Aspni7_172914_Aspni5.e_gw1.2.1410.1

TGELPESLLKPFFSSLLATPEHLSHYENRAGVLAGATTIALSQAASLADVCTTSHAKSSL
PSSDEIEGLEIDTSSITAGAVYNASTSGSYFFPAASYSYCNVTLSYTRPGLDTPFLLELWL
PAPEDYQNRWLSTGGGGFAINSGSSSLPGGVMYGAAAGITDGGFGGFDTQFDAVFLEA
NGTINYEALYSFGYLAHHELTVVGKAFSRNFYSVGDSKIYAYYQGCSEGGREGWSQV
QRYADQWDGAVIDAPAFRYGQQQVNHLYPDVVEHTLGYYPPTCELEKIVNLTIAACD
ALDGRVDGVVARTDLCKLHFNVNSTVGAPYSCPASTSTTVLKKRTSMSNTTPAQNGT
VSAIGAAAASKMLDGLRTLDGRRAYIWYQPSSTFDDAQTQYNYDTGKWELDITSLGG
EWVARFLELRDADNLSSLDNVTYDTLKEWMELGWQRYEDSLQTTWPDLTPFQSAGG
KIIHVHGESDPSIPTGSSVHYHESVRQIMYPGMSFNESSEALNEWNRLYLIPGAAHCASS
TTQPNGPFPQTTLETLIQWVEGGVFPTRLNATVLAGEREGEQQQLCAWPLRPLWKNNG
TVMDCVFDEESYKTWVYDFDAYKLPLY

>TAN_ABX89592_*Aspergillus_niger*

MRQHSRSVRCSGSSTANAASLSDVCTVSNVQSALPSNGTLLGIDLIPSAVTANTVTDAT
AGMGSTTTYDYCNVTVTYTHTGKGDQVVVKYAFPAPSDFENRFYVAGGGGFSLSSDA
TGGLEYGAASGATDAGYDAFSYSYDEVVLYGNGFINWDATNMFGYQALGEMTKIGK
PLTQGFYGLSSDQKIYTYYEGCSDGGREGMSQVHRWGDEYDGVIAGARPSALAQQQG
GPRLPGTIGHTMDYYPPPCELEKIVTETISACDPLDGRTDGVVSRTDLCMLNFNLTSLIG
TPYYCAAQNYTSLGFGLRKRKAKGSTTSYQPAQNGTITRRVALAQAYDGLHDSEGKR
AYLSWQIAAELFDADTTYDSTTDSWTLDIPSTGGEYVTKFVQLLNIDNLENLDNVTYDT
LVDWMNTGMIRYIDSLQTTVVDLTDFQKSGGKMIHYHGESDPSIPTASSVHYWQAVR
QAMYPNVTYTDSLKRMADWYQLYLIPGAAHCGTNSLQPGPYPEDNMEIMINWVENG
VKPSRLNATVSSGTYEGETQMLCQCPLVPLWTSASTPPSPPVPMLVRPRCSASGLPVPS
GPATPASLVSMTPSLLLLGTTLLTLSRCPFFK

>TAN_BAA09656_*Aspergillus_oryzae*

ASFTDVCTVSNVKAALPANGTLLGISMLPSAVTANPLYNQSAGMGSTTTYDYCNVTVA
YTHTGKGDKVVIKYAFPKPSDYENRFYVAGGGGFSLSSDATGGLAYGAVGGATDAGY
DAFDNSYDEVVLYGNGTINWDATYMFAYQALGEMTRIGKYITKGFYGQSSDSKVYTY
YEGCSDGGREGMSQVQRWGEEYDGAITGAPAFRFAQQQVHHVFSSEVEQTLDYYPPP
CELKKIVNATIAACDPLDGRTDGVVSRTDLCKLNFNLTSIIGEPYYCAAGTSTSLGFGFS
NGKRSNVKRQAEGSTTSYQPAQNGTVTARGVAVAQAIYDGLHNSKGERAYLSWQIAS
ELSDAETEYNSDTGKWELNIPSTGGEYVTKFIQLLNLDNLSDLNNVTYDTLVDWMNTG
MVRYMDSLQTTLPDLTPFQSSGGKLLHYHGESDPSIPAASSVHYWQAVRSVMYGDKT
EEEALEALEDWYQFYLIPGAAHCGTNSLQPGPYPENNMEIMIDWVENGNKPSRLNATV
SSGTYAGETQMLCQWPKRPLWRGNSSFDCVNDEKSIDSWTYEFPAFKVPVY

>TAN_XP_748839_*Aspergillus_fumigatus*

ASPSTLEELCTVSYLQTVLPSSKFIQGITIDSDSLTTSVVTNSTVSSVDYPTATIDYCNVTL
AYSHDGIDGDRVLLQIWLPAPTDFQNRWLSTGGGGYAINSGTRMLPEGIIYGAASGLTD
GGFGGFSVNADSAMLLANGTLNYEALYMFGYKAHRELSLIGKAFTRKVYGMADSEKL
YAYYHGCSEGGREGWSQVQRYGDEWDGAIIGAPAFRWSFQQTQHLFSNIVEKTLDYY
PPPCELEKIVNETIVACDPLDGRTDGVVARTDLCLLHFDLKHVIGKKYSCAASTTAPAQ
SGTVSAKAVEVAKTIINGLHDTQGRRVYFSYQPSAAFDDAQTQFNADTGKWELSINQL
GGKHIALLMNKNSTTLDSLNGVTYDTLKDWIISGMQEYYSTLQTTWPDLTPFHQAGGK

VIHYHGDADFSIPTASSIRYWESVRSTMYGNLSYKAGANALNEWYRLYTVPGAGHCST
NDAMPNGPWPQTNLATMVEWVEKGVTPVTLNATVLQGEYEGETQQLCAWPLRPLW
KNKGKTLNCVYDQASINSWHYDLDAVPMPVY

>*Pe*FaeA_*Pleurotus_eryngii*

QAPQVQVGNTAVIGRSIPEFGQELFGGIPFAEPPVGQLRLSNPVLKTRLGTPTFDASNFG
PACLQSASVPLMSEDCLRINVLRPAGIPTGVVLPVMAWVYGGGFDFGDSSIYNASAIVA
QSVIRGTPVVFVSLNYRLGPLGFPQGVEAQKRGALNLGLKDQLAALEWVQANIGLFGG
DKSKVTIFGQSAGSISLSILFLNSNIKRLARAAIFESGFTATSLNFPASHREQSWANFVKD
VPQCASTAGSKDTFSCLRSDSIDEATLLKAGSLADDQSGELFAWDPTIDGPGGILPDIPS
KLLARGQFARLPFIAGTVLDEGTTFTPKFITTEDQIRQSIIANFTPSPFGPAVLAKSAETIL
QLYPDVPALGSPFGTGNETFGLSSQYKRAAAIFGDVSFQSQRRFWIQTLSKAGLKTFGY
LFADPQSSDPVNGVPHASEIPYVYGALGILGGTVTPQALALSRIMVDYWVSFATSLDPN
DGKGLPRPLWTQYTPSNQAIMLLNSTGTTMIPDDYRKKQIDFINSNPAVWHHRRSFST

>*Ps*FAE_*Pleurotus_sapidus*

QAPQVQVGNTAIIGRSIPEFSKEFFGGIPFAEPPVGQLRLNNPVLKTRLDTPTFDAGNYGP
ACLQPDPSVPVISEDCLRINVLRTAGVPAGVLLPVMAWVYGGGFDFGDSSIYNASAIVA
QSALRGTPVVFVSLNYRLGPLGFPKGQEAHDRGALNLGLKDQLAALEWVQANIGFFG
GDKSKVTVFGQSAGSICLSIHFLNSPIQSLARAAIFESGFPATSQNFPAAHREKSWTNFV
KDIPQCASTAGSKDTFSCLRSDTIDQATLLQAGALADDQSGELFAWDPTIDGPGGILPDI
PSKLLAKGKFVRLPFIAGTVLDEATTFTPKTVTTEDQIRQSLIANFTPSPFGPTVLAKDVE
TILKLYPDIPALGSPFGTGNETFGLSSQYKRTAAIFGDLSFQSQRRSWIQTLSKAGVKTFG
YLFTDPQSSNPANGVSHASEIPYVYGAPGIFGGTVTPEAIALSRIMVDYWVSFATSLDPN
DGKGFPRPVWTQYTPSNQAIMLLNSTGTSMIPDDYRKKQIDFINSNPAVWHHRRSF

>jgi_Aspni7_1079885_e_gw1.101.706.1

QNADTPTSAPTVQVRNGTYEGLYNPTYNQDLFLGIPYAQPPVGELRFRPPQPLNTTWTG
TRNATAYYNECIGYGSDDWYWTDVVSEDCLALSVIRPHGIDSSAKLPVVFWMHGGEF
AEGGTRDSRYNLSYIVQQSQEMQSPIIGVTVNYRLSGWGFLYSQEVADEGSANLGLRD
QRHALYWLQENIASFGGDPSRLTIWGQSAGANSVGLHLVAYDGQNDGIFRAGIAESGS
VPSLAAYMSAEDAQPYYDAVVNATNCTGSSNTLTCLREVPTDVLSSIFNSSLVAGAGY
HPVIDGDFLRASGIVNLQTGQFAKTPLLIGTNFDEGTKYAPHGYNTTDQFVSLVQANGT
NYTSALTIASLYPDDPAVGIPGTLQGRPPPSYGYQWKRVAAFLGDLLMHAPRRVTTQW
LAHWNVPAYVYHWNVMTLGPLDGAAHGYEVPFSFHNYDGLGDERGNDSVTWPQLST
MMSRMWVSFINHLDPNYSNMTDIHWPVYTTETPQNMVFDVNVTGLAYVEPDTYRAE
GIAYITSILQSAFNR

>jgi_Aspni7_1169907_gm1.636_g

STLDTSNTPTIKRADAGSNTSSIPTATLNNTVFVGRSLPEFEQELFLGIKFADEPVRFTPST
LKTVYRANDSDNGVYHASSASGLQTSSGTVLYNATEYGYDCPGYGSDETELAEEGYA
RFDENCMNLNIIRPKREKEDELLPVMIWIFGGGWVQGATADPRYNMSYIVRQGALNDK
PVLGVSINYRVAAFGFLDSVEVMESGNTNLGLRDQRVAMHWVKQNIKAFGGDPDKITI
WGESAGAYSVGAHLVTNDGDNEGLFRAAIMESGNAVGPPYNGTDWYQPMYDQIVNA

TNCTTSSNTLQCLREVPFSTIYTAADIGLEWFATIDGTFIKEYPQISITEGRFAKVPILHGT
NTDEGVSFGTTGVNTDAEAIQQLMASKRWVLNETQATTLLSHYPNISALGCPYGWGN
TTWPKLGYEYKRYESMAGDLCMVAPRRLLSQKMKEYEEQVFAYRWDVAALNDSSTI
GVAHFAEIPFVFANPVQNITPLGSDPARLELGNLAARMWTAFVTDLDPNGHGVSGIPH
WPKYNLTDPRDFVFRLPRDGSYVEKDTFRTGGIDYINTIVR

>GE_3PIC_*Hypocrea_jecorina*

QQTSGAGGATCSALPGSITLRSNAKLNDLFTMFNGDKVTTKDKFSCRQAEMSELIQRY
ELGTLPGRPSTLTASFSGNTLTINCGEAGKSISFTVTITYPSSGTAPYPAIIGYGGGSLPAP
AGVAMINFNNDNIAAQVNTGSRGQGKFYDLYGSSHSAGAMTAWAWGVSRVIDALEL
VPGARIDTTKIGVTGCSRNGKGAMVAGAFEKRIVLTLPQESGAGGSACWRISDYLKSQ
GANIQTASEIIGEDPWFSTTFNSYVNQVPVLPFDHHSLAALIAPRGLFVIDNNIDWLGPQ
SCFGCMTAAHMAWQALGVSDHMGYSQIGAHAHCAFPSNQQSQLTAFVQKFLLGQST
NTAIFQSDFSANQSQWIDWTTPTLS

>GE_4G4G_*Sporotrichum_thermophile*

SMAPMNHIFERQDTMVHLTSALLVAGAAFAAAAPMNHIFERQDTCSVSDNYPTVNSA
KLPDPFTTASGEKVTTKDQFECRRAEINKILQQYELGEYPGPPDSVEASLSGNSITVRVT
VGSKSISFSASIRKPSGAGPFPAIIGIGGASIPIPSNVATITFNNDEFGAQMGSGSRGQGKF
YDLFGRDHSAGSLTAWAWGVDRLIDGLEQVGAQASGIDTKRLGVTGCSRNGKGAFIT
GALVDRIALTIPQESGAGGAACWRISDQQKAAGANIQTAAQIITENPWFSRNFDPHVNSI
TSVPQDHHLLAALIVPRGLAVFENNIDWLGPVSTTGCMAAGRLIYKAYGVPNNMGFSL
VGGHNHCQFPSSQNQDLNSYINYFLLGQGSPSGVEHSDVNVNVAEWAPWGAGAPTLA
LEQKLISEEDLNSAVDHHHHHH

>GE_AFM93784_*Phanerochaete_carnosa*

LPQTSSKEAQSFGCSTPANIPFNDDKLPDPFLFNDGTPVRSLTDWSCRRQQLASLIQGYE
AGTLPPKPPIVTSTFSQNGLTGNLTVTAGFPGNTTTFSSPVTFPNGTVPTEGWPLLIAYSG
LSIPIPDGIAVLTYDNSAIGEQNDQTSRGVGQFFDVYGHNATASAMSAWVWGVSRIIDV
LEVTPAAHVNTAKIAVTGCSRDGKGALMAGAFEERIALTIPQESGSGGDTCWRLSKFE
QDSGDVVQQATEIVQENVWFSTNFDNFVFNISVLPYDHHSLAGLIAPRPMISYENTDFE
WLSPLSGFGCMTAAHPIWEAMGVPDNHGFVQVGNHSHCEFPSDLNPTLFAFFDKFLLG
KEANTTIFETNEVFNGTVWNPSQWINWTTPTLSH

>*Coprinopsis cinerea*

ASTPGCGRSPPSSGTKNIGNRQYILQVPANYDPNRPYKLIFGFHWLGGNMNNVAPGYY
GLRALANESAIFVAPNGLDAGWANTGGRDTTFVNQMLTELKNSMCVDDTQIFATGFS
YGGGMSYSLACSNPRDFRAVSVIAGAQLSGCSGGETRVAYLGIHGVVDSVLNVSAGRQ
LRDRYLRVNGCQSKNAPEPPRGPDSNYMKTEYSCAAGYPVWWIAHGGDHVAQPSGQ
NWMATQTWDFWTRAIYEGVENPNPPTNPPPTQPTQPPTQPTNPPPGNCAARYGQCGG
QGWNGPSCCQSGASCQVVNQWYHQCL

Appendix B

DNA sequence of *As*FaeE fused with N-terminal His tag

TCAAGTGAATTCCACCATCACCATCACCATGCCACCTTGTCACAAGTCCTTGACTTC
GGCAACAATCCCGGCGATAACGAGATGTGGATATATGTACCAGATCAACTTGCCGC
TAACCCTGCCGTGATCGTGGCTTTGCACGGATGCCTTGGCTCCGCTGAGGGCTACTA
CTCCGAAGTTCAGGATTTGCCACCTGCCGCTGACGAAAATGGTTTCATATTAGTGTA
TCCCGGTTCAAACGACGATTTCCACTGCTGGGACGTGGCAACGGCAGAATCATTGA
CGCACGATGGAGGTTCAGACAGTAGATCTATTGTCAATATGGTACAATACACCCTA
GATAAGTACAGTGGAGATTCATCAAAAGTATTTACAACTGGCAGTTCCTCAGGTGC
CATGATGTCTTTGGTTCTTGCCGCTGCCTACCCCGACGTCTTCTCTGGAGTAGCCGC
TTATTCCGGAGTTCCATATGGTTGTTTACGTGGCTCTCCAGGCTCTTCCCCCTTCACT
GCTGATCAGGCATGTGCCAATGGAGAGGTTAGTAGAACTGCTCAGGAATGGAAAG
ACGAGGTTAAAATGGCTTGGCCCGGCTACAATGGAACTTACCCTAAAGTTCAGGTT
TGGCATGGAACCGCCGATTCCGTCATTTCACCAAATAATTTTGACGAAGAGGTAAA
GCAATGGTCCGCCGTCTTTGGCGTCAATGTGACAAAGGAGGAGCAAGACTCTCCAC
TTGACGGATATACGCGTAGTATATTCGGCGATGGATCTCACTTCGAGGCTTACTTGG
CAGAAGGCGTGGGCCACGTAGTGCCTACCCAGGTCGATTCAACGCTTAGATGGTTC
GGCCTAATCTGAGCGGCCGCTCACG

Appendix C

Amino acid sequences of FAEs in CE1 used for phylogenetic analysis in Chapter 3

>*Aa*FaeD_jgi|Acral2|1082309_*Acremonium_alcalophilum*

QNTAGCGQNPPSSGVKSINVGGMNREYILQLPNNYDPNKGHMLIFGLHWLSGSMHDV
HPNYYGLRQLAGNNAIFISPNGINNGWANDGGRDVNFIDAILQQVRSQLCINDSQIFAT
GFSFGGGMSYALGCARANVFRAIAPIAGAQISGCSGGTSPIAFLGIHGTNDDVLPIAMGR
QVRDRFLQNNGCQPKNAPEPGWGQGPIKTEYSCQPNYPVTWIAFSGGHDPNQSFVGRE
IWDFFSQFGGTNPGPSPPPTSPPPTNPPPTNPPPTNPPPTNPPPSGGCSAMWGQCGGMGW
NGPQCCSQGTCQQQNPWYSQCL

>*Cl*FaeA1_XP_003666339.1_*Chrysosporium_lucknowense*

ASAGCGKAPPSSGTKSMTVNGKQRQYILQLPNNYDANKAHRVVIGYHWRDGSMNDV
ANGGFYDLRSRAGDSTIFVAPNGLNAGWANVGGEDITFTDQIVDMLKNDLCVDETQFF
ATGWSYGGAMSHSVACSRPDVFKAVAVIAGAQLSGCAGGTTPVAYLGIHGAADNVLP
IDLGRQLRDKWLQTNGCNYQGAQDPAPGQQAHIKTTYSCSRAPVTWIGHGGGHVPDP
TGNNGVKFAPQETWDFFDAAVGAAGAQSPMT

>*Cl*FaeA2_AEP33617.1_*Chrysosporium_lucknowense*

TPSPGCGKTPTLITDGSATTPLTLTSNGKTRRFYVKLPDDYDNSHPYRLIFALHALGGTA
QQVTTGTGGYLPWYGIPDLAANDTVGAVYVAPDGLNNGWANQGGEDVAFLEAVME
TVEQDVCVDRDLRFSTGFSYGAAMSYTLACALGRRIRAVAVLSGSPVISGGCAGAGSG
ASEPVAYYGQHGMSDPVLPVAGGREMRDHFVRTNGCDAGRGPPREPARGSGTHVKT
VYDGCDPDYPVVWNAFDGDHTPQPVDRGATTTFSAVETWEFFSQFK

>*An*FaeC_XP_001395336.1_*Aspergillus_niger*

ANSPGCGKNPTLANGVHQINGREYTLKIPDDYDANNPYHLIFGLHWRGGNMDNVVSG
DSIQPWYGLESRAQGSAIFIAPNGLNAGWANTNGEDVAFIDAIMEQVESDLCVDQSSRF
ATGFSWGGGMSYSLACSRAKEFRAVSVLSGGVISGCDGGNDPIAYLGIHGINDPVLPFD
GGVELAERFVGNNGCQPASIEKPQSGSNGWKRTDFYGCSKPVSFIAYDGGHDGAPLGV
QSSLAPDATWEFFMAA

>*Anid*FAEC_XP_662871.1_*Aspergillus_nidulans*

ANSPGCGKQPTLTNGVNQINGREYVLKIPDGYDPSKPHHLIFGLHWRGGNMYNVVNG
DSIQPWYGLEARAQGSAIFVAPNGLNAGWANTNGEDVAFIDAIMEQVEDDLCVDQAS
RFATGFSWGGGMSYALACARAAEFRAVSVLSGGLISGCDGGNDPIAYLGIHGINDPVLP
LDGGVTLANTFVSNNGCQPTDIGQPASGSGGSVRTDFSGCSHPVSFIAYDGGHDGAPLG
VGSSLAPDATWEFFMAA

>*As*FaeC_jgi|Aspsy1|154482_*Aspergillus_sydowii*

ANSPGCGKQPTLASGVQYINDREYILKIPDNYDPAKPHHLVFGLHWRGGNMNSVVDG
QSVEPWYGLETRAQGSAILVAPNGRDAGWANPNGEDVALIDAIIEQVEGDLCVDQSSR
FSTGFSWGGGMSYALACARANEFRAVSVLSGGVISGCDGGEDPIAYLGIHGINDPVLPF
DGGVTLAQQFAANNGCSPTDIGRPDAGSAGSVRTDFTGCSRPVSFVAYDGGHDAAPLG
VGNPLAPDATWEFFMAA

>*Nc*FaeD_XP_956228.1_*Neurospora_crassa*

APSSGCGKGPTLRNGQTVTTNINGKSRRYTVRLPDNYNQNNPYRLIFLWHPLGSSMQKI
IQGEDPNRGGVLPYYGLPPLDTSKSAIYVVPDGLNAGWANQNGEDVSFFDNILQTVSD
GLCIDTNLVFSTGFSYGGGMSFSLACSRANKVRAVAVISGAQLSGCAGGNDPVAYYAQ
HGTSDGVLNVAMGRQLRDRFVRNNGCQPANGEVQPGSGGRSTRVEYQGCQQGKDVV
WVVHGGDHNPSQRDPGQNDPFAPRNTWEFFSRFN

>*Cg*Fxe1_XP_001224572.1_*Chaetomium_globosum*

ATPSKGCGNAPKLITDAAATTPLKLTSNGKNREFYVKLPANYDNTHPYRLIFTLHALGG
NAQQVTVGTGGYLPWYGIPALANDTLGAVYIAPNGLNNGWANQGGEDITFLKSVMDT
VEADLCIDQNLRFSTGFSYGAAMSYSLACSLGKQIRAVAVLSGNPQISGCTGGTEPVAY
YGQHGTKDSVLPIAGGRQMRDRFVKNNGCAAKQPAEPAAGGKHIKTAYEGCGEYPVV
WNAFDGDHTPQPVDPGTSGTFSAVESWAFFEQFT

>*Cs*Fae1_*Corynascus_sepedonium*

ASAGCGKNPPTSGTKSMTVNGKQRDYILQVPNNYDSSKAHRVVIGYHWLDGSMNDV
ANGGFYDLQSRAGDSTIFVAPNGLNRGWANQGGEDITFTDQLVEMLLNDLCVDEEQFF
ATGWSYGGAMSHSVACSRPDVFKAVAVISGAQLSGCNGGATPVPYLGIHGAADNVLPI
DLGRQLRDKWLQTNGCTSKNAQDPSAGQQAHIKTEYSCSNAPVVWIGHGGGHVPDPT
GNNGVKFAPEETWNFFNAAV

>*Cs*Fae2_*Corynascus_sepedonium*

ASAGCGKAPPSSGTKSMTVNGKQRQYILQLPSNYDSSKAHRVVIGYHWRDGSMNDVA
SGGFYGLRSLAGDSTIFVAPDGLNKGWANSGGEDITFTDQIVAMLKNDLCVNEGEFFA
TGWSYGGAMSHSAACSRPDVFKAVAVIAGAQLSGCSGGTTPVAYLGIHGAADDVLGIS
LGRQLRDKWLQTNGCASKSAPEPSSGQQSHIKTTYSCSLAPVTWIAHGGGHVPDPSGS
NGVKFAPGETWSFFNAAVSGGGGGPGTSAPASTTPTTSPGTQPTTNPTNCAPMYGQCG
GNGWTGATCCQSGSTCRYSNDWYSQCL

>*Ct*FaeB_*Chaetomium_thermophilum*

ASAGCGKAPPSSGVKTVTVNGRQRQYILQLPNNYDRNKPHRVVFGFHWLGGSMNDV
ANNNFYDLRSRGASNTIFVAPNGLDAGWANTGGQDIAFVDQIVASLKNDLCVDEGQFF
ATGFSYGGAMSHSVACSRPNVFKSVSVIAGALLSGCDGGNTPVAYMGIHGAADNVLPI
AMGRQLRDKWLQTNGCQNKNAPEPSAWNSVNGLPGEGGQQTHIKTVYECSRAPVVW
IATGNGHVGDPTGTGGVKFAPGETWDFGSNPNPTPTTIVTSVIPTPTSVIVQPTQPPANC
SPRWGQCGGQGWQGPTCCEAGTTCQSFNPWYSQCL

>*Al*AXE_OJZ85059.1_*Aspergillus_luchuensis*

SGSLQQVTDFGDNPTNVGMYIYVPNNLASNPGIVVAIHYCTGTGPGYYGDSPYATLSE
QYGFIVIYPSSPYSGGCWDVSSQATLTHNGGGNSNSIANMVTWTISKYGADSSKVFVTG
SSSGAMMTNVMAATYPELFAAATVYSGVSAGCFYSNTNQVDGWNSTCAQGDVITTPE
HWASIAEAMYSGYSGSRPRMQIYHGSIDTTLYPQNYYETCKQWAGVFGYDYSAPEKT
EANTPQTNYETTIWGDSLQGIFATGVGHTVPIHGDKDMEWFGFA

>*Ao*AXE_BAD12626.1_*Aspergillus_oryzae*

RAIHNGRSLIPRAGSLEQVTDFGDNPSNVKMYIYVPTNLASNPGIIVAIHYCTGTAQAYY
QGSPYAQLAETHGFIVIYPESPYEGTCWDVSSQATLTHNGGGNSNSIANMVTWTTKQY
NADSSKVFVTGTSSGAMMTNVMAATYPNLFAAGVAYAGVPAGCFLSTADQPDAWNS
TCAQGQSITTPEHWASIAEAMYPDYSGSRPKMQIYHGNVDTTLYPQNYEETCKQWAG
VFGYNYDAPESTESNTPEANWSRTTWGPNLQGILAGGVGHNIQIHGDEDMKWFGFTN

>*An*AxeA_CAK46215.1_*Aspergillus_niger*

HVAKRSGSLQQITDFGDNPTGVGMYIYVPNNLASNPGIVVAIHYCTGTGPGYYSNSFYA
TLSEQYGFIVIYPSSPYSGGCWDVSSQATLTHNGGGNSNSIANMVTWTISEYGADSKKV
FVTGSSSGAMMTNVMAATYPELFAAGTVYSGVSAGCFYSDTNQVDGWNSTCAQGDV
ITTPEHWASIAEAMYPGYSGSRPKMQIYHGSVDTTLYPQNYYETCKQWAGVFGYDYS
APESTEANTPQTNYETTIWGDNLQGIFATGVGHTVPIHGDKDMEWFGFA

>*Ac*Axe1_EAW06435.1_*Aspergillus_clavatus*

ASVLESRSSALLPRAGSLQQVTNFGDNPTNVGMYIYVPNNLASNPGIIVAIHYCTGTAE
AYYNGSPYAKLAEKHGFIVIYPESPYQGKCWDVSSRASLTHNGGGNSNSIANMVKWTI
KKYKTNTSKVFVTGSSSGAMMTNVMAATYPDMFAAGVVYSGVAAGCFMSNTNQQA
AWNSTCAHGKSIATPEAWAHVAKAMYPGYDGPRPRMQIYHGSADTTLYPQNYQETC
KEWAGVFGYDYNAPRSVENNKPQANYKTTTWGKELQGIYATGVGHTVPINGDRDMA
WFGFAK

>*Mt*Axe3_ADZ98864.1_*Myceliophthora thermophila*

HPVFDELMRPTAPLVRPRAALQQVTNFGSNPSNTKMFIYVPDKLAPNPPIIVAIHYCTGT
AQAYYSGSPYARLADQKGFIVIYPESPYSGTCWDVSSRAALTHNGGGDSNSIANMVTY
TLEKYNGDASKVFVTGSSSGAMMTNVMAAAYPELFAAGIAYSGVPAGCFYSQSGGTN
ARNSSCANGQINSTPQVWAKMVFDMYPEYDGPRPKMQIYHGSADGTLRPSNYNETIK
QWCGVFGFDYTRPDTTQANSPQAGYTTYTWGEQQLVGIYAQGVGHTVPIRGSDDMAF
FGL

>*Ct*AxeA_ADZ98864.1_*Chaetomium thermophilum*

HPVFDELMRPTAPLVRPRAALQQVTNFGSNPSNTKMFIYVPDKLAPNPPIIVAIHYCTGT
AQAYYSGSPYARLADQKGFIVIYPESPYSGTCWDVSSRAALTHNGGGDSNSIANMVTY
TLEKYNGDASKVFVTGSSSGAMMTNVMAAAYPELFAAGIAYSGVPAGCFYSQSGGTN
ARNSSCANGQINSTPQVWAKMVFDMYPEYDGPRPKMQIYHGSADGTLRPSNYNETIK
QWCGVFGFDYTRPDTTQANSPQAGYTTYTWGEQQLVGIYAQGVGHTVPIRGSDDMAF
FGL

>*As*FaeE_XP_040699905.1_*Aspergillus_sydowii*

ATLSQVLDFGNNPGDNEMWIYVPDQLAANPAVIVALHGCLGSAEGYYSEVQDLPPAA
DENGFILVYPGSNDDFHCWDVATAESLTHDGGSDSRSIVNMVQYTLDKYSGDSSKVFT
TGSSSGAMMSLVLAAAYPDVFSGVAAYSGVPYGCLRGSPGSSPFTADQACANGEVSRT
AQEWKDEVKMAWPGYNGTYPKVQVWHGTADSVISPNNFDEEVKQWSAVFGVNVTK
EEQDSPLDGYTRSIFGDGSHFEAYLAEGVGHVVPTQVDSTLRWFGLI

>*Cl*FaeB1_API68922.1_*Chrysosporium_lucknowense*

ASLQPVTNFGDNPTGLQMYVYVPDKVAVSPAIIVALHPCGGSAQGWYSQTRLPSYADQ
LGFILIYAGTTKMSNCWDVQNPASLTHNGGGDAGGIVSMVKYALKQYNGDASRVYV

MGGSSGAMMTNVLAGSYPDVFEAGAAFSGVAHACFLGADSATPFSPNQTCAQGRIQR
SAREWGDLVRNSFPAYDGRRPRMQIFHGNADFLVHPECAHQALAQWADVLGLQLTQT
NKGVPSAEYTQEVYGDGTQLQGFFGDGVGHIAPVNEPVMLRFFGLMN

>*Cl*FaeB2_XP_003667253.1_*Chrysosporium_lucknowense*

ASLQEVTEFGDNPTNIQMYIYVPDQLDTNPPVIVALHPCGGSAQQWFSGTQLPSYADD
NGFILIYPSTPHMSNCWDIQNPDTLTHGQGGDALGIVSMVNYTLDKHSGDSSRVYAMG
FSSGGMMTNQLAGSYPDVFEAGAVYSGVAFGCAAGAESATPFSPNQTCAQGLQKTAQ
EWGDFVRNAYAGYTGRRPRMQIFHGLEDTLVRPQCAEEALKQWSNVLGVELTQEVSG
VPSPGWTQKIYGDGTQLQGFFGQGIGHQSTVNEQQLLQWFGLI

>Fae7262_AXF48686.1_*Talaromyces_wortmannii*

ASLTQVNNFGDNPGSLQMYIYVPNTLASKPAVIVAMHPCGGSATEYYGMYDYHSPAD
QYGYILIYPSATRDYNCFDAYSSSSLTHNGGSDSLSIVNMVKYVISTYGADSSKVYMTG
SSSGAIMTNVLAGAYPDVFAAGSAFSGMPYACLYGAGAADPIMSNQTCSQGQIQHTGQ
QWAAYVHNGYPGYTGRYPRLQMWHGTADNVISYADLGQEISQWTTVMGLSFTGNQT
NTPLSGYTKMVYGDGSQFQAYSAAGVGHFVPTDVSVVLDWFGITSGTTTTTTSKTTSA
TTSTTSSAPSSTGGCTAAHWAQCGGIGYTGCTACVSPYTCQKSNDYYSQCL

>*Ct*FxeA_*Chaetomium_thermophilum*

ASLQQVWNFGSNPTNINMHIYVPDRLANKPPIIVALHPCGGNAQQWFSGTRLPQYADR
HGFILIYPSTPHMSNCWDVQNPASLTHGAGGDALGIVSMVNYAIDRYGADRDRVYVM
GFSSGGMMTNVLAGSYPDVFEAGAAYSGVPHACFLGAPAATPFSPNQTCAQGLQKTPE
EWGNLVRNSYPGYNGRRPRMQITHGLNDWLVRPQCAYESLKQWSNVLGLQLTRQVT
SGQWTQHIYGDGTKLVGYLGQGIGHEPAVNEEQLLRFFGIIN

>*Cs*Fxe1_*Corynascus_sepedonium*

ASLQEVTNFGSNPSNIQMYIYVPDQLAANPPVIVALHPCGGNAQQWFGGTQLPSYADS
HGFILIYPSTSHMSNCWDIQNPDTLTHGQGGDALGVVNMVNYALETYSGDSSRVYAM
GFSSGGMMTNQLAGSYPDVFEAGAAYSGVAFGCSAGAESATPGSANQTCAQGLQHTE
QEWGDFVRNAYPGYTGRRPRMQIFHGLADTLVRPQCAEEALKQWSNVLGVEFTQEVS
GVPSAGWTQKIYGDGTQLQGFFGQGIGHQSTVNEQQLLEWFGLL

>*Pa*FaeD_jgi|Podan2|8825_ *Podospora_anserina*

APSAPGANDIPRDVLTQISSWGSNPTNLQLHLYAPSNLTGKPAIILALHGCFGSGPGHAE
MTSQFQTLSSSRNFVVLYPSSINDNNCWDVASPASLTRDGGGDSTGLATIVRWATTTFN
ADPKKVFITGSSSGCMMSNVMASAYPDLFSAVSCYSGVPAGCLAGSPGSSPISADQTCA
NGGIRKTGEEWAEVVRGMAPLPEGKGKGKGKGKGKGNQGWKYPKVATWHGDNDFF
VNYFHNFEEQLKQWGAIHGVEFTRNETNVPAAGYTKMVYGDGTKLVGYSASGVGHV
VPQFEAVDLEWFGL

>*Am*CE1_ADZ47894.1_*Anaeromyces_mucronatus*

MSKLQISNTCPDKYRTKQEGVEYPTAKKITYYSKVTETERKMNVILPVGYDENKKYPV
VYYLHGLMSYEDSMLEDDSTLAIPTNLLKEGRAKEMIIVLPDVYAPKPGTAVTPDFNPE
YYKGYDNFINELIEVIMPYMEEHYSILTGRENTALCGFSMGARTSLYIGYMRSDLIGYV
GAFAPAPGITPGEDSFSGKHEGLISEDEFRAEIQPIVSLIDCGTNDSVVGQFPKSYHEILTR
NNQEHIWFEVPGADHDWNAISAGFYNFIQTTFGALNN

>alpha/beta-hydrolase_ORX42573.1_*Piromyces_finnis*

MSKLKVVNTCPSDAITKQGGVEYPTPKKVTYHSKTTNTDRPLNIILPSGYSESKKYPVL
YLLHGIMGNEDSMLETGMGSITIPANLAKSGKSKEMIIVLPNEYAPEGGKEVEPAFNQE
YFDGYDNFINDLVDSIMPYIESNYSVLTGRENTAICGFSMGGRNSLYIGFKRSDLFGYVG
AFSPAPGVTPGTDFSGYHKGLFTEDSFRAQVAPIVTLISCGTSDSVVGQFPKSYHEILTR
NNQEHIWFEVPGADHDGTAVSAGFYNFISAAFGALN

>*Um*EstD_XP_011391122.1_*Ustilago_maydis*

MTLTKKSCNKVFNGTLTKYALTSTSLGSLETAINVFLPSCASESTRVPVLYYLSGLTCTE
DNAAQKGGFFGAAEEEGIAIVFPDTSPRGAGVQGEDDSYDFGTGAGFYLNATKQPWSK
HYNMYDYIVKELPAILSHHLPIDTSRSSIFGHSMGGHGALTLYLKNRSQYRSASAFSPIC
NPTQCAWGEKAFNGYLENGLHDGAQYDATLLLKSTDIQPDILIDSGTADDFYKQRQLL
PENFETVANEKRFHNVTVRLHDGYDHSYYFISTFASEHVRYHAKFLKQ

>*Sc*CE1_XP_003038790.1_*Schizophyllum_commune*

MSVKLETVSTNKTFEGTLTKYKFKSAALGGLDAQFNLFIPPIASKHKVPVLFYLAGLTC
TEDNGAQKGGFLGVAATQGIAIIFPDTSPRGAGIEGEDADWDFGTGAGFYVNATNPKY
SKHYNMLTHVTLELPQVIEAAGLPIDFKRQSVMGHSMGGHGALTMYLASQTKQFRSA
SAFAPIANPTKCPWGEKAFSGYLQGGTEEAKATYDATELVAKHKDPVHILIDYGTGDN
FYKQGQLLPENFLKAARDAGYDEVQVRVRSQDDYDHSYYFISTFASDHVHFHANFLKE
A

>jgi|Botbo1|29930|fgenesh1_kg.8_#_299_#_Locus4402v1rpkm38.46

MSFQHAALEELSRNHAFAGYIVKYKYKSKTLGDLSTQFNIFLPKEAESASVPVLYYLAG
LTCTEDTGAWKGGFLRDAAEEGIALVFPDTSPRGANVPGEDDSWDFGTGAGFYLNATE
PKFSKHYNMFSFITEELPEVLKSSGLPIDPHRQSIFGHSMGGHGAITLHLLKPGQYRSAS
AFSPILNPTQAPWGHKAFGGYLAGGTEEGKKWDATELISTAKGRRLNIRIDYGDADQF
YQQKQLLPENFVAAARAAGFGEDDVQVQSHSGYDHSYYFISTFGREHIQHHAKYLKA

>jgi|Dacsp1|95872|estExt_fgenesh1_pm.C_140217

MPVTKVKTNKCFSGHLTKFSAHSESYGLETKFNVFVPEGEGPFSVLYYLAGLTCDEDT
GAMKGTFESPAAKHGIALVFPDTSPRGSNSPNDTLSWDFGTGAGFYIDASHPSYAAYQ
GYTYVTHELPELLRETGLALDLERQGIMGHSMGGHGALTLYLKSGQYRSCSAFAPITNP
RLCPWGKKAFAGYLKGGVEEGKEWDATELLGLSRVKEGLRILFDWGDADEFYLSGQL
LPENFLNKVKELGVEGVQGRKREGYDHSYYFIMSFAEEHVAFHAKWLKA

>jgi|Jaaar1|169115|estExt_Genewise1Plus.C_2_t30083

MSLEKLSSNKAFHGDLIKYKFKSTALGGLDSQFNLFLPANAAQGKVPVLFYLSGLTCTE
DNGAQKGAFLKDAADEGIAIIFPDTSPRGAGIEGEDTDWDFGTGAGFYVDATNSKYAK
HYNMATHVSVELPQVIEAAVIPIDFSRQSIMGHSMGGHGALTLYLKNLKQYRSASAFAP
ISNPTKCPWGEKAFKGYLQGGLEEAAPKYDATELIAKVKENVHILIDYGTADNFYKQG
QLLPENFLKAAREAGHDEVQVRVRPQEGYDHSYFFISTFAADHVHFNANFLKA

>jgi|Fomme1|120874|estExt_Genewise1.C_5_t10463

MAVSAELKEEAKNKVFEGTLTKYSFKSEALGGLKANFNLFLPANASASAKVPLLTYLA
GLTCTEDNGAQKGGFFKDAASEGIAILFPDTSPRGAGVEGEDDDWDFGTGAGFYLDAT

ATKYAKNYRMLTHVTKELPSVLESSGIPIDLSRQSIFGHSMGGHGALTIYLASPPGTYRS
ASAFAPISNPTKCPWGEKAFKGYLAGGLEEASKRYDATELIRGVKEPVHILIDYGTGDN
FYKQGQLLPENFLKAAREAGHDEVEIRVRSQEDYDHSYYFISTFAPDHVHFHANFLKA

>jgi|Bjead1_1|34143|fgenesh1_kg.1_#_945_#_Locus732v1_medCvg1115.6s

MSLEKLSSNKAFEGELIKYKFKSEVLGGLNTQFNLFLPANISQGKVPVLVYLSGLTCTE
DNAAQKGNFLGPAASEGIAILFPDTSPRGAGIEGEADSWDFGVGAGFYVDATNPKYAK
HYNMLTHVTVELPTVIQAAGIPVDFSRQSIFGHSMGGHGALTIYLNSLRKGSKAYRSVS
AFAPVSNPVNCPWGQKAFTGFLQGGIEEAKEQYDATELIKKLDGKPVNVLIDYGTGDN
FYKQGQILPENFLKAARDTGYDEVQVRVRRQEGYDHSYYFMSTFGADHIHFHANFLK
A

>jgi|Glotr1_1|68109|estExt_Genewise1Plus.C_00001_t30120

MALETLSANKSFGGVLTKYKFKSAALGGLDTQFNLFLPANASQGKVPVLFYLAGLTCT
EDNGAQKGGFLRDAADEGIAIIFPDTSPRGAGIEGEDKDWDFGVGAGFYLNATNPKFA
KHYNMATHITLELPEVIEAAGIPIDFSRQSIFGHSMGGHGALTLYLAALSKGSKSRPFRS
ASAFSPIANPTKCPWGQKAFNGYLQGGVDEAAPLYDATEMITKIKDNVHILIDYGTGD
NFYKQGQLLPENFLKAARDAGHDEVQVRVRPQDGYDHSYYFISTFGPDHVHFHANFL
KA

>jgi|Aspcl1|4279|7000001156851589

AGSAGCEKPLPKHHGAGGSYPTDFTTSDGTKRSYIIHIPSSYDVNQAAPLIFSFHGRSKT
AENQEALSQFSNEEWNPNAIAVYPQGLDKQWQGDPASSGVDDIAFTMEMLDHFEQRY
CIDSSRVYAAGKSNGGGFTDLLACDATASRRIAAFAPVSGAYYQDVSEDACDAETVPI
DCSPGRNPIPVIEFHGTADKTIPYAGGGRRGECLPSVPHFVREWSKRDGFGLQNKTTNF
YDGNVQKYQYGRGNALGTVTHYRIKGLGHDWPSVAPNSDNPDGTYLNATPIIMDFFG
RWTLP

>jgi|Aspor1|1205|AO090005000945_mRNA

AGSAGCGQPLPPSQNPGGSSYGVNFTLSAGTQRFYRIHIPSNYNVNTPTPLIFSFHGRGK
TAESQEKLSQFSNEDWNPDAIAVYPQGLNKEWQGDPHSKDVDDIAFTMEMLDYFQEK
FCIDSTRVYAAGKSNGGGFTNLLACDPTASTRIAAFAPVSGAYYQDVSEEACHPTTVPI
KCSPGRPSIPILEFHGTADKTIPYGGGGRRGECLPSIPHFVREWSKRLGYGLHNTTKELY
DGNVQEYQYGNGDNRGIVTHYRIGGLGHDWPSVRPNSDNPNGTYLDATPIIVKFFKQW
VLPGSSEDNQSEAR

>jgi|CocheC5_3|1100730|e_gw1.7.316.1

CDKLLPDGVLPGESTSLTIHSESGVTPRGYRLHIPTSYEKNTPVPLILSFHGRGKTAEYQE
ALSQFSNASYGFEGIAVYPEGVPNKKGTQQFQGDPDAPKSIDDVAFTLELLDHLETTYC
IDTSRIYATGKSNGGGFTGILACDAEATKRIAAFAPVSGAFYLDANQQPPPCNPSRKPVP
LMEFHGYRDTTIPYKGGINTRGNANSSNVVTYVNDWAERNGYDACVNTTTYLCDGER
KVARHSWDDVVVHYNYTNIEHDWPSSFPNGDTDDLLTCEEAEATSVILEWFKKWTI

>jgi|Botci1|13853|BC1T_02367

MVSLKNAIHAWSAHNLLRSVTNSREDPSCTKAIPSDITIGQPKNISITSSDVHRSYLIVVP
PLYTSQPSTPVIFSFHGGHRNASQQLALDQISNPEFNDFAITIYPQGVKGKWEGNPGNTA
NDTQLVSDIIDSLDKTYCIDNKRIWATGKSDGGGFCNTLACHPIMSTKIAAFAPVSGAFY

IDTQPCFPNNVTIPCSPGRLKIPMLEFHGGDDSTIRYYGQANRSNQCIPTVPHWVHQWA
LRDELGLKNVTTDLTNDTVVYSYGKGEDEGLVQHVFDRKLGHDWPSTVLNADLIGHG
ERPASFDATPMILDFFQNHILP

>jgi|Stano2|2954|SNOG_09365.3

LSIRDTKSGCGSALPEGLVPGKSKNLTLSSNSGTPTRKYRLHLPANYDGTKKLPVILSFH
GRTQDAKYQEKLSQFSNASYGFEGISVYPQGVNFITDDGKEVPQWQGDPDADPKISDV
KFTLELIDHLESTFCIDKSRIYAAGKSNGGGFTGLLACDPVASTRIAAFAPVSGAFYLNK
TTQQLPDCKPGRLPIPIAEFHGIEDKTIKYKGALNSRKNANSTNIPAYVNAWADRDHFK
ADQNVTSTLCSGKKSVTKHTWGDETVVHYAYKNLEHDWPSSFFNLDTDDKTSLLTCK
DAEATSLILDWFARWTL

>jgi|Trire2|72072|kg2.C_scaffold_19000042

AAIDTVKRASAGCSRTHDWAGQTKEFSFASSGGTRSYRIHLPSNYQPGSPKPLIIAYHGS
GDNPANFEKTTRFSDPSVNPNMIVVYPAGINGNWQGPTYATPGVSDKVFTTDLVNRIK
DNYCVDEARVYATGHSNGGGFVNTLACSPDHGGQFAAFAPVSGAMYTDVNGNDNCH
PARSPLPMFETHGTADTTIPYNPTGPGRGGPLPSIPDWLSRWQGRNQCTTSTTTDLAAG
VHDTRWKCAGIDSLLRHVKIDGAGHAWPGPGSQIDISPQVIQFLSVNFKP

>jgi|Glalo1|12472|GLAREA_08027m.01

TQSPGCGKPLPKAQQPAGGASHKVQFKQTNGTPRTYLIHIPKGYSSDKPAPLIFSFHGRG
KTASSQEELSQFSNEKWNPDGIAVYPQGIENAWQGAPYSKGVDDIEFVTDMIDHITDRY
CVNPKRIYAAGKSNGGGFTGTLACSALSSKIAAFAPVSAAFYIPLPKDTSVCEPEFIDIPC
KPSRSVPILEFHGSQDDTIAYAGGARSGECLPSVARWVKAWSVRDGLGTCANETVLYG
GKVRKWEYGGGKVTGYLTEGLGHAWPSLEANGDNPNGTYFDATPVIMEWFGRWSL

>*An*FaeA_CAA70510.1_*Aspergillus_niger*

ASTQGISEDLYNRLVEMATISQAAYADLCNIPSTIIKGEKIYNAQTDINGWILRDDTSKEI
ITVFRGTGSDTNLQLDTNYTLTPFDTLPQCNDCEVHGGYYIGWISVQDQVESLVKQQAS
QYPDYALTVTGHSLGASMAALTAAQLSATYDNVRLYTFGEPRSGNQAFASYMNDAFQ
VSSPETTQYFRVTHSNDGIPNLPPADEGYAHGGVEYWSVDPYSAQNTFVCTGDEVQCC
EAQGGQGVNDAHTTYFGMTSGACTW

>*At*FAE1_XP_001217493.1_*Aspergillus_terreus*

ASTQGISEDLYNRLVEMATISQAAYANMCNIPSTITVGEKIYNAQTDINGWVLRDDSTK
EIITVFRGTGSDTNLQLDTNYTLTPFSTFSECSGCEVHGGYFIGWSSVQDQVMSLVKEQ
ADQYPDYTLTVTGHSLGASMATLAAAQLSGTYDNITLYTFGEPRSGNEAFASYMNDKF
TATSADTTKYFRVTHSNDGIPNLPPAEQGYVHGGVEYWSVDPYSAQNTYVCTGDEVQ
CCEAQGGQGVNDAHTTYFGMTSGACTW

>*Ao*FaeA_AHZ18111.1_*Aspergillus_oryzae*

AITQGISEGTYSRIVEMATISQAAYANLCNIPSTITSAGKIYNAETDINGWVLRDDSRQEII
TVFRGTGSDTNLQLDTNYTQAPFDTLPQCSGCAVHGGYYVGWISVKDQVEGLVQQQA
SQYPDYSLVITGHSLGASMAAITAAQLSATYNNITVYTFGEPRTGNQAYASYVDETFQA
TNPDATKFYRVTHTNDGIPNLPPTSQGYVHHGTEYWSVEPHGPQNMYLCLGDEIQCCE
AQGGQGVNDAHVTYFGMASGACTW

>*Ps*FaeA_OKO98639.1_*Penicillium_subrubescens*

VTRGISDTVYSRLIEMATISQAAYANLCKIPTTITTVQKIYNAQTDINGWVLRDDSRKEII
TVFRGTGSDVNLQLDTNYTLASFQTNPKCAGCSVHGGYYLGWVSVKDQVESLVQTQA
SKYPDYALTVTGHSLGASIAAITAAQLSATYNGVKLYTFGEPRTGNQAFASYINKQFQS
TDPNTTQYFRVTHANDGIPNLPPTAQGYAHSGVEFWSVDPYSAYNTYICTGTQDQCCE
AQGGQGVNAAHFTYFGMSSGACTW

>*Pc*FaeA_AGR85377.1_*Penicillium chrysogenum*

AVTKGVSENVYNRLVEMATISQAAYANLCKIPSTITTVGKIYNAQTDINGWVLRDDSR
KEIITVFRGTGSNVNLQLDTNYTPAAFATNPLCKGCFVHGGYYKGWLSVKDQVESLVH
QQASQHQDYALTVTGHSLGASMAAITAAQLAATYSNVNLYTFGEPRTGNQAFASYMN
DHFHASSPATTRYFRTTHTNDGIPNLPPTSQGYVSSGIEYWSVEPHNAHNVFVCTGNQN
QCCEAQGGQGVNAAHVTYFGMSSGACSW

Appendix D

DNA sequence of *Aa*FaeD with N-terminal His tag

TCAAGTGAATTCCACCATCACCATCACCATCAAAACACAGCAGGTTGCGGCCAAAA
CCCACCTAGTTCAGGCGTCAAAAGTATTAACGTCGGTGGCATGAATAGAGAATACA
TCTTGCAGCTTCCCAATAACTACGACCCAAACAAGGGCCACATGTTAATCTTCGGC
CTACACTGGCTATCCGGCAGTATGCACGACGTTCATCCTAATTATTACGGTCTTCGT
CAGTTGGCTGGAAATAATGCAATATTTATATCTCCTAATGGCATTAACAACGGCTG
GGCTAATGATGGCGGCAGGGACGTAAACTTCATTGATGCAATATTGCAACAGGTAA
GATCTCAACTTTGCATCAACGATTCTCAAATTTTTGCTACCGGATTCTCTTTCGGAG
GTGGAATGTCCTATGCCTTAGGCTGCGCCCGTGCCAATGTCTTCAGAGCCATTGCAC
CAATCGCCGGTGCACAAATAAGTGGATGCAGTGGCGGCACCTCACCTATCGCCTTC
CTGGGTATCCACGGAACTAACGACGACGTTCTGCCAATAGCAATGGGAAGACAGGT
TAGGGATAGGTTTCTACAAAATAATGGCTGTCAGCCCAAGAATGCACCAGAGCCTG
GCTGGGGTCAAGGACCCATAAAAACAGAGTATTCCTGCCAGCCCAACTATCCCGTA
ACATGGATCGCATTTTCAGGAGGTCACGATCCTAACCAGAGTTTTGTCGGAAGGGA
GATCTGGGACTTTTTTAGTCAGTTCGGTGGTACAAACCCAGGTCCTTCTCCCCCACC
TACGTCACCACCCCCAACCAACCCCCCACCTACGAACCCTCCACCCACCAATCCCC
CTCCAACGAATCCCCCACCCAGTGGAGGATGTTCTGCAATGTGGGGTCAATGCGGA
GGCATGGGCTGGAACGGACCACAATGCTGTAGTCAAGGCACCTGTCAACAGCAGA
ACCCCTGGTACTCCCAGTGCCTGTGAGCGGCCGCTCACG

Appendix E

DNA sequence of *Tl*XYN with N-terminal FLAG tag and C-terminal G4S linker

GAAGCTGAATTCGACTACAAAGATGATGACGACAAATTTCCAGCTGGAAACGCTAC
TGAATTAGAAAAGCGTCAAACCACGCCTAATTCCGAGGGCTGGCATGACGGATATT
ACTATTCTTGGTGGTCTGATGGAGGTGCACAAGCAACCTATACAAATCTTGAGGGT
GGTACTTACGAGATCTCTTGGGGCGACGGTGGAAATCTTGTCGGAGGTAAGGGCTG
GAACCCCGGACTGAATGCCCGTGCCATCCACTTTGAAGGCGTGTACCAACCCAATG
GCAATAGTTACTTAGCCGTGTATGGCTGGACCAGGAATCCACTGGTCGAATACTAT
ATCGTCGAAAATTTCGGTACTTACGACCCTTCCTCTGGTGCAACCGACTTGGGTACG
GTGGAATGTGATGGATCAATATATAGACTTGGAAAGACCACCAGGGTGAACGCCCC
CTCTATCGATGGTACACAAACCTTCGATCAATACTGGTCAGTTCGTCAGGACAAGA
GAACATCCGGCACCGTACAGACCGGTTGTCACTTTGATGCATGGGCCAGAGCCGGT
TTGAACGTTAACGGCGACCACTATTACCAGATTGTAGCCACTGAAGGTTATTTTTCC
TCTGGATACGCTAGGATCACGGTAGCAGATGTCGGTGGTGGCGGAGGAAGTGGTG
GAGGAGGTAGTGGCGGCGGAGGCTCTACTAGTCAATTT

DNA sequence of *Aa*FaeD with N-terminal His tag

TCAAGTGAATTCCACCATCACCATCACCATCAAAACACAGCAGGTTGCGGCCAAAA
CCCACCTAGTTCAGGCGTCAAAAGTATTAACGTCGGTGGCATGAATAGAGAATACA
TCTTGCAGCTTCCCAATAACTACGACCCAAACAAGGGCCACATGTTAATCTTCGGC
CTACACTGGCTATCCGGCAGTATGCACGACGTTCATCCTAATTATTACGGTCTTCGT
CAGTTGGCTGGAAATAATGCAATATTTATATCTCCTAATGGCATTAACAACGGCTG
GGCTAATGATGGCGGCAGGGACGTAAACTTCATTGATGCAATATTGCAACAGGTAA
GATCTCAACTTTGCATCAACGATTCTCAAATTTTTGCTACCGGATTCTCTTTCGGAG
GTGGAATGTCCTATGCCTTAGGCTGCGCCCGTGCCAATGTCTTCAGAGCCATTGCAC
CAATCGCCGGTGCACAAATAAGTGGATGCAGTGGCGGCACCTCACCTATCGCCTTC
CTGGGTATCCACGGAACTAACGACGACGTTCTGCCAATAGCAATGGGAAGACAGGT
TAGGGATAGGTTTCTACAAAATAATGGCTGTCAGCCCAAGAATGCACCAGAGCCTG
GCTGGGGTCAAGGACCCATAAAAACAGAGTATTCCTGCCAGCCCAACTATCCCGTA
ACATGGATCGCATTTTCAGGAGGTCACGATCCTAACCAGAGTTTTGTCGGAAGGGA
GATCTGGGACTTTTTTAGTCAGTTCGGTGGTACAAACCCAGGTCCTTCTCCCCCACC
TACGTCACCACCCCCAACCAACCCCCCACCTACGAACCCTCCACCCACCAATCCCC
CTCCAACGAATCCCCCACCCAGTGGAGGATGTTCTGCAATGTGGGGTCAATGCGGA
GGCATGGGCTGGAACGGACCACAATGCTGTAGTCAAGGCACCTGTCAACAGCAGA
ACCCCTGGTACTCCCAGTGCCTGTGAGCGGCCGCTCACG

DNA sequence of *Sc*SED1 from *Saccharomyces cerevisiae* strain Cen-PK2-1C

CAATTTTCCAACAGTACATCTGCTTCTTCCACCGATGTCACTTCCTCCTCTTCCATCT
CCACTTCCTCTGGCTCAGTAACTATCACATCTTCTGAAGCTCCAGAATCCGACAACG
GTACCAGCACAGCTGCACCAACTGAAACCTCAACAGAGGCTCCAACCACTGCTATC
CCAACTAACGGTACCTCTACTGAAGCTCCAACCACTGCTATCCCAACTAACGGTAC

CTCTACTGAAGCTCCAACTGATACTACTACTGAAGCTCCAACCACCGCTCTTCCAAC
TAACGGTACTTCTACTGAAGCTCCAACTGATACTACTACTGAAGCTCCAACCACCG
GTCTTCCAACCAACGGTACCACTTCAGCTTTCCCACCAACTACATCTTTGCCACCAA
GCAACACTACCACCACTCCTCCTTACAACCCATCTACTGACTACACCACTGACTACA
CTGTAGTCACTGAATATACTACTTACTGTCCAGAACCAACCACTTTCACCACAAACG
GTAAGACTTACACCGTCACTGAACCAACCACATTGACTATCACTGACTGTCCATGC
ACCATTGAAAAGCCAACAACCACATCAACCACCGAATACACTGTAGTCACTGAGTA
CACTACTTACTGTCCAGAACCAACCACTTTCACCACAAACGGTAAGACTTACACCG
TCACTGAACCAACCACTTTGACTATCACTGACTGTCCATGTACTATTGAAAAGAGC
GAAGCCCCTGAGTCTTCTGTCCCAGTTACCGAATCTAAGGGCACTACCACCAAAGA
AACAGGTGTTACTACCAAACAAACCACAGCCAACCCAAGTCTAACCGTCTCCACAG
TCGTCCCAGTTTCATCCTCTGCTTCTTCTCATTCCGTTGTCATCAACAGTAACGGTGC
TAACGTCGTCGTTCCAGGTGCTTTAGGTTTGGCTGGTGTTGCTATGTTATTCTTATAA

Amino acid sequence of XYN-SED fusion gene comprising N-terminal FLAG tag (red), *Tl*XYN, G4S linker (underlined), and C-terminal SED1 anchoring domain (blue)

DYKDDDDKFPAGNATELEKRQTTPNSEGWHDGYYYSWWSDGGAQATYTNLEGGTYE
ISWGDGGNLVGGKGWNPGLNARAIHFEGVYQPNGNSYLAVYGWTRNPLVEYYIVENF
GTYDPSSGATDLGTVECDGSIYRLGKTTRVNAPSIDGTQTFDQYWSVRQDKRTSGTVQ
TGCHFDAWARAGLNVNGDHYYQIVATEGYFSSGYARITVADVGGGGGSGGGGSGGG
GSTSQFSNSTSASSTDVTSSSSISTSSGSVTITSSEAPESDNGTSTAAPTETSTEAPTTAIPT
NGTSTEAPTTAIPTNGTSTEAPTDTTTEAPTTALPTNGTSTEAPTDTTTEAPTTGLPTNGT
TSAFPPTTSLPPSNTTTTPPYNPSTDYTTDYTVVTEYTTYCPEPTTFTTNGKTYTVTEPTT
LTITDCPCTIEKPTTTSTTEYTVVTEYTTYCPEPTTFTTNGKTYTVTEPTTLTITDCPCTIE
KSEAPESSVPVTESKGTTTKETGVTTKQTTANPSLTVSTVVPVSSSASSHSVVINSNGAN
VVVPGALGLAGVAMLFL

Amino acid sequence of FAE-SED fusion gene comprising N-terminal His tag (green), *Aa*FaeD catalytic domain, proline-rich linker (underlined), and SED1 anchoring domain (blue)

HHHHHHQNTAGCGQNPPSSGVKSINVGGMNREYILQLPNNYDPNKGHMLIFGLHWLS
GSMHDVHPNYYGLRQLAGNNAIFISPNGINNGWANDGGRDVNFIDAILQQVRSQLCIN
DSQIFATGFSFGGGMSYALGCARANVFRAIAPIAGAQISGCSGGTSPIAFLGIHGTNDDV
LPIAMGRQVRDRFLQNNGCQPKNAPEPGWGQGPIKTEYSCQPNYPVTWIAFSGGHDPN
QSFVGREIWDFFSQFGGTNPGPSPPPTSPPPTNPPPTNPPPTNPPPTNPPPSGGTSQFSNSTS
ASSTDVTSSSSISTSSGSVTITSSEAPESDNGTSTAAPTETSTEAPTTAIPTNGTSTEAPTTA
IPTNGTSTEAPTDTTTEAPTTALPTNGTSTEAPTDTTTEAPTTGLPTNGTTSAFPPTTSLPP
SNTTTTPPYNPSTDYTTDYTVVTEYTTYCPEPTTFTTNGKTYTVTEPTTLTITDCPCTIEK
PTTTSTTEYTVVTEYTTYCPEPTTFTTNGKTYTVTEPTTLTITDCPCTIEKSEAPESSVPVT
ESKGTTTKETGVTTKQTTANPSLTVSTVVPVSSSASSHSVVINSNGANVVVPGALGLAG
VAMLFL

List of Publications & Presentations

List of Journal papers that constitute the thesis:

1. **Apisan Phienluphon**, Keiko Kondo, Bunzo Mikami, Takashi Nagata, & Masato Katahira. (2023). Structural insights into the molecular mechanisms of substrate recognition and hydrolysis by feruloyl esterase from *Aspergillus sydowii*. International Journal of Biological Macromolecules, 253, 127188.
2. **Apisan Phienluphon**, Keiko Kondo, Bunzo Mikami, Kenneth Sze Kai Teo, Kaori Saito, Takashi Watanabe, Takashi Nagata, & Masato Katahira. (2024). Structure-based characterization and improvement of an enzymatic activity of *Acremonium alcalophilum* feruloyl esterase. ACS Sustainable Chemistry & Engineering (In press)
3. **Apisan Phienluphon**, Keiko Kondo, Hiroyuki Okano, Takashi Watanabe, Takashi Nagata, & Masato Katahira. (2024). Boosting sugarcane trash decomposition: synergistic action and proximity effect of xylanase and feruloyl esterase co-displayed on the cell surface of *Pichia pastoris* (*Komagataella phaffi*). Sustainable Energy & Fuels (In press)

List of other Journal papers:

1. Jantima Arnthong, Piyada Bussadee, **Apisan Phienluphon**, Pacharawan Deenarn, Kan Tulsook, Sa Ngapong Plupjeen, Chatuphon Siamphan, Chakrit Tachaapaikoon, Verawat Champreda, & Surisa Suwannarangsee. (2022). Overexpression of *LAS21* in Cellulase-Displaying *Saccharomyces cerevisiae* for High-Yield Ethanol Production from Pretreated Sugarcane Bagasse. Fermentation, 8(11), 652.
2. Jantima Arnthong, Jatupong Ponjarat, Piyada Bussadee, Pacharawan Deenarn, Parichat Prommana, **Apisan Phienluphon**, Salisa Charoensri, Verawat Champreda, Xin Qing Zhao, & Surisa Suwannarangsee. (2022). Enhanced surface display efficiency of *β*-glucosidase in *Saccharomyces cerevisiae* by disruption of cell wall protein-encoding genes *YGP1* and *CWP2*. Biochemical Engineering Journal, 179, 108305.

3. **Apisan Phienluphon**, Wuttichai Mhuantong, Katewadee Boonyapakron, Pacharawan Deenarn, Verawat Champreda, Duangdao Wichadakul, & Surisa Suwannarangsee. (2019). Identification and evaluation of novel anchoring proteins for cell surface display on *Saccharomyces cerevisiae*. Applied Microbiology and Biotechnology, 103(7), 3085–3097.

List of presentations (○Presenter):

1. ○**Phienluphon A**, Kondo K, Nagata T, & Katahira M. Construction of yeast co-displaying xylanase and feruloyl esterase for an efficient biomass degradation. The 14th International Symposium of Advanced Energy Science, Aug. 30th–Sep 1st, 2023, Kyoto, Japan. (**Poster**)
2. ○**Phienluphon A**, Kondo K, Nagata T, & Katahira M. Co-displaying of xylanase and feruloyl esterase on the *Pichia pastoris* cell surface. The 73nd Annual Meeting of the Japanese Wood Society Research Presentation, Mar 14th–16th, 2023, Fukuoka, Japan. (**Oral**)
3. ○**Phienluphon A**, Kondo K, Mikami B, Nagata T, & Katahira M. Effect of site-directed mutagenesis of *Aspergillus sydowii* feruloyl esterase on substrate preference. The 2022 Annual Meeting of The Japan Society for Bioscience, Biotechnology, and Agrochemistry (**Online**), Mar 15th–18th, 2022, Kyoto, Japan. (**Oral**)
4. ○**Phienluphon A**, Kondo K, Nagata T, & Katahira M. Cell surface display of fungal feruloyl esterase in *Pichia pastoris*. The *72nd* Annual Meeting of the Japanese Wood Society Research Presentation (**Online**), Mar 15th–17th, 2022, Nagoya, Japan. (**Oral**)
5. ○**Phienluphon A**, Kondo K, Nagata T, & Katahira M. Screening and heterologous expression of fungal feruloyl esterases, a lignin-carbohydrate complex degrading enzymes, in *Pichia pastoris* X-33. The 2021 Annual Meeting of The Japan Society for Bioscience, Biotechnology, and Agrochemistry (**Online**), Mar 18th–21st, 2021, Sendai, Japan. (**Oral**)

Acknowledgments

I would like to express my heartfelt gratitude to the individuals who have played a vital role in the completion of this dissertation. Their unwavering support and guidance have been invaluable throughout this journey.

First and foremost, I would like to express my sincere gratitude to my major advisor, Prof. Masato Katahira, whose benevolent supervision, invaluable counsel, and unwavering encouragement have been instrumental in shaping the trajectory of this research journey. I am deeply grateful to Assoc. Prof. Takashi Nagata for his generous support and insightful guidance, which significantly contributed to the success of this research. Prof. Keiko Kondo and Prof. Yudai Yamaoki merit special recognition for their patient instruction and invaluable advice, which have been consistent pillars throughout my research within this laboratory. Furthermore, I am indebted to Prof. Bunzo Mikami for his tireless encouragement and instruction throughout my works on the enzyme structures. Their collective mentorship not only enhanced my experimental skills but I also learnt soft and hard skills, fostering effective verbal and non-verbal communication. I would like to extend my thanks to Prof. Kaori Saito for her support and mentorship for my work and to Dr. Hiroyuki Okano for his invaluable assistance in the meticulous preparation of biomass. Their collective contributions have profoundly shaped my academic and research experience, for which I am deeply grateful.

My heartfelt appreciation also extends to all members of the Katahira laboratory including Keisuke Kamba, Sadat Mohamed Rezk Khattab, Chaithanya Pudushery Kunjiveedu, Wei Hsun Tu, Kenneth Teo Sze kai, Nesreen Hamad Abdel Gawwad Hamad, Mohammed Mahmoud Afaf Sobhi Eladl, Tomoki Sakamoto, Mohamed Mahmoud Omar Sobhi Eladl, Mohammed Oksh Mohammed Mousa, Yusei Yagi, Yuka Yonezawa, Yuta Sakai, Ryosuke Ono, Shohei, Takami, Ningxin Wang, Chuyu Zheng, Chihiro Nakayama, Weile Dai, Aika Kawakami, Taiki Morita, Soichiro Matuso, Chenchen Ji, Eri Oga, Kentaro Tozawa, Taisei Masunaga, Makoto Shimizu, and Yuto Goto, as well as the Thai student community in Kyoto University, including Nittikarn Suwanawat, Panith Adulsiriswad, Anucha Koedtruad, and Chuaychob Surachada. My student life was enriched by your presence, and I will forever cherish the wonderful memories we created

together. These moments of joy and friendship were invaluable in pushing me through the challenges of scientific research, and I am truly grateful for your support and friendship.

Special thanks are also due to our dedicated secretary, Naomi Murakami, whose unwavering support was a constant source of help and friendship. Her assistance exemplified her dedication to my success and I am truly grateful for her kindness. Additionally, I extend my appreciation to the GSES administration officer for their consistent communication and assistance throughout my academic journey. Your support has been invaluable, and I am thankful for your contributions to my research endeavors.

I would like to extend my sincere gratitude to Dr. Surisa Suwannarangsee and Dr. Verawat Champreda for their extensive encouragement and invaluable guidance during my tenure at BIOTEC, NSTDA, Thailand. Their support was instrumental in shaping my academic journey. Additionally, I am deeply appreciative of Prof. Takashi Watanabe, who recognized the compatibility of my expertise with the Katahira laboratory. With their precise guidance during the doctoral course admission process, I was afforded the opportunity to pursue my future career and embark on my doctoral studies at Kyoto University, Japan.

I want to extend a special thanks to Dr. Tiwaporn Nualkaew, my beloved girlfriend, for her unwavering support, constant presence, and invaluable assistance, day and night. Her encouragement has been a source of strength throughout my journey, and I am profoundly grateful for her love. Additionally, I would also like to extend my thanks to all my friends in Thailand and Japan for their support and warm wishes. Your presence and friendship were truly invaluable to me.

Last but not least, my deepest gratitude goes to my beloved family, Kanyarat Phienluphon, Naisan Phienluphon, Thirawat Phienluphon, and Warisa Phienluphon, for their boundless love, unwavering encouragement, and profound understanding that has enriched my life in immeasurable ways. Their support has been a steadfast pillar in my life, providing both the emotional sustenance and unwavering belief in my abilities that have enabled me to pursue my academic and personal aspirations. I consider myself

incredibly fortunate to have such a loving and supportive family, and I cherish the invaluable role they have played in shaping who I am today.

I consider myself the luckiest person to have such an incredible family in my life. The love and support you've given me are beyond measure, and I hope to find ways to express my gratitude and compensate you for all that you have done for me.

Apisan Phienluphon

Bioenergy Research Section

Department of Fundamental Energy Science

Graduate School of Energy Science

Kyoto University

Dec. 2023