Supporting Information for

Selective cleavage of lignin β-*O*-4 aryl ether bond by β-etherase of the white-rot fungus *Dichomitus squalens*

Mila Marinović[a], Paula Nousiainen[b], Adiphol Dilokpimol[c], Jussi Kontro[b], Robin Moore[b], Jussi Sipilä[b], Ronald P. de Vries[a,c], Miia R. Mäkelä[a], Kristiina Hildén[a*]

[a]Division of Microbiology and Biotechnology, Department of Food and Environmental Sciences, University of Helsinki, Viikinkaari 9, Helsinki, Finland

[b]Department of Chemistry, Laboratory of Organic Chemistry, University of Helsinki, Finland

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

*Corresponding author: email: kristiina.s.hilden@helsinki.fi

Contents:

Table S1. Putative GST encoding sequences of *D. squalens* with corresponding JGI protein IDs (http://genome.jgi-psf.org/Dicsq1/Dicsq1.home.html) identified in this study.

References

File S1. Materials and methods for cloning, heterologous expression and purification of Ds-GSTs.

Cloning of Ds-GSTs

The genes encoding *D. squalens* Ds-GST1 and Ds-GST2, and *Sphingobium* sp. SYK-6 LigF, used as a positive control, were codon-optimized for *E. coli*, synthesized and cloned into a pET-24a(+) vector (GenScript, USA). Ds-GST3 and Ds-GST4 were amplified from the *D. squalens* cDNA library with primer pairs designed according to the *D. squalens* LYAD-421 SS1 v1.0 genome sequence (Ds-GST3 forward: 5'-CGGAATTCCCATATGCCGGAGCCGATCATATTCTACGAC-3', *Nde*I restriction site underlined; Ds-GST3 reverse: 5'- CGGAATTCGGATCCTCAGAGCTCGACGACCGACCCTTC-3', *Bam*HI restriction site underlined; Ds-GST4 forward: 5'-CGGAATTCCCATATGTCAGACCCCATCGTTCTATACGACATACC, *Nde*I restriction site underlined, Ds-GST4 reverse: 5'-CGGAATTCGGATCCTTAGACCTTCACCTCGGCGTACACCTCC-3', *Bam*HI restriction site underlined). A 25 µl PCR reaction mixture contained 0.5 µL *D. squalens* cDNA template, 5 µL HF buffer, 0.4 µM dNTP-mix (Thermo Fischer Scientific), 2 µM primers, 0.75 µL DMSO, and 0.5 U Phusion High-Fidelity polymerase (Thermo Fischer Scientific). PCR was performed with initial denaturation at 98°C for 30 s; followed by 45 cycles of (1) denaturation at 98°C for 10 s, (2) annealing at 70°C for 30 s, (3) elongation at 68°C for 1:45 min; and final extension at 68°C for 8 min. Amplification products were ran on 1% agarose gel, the correct size fragments were cut off and purified by using Wizard® SV Gel and PCR Clean-Up System (Promega). The fragments were digested with *Nde*I and *Bam*HI (Promega), ligated into pET28a (T4 DNA ligase, Thermo Scientific) and transformed into *E. coli* DH5α. Selected transformants were sequenced (Macrogen, The Netherlands) and sub-cloned into *E. coli* Rosetta^{TM}(DE3)pLysS.

Heterologous expression and enzyme purification

For protein expression, *E. coli* cultures with optical density of 0.6 at 600 nm were induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Sigma-Aldrich) and incubated for 24 h (250 rpm, 12°C). The cells were harvested by centrifugation (15 min, 7,000 × *g*, 4°C) and incubated overnight at –20°C. The cells were resuspended in 30 mL Bugbuster™ protein extraction reagent (Novagen) and centrifuged (30 min, 16,000 × *g*, 4°C) to separate soluble protein fraction (supernatant) from insoluble fraction (cells).

Filtered culture supernatants (0.22 µm, Sterile Syringe Filter, VWR International) were purified chromatographically with ÄKTA explorer apparatus (GE Healthcare) with Q Sepharose FF and Mono Q 5/50 GL columns (GE Healthcare), respectively. A linear gradient of 0 – 1 M NaCl in 10 mM sodium-acetate buffer, pH 8.5, was used for protein elution. Protein containing fractions were pooled, concentrated and dialyzed into 20 mM HEPES-buffer, pH 7.0, and analyzed by SDS-PAGE (Mini-PROTEAN® TGX™ Precast Protein Gels, Bio-Rad) and stained with PageBlue Protein Staining Solution (Thermo Scientific). Protein concentration was determined by Pierce™ BCA protein assay kit according to manufacturer's instructions (Thermo Fischer Scientific).

File S2. Materials and methods for the synthesis of lignin model compounds.

2-(2-methoxyphenoxy)-3-oxo-3-(4-hydroxy-3-methoxyphenyl)propan-1-ol (**1**) and 2-(2-methoxyphenoxy)-3-(4-hydroxy-3-methoxyphenyl)propan-1,3-diol (**2**)[1,2] were used as lignin model compounds. The products were purified with silica flash chromatography (Biotage SP4, Biotage AB, Sweden) and characterized by NMR and ESI-TOF (Bruker MicroTOF). The purity of the products was determined by Agilent 1260 Infinity HPLC. NMR-spectra were recorded using a Varian Inova 500

spectrometer in $CDCl_3$. The purified products were dissolved in freshly distilled 1,4-dioxane (Fischer Scientific) before use in enzyme assays.

Guaiacyl dehydrogenation polymer (G-DHP), was prepared from coniferyl alcohol[3], which was synthesized from methyl ferulate by reduction with DIBAL-H[4]. The G-DHP was oxidized in the presence of *Trametes villosa* NS-51002 laccase (Novozymes A/S) with methyl syringate (MeS, synthesized from syringic acid (Sigma-Aldrich) and methanol) as a mediator to provide α-carbonyl functional groups in G-DHP[5]. G-DHP in cellulose (0.5 g) was mixed with 2.5 mL of dioxane (20% (v/v)) in 50 mM succinic acid buffer, pH 5.5, supplied with 0.9 U/ml NS-51002 laccase and 0.38 mM MeS. The mixture was stirred for 24 h, after which the enzymatic reaction was stopped by adding 1mg/mL (2.5 mL) sodium azide solution[6]. After concentrating on rotary evaporator under reduced pressure, the laccase-treated G-DHP was washed with water and ethyl acetate (Fischer Scientific) to remove salts, laccase and mediator together with oligomeric lignin fragments, and finally extracted with 1,4-dioxane and dried under reduced pressure.

File S3. Materials and methods for chiral chromatography

Enzymatic reaction mixtures were evaporated using 35°C vacuum oven (Thermo Scientific) until dry, and subsequently dissolved in 0.1% TFA in 65:35 (v/v) hexane/isopropanol (Fischer Scientific). Analysis of compound (**1**) and β–etherases-derived residual substrates were conducted *via* chiral chromatographic separation using Agilent 1200 HPLC (Santa Clara, CA, USA) equipped with Chiracel OD-H column (4.6 mm × 250 mm, 5 µm; Diacel Corporation) at 20°C and detection wavelength 280 nm. A mobile phase consisting of 0.1% TFA in 65:35 (v/v) hexane/isopropanol was used at the flow rate of 0.6 mL/min. Sample injection volume was 5 µL.

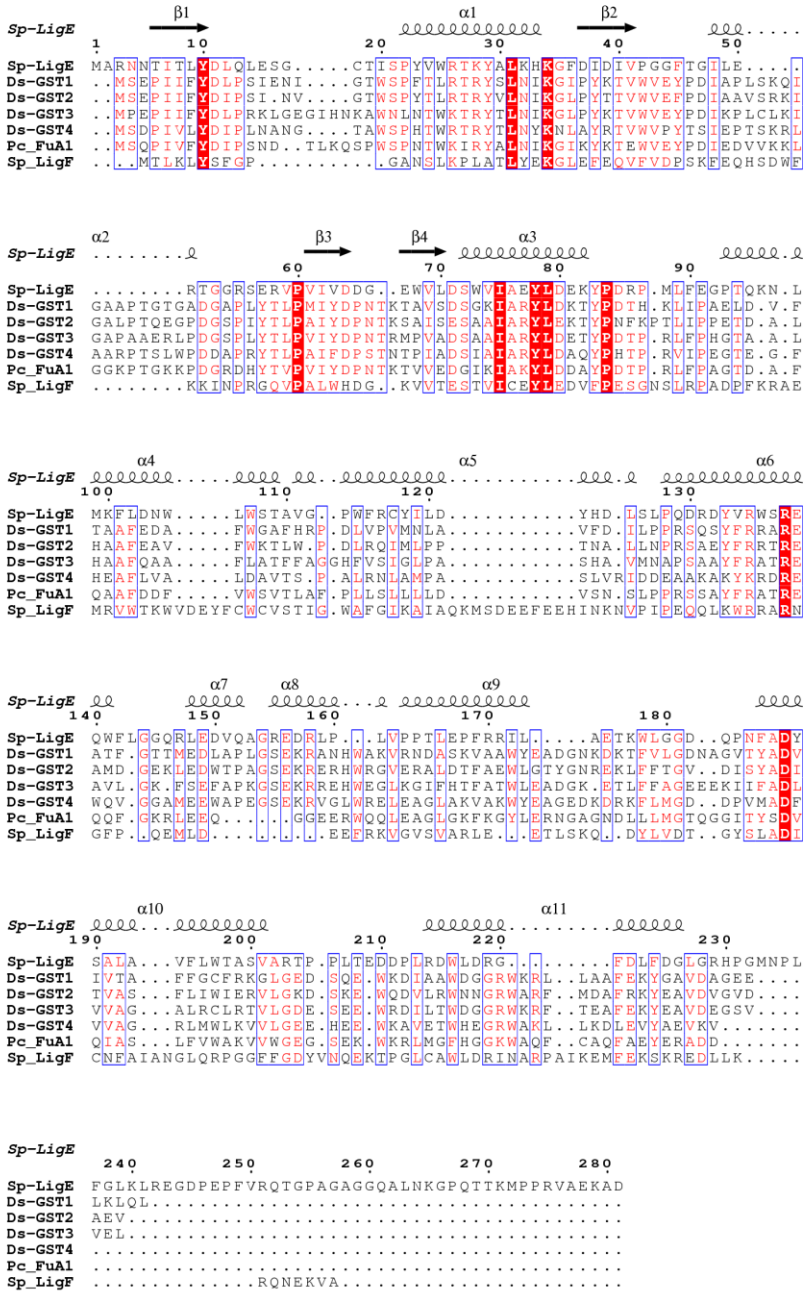

Figure S1. Alignment of Ds-GST1, Ds-GST2, Ds-GST3 and Ds-GST4 from *D. squalens* with homologs from *Sphingobium* sp. SYK-6 (LigE and LigF), and *P. chrysosporium* (Pc-FuA1).

Conserved amino acid residues are highlighted in red. LigE secondary structure elements, α-helixes and β-sheets, are shown above the sequence alignment. JGI protein ID numbers are: Ds-GST1, 82843; Ds-GST2, 56311; Ds-GST3, 102613; Ds-GST4, 61773 and Pc-FuA1, 5118. NCBI accession numbers

are: LigE, BAA02032 and LigF, WP_014075191. The multiple-sequence alignment was generated by ESPript[7]. The amino acid sequence similarity between the four Ds-GSTs ranged from 42% to 55%. Despite the low amino acid sequence similarity, 17-25%, between Ds-GSTs and *Sphingobium* sp. SYK-6 LigE and LigF, the key amino acid residues, e.g. those responsible for non-covalent anchoring of γ-glutamyl (Asp-71 and Ser-72) and glycyl (Tyr-133 and Arg-138) GSH moieties in LigE[8], are well conserved in Ds-GST1 and Ds-GST3. In Ds-GST2 and Ds-GST4, these residues are replaced by Glu and Lys, respectively. Sequence and crystal structure analyses have revealed a conserved Ser residue near the thiol of the bound GSH in the active site of LigE, LigF and *P. chrysosporium* GSTFuA1 (Ser-21, Ser-14 and Ser-22, respectively) that may have a role in the binding of GSH[8,9]. This residue was also present in Ds-GST1 (Ser-20), Ds-GST2 (Ser-19) and Ds-GST4 (Ser-20). In Ds-GST3, Asn-24 was detected instead of Ser. Surrounding residues, Trp-20 and Pro-22 in LigE, are considered to be a characteristic feature in the GSTFuA class enzymes[10], and they were conserved in the studied Ds-GSTs, except Ds-GST3 in which Pro was replaced by Leu-25 residue.

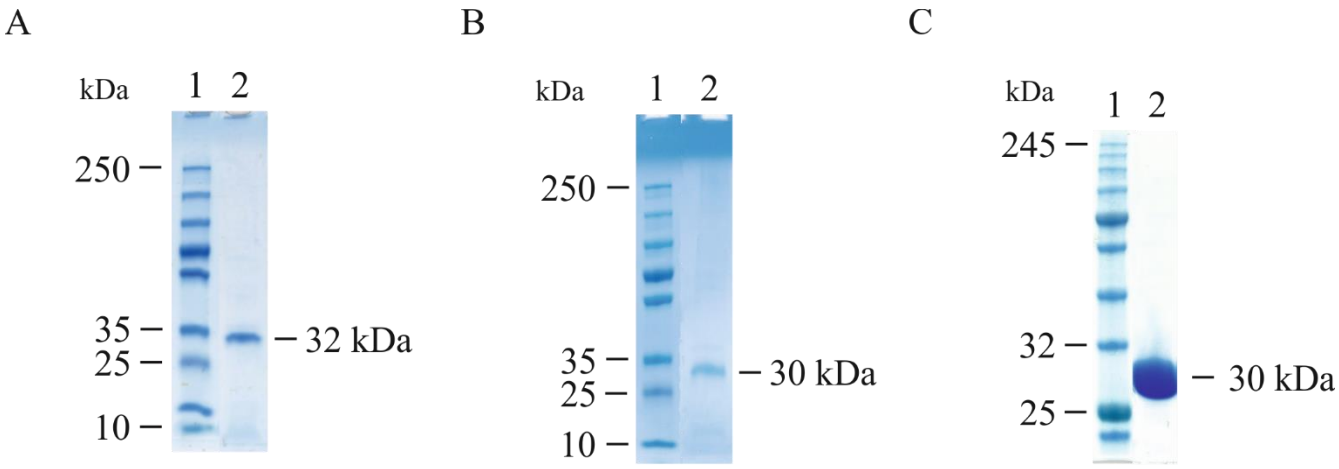

Figure S2. SDS-PAGE analysis of chromatographically purified recombinant Ds-GSTs. A) Lane 1: molecular mass standard, lane 2: Ds-GST1. B) Lane 1: molecular mass standard, lane 2: Ds-GST2. C) Lane 1: molecular mass standard, lane 2, Ds-GST3.

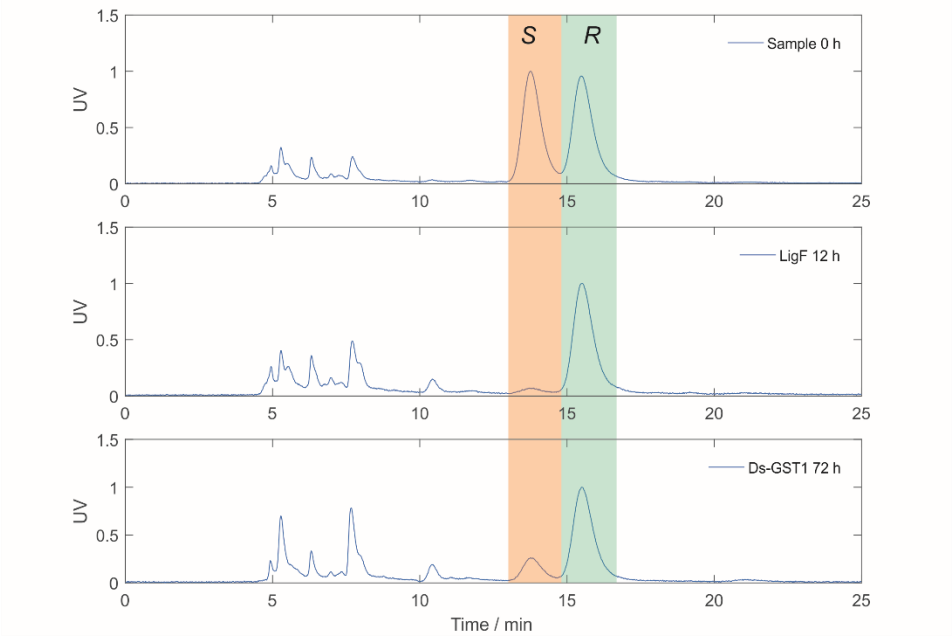

Figure S3. Chiral chromatography of a racemic lignin model compound (1) at 0 h with LigF and Ds-GST1 catalyzed enzymatic reactions. A) Racemic lignin model compound (**1**) exhibiting *S*- and *R*-enantiomers at 0 h. Stereoselective degradation of *S*-enantiomers by B) positive control LigF and C) Ds-GST1.

Table S1. Putative GST encoding sequences of *D. squalens* with corresponding JGI protein IDs (http://genome.jgi-psf.org/Dicsq1/Dicsq1.home.html) identified in this study.

#	Name	JGI protein ID
1	Ds-GST1	82843
2	Ds-GST2	56311
3	Ds-GST3	102613
4	Ds-GST4	61773
5		101677
6		136582
7		69238
8		110722
9		174906
10		76710
11		70508
12		107335
13		135184
14		137430
15		167503
16		169668
17		170658
18		171068
19		174286
20		71278
21		183150
22		65248
23		161976
24		169691
25		172112
26		112982
27		148760
28		111591
29		62518
30		137388
31		181982
32		75740
33		46340
34		51898
35		91147
36		100284

37	129051
38	134414
39	173218
40	105731
41	83422
42	49777

References

1. Adler, E.; Lindgren, B. O.; Saedén, U., The β-guaiacyl ether of α-veratrylglycerol as a lignin model. *Svensk Papperstidn.* **1952**, *57*, 245-254.

2. Sipilä, J.; Syrjänen, K., Synthesis and ^{13}C NMR spectroscopic characterization of six dimeric arylglycerol β-aryl ether model compounds representative of syringyl and *p*-hydroxyphenyl structures in lignins. On the Aldol reaction in β-ether preparation. *Holzforschung* **1995**, *49*, 325-331.

3. Hofrichter, M.; Vares, K.; Scheibner, K.; Galkin, S.; Sipilä, J.; Hatakka, A. Mineralization and solubilization of synthetic lignin by manganese peroxidases from *Nematoloma frowardii* and *Phlebia radiata*. *J. Biotechnol.* **1999**, *67*, 217-228.

4. Quideau, S.; Ralph, J. Facile large-scale synthesis of coniferyl, sinapyl, and *p*-coumaryl alcohol. *J. Agric. Food Chem.* **1992**, *40*, 1108-1110.

5. Nousiainen, P.; Kontro, J.; Manner, H.; Hatakka, A.; Sipilä, J. Phenolic mediators enhance the manganese peroxidase catalyzed oxidation of recalcitrant lignin model compounds and synthetic lignin. *Fungal Genet. Biol.* **2014**, *72*, 137-149.

6. Ürek, R. Ö.; Pazarlioğlu, N. K. Purification and partial characterization of manganese peroxidase from immobilized *Phanerochaete chrysosporium*. *Process Biochem.* **2004**, *39*, 2061-2068.

7. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320-W324.

8. Helmich, K. E.; Pereira, J. H.; Gall, D. L.; Heins, R. A.; McAndrew, R. P.; Bingman, C.; Deng, K.; Holland, K. C.; Noguera, D. R.; Simmons, B. A.; Sale, K. L.; Ralph, J.; Donohue, T. J.; Adams, P. D.; Phillips, G. N. Structural basis of stereospecificity in the bacterial enzymatic cleavage of β-aryl ether bonds in lignin. *J. Biol. Chem.* **2015**, *291*, 5234-5246.

9. Mathieu, Y.; Prosper, P.; Favier, F.; Harvengt, L.; Didierjean, C.; Jacquot, J. P.; Morel-Rouhier, M.; Gelhaye, E. Diversification of fungal specific class A glutathione transferases in saprotrophic fungi. *PLoS One* **2013**, *8*, e80298.

10. Mathieu, Y.; Prosper, P.; Buee, M.; Dumarcay, S.; Favier, F.; Gelhaye, E.; Gerardin, P.; Harvengt, L.; Jacquot, J. P.; Lamant, T.; Meux, E.; Mathiot, S.; Didierjean, C.; Morel, M. Characterization of a *Phanerochaete chrysosporium* glutathione transferase reveals a novel structural and functional class with ligandin properties. *J. Biol. Chem.* **2012**, *287*, 39001-39011.