UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS QUIMICAS

TESIS DOCTORAL

Resurrection of ancestral ligninolytic peroxidases

Resurrección de peroxidasas ligninolíticas ancestrales

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Iván Ayuso Fernández

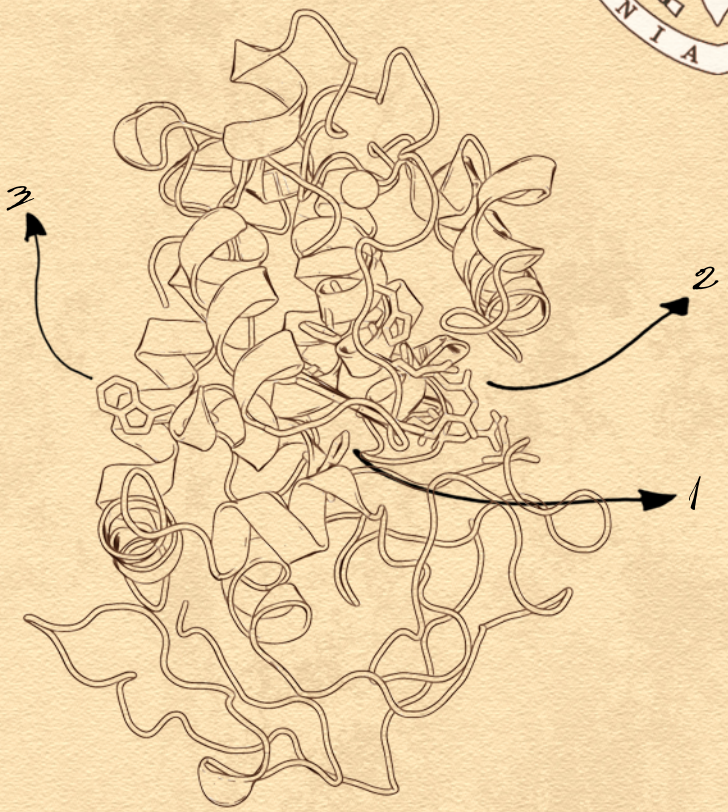
Directores

Ángel T. Martínez Ferrer
F. Javier Ruiz Dueñas

Madrid

UNIVERSIDAD COMPLUTENSE MADRID

Facultad de Ciencias Químicas
Departamento de Bioquímica y
Biología Molecular



Doctoral Thesis

Resurrection of ancestral ligninolytic peroxidases

Resurrección de peroxidasas ligninolíticas ancestrales

Iván Ayuso Fernández

Supervisors
Ángel T. Martínez Ferrer
F. Javier Ruiz Dueñas

Madrid, 2019

UNIVERSIDAD COMPLUTENSE MADRID

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña. Iván Ayuso Fernández,
estudiante en el Programa de Doctorado de Bioquímica, Biología Molecular y Biomedicina,
de la Facultad de Ciencias Químicas de la Universidad Complutense de Madrid, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

RESURRECCIÓN DE PEROXIDASAS LIGNINOLITICAS ANCESTRALES /
RESURRECTION OF ANCESTRAL LIGNINOLYTIC PEROXIDASES

y dirigida por: D. Ángel T. Martínez y D. F. Javier Ruiz Dueñas

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

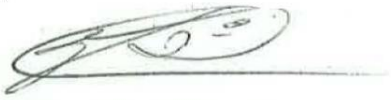
Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 10 de junio de 2019

Fdo.: Iván Ayuso

Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

V°B° Director Tesis
Fdo



Francisco Javier Ruiz Dueñas

V°B° Director Tesis
Fdo

Ángel Tomás Martínez Ferrer

V°B° Tutora
Fdo

Ana Isolina Saborido Modia

Madrid a 10 de junio de 2019

Universidad Complutense de Madrid

Facultad de Ciencias Químicas

Departamento de Bioquímica y

Biología Molecular

Doctoral Thesis

Iván Ayuso Fernández

Resurrection of ancestral ligninolytic peroxidases

Resurrección de peroxidasas ligninolíticas ancestrales

Supervisors:

Dr. Ángel T. Martínez Ferrer
Profesor de Investigación, CSIC

Dr. Francisco Javier Ruiz Dueñas
Científico Titular, CSIC

Madrid, 2019

Front cover: artistic representation of the first versatile peroxidase that appeared in Polyporales evolution, named AVPd in this thesis.

Life, uh, finds a way
Dr. Ian Malcolm
Jurassic Park

A mi madre, Julia
Descansa en paz

Agradecimientos

La presente tesis doctoral se ha realizado en el Centro de Investigaciones Biológicas y durante una estancia breve en la Universidad de California en Irvine. Ha sido financiada por una beca de Formación del Personal Investigador (FPI, Ref BES-2012-053513), una beca para estancia breve (Ref EEBB-I-15-10267) así como por los proyectos europeos "Optimized oxidoreductases for medium and large scale industrial biotransformations" (INDOX, Ref KBBE-2013-613549) y "New enzymatic oxidation/oxyfunctionalization technologies for added value bio-based products" (EnzOx2, Ref H2020-BBI-PPP-2015-720297) y del Plan Nacional "Nuevas enzimas oxidativas para una industria sostenible" (NOESIS, Ref BIO2014-56388-R) y "Genomas de basidiomicetos para las biorrefinerias de lignocelulosa" (GENOBIOREF, Ref BIO2017-86559-R).

La escritura de la tesis, por mucho que uno se empeñe en hacerla metódica y al amparo de la Ciencia, acaba siendo algo muy personal e íntimo. Por ello, en esta sección quiero reflejar un sincero agradecimiento a aquellas personas que por un motivo u otro han tenido a bien estar cerca de mí durante su desarrollo. Al fin y al cabo la tesis no es del que la hace, sino de todo su entorno.

Lo primero agradecer a Ángel Martínez haberme acogido en este grupo y haberme guiado de forma magistral durante todos estos años, especialmente a la hora de exprimir los resultados y sacar las publicaciones adelante. También por tu resolución y gestión de aspectos accesorios de la Ciencia (un mal necesario), y por tu humor irónico. También quiero agradecer a Javier Ruiz Dueñas su guía cuando estaba más perdido y su paciencia cuando he sido más rebelde. Tu amabilidad como persona y tus insondables conocimientos sobre peroxidasas han hecho esta tesis posible, aparte de haberme concedido la beca FPI, claro. Muchas gracias, Javi.

Muchas gracias al resto de jefas de este grupo, por su apoyo a lo largo de los años, las conversaciones sobre ciencia y por preocuparse por mí. Gracias a Susana, a Marta Pérez, Ali y María Jesús, porque los funguitos no serían lo mismo sin vosotras.

Los agradecimientos a mis compañeros de poyata deberían ocupar un capítulo entero, y no porque seamos muchos. Hacéis que dedicarse a un trabajo como este, a veces ingrato y otrora frustrante, no solo sea posible sino que incluso apetezca.

Gracias a Ana, por tu amistad y sinceridad, por tus consejos certeros, y por ser tan maña que ni tus telómeros se han atrevido a acortarse. A David, por las buenas conversaciones independientemente del tema y por tu espíritu curioso. Gracias también a Pablo, cuya amistad con David me obliga a saltarme el orden alfabético. Gracias por las épicas historias de mayor o menor veracidad y por los duelos de *Magic*. Gracias a Gonzalo, por odiar todo, lo que hace que signifique más que te gustase una de mis charlas. Y por los adjetivos cada vez más locos. Gracias a Isa Pardo, porque siempre has sido uno de los faros que he seguido como ejemplo, por ser una gran amiga y por tu paciencia ante las absurdas aventuras que siempre te proponíamos Carlos y yo. Gracias a Isa V, por el drama (por supuesto), por los rayos, por todos los planes a lo largo de estos años y porque en el tiempo que coincidimos en el CIB, el sol brillaba más. Gracias a Jorge, por tu sabiduría e innumerables consejos, casi tantos como cañas contigo, y por demostrar que

recorrer el camino de la Ciencia por la vía espartana es posible. Gracias a Juan, por ser el mejor compañero de piso del mundo, por aguantarme en casa en esta última etapa de muchas quejas, por estar siempre dispuesto a un plan y liarlo hasta su límite. Gracias por tu humor y tus anécdotas geniales, y en definitiva por tu amistad a lo largo de estos años. Gracias a Juanan, por las conversaciones sobre cualquier videojuego, y por creer en mí más que yo mismo. Gracias a Lola, porque eres la mejor persona posible, por enseñarme en mis primeros meses y no parar de enseñarme nunca, y por descubrirnos los manolitos (los de mantequilla tallada). Gracias a Manuel, porque su presencia eleva la conversación, por demostrar que los trasgos *a veces* ganan a los elfos, por ser el mejor DM que jamás existirá y por ser una de las mejores amistades que puedo tener. Sé que demostrarás que el *hygge* es una mamarrachada vikinga. Gracias a María Molina, por sus tartas de queso y su simpatía contagiosa! Gracias a Minipiticli (aka Alberto), por las anécdotas y las risas, y por traer torreznos sorianos en tu cumpleaños. Gracias a Rash y a Ander, mi dúo cómico favorito, por ser tan geniales sin pretenderlo. Sabía que os iría bien. Gracias a Vero, un ángel PhD, porque tu infinita bondad nos hacía a todos mejores, y por enseñarme tanto.

Special thanks to Dr. Poulos and his team in California, for hosting me overseas. Thanks to Georges, Alec and Dillon for the great times there.

Gracias a los jinetes con los que más he recorrido esta travesía, porque con vosotros cabalgaría a la muerte (bueno no nos pasemos). Gracias a Felipe, la mezcla perfecta de pillería y mejor persona del planeta. Gracias por todos los planes, cervezas y viajes juntos, que se dice pronto pero han sido muchos, por ser un compañero de trabajo perfecto, y por ser uno de mis mejores amigos. Gracias a Joan, el mejor batería del mundo, por salvar a los *Michael Just Furniture*, por intentar mejorarme en la música a pesar de mi torpeza, por las conversaciones intensas y no tanto, y por ser la roca que me mantiene cuerdo. Y gracias a Carlos. Has sido la persona que más ha influido en mi vida a miles de niveles (Scaruffi te daría un 9.5). Las aventuras contigo fueron las mejores (publicamos el disco contra todo pronóstico!) y aún nos quedan muchas cosas por hacer. Tenemos que tocar en el Primavera con la rumba congoleña *noise power pop*, resolver misterios, visitar la tumba de Townes van Zandt y montar la trattoria, por ahora. Gracias hamijo, porque encontrar a tu alma gemela no es fácil, y tú eres mucho más.

Gracias a Lidem, que merece un capítulo ella sola (si no digo esto me mata). Esta tesis ha sido posible a pesar de ti, pero sin ti no habría tenido sentido. Has hecho mejores todos y cada uno de los días, y no sé si en el CIB, pero en mi vida sí eres *estructural*. Gracias por alimentarme a lo largo de los años, por los planes (a pesar de los juegos de mesa) y viajes, y por tu psicología extrema. No me olvides, por fa.

Gracias a mi familia, por su incondicional apoyo y paciencia a lo largo de estos años. Gracias a mi padre, porque se echó el mundo a los hombros cuando tocó. Si yo estoy aquí es por ti, así que muchas gracias, papa. Gracias a mi hermano, por su apoyo continuo a lo largo de toda mi vida, y por ser tan bueno siempre.

Gracias a Marisa. Porque no imagino la vida sin ti, y porque has hecho más por esta tesis que yo mismo. Gracias por hacerme reír más que nadie, por los cientos de memes, los cientos de planes (locos o de seta *mácsima*, no hay término medio) y porque contigo el viaje más desastroso imaginable ha sido el mejor momento de mi vida. Gracias por la continua música de ida y vuelta, por componer canciones que no son demasiado punk (o Nirvana…), y por descubrirme tantísimas cosas. Gracias por todo, Marisota.

Table of Contents

Abreviaturas/ *Abbreviations* ..... **i**
Resumen / Summary **1**
Introduction **11**
1. Plant biomass and its use in industry 13
1.1 Fungi in forest ecosystems 13
1.2 Lignin biosynthesis and structure 14
1.3 Importance of lignin in the bio-based industry 17
2. Wood rotting fungi 19
2.1 Fungal strategies for wood degradation 19
2.2 Ligninolytic peroxidases 21
2.2.1 Common features of PODs and their classification 21
2.2.2 Global catalytic cycle of PODs 25
2.2.3 Redox properties of PODs 26
3. Evolution of wood rotting fungi and land plants 28
3.1 Origin of wood-rotting fungi 28
3.2 Polyporales evolution 31
3.3 Lignin evolution was essential for land colonization 32
4. Enzyme resurrection as a method to study evolution 34
4.1 Origins of ancestral sequence reconstruction 34
4.2 Uncertainty in ancestral sequence reconstruction 36
4.3 Biotechnological potential of ancestral enzymes 39
4.4 Resurrection of ligninolytic peroxidases 41
Objetivos / Objectives 43
Materials & Methods 49
Chapter 1 Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date 61
General abstract 63
Abstract 64
Background 64
Results 66
Reconstruction of ancestral sequences from Polyporales genomes 66
In silico analysis of ancestral heme pocket and substrate-oxidation sites 70
Kinetic properties of resurrected peroxidases 71
Stability of ancestral peroxidases 74
Discussion 76
Conclusions 78
Methods 79
Phylogenetic analysis 79
Ancestral sequence reconstruction 79
Protein modeling 80
E. coli expression 81
Steady-state kinetics 81
pH and temperature stability 81

Funding 82
Acknowledgements 82
Additional material 82
Additional file 1: Supplemental Figures S1-S4 84
Chapter 2 Evolutionary convergence in lignin degrading enzymes 89
General abstract 91
Abstract 92
Significance statement 92
Introduction 93
Results 94
Ancestral Sequences of Polyporales Peroxidases. 94
Structural Comparison of two Peroxidase Lineages 95
Reaction Kinetics and Convergent Evolution of Ancestral Peroxidases 98
Stability Comparison in the two Peroxidase Lineages 101
Discussion 102
Materials and methods 105
Acknowledgements 105
Suporting Information 106
Supporting Materials and Methods 106
Supporting Figures and Tables 108
Chapter 3 Redox potential increased during the evolution of enzymes degrading recalcitrant lignin 115
General abstract 117
Abstract 118
Results & Discussion 118
Experimental Section 126
Acknowledgements 126
Suporting Information 128
Supplementary materials and methods 128
Supplementary tables 134
Supplementary figures 140
Chapter 4 Peroxidase evolution in white rot fungi follows wood lignin evolution in plants 147
General abstract 149
Abstract 150
Significance statement 150
Introduction 151
Results 152
Evolution of Polyporales lignin-degrading peroxidases 152
Kinetics of lignin oxidation through peroxidase evolution 154
Steady-state treatment of lignins 156
Polyporales specialization for plant hosts 158
Discussion 158
Acknowledgements 160
Materials and methods 160
SI Appendix 162

Supporting Materials and Methods 162
Supporting Tables and Figures 167
General discussion 173
Conclusiones / Conclusions 181
Bibliography 187
Introduction references 189
Materials & Methods references 198
Chapter 1 references 201
Chapter 2 references 206
Chapter 2 Supporting Information references 210
Chapter 3 references 211
Chapter 3 Supporting Information references 215
Chapter 4 references 217
Chapter 4 Supporting Information references 220
General discussion references 222

Abreviaturas/ *Abbreviations*

ABTS	2,2'-azinobis (3-etilbenzotiazolin-6-sulfonato) / *2,2'-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid)*
ALiP	Lignina peroxidasa ancestral/ *Ancestral lignin peroxidase*
ASR	Reconstrucción de secuencias ancestrales/ *ancestral sequence reconstruction*
AVPb/d	Peroxidasa versátil ancestral del clado B o D/ *Ancestral versatile peroxidase of clade B or D*
CaD/CaB	Ancestro común de las peroxidasas del clado D o B/*Common ancestor of clade D or B peroxidases*
CaPo	Ancestro común de las peroxidasas de Polyporales/*Common ancestor of Polyporales peroxidases*
CI	Compuesto I / *Compound I*
CII	Compuesto II/ *Compound II*
COMT	Ácido cafeico/5-hidroxiconiferaldehído O-metiltransferasa/ *Caffeic acid/5-hydroxyconiferaldehyde O-methyltransferase*
CT	Bandas de transferencia de carga/ *Charge transfer bands*
DMP	2,6-dimetoxifenol / *2,6-dimethoxyphenol*
DyP	Peroxidasa decolorante de tintes/ *Dye-Decolorizing peroxidase*
E$^{o'}$	Potencial de reducción estándar/ *Standard reduction potential*
F5H	Ferulato-5-hidroxilasa/ *Ferulate 5-hydroxilase*
GP	Peroxidasa genérica/ *Generic peroxidase*
HSQC	*Heteronuclear single-quantum correlation*
IPTG	Isopropil-ß-D-1-tiogalactopiranósido Isopropil tio-ß-D-galactósido / *Isopropyl β-D-1-thiogalactopyranoside*
k_1 /k_2 /k_3	Constante de velocidad de primer orden para la formación del C-I/reducción del C-I/ reducción del C-II / *First-order rate constant for CI formation /CI reduction/ CII reduction*
k_{app}	Constante de velocidad aparente de segundo orden/ *Apparent second-order rate constant*
k_{cat}	Constante catalítica/ *Catalytic constant*
k_{cat} / K_m	Eficiencia catalítica/ *Catalytic efficiency*
K_D	Constante de disociación/ *Dissociation constant*
K_m	Constante de Michaelis/ *Michaelis constant*
k_{obs}	Constante de velocidad de pseudo-primer orden/ *Pseudo first-order rate constants*
LiP	Lignina peroxidasa / *Lignin peroxidase*
LRET	Transferencia electrónica de largo recorrido/ *Longe range electron transfer*
ML	*Maximum likelihood*
MnP	Manganeso peroxidasa/ *Manganese peroxidase*
NMR	Resonancia paramagnética nuclear/ *Nuclear paramagnetic resonance*
PC-LiPA (LiPH8)	Lignina peroxidasa A (isoforma H8) de *Phanerochaete chrysosporium* / *Lignin peroxidase A (isoform H8)* from Phanerochaete chrysosporium
PDB	*Protein Data Bank*
POD	Peroxidasas fúngicas de clase II/ *Class-II fungal peroxidases*
RB5	*Reactive Black 5*

RS	Estado de reposo/ *Resting state*
SEC	Cromatografía de exclusión molecular/ *Size-exclusion chromatography*
T_{50}	Temperatura a la que la actividad es mitad de la máxima/ *Temperature at which the activiy is half the initial*
T_m	Temperatura del punto medio en la transición al estado desnaturalizado/ *temperature Temperature at the midpoint of the unfolding transition in the thermal melting profiles*
TV-VP2	Peroxidasa versátil 2 de *Trametes versicolor*/ *Versatile peroxidase 2 from* Trametes versicolor
VA	Alcohol veratrílico (alcohol 3,4-dimetoxibencílico)/ *Veratryl alcohol*
VP	Peroxidasa versátil/ *Versatile peroxidase*

Resumen / Summary

Resurrección de peroxidasas ligninolíticas ancestrales

Antecedentes

La lignina es uno de los biomateriales más abundantes sobre la Tierra, y su degradación es un problema importante tanto a nivel biológico como a nivel industrial. En la naturaleza, el reciclado de la lignina es esencial en el ciclo del carbono, siendo los hongos de la podredumbre blanca los principales organismos capaces de mineralizar este polímero. A nivel industrial, el uso de dichos hongos o de su maquinaria enzimática se estudia para un mejor aprovechamiento de la lignocelulosa en la industria de la biorefinería.

Para la degradación de la lignina los hongos de la podredumbre blanca secretan diversos tipos de peroxidasas: i) manganeso peroxidasas (MnPs), que poseen un sitio de unión de Mn^{2+} para la oxidación de dicho catión metálico a Mn^{3+}, que puede ser quelado y difundir en la madera, oxidando la parte fenólica de la lignina; ii) lignina peroxidasas (LiPs), que poseen un triptófano catalítico en su superficie donde la lignina se puede oxidar directamente; y iii) peroxidasas versátiles (VPs), que combinan en su estructura los dos sitios de oxidación anteriores

El origen de la degradación fúngica de la madera se estableció en el Carbonífero asociado a la producción de las primeras peroxidasas ligninolíticas y, junto a otros factores geoquímicos, contribuyó al fin de la acumulación de carbón fósil. Dicho evento es el punto de partida de esta tesis, cuyo principal objetivo es analizar la posterior evolución de las peroxidasas ligninolíticas usando la resurrección de enzimas ancestrales.

Objetivos

i. Reconstrucción de secuencias de peroxidasas fúngicas usando la filogenia de las peroxidasas ligninolíticas de Polyporales y el software PAML. Se evaluarán las secuencias de interés y los modelos tridimensionales de sus estructuras para seleccionar aquellas enzimas ancestrales que se estudiarán en detalle.
ii. Resurrección de las enzimas ancestrales seleccionadas para su posterior caracterización bioquímica. Para ello se usarán sustratos modelo y se estudiarán tanto sus constantes cinéticas como su estabilidad.
iii. Puesta a punto de técnicas para medir y evaluar el potencial redox de las peroxidasas ligninolíticas a lo largo de su evolución, incluyendo espectrometría de flujo detenido (*stopped-flow*) en equilibrio redox, titulación espectroelectroquímica y NMR de proteínas.
iv. Estudio de la actividad de las peroxidasas ancestrales y actuales usando lignosulfonatos de gimnospermas y angiospermas como modelos de lignina soluble en agua. Para ello se realizarán estudios cinéticos de los estados transitorios de las enzimas mediante espectrofotometría de

stopped-flow, y análisis de NMR bidimensional (2D-NMR) de los productos de reacción en los que se evaluará la modificación de los polímeros en tratamientos prolongados.

Resultados

Capítulo 1: *Recreación experimental de la evolución de enzimas degradadoras de lignina desde el Jurásico hasta hoy*

En este capítulo se presenta el primer estudio de resurrección de enzimas ligninolíticas realizado y ya recogido en la literatura. Usando PAML y la filogenia de las peroxidasas de clase II de los Polyporales (donde se incluyen la mayoría de hongos degradadores de la madera) se obtienen ancestros en el linaje que conduce a las enzimas ligninolíticas más eficientes que existen actualmente, las LiPs. Usando sustratos modelo simples, se caracterizan los distintos sitios de oxidación en las enzimas y se comprueba cómo han ido cambiando a lo largo de la evolución, explorando nuevas estrategias para la oxidación de lignina. Así mismo, se analiza la estabilidad de las enzimas ancestrales y actuales, viendo que su estabilidad a pHs ácidos, donde actúan en la naturaleza, aumenta cuando aparece el triptófano catalítico para la oxidación directa de la lignina.

Capítulo 2: *Convergencia evolutiva en enzimas degradadoras de lignina*

En este segundo capítulo se analiza otro linaje evolutivo que lleva a un clado de enzimas diferenciado y distante con respecto a las LiPs estudiadas en el capítulo anterior. En este clado se incluyen VPs actuales, que también poseen el sitio de oxidación directa de la lignina. Para evaluar si la aparición de dicho sitio se debió a eventos de duplicación desde un ancestro que lo contenía en su estructura, o se produjo como consecuencia de un proceso de convergencia evolutiva apareciendo dos veces en distintos linajes, se resucitan los ancestros correspondientes. Los resultados muestran que se produjo convergencia evolutiva en hongos degradadores de madera para la oxidación directa de la lignina debido a la producción de enzimas similares durante la evolución. Se muestra también que durante la evolución de las LiPs se produjo un aumento de carga negativa alrededor del triptófano, lo que podría explicar su mayor eficacia oxidando sustratos modelo. Los resultados también muestran que se produjo convergencia en las propiedades de las enzimas de ambos linajes, siendo más estables a pHs ácidos tras la aparición del triptófano catalítico.

Capítulo 3: *El potencial redox de las enzimas que degradan lignina aumentó durante la evolución*

Las peroxidasas ligninolíticas se caracterizan por su alto potencial redox. Esto les permite oxidar múltiples sustratos entre los que se encuentra la lignina. En este capítulo se analiza cómo cambió el potencial redox de estas enzimas a lo largo de la evolución. Para calcular los potenciales de reducción

de las enzimas ancestrales y actuales se utilizan medidas espectrofotométricas de *stopped-flow* en equilibrio redox y titulación espectroelectroquímica, obteniéndose los valores de potencial redox de todas las parejas del ciclo catalítico y del par Fe^{2+}/Fe^{3+}. Los resultados muestran que se produjo un aumento del potencial redox en las peroxidasas ligninolíticas a lo largo de la evolución, un hecho que se ha podido correlacionar con el desplazamiento observado de la señal correspondiente a la histidina proximal usando NMR de proteínas. Esto indica que un cambio estructural en el entorno proximal del hemo, con alteraciones en la geometría del enlace histidina proximal – hierro del hemo, estaría implicado en la modificación del potencial redox a lo largo de la evolución.

Capítulo 4: La *evolución de las peroxidasas de los hongos de podredumbre blanca sigue la evolución de la lignina de la madera en plantas* En este último capítulo se caracteriza el linaje de las LiPs usando dos tipos de lignina soluble en agua: lignosulfonatos de gimnospermas y de angiospermas. Mediante espectrofotometría de *stopped-flow* se obtienen las constantes de estado transitorio para la oxidación de lignina, y usando 2D-NMR se analiza cómo se modifica el polímero a lo largo de la evolución en base a tratamientos prolongados de ambos lignosulfonatos con enzimas actuales y resucitadas. Tras la calibración temporal de la filogenia de las peroxidasas, se ha podido concluir que hubo un cambio en la preferencia de sustrato cuando apareció el triptófano catalítico en la superficie de las peroxidasas ligninolíticas. De este modo, se pasó de una mejor oxidación de lignina de gimnospermas, más basal, a una preferencia por la lignina de angiospermas, más reciente y compleja, coincidiendo aproximadamente con el origen de este tipo de plantas en la evolución.

Conclusiones

Usando la resurrección de enzimas ancestrales se ha analizado la evolución de las peroxidasas ligninolíticas de los hongos degradadores de la madera. En esta tesis se demuestra que durante la evolución de estas enzimas se exploraron varias estrategias para la degradación de lignina, convergiendo en linajes distintos la capacidad de oxidar directamente este polímero en un triptófano expuesto en la superficie de la proteína. Con esta adquisición, junto con su estabilización a pHs ácidos, el incremento en la carga negativa alrededor del triptófano y el aumento del potencial redox a lo largo de la evolución, las enzimas adquirieron las propiedades únicas que tienen hoy día. Así mismo, empleando ligninas solubles se muestra el cambio en la preferencia de sustrato experimentado por las peroxidasas ligninolíticas, de lignina de gimnospermas a lignina de angiospermas, coincidente con la aparición del triptófano catalítico.

Resurrection of ancestral ligninolytic peroxidases

Background

Lignin is one of the most abundant biomaterials on Earth, and its degradation is an important biological and industrial problem. In nature, lignin recycling is essential in the carbon cycle, being white-rot fungi the main organisms able to mineralize this polymer. In the industry, the use of these fungi or their enzymatic machinery is studied for a better use of lignocellulose in the biorefinery context, as a green alternative to chemical (or physical) methods.

For lignin degradation, white-rot fungi secrete different types of peroxidases: i) manganese peroxidases (MnPs), which have a Mn^{2+}-binding site for oxidation of the metal cation to Mn^{3+}, whose chelates diffuse in the wood and are able to oxidize the phenolic moiety of lignin; ii) lignin peroxidases (LiPs), which have a catalytic tryptophan in their surface where lignin is oxidized directly; iii) versatile peroxidases (VPs), which combine both oxidation sites in their structure.

The origin of fungal wood degradation was established in the Carboniferous period associated to the production of the first ligninolytic peroxidases and, together with other geochemical factors, contributed to the end of coal accumulation. This evolutionary event is the starting point of this thesis, where the posterior evolution of ligninolytic peroxidases is analyzed using ancestral enzyme resurrection.

Aims

i. Ancestral sequence reconstruction of fungal peroxidases using the phylogeny of Polyporales peroxidases and the PAML software. The sequences of interest and their molecular models will be evaluated to select those ancestral enzymes that will be studied in detail.

ii. Resurrection of the selected ancestral enzymes for their biochemical characterization. Model substrates will be used and both the kinetic constants and stability parameters will be obtained.

iii. Optimization of reduction potential measurements to evaluate the redox potential of ligninolytic peroxidases through evolution, including stopped-flow spectrophotometry at redox equilibrium, spectroeloectrochemical titration and protein NMR.

iv. Characterization of ancestral and extant peroxidases using lignosulfonates from gymnosperms and angiosperms as water-soluble lignin models. Stopped-flow spectrophotometry will be used to analyze transient-state kinetics, and 2D-NMR to evaluate the modification of the polymer in long-term treatments.

Results

Chapter 1. *Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date*

In this chapter is presented the first resurrection study of ligninolytic enzymes, already published in the literature. Using PAML and the phylogeny of class-II peroxidases from Polyporales (where most wood-rotting fungi are included) the ancestors in the lineage leading to the most efficient ligninolytic enzymes, LiPs, are obtained. Using simple model compounds, the different oxidation sites and their change through evolution are characterized, showing the exploration of new strategies for lignin oxidation. Likewise, the stability of ancestral enzymes is analyzed, showing that their stability to acid pHs, at which ligninolytic peroxidases act in nature, increases when the catalytic tryptophan for direct lignin oxidation appears.

Chapter 2. *Evolutionary convergence in lignin degrading enzymes*

In the second chapter a different lineage leading to a distant clade from LiPs is analyzed. This clade includes extant VPs, which also have the direct lignin oxidation site. To evaluate if the appearance of this site is due to duplication events of an ancestor already containing it in its structure, or due to evolutionary convergence appearing twice in different lineages, the ancestors of interest are resurrected. The results show that convergence took place in wood-rotting fungi evolution for a direct lignin oxidation, due to the production of similar types of enzymes through time. It is also showed that during LiPs evolution, the catalytic tryptophan environment became more negative, which could explain their higher efficiency oxidizing model substrates. The results also show convergence in the stability properties of the enzymes in both lineages, being more stable to acidic pHs after the appearance of the catalytic tryptophan.

Chapter 3. *Redox potential increased during the evolution of enzymes degrading recalcitrant lignin*

Ligninolytic peroxidases are characterized by their uniquely high reduction potential. This allows them to oxidize multiple substrates including lignin. In this chapter, it is evaluated how the redox potential changed through evolution. To calculate reduction potentials of ancestral and extant enzymes, stopped-flow spectrophotometric measurements at redox equilibrium and spectroelectrochemical titration are employed, obtaining the values of reduction potential of all the catalytic cycle couples and the Fe^{2+}/Fe^{3+} pair. The results show that there is an increase in the redox potential of ligninolytic peroxidases through evolution, a fact correlated with the displacement of the signal of the proximal histidine observed using protein NMR. This indicates that an structural change in the heme proximal side, with alterations in the geometry of the proximal histidine –

heme iron bond, would be involved in the modification of redox potential in evolution.

Chapter 4. *Peroxidase evolution in white-rot fungi follows wood lignin evolution in plants*

In the last chapter the linage to LiPs is characterized using two types of water-soluble lignin: lignosulfonates from gymnosperms and angiosperms. With stopped-flow spectrophotometry the transient-state kinetic constants in lignin oxidation are obtained and using 2D-NMR it is evaluated how the polymer is modified through evolution in long-term treatments of both lignosulfonates using extant and resurrected enzymes. After time calibration of the peroxidases phylogeny, it has been concluded that there was a switch in the substrate preference when the solvent exposed catalytic tryptophan appeared in the surface of ligninolytic peroxidases. This way, there was a change from a better oxidation of gymnosperm lignin, more basal, to a preference for angiosperm lignin, more recent and complex, coincident approximately with the origin of this type of plants in evolution.

Conclusions

Ancestral enzyme resurrection allowed the analysis of ligninolytic peroxidase evolution in wood-rotting fungi. In this thesis, it is demonstrated that during evolution these enzymes explored several strategies for lignin degradation, converging in distant lineages the ability to oxidize the polymer directly in a solvent-exposed tryptophan. With this acquisition, together with the stabilization to acidic pHs, the increase in the negative charge surrounding the tryptophan and the boost of reduction potentials through evolution, the enzymes acquired the unique properties they have today. Likewise, the use of soluble lignin shows a change in the substrate preference from gymnosperms to angiosperms wood lignin, concomitant with the rise of the catalytic tryptophan.

Introduction

1. Plant biomass and its use in industry

1.1 Fungi in forest ecosystems

Fungi are a central part of forest and other land ecosystems, associating to all other life forms and interconnecting them (Dighton and White 2017). In the global picture of the forest they can be roughly defined by the functions they perform, which are multiple and relevant to other organisms. This way, fungi can exist as pathogens, regulating population dynamics by killing plants. This may generate gaps in the forest, which can be opportunities to other colonizers (increasing biodiversity and also contributing to the accumulation of dead wood) (Holah et al. 1993; Jarosz and Davelos 1995). Fungi can also have a role as symbionts, either in mycorrhizal associates or in lichens. Mycorrhizal organisms are part of the absorption components of plant roots, increasing the uptake of water and nutrients in exchange of carbon, being this association essential to most terrestrial plants (Hoysted et al. 2018; Balestrini and Lumini 2018). In lichens, one fungal organism (usually an ascomycete) is associated to one cyanobacteria species (and sometimes a third component, a basidiomycete yeast, as recently discovered) (Spribille et al. 2016) for a mutual beneficial relationship, with an important role in the total primary productivity of forest. But, it is as decomposers that fungi drive the global carbon cycle in the forest. Wood and litter decay fungi recycle carbon stored in plant biomass and make it available for other organisms. In this way, lignin degrading fungi expose cellulose and hemicelluloses for a direct consumption by other microorganisms.

The function of fungi as decomposers is essential given the fact that almost 2000 gigatons of carbon are present in terrestrial ecosystems, with 550 gigatons stored in plant biomass (McCarl et al. 2007), mainly in form of wood. This way, saprotrophic fungi of different orders of basidiomycetes participate in the carbon cycle performing an enzymatic “combustion” (Kirk and Farrell 1987) of recalcitrant lignin that enables the use of plant biomass for nutrition of fungi and other microorganisms.

Plant biomass is essentially lignocellulose from woody and nonwoody plants, and represents most of the total carbon fixed in the photosynthesis of land species (Higuchi 1997). Lignocellulosic biomass is the most abundant renewable raw material on Earth, so understanding its biodegradation by fungi is mandatory to try to implement these organisms and their molecular tools in the bio-based industry. Lignocellulose is composed of three main polymers: cellulose (30-50%), hemicelluloses (15-30%) and lignin (10-25%) (Fengel and Wegener 1984; Pettersen 1984). While cellulose and hemicelluloses are sugar polymers more or less structured that serve as direct carbon source for other organisms, lignin is the generic term for different heterogeneous and highly recalcitrant aromatic polymers, with

some remarkable functions. Lignin gives rigidity to plants and protects them against the UV radiation, which was essential in their land colonization, maintains the structural integrity of the cell wall while waterproofing it (enabling transport of water and nutrients through the vascular system) and acts as a shield against the attack of some pathogens (Boerjan et al. 2003). Given the importance of lignin in Nature and in the industry, its biosynthesis and biodegradation has been thoroughly studied in the last decades.

1.2 Lignin biosynthesis and structure

Lignin derives from the radical polymerization of differently substituted phenyl-propylene precursors. There are three basic lignin monomers or monolignols, which result from the methoxy substitutions of the benzene ring: *p*-coumaryl alcohol (forming *p*-hydroxyphenyl or H units), coniferyl alcohol (forming guaiacyl or G units) and sinapyl alcohol (forming syringyl or S units). The proportion of the different units changes between species or even between tissues in the same plant. However, as rule of thumb, gymnosperms have lignins with G units and some H units, while in angiosperms, dicots have lignins with S and G units and trace amounts of H, and monocots have lignins with S and G units and higher proportion of H units than in dicots (Vanholme et al. 2010).

The precursor of the three monolignols is the aromatic amino acid phenylalanine, synthesized in the shikimate pathway. The main biosynthetic route of all of them is summarized in **figure 1**, where it is shown the route to H, G, and S units. It is noteworthy the need of ferulate 5-hydroxilase (F5H) and caffeic acid/5-hydroxyconiferaldehyde O-methyltransferase (COMT) to form S units: this catalyzes the conversion of coniferaldehyde and coniferyl alcohol into 5-hydroxyferulic acid, continuing the route to the syringyl units. F5H is only present in angiosperms, there the appearance of S units only in this type of plants and its absence in gymnosperms (with some notable exceptions where S lignin appears, as reviewed by Novo-Uzal *et al* (2012) and discussed below).

After their biosynthesis, the monolignols are transported to the cell wall by a mechanism still in mystery, although some different models have been proposed (Vanholme et al. 2010) (simple diffusion, Golgi-derived vesicles, using transporters or translocated as their glucosylated forms). There takes place the polymerization, consisting of an oxidative radicalization of the monomers followed by a combinatorial radical coupling.

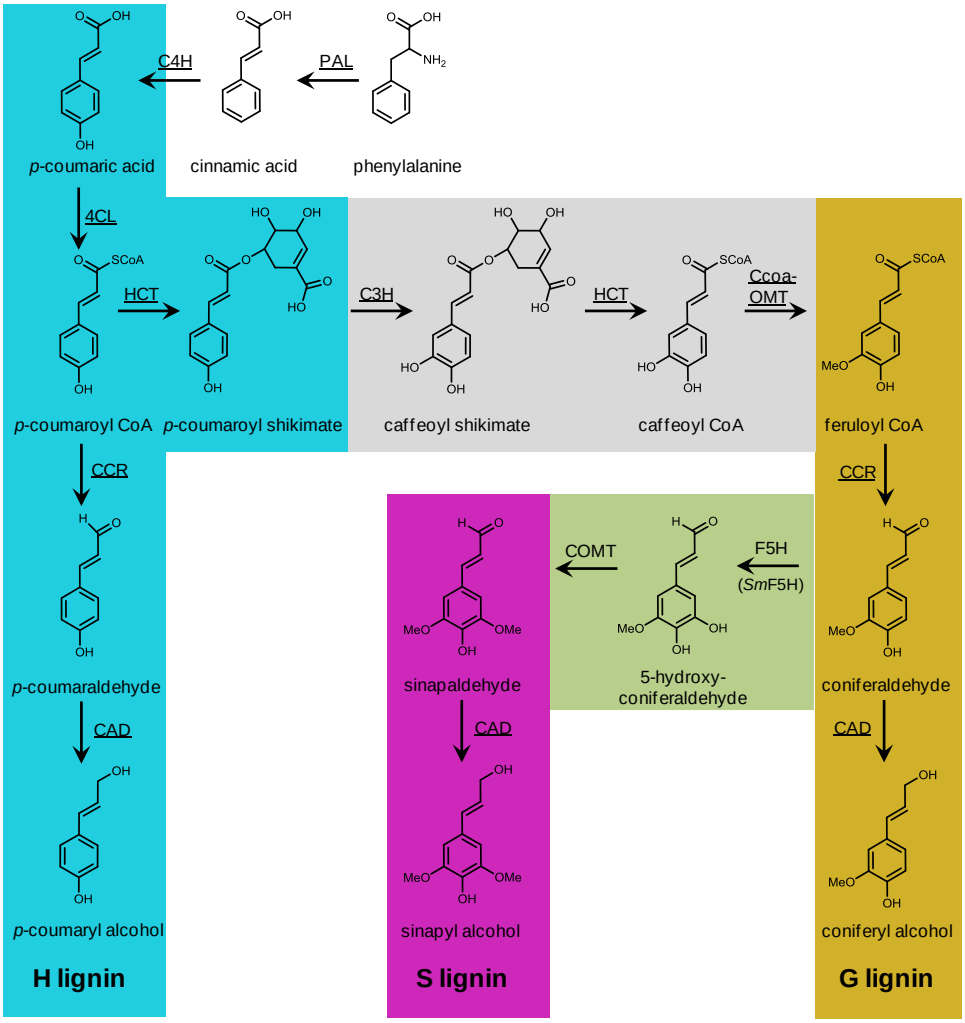


Figure 1. Biosynthetic routes to three monolignols (*p*-coumaryl, coniferyl and sinapyl alcohols) precursors of H, G and S lignin units, respectively. The code of colors indicates 4-hydroxycinnamyl (blue), 3,4-dihydroxycinnamyl (grey), 3-methoxy-4-hydroxycinnamyl (yellow), 3-methoxy-4,5-dihydroxycinnamyl (green) and 3,5-dimethoxy-4-hydroxycinnamyl (pink) intermediate compounds. Enzymes included: **PAL**, phenylalanine ammonia-lyase; **C4H**, cinnamate 4-hydroxylase; **4CL**, 4-coumarate:CoA ligase; **C3H**, *p*-coumarate 3-hydroxylase; **HCT**, *p*-hydroxycinnamoyl-CoA:quinate/shikimate *p*-hydroxycinnamoyl-transferase; **CCoAOMT**, caffeoyl-CoA *O*-methyltransferase; **CCR**, cinnamoyl-CoA reductase; **F5H**, ferulate 5-hydroxylase; **COMT**, caffeic acid *O*-methyltransferase; **CAD**, cinnamyl alcohol dehydrogenase. Adapted from Vanholme et al (2010).

The first polymerization step is performed by peroxidases or laccases present in the cell wall, which extract one electron of the monolignols, generating relatively stable radical monomers that may couple with other radical monomers to form dimers. This second step is combinatorial, meaning that the structure of the growing polymer is largely dependent of the ratio of the monomers and the conditions in the cell wall (Ralph et al. 2004). The dimer has to be oxidized again before it can couple to other radical monomer, being this way of polymerization called endwise coupling. It is also possible the union of lignin oligomers, which is especially common

in G-lignins. In the final structure of the macromolecule there will be different types of ether and C-C bonds, being the most abundant the ß-O-4', followed by β-β', β-5' and 4-O-5' (**figure 2**).

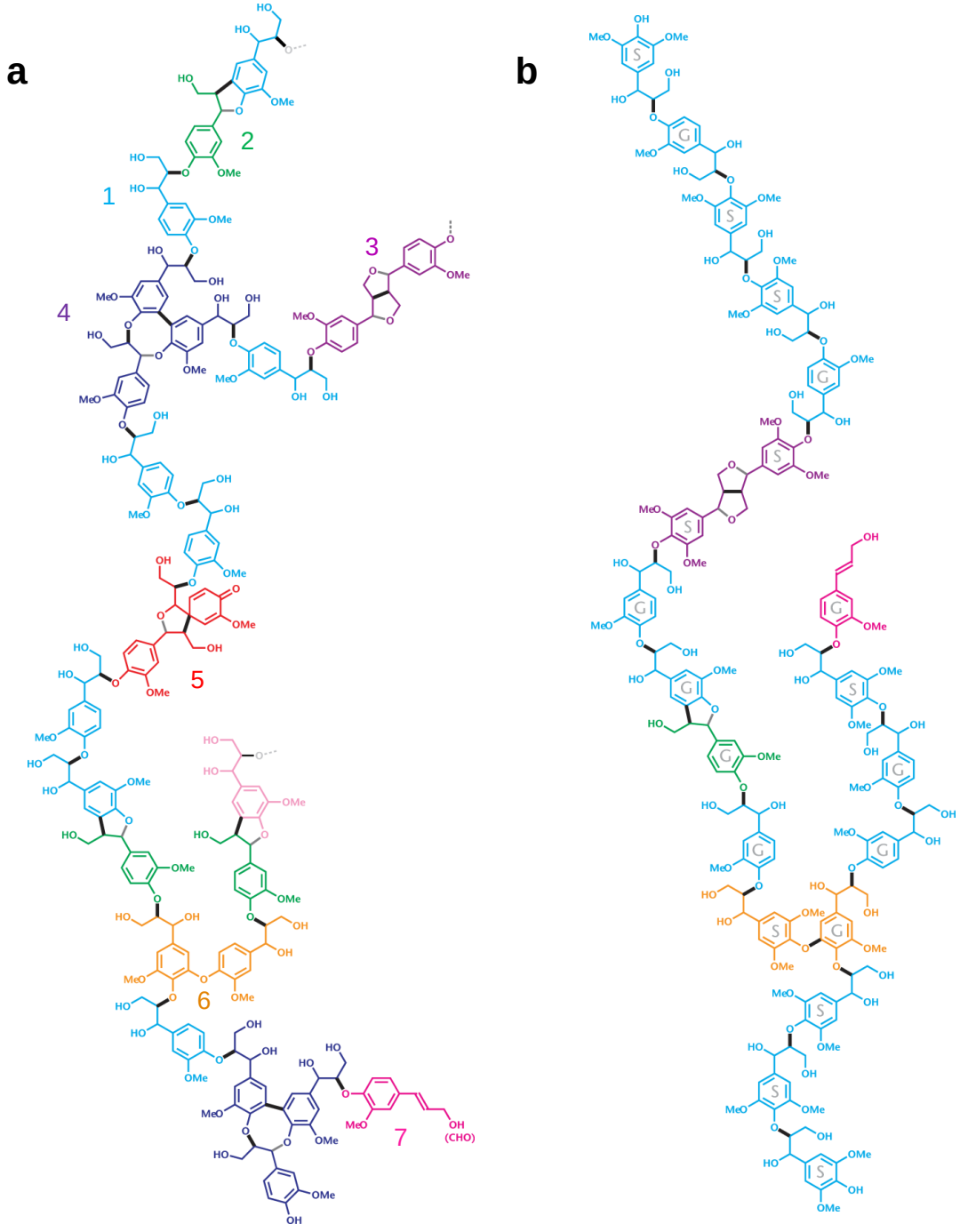


Figure 2. Structural models of gymnosperm (spruce) G-lignin (a) and angiosperm (poplar) S-G lignin (b) deduced using NMR and other techniques**.** The variety of bonds results in the following substructures, shown using a color code: **1)** β-O-4' ether (cyan); **2)** Phenylcoumaran (green); **3)** Resinol (purple); **4)** Dibenzodioxocin (dark blue); **5)** Spirodienone (red); **6)** 4-O-5' ether (orange); and **7)** Cinnamyl alcohol end-group (magenta). Adapted from Rencoret (2008).

Aside from the basic monolignols described above, there are other less common units such as sinapyl *p*-coumarate, sinapyl acetate or sinapyl *p*-hydroxybenzoate. This etherification of monolignols would allow the presence of some highly acylated lignins in angiosperms, being a way to control the structure and some properties of lignin (Martínez et al. 2008). Together with the not so clear specificity of peroxidases and laccases oxidizing phenols, which would allow the control of lignin structure, the engineering of plants with *a la carte* lignins is one of the most exciting aspects in lignin biotechnology (Ralph et al. 2004; Abramson et al. 2010).

1.3 Importance of lignin in the bio-based industry

Lignocellulosic biomass is one of the main renewable materials in the world. Humanity has used it for centuries as elements for construction or as source for heat. Biomass can be also stored as fossil in the form of coal, whose massive accumulation ended with the appearance of the first ligninolytic fungi at the end of the Carboniferous (Floudas et al. 2012). However, the discovering and intense use of coal (and petroleum) since the industrial revolution caused a major unbalance in the global carbon cycle. The carbon stored as fossil, and hence out of the cycle, has been transferred massively to it, altered since then. The accumulation of CO_2 that cannot be assimilated in the cycle, together with other factors, produced the well known and possibly irreversible green-house effect. This led to the urge to search alternatives to coal and petroleum, whether in producing fuels or in its derivates.

Fossil fuels are finite and not really an option on the long run to satisfy the global demand for combustibles (Ragauskas et al. 2006). For that reason, during the last decades the concept of biorefinery to substitute the petroleum refinery is attracting a considering amount of dedication and investment (Martínez et al. 2009). In this sense, the biorefinery would imply obtaining the same type of products as the classic petroleum refinery but with the use of biomass as the starting raw material (Ragauskas et al. 2014). The work flux is summarized in **figure 3**. Starting with natural (or engineered as previously said) lignocellulose, a first step is the separation of the components of the biomass (cellulose, hemicelluloses, lignin and minor components). Cellulose has an important market in paper industry, and together with hemicelluloses would be used to obtain added-value chemicals and biofuels (bioethanol) by fermentation in the biorefinery (Martínez et al. 2009).

Lignin would imply much more effort, and currently is underused in the industry. As said above, lignin is a highly heterogeneous and recalcitrant polymer, properties that make it very difficult to work with. For that reason, lignin has been directly burned as mere combustible for heat and power

after their extraction in the biomass industry. Today, important advances in lignin technology have been made, both trying to make lignin rentable by its valorization and in its degradation for a more sustainable process working with biomass.

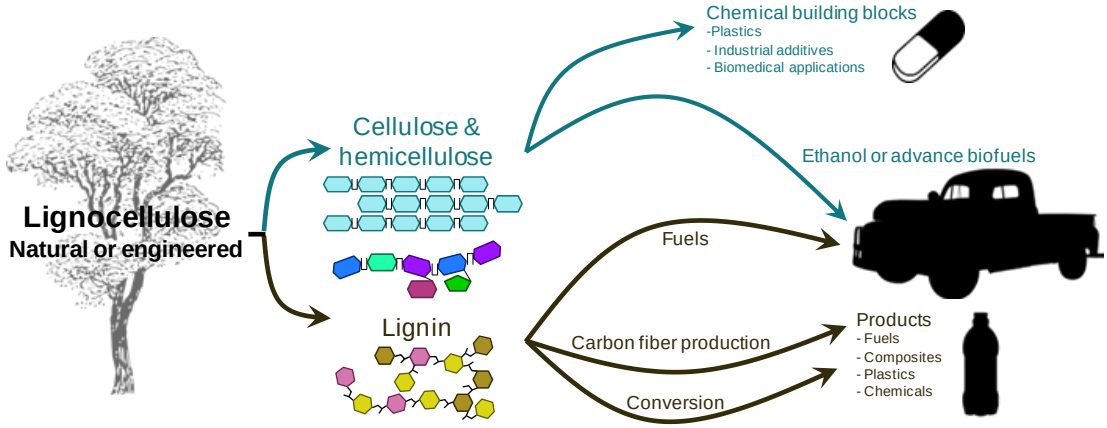


Figure 3. Schematic representation of the lignocellulosic biorefinery. The different components of the biomass are separated and used to obtain fuels and different added value products. Adapted from Ragauskas et al (2014).

Lignin valorization is a challenge that requires a multidisciplinary approach (Beckham et al. 2016). Treating lignin as a valuable feedstock allows its use to obtain different types of materials (carbon fiber), fuels (production of gasoline-range aromatics) and different aromatic chemical compounds derived from it (Ragauskas et al. 2014). In order to improve lignin valorization, the efforts are being focused on several aspects: i) its fractionation and use as substrate for microbial utilization; ii) the host selection for biological funneling of the aromatic compounds as carbon source; iii) understanding of ligninolytic enzymes-microbes relationship for a better advantage of their synergy; iv) expansion of enzyme substrate specificity of the aromatic microbes chosen; v) metabolic engineering to lead the flux of metabolite productivity to the chosen compounds; and vi) the integration of all processes (Beckham et al. 2016).

Additionally to lignin valorization, there have been great efforts dedicated to a more sustainable degradation of lignin in the industry. Lignin is one of the main factors to overcome biomass recalcitrance and therefore to low the cost of operations separating biomass components in a biorefinery (Himmel et al. 2007). During the last years, the investigation of lignin biosynthesis allowed the design of special trees with modified lignin, either for an easier removal of the polymer or for an improved lignin valorization (Ralph et al. 2004; Vanholme et al. 2008). However, these trees with altered lignin can also result in altered growth and other related environmental problems. Together with this choice, other options have been considered. Among them, biological degradation and better biocatalysts for lignin degradation are well

studied. The main organisms that degrade lignin in nature are fungi, being the so-called white rot basidiomycetes the principal decomposers, using an array of enzymes to perform the enzymatic mineralization of lignin (Hammel and Cullen 2008; Martínez et al. 2018).

2. Wood rotting fungi

2.1 Fungal strategies for wood degradation

There are different types of organisms able to degrade wood to some extent, but only wood-rotting fungi are able to mineralize lignin properly, exposing the biomass sugars. Wood rotting-fungi are principally saprotrophic Agaricomycetes (a class within the phylum Basidiomycota), which can be classified taking into account the type of wood degradation they perform. There are mainly three types of wood degradation by fungi: soft, brown and white rot, depending of the type of wood residue that these fungi leave after its decomposition (Martínez et al. 2005).

Soft-rot fungi are Ascomycetes and Deuteromycetes that are able to degrade wood sugars, but barely affect lignin. Their limited action leaves a soft consistency wood residue in wet environments (brown and crumbly in dry environments). These organisms act mainly on hardwoods, i.e. the wood from dicot trees.

Brown-rot fungi include only a 7% of wood-rotting Basidiomycetes, being some representative species *Postia placenta, Gloeophyllum trabeum, Laetiporus sulphureus or Piptoporus betulinus*. These organisms attack selectively hemicelluloses and cellulose of wood distanced from the hyphae, leaving a brown residue (hence their name) with modified lignin that breaks in cubical pieces (**figure 4a**)**.** Brown-rot fungi do not produce peroxidases to act on lignin units, and developed an alternative way to attack wood based on Fenton chemistry. This way, they produce radicals able to diffuse and oxidize cellulose, leaving a lignin residue with untouched aromatic rings but demethylated to some extent (Martínez et al. 2011; Sigoillot et al. 2012). These organisms attack mainly softwood, i.e. the wood from gymnosperms, and occasionally hardwood. As demonstrated from genomic studies, brown-rot fungi are descendants of the white-rot lineage, losing the ligninolytic peroxidases that these other organisms have (Ruiz-Dueñas et al. 2013).

White-rot basidiomycetes are the most frequent wood-rotting organisms and the only ones able to completely mineralize lignin (Kirk and Farrell 1987). These organisms leave, after extensive lignin degradation, the white wood residue that names them (**figure 4b**). They do not use lignin as the only source of carbon, and also degrade cellulose and hemicelluloses for their consumption. Therefore, white-rot fungi are able to modify all the major components of wood. White-rot fungi are mainly Basidiomycetes,

being the most iconic and well studied *Phanerochaete chrysosporium* (Tien and Kirk 1984; Martinez et al. 2004), although other species have brought the attention because of their industrial applications, such as *Irpex lacteus* (Salvachúa et al. 2013), *Ceriporiopsis subvermispora* (Fernández-Fueyo et al. 2012) or *Pleurotus spp.* (Camarero et al. 1998; Cohen et al. 2002). There are two types of white-rot decay (Martínez et al. 2005; Sigoillot et al. 2012)

- *Simultaneous decay*: as it name suggest this type of decay is characterized by the concurrent degradation of cellulose, hemicelluloses and lignin. The anatomical features in this decomposition imply a progressive degradation of the cell wall from the lumen an also the appearance of erosions associated to hyphae. Although Basidiomycetes are the main organisms included in this type of decay, some Ascomycetes such as *Xylaria hypoxylon* are also capable of this degradation.
- *Selective decay*: in this case there is a preferential degradation of lignin (and hemicelluloses) in a first attack, often followed by degradation of cellulose. These fungi attack the middle lamella and the secondary wall by a diffusion mechanism (not direct contact with hyphae), and radial cavities appear in the cell wall.

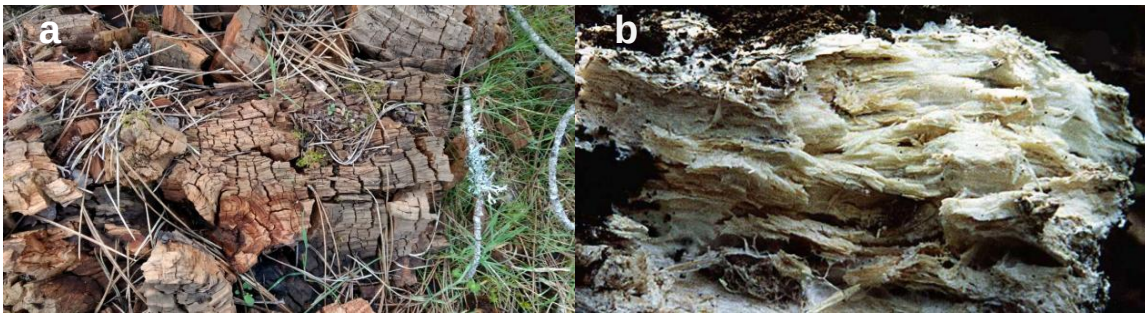


Figure 4. Wood after being attacked by brown rot **(A)** or white-rot **(B)** fungi.

White-rot fungi developed the degradation of lignin with the use of specialized enzymatic machinery to expose the biomass sugars and use them as nutrients. The enzymes involved in lignin degradation are laccases, peroxidases and H_2O_2 generating oxidases.

Laccases are multicopper oxidases that use oxygen as electron acceptor generating H_2O as byproduct. Due to their "medium" redox potential (~0.8 V) their direct action on lignin is limited to the oxidation of the phenolic moiety of the polymer. However, laccases are able to oxidize natural redox mediators that, once oxidized, diffuse and oxidize other substrates allowing the indirect action of laccases over other substrates, expanding the range of compounds they would act on directly (Cañas and Camarero 2010).

Ligninolytic peroxidases belong to the class II (PODs) of the of peroxidase-catalase superfamily (Zámocký et al. 2015). They use hydrogen peroxide as electron acceptor to oxidize a wide variety of substrates, including the direct oxidation of lignin. In the fungal ligninolytic context, hydrogen peroxide is produced by several extracellular oxidases such as aryl-alcohol oxidase, methanol oxidase, pyranose 2-oxidase and glyoxal oxidase (Ferreira et al. 2015). The characteristics of ligninolytic peroxidases are further explained below.

2.2 Ligninolytic peroxidases

2.2.1 Common features of PODs and their classification

Fungal PODs are a group of heme-proteins secreted by white-rot fungi that include four types of enzymes: manganese peroxidases (MnPs), versatile peroxidases (VPs), lignin peroxidases (LiPs), and generic peroxidases (GPs). MnPs, VPs and LiPs are high redox-potential enzymes (>1.2 V) that perform the direct or indirect oxidation of lignin, in addition to other aromatic compounds. GPs however are fungal PODs similar in catalytic properties to plant peroxidases, the latter oxidizing low-redox potential monolignols (Martínez et al. 2009).

The common structural features of PODs include a globular folding with 11-12 α-helices distributed in two domains surrounding the central cavity where the heme cofactor is located (**figure 5**). This arrangement "divides" the enzyme in two commonly referred moieties: the heme-proximal and distal sides. These sides take their names from the position of two axial residues: proximal histidine acting as heme-iron ligand and axial histidine helping activation by peroxide. PODs are stabilized by two structural Ca^{2+} ions, one in the proximal side and one in the distal side, and 4-5 disulfide bonds. The cofactor is a type-b heme (ferriprotoporphyrin IX*)* with a ferric ion (Fe^{3+}) coordinated with the four nitrogens of the porphyrin and the ε nitrogen of the proximal histidine (His176 in *P. chrysosporium* LiP A, PC-LiPA) (Martínez 2002). The sixth coordination position of the Fe^{3+} is available for the H_2O_2 to bind, starting the catalytic cycle. Additionally, all PODs have conserved residues including the proximal histidine mentioned above. Continuing the residue numbering of PC-LiPA, these conserved residues on the proximal side of the heme are Asp238, Ser177, which establish hydrogen bonds with His176 (Piontek et al. 1993) and Phe193, with a proposed role in the stability of PODs (Kishi et al. 1997). On the distal side the conserved residues are His47, Arg43 (both participating in reactivity with H_2O_2, (Erman et al. 1993; Vitello et al. 1993; Hiner et al. 2002)), Asn84 and Phe46, both establishing hydrogen bonds with the distal His47.

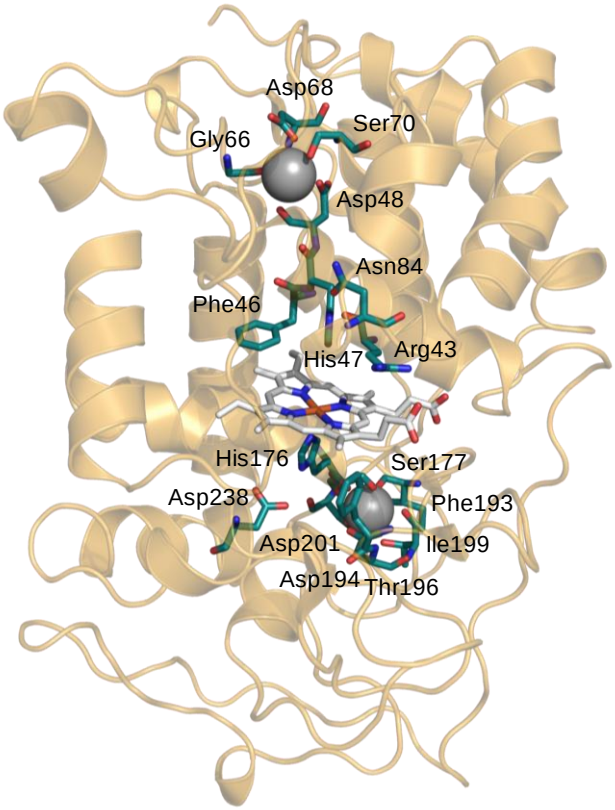


Figure 5. Conserved structure of fungal ligninolytic peroxidases. The heme cofactor is located in the center of the structure, dividing the enzyme in the proximal side, where a proximal histidine acts as the fifth ligand of heme iron (bottom) and distal side, where a distal histidine contributes to activation by H_2O_2 (top). The amino acid numbers refer to the sequence of LiP A from *P. chrysosporium* (PDB 1b82).

Although PODs have numerous conserved amino acids, they differ in the oxidation sites they have (**figure 6**) (Martínez 2002):

- *Manganese binding site*. Mn^{2+} is naturally present in wood. It has been demonstrated that the cation upregulates the production of MnPs, although it inhibits the production of LiPs and VPs (Bonnarme and Jeffries 1990; Ruiz-Dueñas et al. 1999). It can be oxidized by peroxidases to Mn^{3+} in a small channel that leads to one of the heme propionates. The Mn^{2+}-binding site is formed by three acidic residues, conserved among MnPs (Glu39, Glu35 and Asp179, following the reside numbering of MnP6 from *C. subvermispora*). Although the Mn^{2+}-binding site is a signature amino acid triad of fungal MnPs, the oxidation of Mn^{2+} has been described also in dye-decolorizing peroxidases, both fungal (Fernández-Fueyo et al. 2015), and bacterial (Brown et al. 2012). Additionally, there has been described "atypical" MnPs, fungal PODs that lacks one of the canonical acidic amino acids of the triad of MnPs (Hildén et al. 2014). Mn^{3+} is a highly reactive metal cation (E^{o}" = 1.54 V) that is unstable in aqueous solutions.

White-rot fungi secrete organic acids which associate to Mn^{3+} forming chelates, more stable but with a lower reduction potential ($E^{o'}$ = 0.9-1.2 V). These Mn^{3+} chelates are diffusible, and they oxidize the phenolic moiety of lignin (Gold et al. 2000).

- *Surface tryptophan*. In order to oxidize the bulky lignin polymer, unable to access the heme-channel, some ligninolytic PODs evolved acquiring a radical-forming surface residue that allows them to oxidize high redox potential substrates. This residue is Trp171 (numbering of PC-LiPA) that during the catalytic cycle of PODs is oxidized to an amino acidic radical by a mechanism called long-range electron transference (LRET). This way, one of the oxidation equivalents of the heme during the cycle is transferred to the surface residue, which is then able to oxidize the non-phenolic moiety of lignin and many other compounds, such as veratryl alcohol (VA, a fungal metabolite commonly used as non-phenolic lignin model) or recalcitrant industrial dyes (as reactive black 5, RB5) (Sáez-Jiménez et al. 2015b; 2016). It is important to consider not only the surface tryptophan, but also its surroundings. It has been demonstrated that differences of charge in the environment of the catalytic Trp are responsible for substrate preferences among different types of PODs. For example, VPs are able to oxidize RB5, while LiPs are not, and the main difference is a more positive charge surrounding the Trp in VP than in LiP. On the other hand, the better oxidation of lignin-model compounds by LiPs can be explained taking into account that they have a more acidic tryptophan environment that would stabilize the cationic radicals of these compounds, promoting the reaction (Ruiz-Dueñas et al. 2008).
- *Heme channel*. This aperture to the solvent allows the peroxide to access the heme and oxidize it, starting the catalytic cycle. The channel is conserved in all PODs, although its size varies between them. Also it allows the entrance of small molecules that can be oxidized in direct contact with heme, such as simple phenols (e.g 2,6-dimethoxyphenol). It has been demonstrated that enlarging the heme channel in enzymes with a small channel allows the oxidation of these type of substrates (Morales et al. 2012) but with the drawback of inactivation of the enzyme due to the generation of radicals. This considered, it is possible that ligninolytic PODs evolved narrowing the heme channel to avoid oxidative damage.

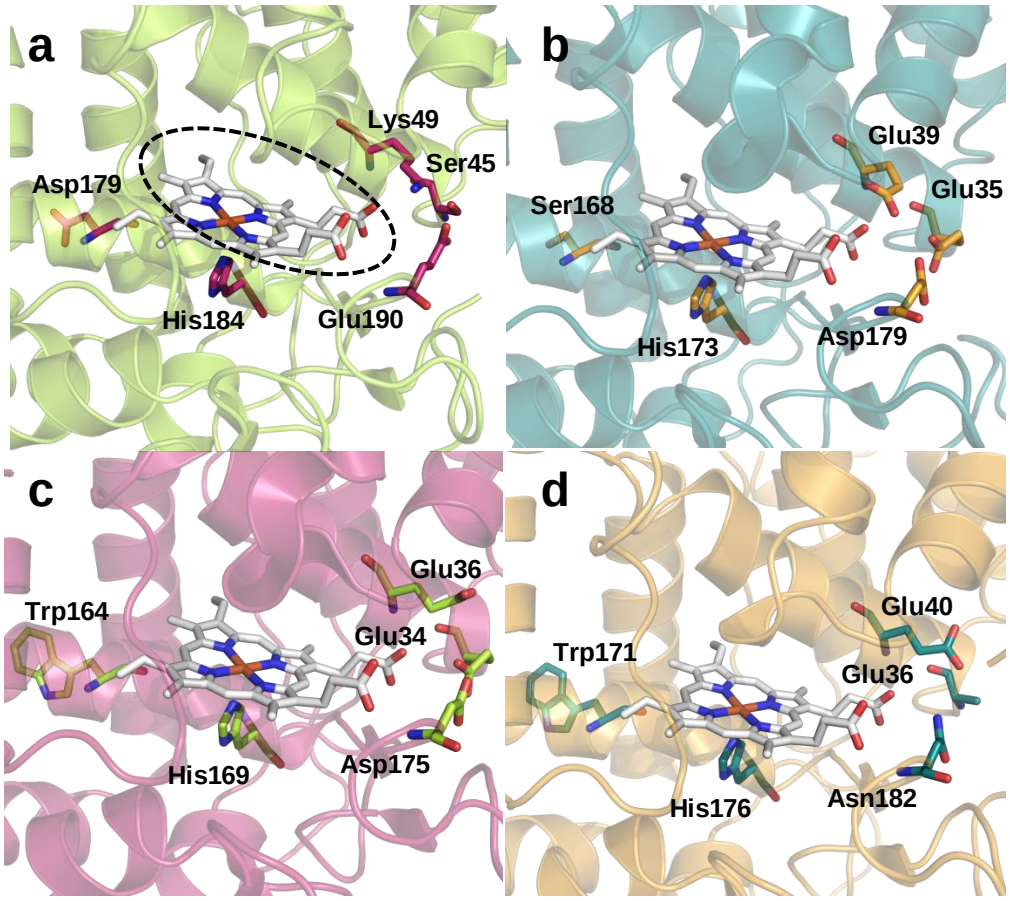


Figure 6. Different types of fungal peroxidases considering their oxidation sites. **A**: GP (*Coprinopsis cinerea* PDB 1arx), which oxidizes low redox-potential substrates at the main heme access channel (circled). **B:** MnP (*P. chrysosporium* PDB 1yyd), that in addition has a Mn^{2+}-binding site formed by Glu35, Glu39 and Asp179. **C:** VP (*Pleurotus eryngii* PDB 2boq), which aside from the Mn^{2+}-binding site, formed by Glu36, Glu40 and Asp175, also has a solvent-exposed tryptophan (Trp164) for direct oxidation of the bulky lignin polymer. **D**: LiP (*P. chrysosporium* PDB 1b82), which only has the solvent-exposed tryptophan (Trp171). Adapted from Ruiz-Dueñas et al (2009).

Taking the oxidation sites into account, PODs can be classified into four types of enzymes (**figure 6**). Generic peroxidases (GPs) only have the heme access channel, which is present in the other three types of PODs, and they can only oxidize low redox potential substrates (similar to the plant peroxidases). Manganese peroxidases (MnPs) additionally have the Mn^{2+}-binding site, and are differentiated in long- and short-MnPs (depending on the length of their C-terminal tail), with different catalytic and stability properties (Fernández-Fueyo et al. 2014). Lignin peroxidases (LiPs) have a narrowed heme channel, but they have the solvent-exposed tryptophan able to oxidize lignin directly (and other high redox-potential substrates) (Sáez-Jiménez et al. 2015b; 2016). And finally, versatile peroxidases (VPs) that combine all the oxidation sites described (Ruiz-Dueñas et al. 2009).

2.2.2 Global catalytic cycle of PODs

All PODs have the same global catalytic cycle, with exceptions for the enzymes with catalytic Trp (**figure 7**). This cycle has three consecutive oxidation/reduction steps being the initial reaction the oxidation of the resting state (RS, with a Fe^{3+} heme) enzyme by H_2O_2, extracting two electrons from the cofactor. This generates the Compound I (CI) with two oxidizing equivalents: one in an oxy-ferryl cation (Fe^{4+}=O) and the other in a porphyrin radical. Alternatively, CI of LiPs and VPs transfers one of these equivalents forming a solvent-exposed tryptophanyl radical being able to directly oxidize the bulky lignin polymer, which is unable to access the buried cofactor (Sáez-Jiménez et al. 2015a). CI is reduced to Compound II (CII) with the concomitant one-electron oxidation of an external substrate molecule. CII only has one oxidation equivalent (Fe^{4+}-OH or Fe^{4+}=O, in equilibrium with the tryptophanyl radical in LiPs and VPs), and is further reduced to the RS by oxidation of another molecule of substrate, closing the cycle.

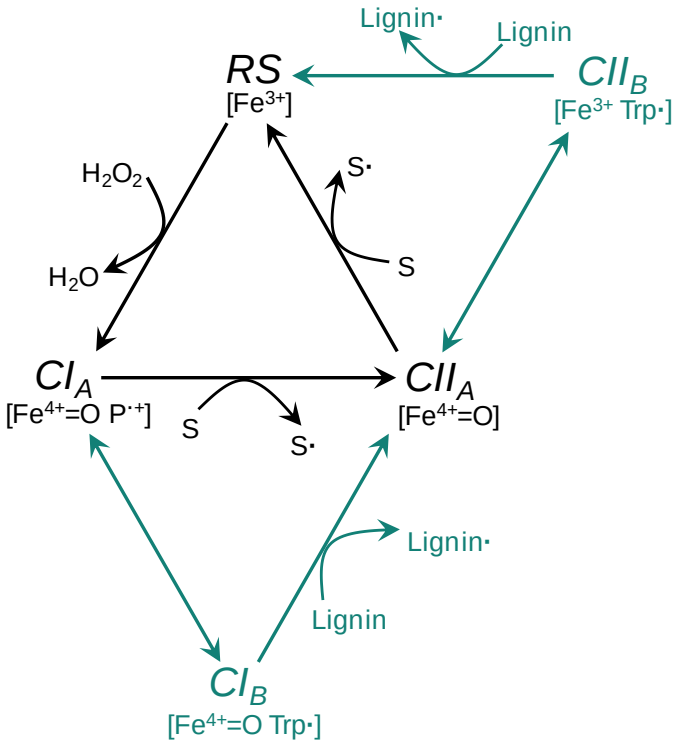


Figure 7. Global catalytic cycle of fungal ligninolytic peroxidases. In black it is represented the common cycle initiated with the oxidation of the RS enzyme by H_2O_2 to form CI. Two consecutive one-electron oxidations of substrate (S) lead to CII and RS. In blue is showed the expanded cycle for VPs and LiPs, with the displacement of one oxidation equivalent to the solvent-exposed tryptophan, able to attack lignin directly. Adapted from Sáez-Jiménez et al (2015b).

Each of the above catalytic intermediates of the cycle is characterized by different spectral features (**figure 8**), allowing the study of the catalytic cycle spectroscopically (Andrawis et al. 1988; Wariishi et al. 1991). Starting with the RS, it presents the typical Soret band of heme-peroxidases at ~410 nm (varying a few nm between enzymes) and two charge-transference bands, CT2 at ~505 nm and CT1 at ~560 nm. This spectrum reflects a heme-peroxidase with a Fe^{3+} penta-coordinated (sixth position occupied by a water molecule and available for H_2O_2) and in its high-spin form.

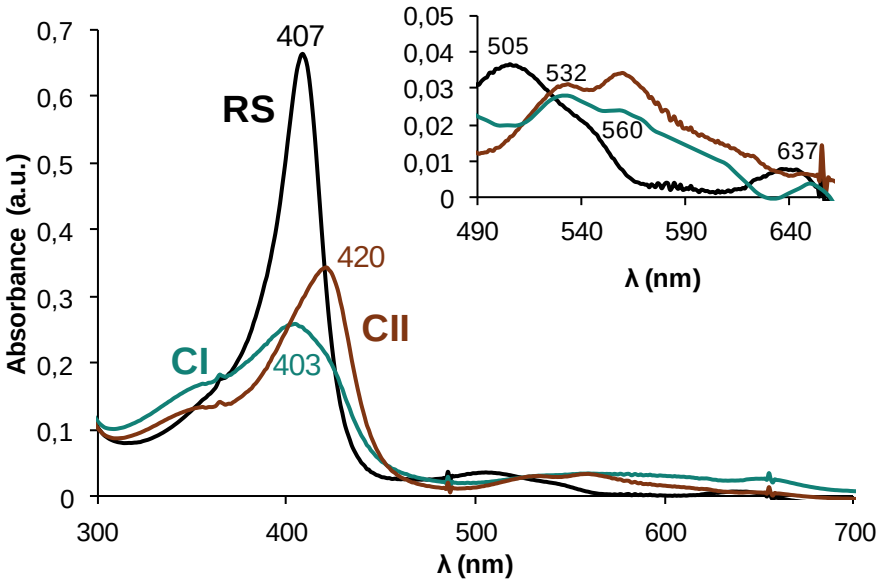


Figure 8. UV-visible spectra of the different intermediates in the peroxidase catalytic cycle. Black, resting state. Blue, compound I. Brown, CII.

As said above, the reaction with H_2O_2 leads to CI formation, whose spectrum differs from RS. In this case, the Soret band is displaced to ~400 nm, and two new maxima appear at ~530 nm and ~ 560 nm. CII is formed after one-electron reduction of the enzyme and its spectra is characterized by a maximum at ~420 nm and additional maxima at ~530 nm and ~550 nm.

The differences in spectra between the three intermediates of the catalytic cycle of ligninolytic PODs is especially useful to study its mechanism. It is possible to analyze all the semi-reactions of the cycle and measure the transient state kinetics of each reduction step. Using stopped-flow spectrophotometry it is possible to address complex reactions, such as the oxidation of complex lignin models as lignosulfonates, and to evaluate the fine aspects (transient-state constants, etc) of PODs mechanism (Sáez-Jiménez et al. 2016). Moreover, the spectral differences between RS, CI and CII allow the calculation of the formal reduction potential ($E^{o'}$) of these enzymes (Arnhold et al. 2001), one of the most elusive aspects of PODs.

2.2.3 Redox properties of PODs

The thermodynamic and kinetic properties of the catalytic cycle are heavily influenced by the $E^{o'}$ between the species mentioned above. Therefore, the determination of $E^{o'}$(CI/RS), $E^{o'}$(CI/CII) and $E^{o'}$(CII/RS) is essential to know how these enzymes work (Battistuzzi et al. 2010), and knowing the structural elements that confer that redox properties will allow a better understanding of PODs in general. Moreover, if the molecular determinants

of higher E°' values are known, the engineering of especially high redox-potential PODs will be possible.

During the last decades investigation of the redox properties of not only PODs, but heme-peroxidases in general, has been extensive (Battistuzzi et al. 2010). The combination of multiple techniques with site directed mutagenesis to measure different E°' has been useful to elucidate some of the structural determinants for the redox properties. The most used technique is the measurement of E°'(Fe^{3+}/Fe^{2+}) by redox titration employing redox mediators. Except for human myeloperoxidase (Ikeda-Saito and Prince 1985), all the E°'(Fe^{3+}/Fe^{2+}) measured in the literature show a negative E°'(Fe^{3+}/Fe^{2+}), which ensures that at physiological pH heme-peroxidases have a catalytic Fe^{3+} ready to start the cycle. However, caution is needed to infer kinetic parameters using only these data: although is generally agreed that E°'(Fe^{3+}/Fe^{2+}) influences the other redox pairs, Fe^{2+} does not belong to the catalytic cycle of peroxidases and only Fe^{3+} is efficiently oxidized by H_2O_2 (Raven and Dunford 2016).

Other methodologies are more precise, although not always a choice. Perhaps the most accurate is cyclic voltammetry, with which is applied a potential difference on a protein immobilized in an electrode, and that later will be switched to the opposite direction (Oyadomari et al. 2003). This type of potentiometric measurement shows the transition between redox species as peaks in the voltammogram obtained.

Another powerful technique is the use of stopped-flow spectrophotometry to measure directly the concentration of the transient (and unstable) redox species of the cycle in equilibrium (Sørlie et al. 2000; Arnhold et al. 2001). Taking advantage of the spectral differences between RS, CI and CII it is possible to calculate the concentrations of redox pairs in equilibrium, and using the Nerst equation, infer the E°' of each pair.

All considered, it is still unclear how PODs shape their redox power. What seems obvious is that mutations of amino acids close to the heme (either on the proximal or the distal side) affect more or less the different E°' measured (Battistuzzi et al. 2010). To try to determine if some of the conserved amino acids in PODs is critical to tune the redox power, ^{1}H-NMR was used to intensively investigate the closest amino acids to the heme Fe^{3+} (Banci et al. 1991; 1992; 1995). These studies have shown that the proximal His, coordinated in the fifth position of the Fe^{3+}, seems essential to regulate the reduction potential of PODs (**figure 9**).

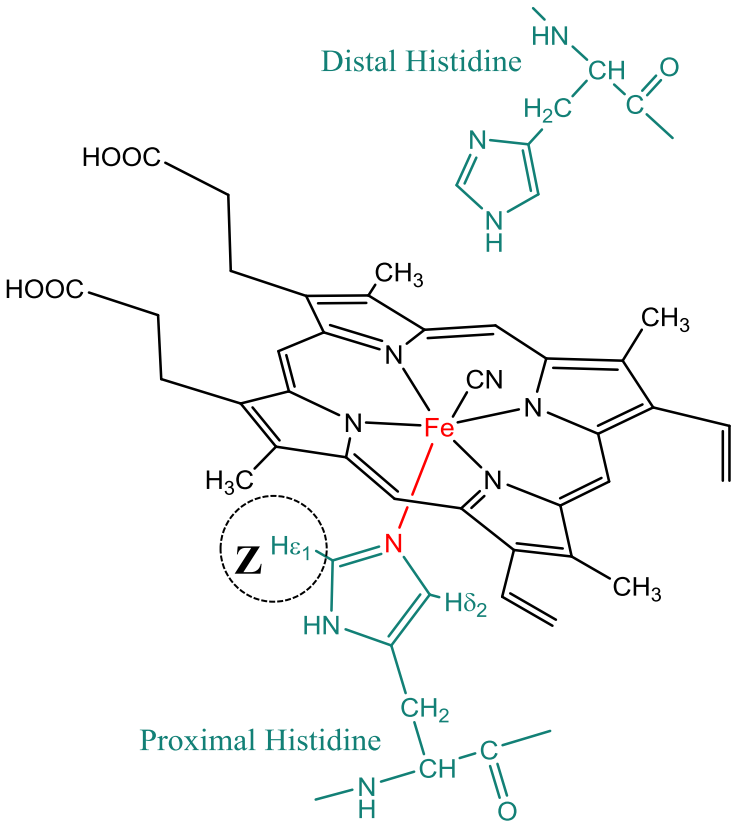


Figure 9. Chemical structure of the heme cofactor and the distal and proximal histidines. The proximal histidine is coordinated in the fifth position of the iron (red bond). Using ^{1}H-NMR of the cyanide adducts of heme-peroxidases it is possible to explore the chemical shift of several atoms close to the iron. One of the most studied signals is the one labeled as Z. The displacement of this signal indicates properties of the His-Fe bond, and has been correlated with the reduction potential of these enzymes (Banci et al. 1991).

3. Evolution of wood rotting fungi and land plants

3.1 Origin of wood-rotting fungi

There is a strong consensus that fungi (and nucleariids) are the sister group of holozoans within the clade Ophistokonta (Baldauf 2008), and their divergence is usually estimated in the Mesoproterozoic/early Neoproterozoic era (~1000 mya) (Heckman et al. 2001; Parfrey et al. 2011). The last common ancestor of fungi is considered to have been non-filamentous and aquatic, with flagellated spores (James et al. 2006). However, some studies point out to an earlier and different origin of fungi that would imply a revisitation of this phylogeny (Butterfield 2005; Bengtson et al. 2017). Specially controversial is the work by Bengtson and coworkers, given the fact that they found fungus-like filamentous fossils originated in the early Paleoproterozoic era (~2400 mya), an origin considerably older (and distinct) than that previously thought. Additionally, it has been demonstrated that

complex multicellularity in fungi appeared in different taxa as a convergent evolutionary adaptation, so the roots of these organisms have to be further investigated in detail (Nagy et al. 2018).

More clear is the origin and further evolution of Dikarya, comprised of Ascomycetes and Basidiomycetes (**figure 10**, bottom). The time-calibration of different phylogenetic trees using fungal fossils establish the origin of Dikarya in the Neoproterozoic era (~750 mya), and the divergence of Ascomycetes and Basidiomycetes in the Cambric period (~500 mya) (Parfrey et al. 2011; Floudas et al. 2012). The evolution of Basidiomycetes shows a rapid and continuous diversification of the different groups that form the order (Zhao et al. 2017). Within Basidiomycetes, the appearance of the class Agaricomycetes, where wood-rotting fungi are included, is estimated to occur at the end of the Carboniferous period (~300 mya) (Floudas et al. 2012; Zhao et al. 2017) with global geological consequences.

The work of Floudas *et al* (2012) shows that the appearance of wood-rotting fungi, together with their production of the first ligninolytic peroxidases, contributed to the end of biomass accumulation in form of coal at the end of the Carboniferous period. However, geoclimatic factors would have also significantly contributed to coal formation under ever-wet tropical conditions, and its decline could also be related to climatic shifts toward drier conditions (Nelsen et al. 2016; Hibbett et al. 2016).

In any case, after the appearance of Agaricomycetes there was a huge diversification of these fungi, generating the great abundance of species observed today. During this speciation, there was a complex evolution in terms of duplication/loss events concerning genes encoding ligninolytic peroxidases: in white rot fungi it is observed a higher ratio of duplication, while brown-rot suffered a clear loss of ligninolytic genes (Ruiz-Dueñas et al. 2013; Nagy et al. 2016a). Although the work of Floudas *et al* (2012) established the common ancestor of Agaricomycetes as a white-rot fungus, recent works involving the genome sequence of undersampled species shows that the ancestor of both white- and brown-rot fungi could have been a soft-rot fungi (Nagy et al. 2016b).

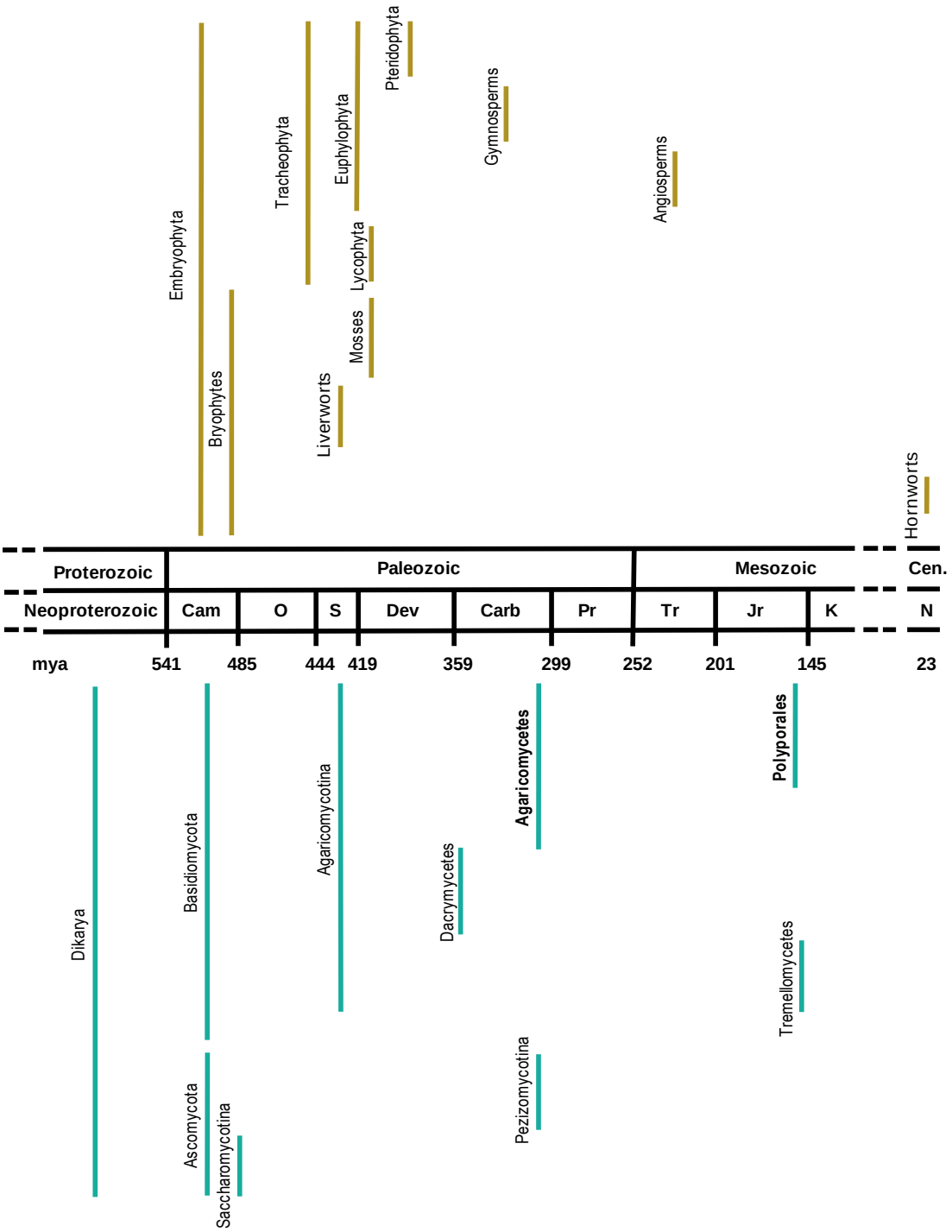


Figure 10. Evolution of land plants (Embryophyta, top) and Dikarya fungi (bottom). **Cam**, Cambric period; **O**, Ordovician; **S**, Sillurian; **Dev**, Devonian; **Carb**, Carboniferous; **Pr**, Permian; **Tr**, Triassic; **Jr**, Jurassic; **K**, Cretacic; **Cen**, Cenozoic era; **N**, Neogene period. Adapted from Floudas et al (2012) and Morris et al (2018b).

3.2 Polyporales evolution

Within Agaricomycetes, the evolution of the order Polyporales, where most wood-rotting fungi are included, has been specially analyzed. Being the main wood-decay fungi, Polyporales have been intensively studied both in basic and applied science. This order contains approximately 1800 species, organized in 18 families (Justo et al. 2017). Concerning lignin degradation, both the first white-rot fungus (*P. chrysosporium*) (Martinez et al. 2004) and brown-rot fungus (*Postia placenta*) (Martinez et al. 2009) whose genome were sequenced are members of Polyporales. Today, in the MycoCosm web of the DOE-JGI (https://genome.jgi.doe.gov/programs/fungi/index.jsf) there are 322 sequenced genomes of Agaricomycotina species, of which 67 are from the order Polyporales.

The information obtained due to genome sequencing and the careful annotation of genes and proteins allowed a deeper understanding of the phylogenetic relationships in Polyporales and their evolution. This way, Binder *et al* (2013) performed the most comprehensive phylogenomic study of Polyporales using a six-gene data set for 373 taxa, establishing the roadmap for further studies that Justo et al. (2017) expanded, with the currently accepted classification of clades within Polyporales. Also, Ruiz-Dueñas *et al* (2013) focused in the annotation of Polyporales peroxidases, which allowed a more detailed study of the different families of heme-proteins, including ligninolytic peroxidases. Concerning lignin-degradation in the evolution of Polyporales, it is clear that white-rot fungi experienced a massive duplication of genes encoding ligninolytic peroxidases, while brown rot suffered a loss of these genes, basing its biomass modification in the Fenton reaction as said above. Interestingly, the production of H_2O_2 needed for peroxidases in white-rot decay or for Fenton's reaction in brown-rot decay, is ensured in evolution with a wide distribution of oxidases that produce the peroxide in both fungal types (Ferreira et al. 2015).

Moreover, using ancestral state reconstruction the evolution of the different oxidation sites in ligninolytic peroxidases has been analyzed (Floudas et al. 2012; Ruiz-Dueñas et al. 2013; Nagy et al. 2016b). Ancestral state reconstruction allows the determination of discrete characters in the nodes of the phylogeny, which make it possible to establish hypothesis for the evolution of the different lineages observed in the phylogeny. Therefore, it has been proposed that the most ancient enzymes were most probably short MnPs. The evolution that lead to the protein clades observed today implied that some enzymes stayed as MnPs, either short or long (enlarging their C-t tail), probably developing their catalytic and stability properties. Other lineages lead to lignin peroxidases, first incorporating the solvent exposed tryptophan and later losing the Mn^{2+}-binding site. Finally, there are VPs in the phylogeny, originated when ancestral short MnPs

incorporated the catalytic tryptophan but kept the Mn^{2+}-binding site through evolution.

Although ancestral state reconstruction is useful to establish evolutionary hypothesis, it does not give information neither of structure nor biochemical properties of ancestral proteins. For that reason, it is necessary a deeper study of evolution using more advanced bioinformatic tools that allows the reconstruction of the whole sequence of amino acids at any node of the phylogeny studied, allowing the further "resurrection" of those proteins in the laboratory. Such technique is known as ancestral sequence reconstruction (ASR), and its advantages will be discussed below.

3.3 Lignin was essential for land colonization

The origin, evolution and land colonization by plants transformed the terrestrial biosphere, and its study is a fascinating field that even today remains controversial. The establishment of plants on land ecosystems had large impact in the biogeochemical cycles, enhancing silicate weathering and increasing carbon burial, what led to mayor perturbations in the global carbon cycle, dropping substantially the concentrations of atmospheric CO_2 during the Paleozoic (Morris et al. 2015) and increasing oxygenesis as well (Lenton et al. 2016). Plants on land also provided new habitats for animals and fungi, as discussed above, while modified soil composition due to their nutrient absorption, influencing rivers and landscapes in general (Gibling and Davies 2012). The development of a clear evolutionary history on early plant evolution to explain its coevolution with Earth Systems is difficult due to the lack of fossil evidence in the first stages of land colonization (Kenrick et al. 2012). However, recent efforts using molecular clock calibration of multiple phylogenetic trees, together with statistical analysis to cover uncertainties of the fossil calibration, establish a clear origin of Embryophyta (land plants) in the middle Cambrian – early Ordovician range (~ 500 mya) and the appearance of Tracheophytes (vascular plants) in the late Ordovician – Silurian interval (~ 440 mya) (Hedges et al. 2018; Morris et al. 2018a; 2018b) (**figure 10**, top).

During land colonization, the pioneer plants needed to develop a series of adaptations to overcome the challenges they were confronted to. They had to protect themselves from the UV-B radiation (280 – 315 nm), because water no longer shielded them, and due to the absence of water they needed also an structural support. In that way, the desiccation stress was also a problem, and eventually they suffered the attack of co-evolving pathogens such as arthropods or fungi (Raven 1984). As discussed above, those are the main functions lignin covers, and the evolution of lignin synthesis was essential in land colonization by plants.

It is considered that one of the main biochemical acquisitions in plants was the ability to deaminate phenylalanine (precursor of lignin synthesis, **figure 2**) and hydroxylate the aromatic ring of subsequent metabolites, allowing the accumulation of simple phenylpropenoids with absorbance maxima in the UV-B range (Weng and Chapple 2010). This important adaptation conferred the early plants (and their spores) the ability to resist the lethal UV radiation during the first stages of land colonization (Lowry et al. 1980). However, plants stayed small during million years due to a lack of structural support. The rise of tracheophytes involved the accumulation of phenylpropenoids units in the form of lignin polymer in the cell wall, allowing the vertical growing and the dominance of land ecosystems. With lignin accumulation, long-distance water transport was ensured due to the waterproofing of vessels. Finally, the complex and random nature of lignin synthesis made of it one of the most difficult polymers to degrade in nature, providing a strong shield against the attack of pathogens.

Today, it is accepted that the routes of synthesis of monolignols appeared gradually due to duplication and neofunctionalization of genes that had different functions in the primary metabolism (Weng and Chapple 2010). This way, the eight core enzymes of the route (underlined in **figure 2**), were recruited from different essential processes in plant metabolism, such as amino acid degradation or beta-oxidation of fatty acids, among others. These side functions of new genes probably generated an ancestral pool of phenylpropenoids with UV absorption properties, giving a clear evolutionary advantage to the first land plants.

It has been suggested that the rise in oxygen concentration during the Ordovician-Silurian facilitated the oxidative coupling mechanism of lignification (Lowry et al. 1980), probably being laccases the first enzymes in evolution that performed this oxidation. The task rapidly change to other enzymes with the rise of class III peroxidases of the peroxidase-catalase superfamily (Zámocký et al. 2015), the most important enzymes involved in the polymerization of lignin. Precisely due to the rise in oxygen concentration in the late Ordovician, the first functions postulated for class III peroxidases were related to protection against oxidative stress, using phenylpropenoids as reducing substrates (Kawaoka et al. 2003). The monolignol radicals eventually would polymerize into lignin, with the concomitant evolutionary advantage that this would grant.

Although the presence of lignin-like compounds in non-vascular plants is well documented (Weng and Chapple 2010; Novo-Uzal et al. 2012), lignin in tracheophytes has been considerably more studied. All tracheophytes synthesize H and G lignins, and the occurrence of S lignins is specific of flowering plants among the organisms (with some exceptions discussed below). This differential synthesis of lignin in angiosperms has been

extensively characterized, genetic and biochemically. As mentioned above, the reactions leading to S units needs two extra enzymes (F5H and COMT), being F5H present only in angiosperms.

However, different studies have demonstrated the presence of S-lignins in other plants, leading to the conclusion that the appearance of its synthesis happened several times in evolution (Weng and Chapple 2010). The most notable example is *Selaginella moellendorffii*, a lycophyte capable of S-lignin synthesis where a F5H-like enzyme (*Sm*F5H) was identified (Weng et al. 2008). This study demonstrated also the different phylogenetic origin of this family of enzymes in lycophytes, pointing to a convergent evolution in the synthesis of S-lignin in plants. This conclusion is reinforced taking into account the localization of S-lignin in angiosperms and *S. moelloendorfii*, with similar tissue specificity: fiber cells in angiosperms (but low in tracheary elements) and stem cortical cells in *S. moelloendorfii* (but not abundant in xylem). Similar tissue localization could mean similar evolutionary advantages. It has been proposed that lignin containing S-units can have better mechanical properties (Li et al. 2001). Also G/S copolymers could represent a better protection against the attack of pathogens due to its greater linkage diversity (Moerschbacher et al. 1990; Menden et al. 2007).

4. Enzyme resurrection to study evolution

4.1 Origins of ancestral sequence reconstruction

In 1859 Charles Darwin established in its book *The Origin of Species* that all organisms on Earth are the progeny of a common ancestor, and the diversity of species observed today is based on descent with modifications. He also proposed that the main force driving evolution is natural selection. Although the first statement was immediately accepted by most scientists, the second is controversial even today. Since then multiple theories attempted to explain evolution, with a long and disputed history between them (Nei 2005). The discussion is still ongoing, confronting the two main views: the neutral theory of molecular evolution *vs* the selectionist proposals. The most rational view encloses multiple perspectives taking into account both genetic and phenotypic diversity in populations. Either way, it is obvious that mutation is the key mechanism under evolution, and the study of genes and proteins is essential to understand how organisms evolve (Nei 2005).

To understand protein evolution is mandatory the construction of phylogenies using their amino acid sequences. The massive sequencing data acquired during the last decades allows the study of the evolution of almost any family of proteins. Their phylogenies not only permits its obvious

classification, but also the formulation of evolutionary hypothesis concerning them, *e.g.* how the different families of a type of enzymes perform different reactions. However, in 1963 Pauling and Zuckerkandle went a step further, and proposed the resurrection of ancestral enzymes to test evolutionary hypothesis. By comparison of the sequences of extant (modern) proteins, they suggested the mathematical inference of their ascendants, reconstructing the sequences of all the nodes that generate the phylogeny of study (Pauling and Zuckerkandl 1963) (**figure 11**). At their time they performed a theoretical study, proposing the sequence of two reconstructed hemoglobins, but pointed that producing and characterizing the ancestral proteins in the laboratory (commonly known as **resurrection**) would be a powerful tool to understand evolution not only at a molecular scale, but with higher-orders of biological interpretation. With this work, they started the exciting field of paleogenetics.

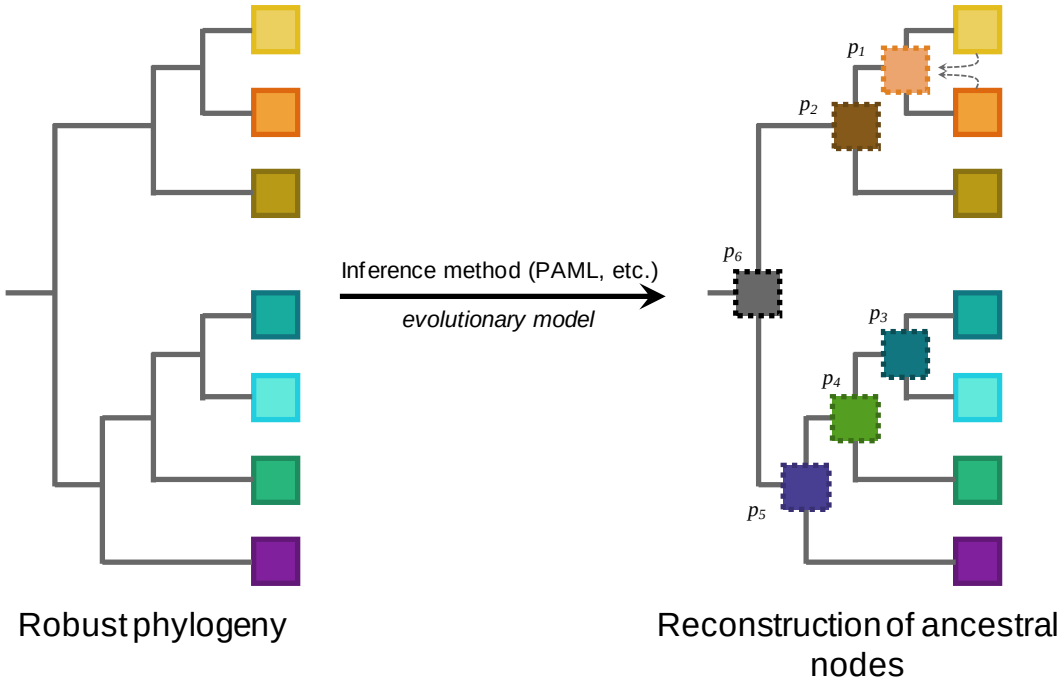


Figure 11. Schematic representation of the reconstruction of ancestral sequences. The most common approach is the use of maximum likelihood methods where a robust phylogeny, together with an evolutionary model and an inference method, is used to obtain the ancestral sequences of each node of the phylogeny. The reconstructed sequences will have an statistic parameter associated to the quality of the reconstruction, named posterior probability (p_n).

However, the first examples of enzymes resurrected in the laboratory did not appear until the 90's because there were not enough sequencing data. These pioneer works showed the resurrection of pancreatic ribonucleases from artiodactyls and bird lysozymes (Stackhouse et al. 1990; Malcolm et al.

1990). The first work resurrecting ribonucleases was especially important: after ~30 years the ideas of Pauling and Zuckerkandle were tested, and the authors connected molecular evolution with organismal evolution and even with planetary biology.

Since then, dozens of examples have been published, showing the power of this technique. Most importantly, there are examples of almost every geological era (**figure 12**) (Gumulya and Gillam 2017), and the collection of resurrected proteins is rich and diverse, covering multiple areas of knowledge. Some works target the origin of life in the Precambrian era (Gaucher et al. 2003; Risso et al. 2013; Akanuma et al. 2013), to illuminate how the first organisms on Earth were. In the field of physiology and behavior, examples as diverse as recreating visual pigments of dinosaurs (Chang et al. 2002) or determining the first primate able to metabolize alcohol (Carrigan et al. 2015) shows that this technique is not only limited to a molecular characterization, but also permits strong biological interpretations. There are also works that investigate the evolution of enzymes, analyzing the changes in their structure that led to the activities and properties they have today. Examples as diverse as determining the specificity of enzymes in secondary plant metabolism (Huang et al. 2012) or to analyze why some kinases are prone to cancer treatment but others are not (Wilson et al. 2015) show the strength of enzyme resurrection.

4.2 Uncertainty in ancestral sequence reconstruction

The multiple examples of evolutionary hypothesis tested using resurrected proteins demonstrate that ASR is very useful to study evolution. However, there are limitations inherent to the method, which must be taken into account carefully for a precise reconstruction. The accuracy of the sequences obtained depends entirely on the statistical methods and models used. Initially, parsimony methods were used because they were easier to employ but their limitations (Sanderson et al. 2000; Krishnan et al. 2004) rapidly made Bayesian and maximum likelihood (ML) approaches more prevalent. ML methods calculate the most probable amino acids for ancestral sequences using a given phylogeny, an evolutionary model and a sequence alignment. However, the true topology of the tree is rarely known, and this uncertainty is not considered in ML methods. Alternatively, Bayesian methods consider the phylogenetic uncertainty by sampling from a distribution of possible trees (Thornton 2004). Today, there is still debate

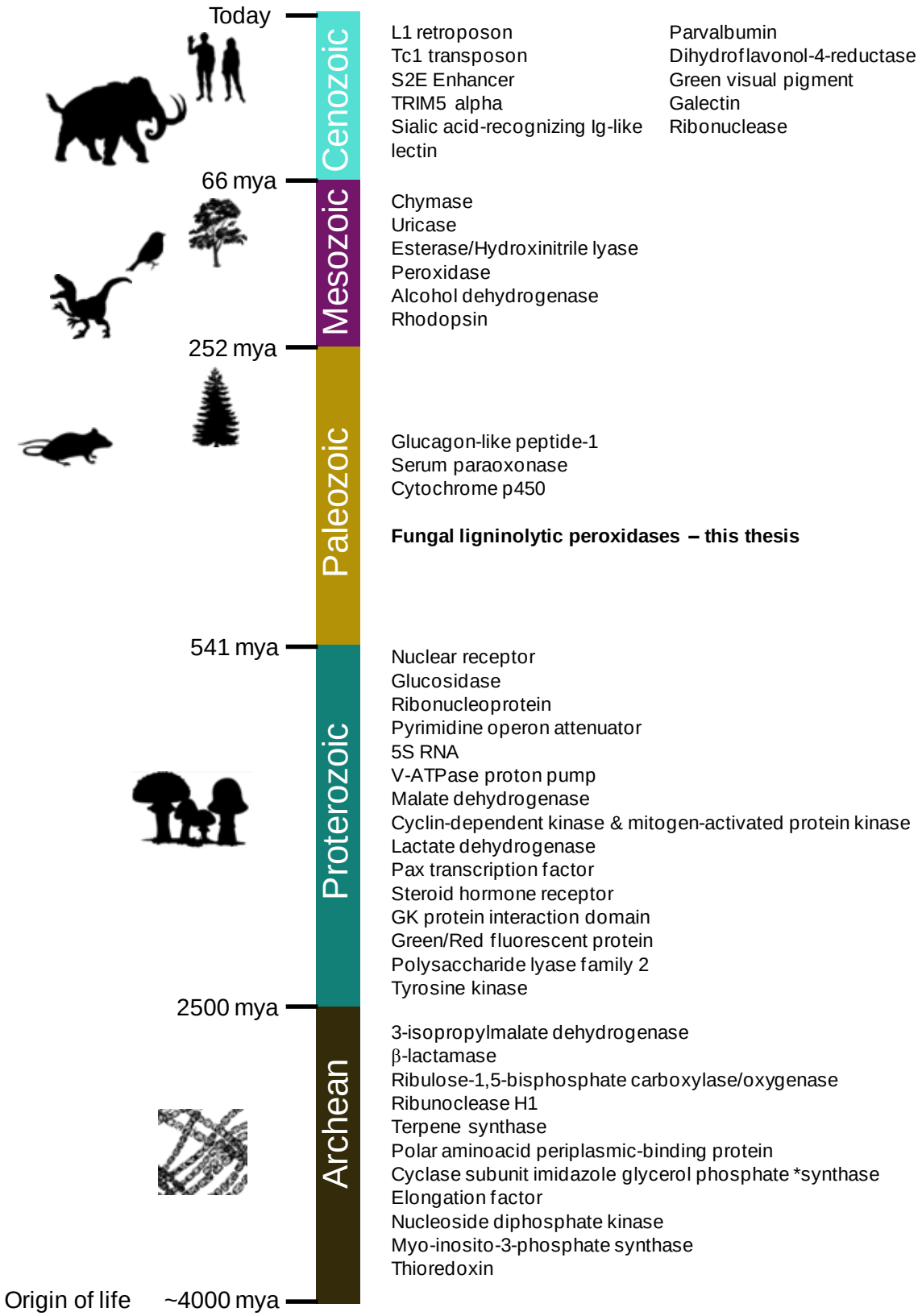


Figure 12. Ancestral proteins resurrected in the different geological eras. The silhouettes represent the different organisms that appeared through evolution. Adapted from Gumulya and Gillam (2017).

concerning whether these two methods is better, and there is not a clear answer for that, but it has been demonstrated that ML approaches are as accurate as the Bayesian ones (Thornton 2004; Hanson-Smith et al. 2010). In any case, ML methods are somehow preferred through the literature, and among all the software implementations, the package PAML (phylogenetic analysis by maximum likelihood) (Yang 2007) is definitely the most popular way for ASR.

Once the statistical method for ASR is chosen, ambiguity in the reconstruction can be managed using three complementary strategies (Liberles 2007):

- Ignoring ambiguity, considering that given the most probable amino acids are obtained, the ambiguity only happens at sites not critical for the biological function.
- Increase the amount of sequences used in the phylogenetic reconstruction. A greater number of sequences would generate more robust trees, and if the divergence is not higher hopefully the ancestral sequences will be less ambiguous. If the goal is reconstructing ancestral sequences with a biotechnological purpose, 50-200 sequences collecting the properties of interest are enough (Cox and Gaucher 2014).
- Obtaining several candidates per ancestral node to cover all alternative reconstructions and deal with ambiguity by sampling in the posterior distribution. The selection of several candidates can be made through a comprehensive study of ambiguous positions, based on structure/function studies on extant enzymes (Chang et al. 2002), or through a complete random generation of alternative ancestors using Monte-Carlo simulations (Risso et al. 2013). Given the fact that only some replacements in crucial sites of the enzyme have large impact on their function, and usually those conserved sites are known prior to the ancestral reconstruction, is easy to analyze the ambiguity in those critical positions.

Finally, characterization of the resurrected enzymes is the confirmation of a correct reconstruction and, if the properties of the enzymes support the biological interpretation, it means that the ASR is robust to the uncertainties in the methods used (Liberles 2007). Ambiguities are usually found in positions that suffered multiple changes through evolution, indicating (not always) that neutral drift is acting at a site. This generally means that the replacement of an amino acid at those sites will not have impact on the general behavior of the enzyme. Of course, all the above has to be carefully analyzed to gain certainty in the reconstruction of the sequences.

4.3 Biotechnological potential of ancestral enzymes

Besides their use to explain important evolutionary issues in the last ~30 years, resurrected proteins often show remarkable properties reflecting their adaptation to ancestral environments. Among the interesting properties they have, ancestral enzymes are in general more stable, have different substrate specificity compared to their descendants, and are often conformationally flexible (Risso et al. 2018).

A noteworthy number of works have reported the increased stability of ancestral proteins, especially to temperature (**figure 13a**). This rise in their thermal stability is especially higher when Precambrian nodes are targeted, a fact that has been discussed to be related to the origin of life in hot environments (Gaucher et al. 2003; 2008; Risso et al. 2013; Akanuma et al. 2013; Nguyen et al. 2017). In these cases, the melting temperatures of ancestral proteins are in the order of tens of degrees higher than their descendants, what makes them suitable for biotechnological applications (Wijma et al. 2013; Risso et al. 2018). Additionally, high stability is important for protein engineering because it contributes to high evolvability (Bloom et al. 2006): it is easier to include mutations that destabilize the enzyme but increase their activity. This is especially important because it may allow to "replay the tape of life" (as Jay Gould dreamt) and direct the evolution of ancestral proteins in the laboratory to obtain new descendants that satisfy different needs (Alcalde 2017).

Another property of ancestral enzymes important in biotechnology is promiscuity (**figure 13b**). In multiple enzyme families it has been already observed generalist ancestors (Siddiq et al. 2017). Promiscuous enzymes are excellent starting points for directed evolution, and combined with the higher stability of ancestral enzymes, ASR becomes an important strategy to obtain more specific reactions in more robust enzymes. It is important to note that during years the generally accepted hypothesis was that ancestral enzymes were promiscuous, catalyzing a wide array of reactions, and later specialized in evolution after gene duplication, refining their activities (Khersonsky and Tawfik 2010; Pandya et al. 2014). However, today it is accepted that this “ancestral promiscuity” is not a rule, and despite the appealing reasons that might be to consider it a general hypothesis (“moonlighting” activities, together with the difficulty for *de novo* generation of active sites and the relatively ease to obtain new functions by optimizing low-level side activities), evolution of enzymes from ancient multifunctional proteins did not always happened this way (Siddiq et al. 2017). However, even in the cases where ancestral enzymes are not promiscuous, they will perform reactions differently than their descendants, opening the possibility

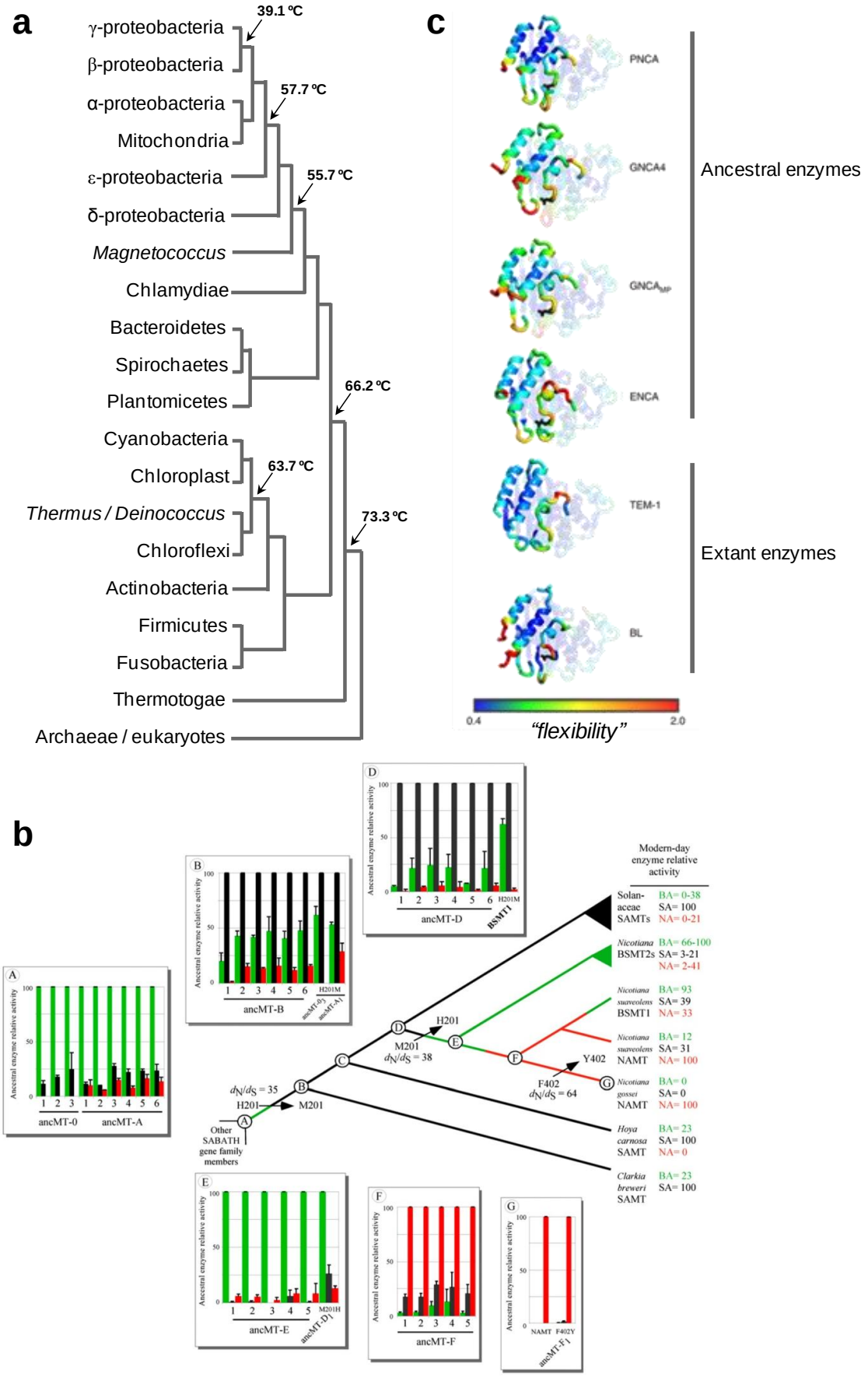

a
39.1 ℃
57.7 ℃
55.7 ℃
66.2 ℃
63.7 ℃
73.3 ℃
γ-proteobacteria
β-proteobacteria
α-proteobacteria
Mitochondria
ε-proteobacteria
δ-proteobacteria
Magnetococcus
Chlamydiae
Bacteroidetes
Spirochaetes
Plantomicetes
Cyanobacteria
Chloroplast
Thermus / Deinococcus
Chloroflexi
Actinobacteria
Firmicutes
Fusobacteria
Thermotogae
Archaeae / eukaryotes
c
PNCA
GNCA4
GNCA_MP
ENCA
TEM-1
BL
Ancestral enzymes
Extant enzymes
0.4
2.0
"flexibility"
b
Modern-day enzyme relative activity
Solan-aceae SAMTs BA= 0-38 SA= 100 NA= 0-21
Nicotiana BSMT2s BA= 66-100 SA= 3-21 NA= 2-41
Nicotiana suaveolens BSMT1 BA= 93 SA= 39 NA= 33
Nicotiana suaveolens NAMT BA= 12 SA= 31 NA= 100
Nicotiana gossei NAMT BA= 0 SA= 0 NA= 100
Hoya carnosa SAMT BA= 23 SA= 100 NA= 0
Clarkia breweri SAMT BA= 23 SA= 100
dN/dS = 35
H201
M201
dN/dS = 38
Y402
F402
dN/dS = 64
Other SABATH gene family members
ancMT-0
ancMT-A
ancMT-B
ancMT-D
BSMT1
ancMT-E
ancMT-F
NAMT F402Y
Ancestral enzyme relative activity

Figure 13 (previous page). Some properties of biotechnological interest in ancestral enzymes. **a**, enhanced thermostability in ancestral proteins has been shown in different families, representing here the increase in melting temperature of ancient bacterial elongation factors (adapted from Gaucher et al (2008)). **b**, promiscuity is present, but not always, in ancestral enzymes. In the example is showed the promiscuity of ancestral enzymes involved in the secondary metabolism of plants, acting on multiple substrates with different specificity during evolution (adapted from Huang et al (2012)). **c**, conformational flexibility is another interesting property in ancestral enzymes. The example shows the differences in flexibility between ancestral and extant beta-lactamases, which allowed the *de novo* design of an active site in ancestral enzymes (adapted from Risso et al (2017)).

for new catalytic properties that could be of biotechnological relevance (Kratzer et al. 2014).

Finally, the conformational dynamics of ancestral proteins are of interest to clarify some of their remarkable properties (**figure 13c**). A higher conformational flexibility is directly related to activity (Tokuriki and Tawfik 2009), and exploring the different conformations of a protein would explain the variety of reactions they can perform, even their promiscuous behavior (Risso et al. 2018). This is supported by the fact that ancestral enzymes often share the same 3D folding as their descendants, even with very different sequences (Risso et al. 2013), but they perform different reactions. Therefore, conformational dynamics could help understanding this interesting behavior. This approach has been used to study the evolution of steroid receptors, showing that promiscuous ancestors have highly flexible pockets, while those more specific are more rigid due to mutations in evolution (Glembo et al. 2012). Again, flexible ancestral proteins become potential scaffolds for protein engineering of new activities within the same structure. This has been tested recently using β-lactamases, generating a *de novo* site for Kemp-elimination activity in ancestral (and conformational flexible) enzymes, a strategy that failed in their more rigid extant counterparts (Risso et al. 2017).

4.4 Resurrection of ligninolytic peroxidases

For all the reasons explained above, the resurrection of ancestral peroxidases to study the evolution of ligninolytic fungi is appealing. Floudas *et al* (2012) established the origin of wood-rotting fungi in the Carboniferous period, associated to the production of the first ligninolytic peroxidases. Later, Ruiz-Dueñas *et al* (2013) expanded the work in the order Polyporales, where most-wood rotting fungi are included. Both works are of outstanding interest, and using ancestral state reconstruction established a series of hypothesis for evolution of lignin-degrading fungi.

In this thesis, the main objective is the ancestral sequence reconstruction of ligninolytic peroxidases for their resurrection and characterization in the laboratory. This will provide ancient proteins to elucidate essential aspects

of fungi evolution such as the concomitant change in catalytic properties with the modification of the oxidation sites, when direct lignin oxidation appeared, their adaptation to acidic environments where they act today or the fine-tuning of their unique high redox-potential required for lignin degradation. This in-detail understanding of fungal evolution is only possible with the use of ASR, but at the same time a new plethora of enzymes will be provided for their engineering or their biotechnological use. Even more, the study of ancestral peroxidases could help to understand essential ecological aspects of ancestral paleo-environments (**figure 14**).

Figure 14. Artistic representation of the Mesozoic landscape, based on the art of Phil Wilson (https://www.deviantart.com/phillustr8r). The use of ancestral enzyme resurrection to obtain ancient ligninolytic peroxidases will allow the study of the evolution of lignin degrading fungi and even some aspects of the ancient ecosystems where these organisms cohabited with other species.

Objetivos / Objectives

Objetivos

El objetivo general de esta tesis doctoral ha consistido en determinar cómo ha sido el proceso evolutivo de las peroxidasas ligninolíticas de basidiomicetos mediante la reconstrucción, resurrección y análisis de enzimas ancestrales. Para poder llevar a cabo este estudio se propusieron una serie de tareas asociadas a diferentes objetivos específicos:

i. Reconstruir secuencias ancestrales, usando la filogenia de las peroxidasas de Polyporales y el software PAML, y analizar *in silico* las enzimas reconstruidas. El análisis de las secuencias de aminoácidos alineadas y de las estructuras 3D de estas enzimas (obtenidas mediante modelado por homología) constituirán la estrategia para evaluar diferentes líneas evolutivas y determinar los ancestros de interés para su posterior resurrección en el laboratorio.

ii. Diseñar los genes de peroxidasas ancestrales de las líneas evolutivas a resucitar y producir, purificar y caracterizar las proteínas que codifican. Caracterizar los distintos sitios de oxidación de las peroxidasas ligninolíticas usando sustratos modelo. Evaluar su estabilidad tanto a temperatura como a pH de cara a futuras aplicaciones biotecnológicas.

iii. Poner a punto la medida del potencial redox de peroxidasas (titulación redox y espectrofotometría de flujo detenido, *stopped-flow*) para analizar como estas enzimas han adquirido a lo largo de la evolución el alto potencial redox que las caracteriza. Emplear NMR de proteínas para evaluar modificaciones en el entorno del cofactor hemo que puedan explicar los posibles cambios en las propiedades redox.

iv. Evaluar la evolución de las peroxidasas ligninolíticas mediante el estudio de reacciones de las enzimas actuales y ancestrales sobre lignina real procedente de angiospermas y gimnospermas usando técnicas de vanguardia como espectrofotometría de *stopped-flow* y 2D-NMR. Los resultados de estos experimentos ayudarán a entender cómo estas enzimas han ido adaptando sus propiedades frente a los distintos tipos de madera, correlacionando su evolución con la de la lignificación en plantas.

Objectives

The main goal of this thesis is to determine how was the evolution of ligninolytic peroxidases in basidiomycetes, using the reconstruction, resurrection and analysis of ancestral enzymes. To that end, a series of tasks corresponding to different specific objectives were proposed:

i. To reconstruct ancestral sequences, using the phylogeny of Polyporales peroxidases and the software PAML, and to analyze *in silico* the reconstructed enzymes. The analysis of the aligned sequences and their 3D structures (obtained by homology modeling) will be the strategy to evaluate different evolutionary lineages and to determine the ancestors of interest for their posterior resurrection in the laboratory.

ii. To design the ancestral peroxidases genes for the resurrection and production of the selected evolutionary lineages, and to purify and characterize the encoded proteins. To characterize the different oxidation sites of the ligninolytic peroxidases using model substrates. To evaluate their temperature and pH stability towards their industrial application.

iii. To optimize the measurement of the redox potential of peroxidases (redox titration and stopped-flow spectrophotometry) to analyze how these enzymes acquired their uniquely high reduction potential. To employ protein NMR for the evaluation of rearrangements in the heme pocket that could explain the hypothetical changes in redox properties.

iv. To evaluate ligninolytic peroxidases evolution with reactions of the extant and ancestral enzymes using real lignin from angiosperms and gymnosperms and employing state-of-the-art techniques such as stopped-flow spectrophotometry and 2D-NMR. The results from these experiments will help to understand how these enzymes adapted their properties against the different types of wood, correlating their evolution with that of plant lignification.

Materials & Methods

Phylogenetic analysis and ancestral sequence reconstruction

Every class-II peroxidase (113 sequences) annotated in the genomes of ten Polyporales (phylum Basidiomycota) species (namely *B. adusta*, *C. subvermispora*, *D. squalens*, *F. pinicola*, *Ganoderma* sp., *P. brevispora*, *P. chrysosporium*, *P. placenta*, *T. versicolor* and *W. cocos*, see file S1) were used in this study, being available at the DOE JGI Mycocosm portal as http://genome.jgi.doe.gov/Bjead1_1/Bjead1_1.home.html (Binder et al. 2013), http://genome.jgi.doe.gov/Cersu1/Cersu1.home.html (Fernández-Fueyo et al. 2012), http://genome.jgi.doe.gov/Dicsq1/Dicsq1.home.html (Floudas et al. 2012), http://genome.jgi.doe.gov/Fompi3/Fompi3.home.html (Floudas et al. 2012), http://genome.jgi.doe.gov/Gansp1/Gansp1.home.html (Binder et al. 2013), http://genome.jgi.doe.gov/Phlbr1/Phlbr1.home.html (Binder et al. 2013), http://genome.jgi.doe.gov/Phchr2/Phchr2.home.html (Ohm et al. 2014), http://genome.jgi.doe.gov/Pospl1/Pospl1.home.html (Martinez et al. 2009), http://genome.jgi.doe.gov/Trave1/Trave1.home.html (Floudas et al. 2012) and http://genome.jgi.doe.gov/Wolco1/Wolco1.home.html (Floudas et al. 2012), respectively.

The amino-acid sequences were aligned using MUSCLE as implemented in MEGA 7 (Kumar et al. 2016). The sequences were tested using ProtTest (Abascal et al. 2005) to determine the evolutionary model that best fits the data for the ML analysis among 64 empirical models of evolution. The ML phylogeny was then constructed using RAxML (Stamatakis et al. 2008), under the Whelan and Goldman (2001) model of evolution using gamma-distributed rate of heterogeneity (gamma shape with 4 rates of categories = 1.33) with empirical amino-acid frequencies and invariant sites (proportion of invariant sites = 0.073) (WAG+I+G+F).

PAML 4.7 package (Yang 2007) was used to obtain the posterior amino-acid probability per site in each ancestor under the WAG model of evolution, and the most probable whole sequences for each of the nodes, using as inputs the ML phylogeny and the MUSCLE alignment previously obtained (PAML reconstructions using the LG and Dayhoff evolution models were also performed for comparison). Marginal reconstruction was selected for the present work. The most probable sequences of the marginal reconstruction were selected and manually corrected for C-terminal and other insertions or deletions according to the sequences of the ancestor progeny.

Time calibration of ancestral nodes in Polyporales phylogeny

To determine the ages of the ancestral nodes, the peroxidase phylogenetic tree was converted in an ultrametric tree using the recognized non-parametric method implemented in PATHd8 (Britton et al. 2007). The

algorithm in PATHd8 only considers the mean path lengths from a node to its descendants and addresses deviations from the molecular clock locally. We used as time constraints those obtained by Floudas *et al.* (Floudas et al. 2012) using Bayesian methods corresponding to: i) the diversification of Pezizomycotina from the rest of the phylogeny (631 mya with 533-762 mya maximum deviations); ii) the internal diversification of *S. nodorium* GPs (403 mya with 276-533 mya maximum deviations); and iii) the diversification of the Antrodia clade peroxidases (formed by *F. pinicola, P. placenta* and *W. cocos*) from the rest of Polyporales peroxidases (399 mya with 446-207 mya maximum deviations).

Protein modeling

Molecular models of the predicted proteins were obtained at the Swiss-Model protein homology modeling server (Guex et al. 2009; Biasini et al. 2014) using related crystal structures, selected using the GMQE parameter, as templates (PDB entries 4BM1, 3FJW, 1QPA and 1B85 corresponding to *Pleurotus ostreatus* MnP4, *Pleurotus eryngii* VPL and *P. chrysosporium* LiPD and LiPA, respectively). All protein models had great quality taking into account the Swiss-Model parameters (good QMEAN and high GMQE). The electrostatic surfaces were computed with PyMOL v1.8 (Schrödinger LLC, http://pymol.org) using default parameters.

***E. coli* expression**

After gene synthesis, the coding sequences of the most probable ancestors selected (common ancestor of Polyporales peroxidases, CaPo, common ancestor of Clade B peroxidaes, CaB, Common ancestor of Clade D peroxidases, CaD, ancestral versatile peroxidase in line D, AVP-d, ancestral versatile peroxidase in line B, AVP-b and ancestral lignin peroxidas, ALiP), plus the two additional CaPo sequences (CaPo-bis and CaPo-tris), and the extant enzymes *Phanerochate chrysosporium* LiP A (PC-LiPA, JGI ID# 2989894) and *Trametes versicolor* VP2 (TV-VP2, JGI ID# 29236) were cloned into pET23b(+) (Novagen). The resulting plasmids were transformed into *E. coli* DH5α for propagation and conservation, and in BL21(DE3)pLysS (ancestors and TV-VP2) or W3110 (PC-LiPA) for expression. With this purpose, cells were grown in Terrific broth until an OD_{600} 0.5-0.6 to reach an adequate expression level, induced with 1 mM isopropyl-ß-D-thiogalactopyranoside, and grown for another 4 h.

The apoenzymes accumulated in inclusion bodies, as reveled by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and were activated *in vitro* (Pérez-Boada et al. 2002; Fernández-Fueyo et al. 2014). After solubilization in 8 M urea the *in vitro* activation conditions of the ancestral proteins and extant TV-VP2 were optimized - including 0.16 M urea, 5 mM

$CaCl_2$, 15 μM hemin, 0.4 mM oxidized glutathione, 0.1 mM dithiothreitol and 0.1 mg/mL of protein in 50 mM Tris-HCl, pH 9.5 - while those reported by Doyle and Smith (1996) where used for PC-LiPA. The active enzymes were purified with a Resource-Q column (GE-Healthcare) using a 0-400 mM NaCl salt gradient, 2 mL/min flow, in 10 mM sodium tartrate, pH 5.5, containing 1 mM of $CaCl_2$.

Steady-state kinetics

Five different substrates were selected for the kinetic characterization of the resurrected peroxidases: **i)** Mn^{2+}, which is oxidized to Mn^{3+} (Mn^{3+}-tartrate complex ε_{238} 6500 $M^{-1} \cdot cm^{-1}$); **ii)** VA, whose oxidation product is veratraldehyde (ε_{310} 9300 $M^{-1} \cdot cm^{-1}$); **iii)** DMP, which dimerizes to coerulignone (ε_{469} 55000 $M^{-1} \cdot cm^{-1}$); **iv)** ABTS, which is oxidized yielding the cation radical (ε_{436} 29300 $M^{-1} \cdot cm^{-1}$); and **v)** RB5, whose disappearance/discoloration after oxidation was measured (ε_{598} 30000 $M^{-1} \cdot cm^{-1}$). These reactions were analyzed using a Thermo Scientific Biomate5 spectrophotometer, at 25 ºC and the optimal pH and H_2O_2 concentration for each enzyme, determined using 50 mM Britton-Robinson (B&R) buffer (pH 2-10) and 2.5 mM ABTS as substrate. The sigmoid kinetic curves obtained for oxidation of DMP and ABTS enabled calculation of two sets of constants for AVP, corresponding to low and high efficiency oxidation sites, as reported for extant VP (Morales et al. 2012).

pH and temperature stability

To study the effect of pH on enzyme stability, the resurrected peroxidases and extant enzymes were incubated in B&R buffer, pH 2-10, at 25 ºC for 4 h. Then, the residual activity was estimated by the oxidation of ABTS (2.5 mM), under the conditions described above. For every enzyme, the activity after 1 min incubation at 25 ºC in pH 5 buffer was taken as 100%, and the percentages of residual activity at the different pH conditions were referred to this value.

To study their thermal stability, the enzymes were incubated in 10 mM acetate, pH 5.5, for 10 min at 5 ºC intervals in the range 25-85 ºC. Residual activity was measured and calculated as described above. Temperature stability was presented as the 10 min T_{50}, i.e. the temperature at which 50% of the activity was lost after 10 min incubation.

The effect of temperature on circular dichroism spectra was addressed studying the changes at 222 nm from 20 ºC to 95 ºC, 30 ºC/h, using a Jasco J-815 spectropolarimeter equipped with a Peltier temperature controller and a thermostated cell holder on a 0. 1 cm path length quartz cell. A final concentration of 6 μM pure enzyme in 10 mM acetate, pH 5.5 was used. T_m

represents the temperature at the midpoint of the unfolding transition in the thermal melting profiles.

Spectro-electrochemical measurement of Fe^{3+}/Fe^{2+} reduction potential

Electronic absorption spectra showing the characteristic Soret band were measured with a Shimadzu UV-2401PC spectrophotometer. Redox titrations were controlled with a BAS-CV27 potentiostat and a FLUKE 77 Series II voltmeter. The potentials were calculated *vs* the standard hydrogen electrode at pH 7 (Tris/HCl 20 mm with 0.2 M of KCl) and 25 ºC for comparison with the literature. For each measurement, 10 µL of enzyme (0.6-1.0 mM concentration) were placed in an *ad hoc* cell,(Moss et al. 1990) provided by Prof. Mäntele, with a 6 µm gold mesh (Buckbee-Mears) as working electrode, a platinum plate as auxiliary electrode and a silver/silver chloride reference electrode whose potential was checked prior and after each experiment. The potentiometric titrations were carried out in the presence of 50 µM of the following redox mediators: methylene blue, 2-hydroxy-1,4-naphtoquinone, anthraquinone-1,5-disulfonate, anthraquinone-2-sulfonate, neutral red and benzyl viologen. To quantify the oxidized and reduced enzyme, we took the absorbance at the 410 nm (Fe^{3+}) and 438 nm (Fe^{2+}) maxima, and adjusted the values to the Nernst equation:

$$[1]\ A_{410} = A_{\max 410}\frac{e^{\frac{(E-E^{O'})nF}{RT}}}{1+e^{\frac{(E-E^{O'})nF}{RT}}}$$

$$[2].\ A_{438} = A_{\max 438}\frac{1}{1+e^{\frac{(E-E^{O'})nF}{RT}}}$$

Stopped-flow experiments for redox-potential calculations

The enzyme in resting state (RS), Compound-I (CI) or Compound-II (CII) states was followed in a stopped-flow rapid spectrophotometry equipment (Bio-Logic) synchronized with a diode array detector (J&M), and the BioKine software. All experiments were made in 100 mM tartrate, pH 3, at 25 ºC. For CI formation, a typical experiment consisted of mixing 4 µM enzyme with different concentrations of H_2O_2 (1-12 molar equivalents till equilibrium) for 3 seconds. To analyze the reduction of CII to RS, a typical experiment started by mixing 8 µM enzyme with enough equivalents (usually 2) of H_2O_2 and 1 equivalent of potassium hexacyanoferrate(II) (ferrocyanide) during 1 second, to ensure total conversion to CII. After that, adding different concentrations of tyrosine, which is oxidized to tyrosinate radical (Tyr·) by the peroxidases, allowed to quantify the reduction of CII to RS, measuring up to 1 min. All experiments were at least triplicates.

Standard reduction potentials of CI/RS, CI/CII and CII/RS couples

The formal reduction potentials ($E^{o\prime}$) of CI/RS and CII/RS were calculated using the above-described stopped-flow spectrophotometry,(Sørlie et al. 2000; Furtmüller et al. 2003) and the $E^{o\prime}$ of CI/CII was inferred. In the two former cases, the reduction potential was determined using the Nernst equation at equilibrium:

[3] $\Delta E^{o\prime} = (RT / nF) \ln K'$

that correlates the difference of reduction potentials between enzyme and substrate with the equilibrium constant K'. R is equal to 8.31 J K^{-1} mol^{-1}, T is set to 298 K, n represents the number of electrons transferred in a single reaction step of the redox couple, and F (the Faraday constant) is 96,485 J V^{-1} mol^{-1}. K' represents the equilibrium constant, and is calculated as follows for the couples CI/RS (equation 4) and CII/RS (equation 5):

[4] $K' = ([H_2O_2][RS])/[CI]$
[5] $K' = ([Tyr^{\cdot}][RS])/([Tyr][CII])$

The different redox species at equilibrium were quantified as follows: i) for the pair CI/RS, the amounts of CI and RS were estimated with their extinction coefficients at 410 nm (see below) at the equilibrium (i.e. when the spectral changes ended, during H_2O_2 addition; see Figure 3A) and the reduction potential for the pair H_2O_2/H_2O at pH 3 ($E^{o\prime}$ = 1.56 V);(Koppenol 1987) and ii) for the CII/RS pair, the equilibrium concentrations of CII and RS (see Figure 3B) were calculated using their extinction coefficients at 410 nm, after incubation with different amounts of tyrosine, and the calculated $E^{o\prime}$ = 1.18 V for the pair $Tyr^{\cdot}$/Tyr.(DeFelippis et al. 1989) The selection of Tyr as substrate was made taking into account that: i) all enzymes of the PC-LiPA lineage were able to oxidize it (data not shown); and ii) it has been previously used to calculate the $E^{o\prime}$ of mammalian peroxidases using a similar approach.(Arnhold et al. 2001; Furtmüller et al. 2003; Furtmüller et al. 2005; Battistuzzi et al. 2010)

At a specific wavelength, the absorbance is an additive measurement of those of the individual components of a mixture. Therefore, using the 410 nm (Soret band) extinction coefficients for RS, CI and CII (given below) the quantification of the different redox pairs at equilibrium is possible using the equations 6 (for equilibrium of CI and RS) and 7 (for equilibrium of CII and RS):

[6] $A_{410} = \varepsilon_{410\text{-}RS}[RS]\, l + \varepsilon_{410\text{-}CI}[CI]\, l$
[7] $A_{410} = \varepsilon_{410\text{-}CI}[CI]\, l + \varepsilon_{410\text{-}CII}[CII]\, l$

where l is the path length of the stopped-flow cuvette.

The RS ε_{410} values were: 138 mM^{-1} cm^{-1} (RS-CaPo), 177 mM^{-1} cm^{-1} (RS-CaD), 149 mM^{-1} cm^{-1} (RS-AVPd), 171 mM^{-1} cm^{-1} (RS-ALiP) and 168 mM^{-1} cm^{-1} (RS-PC- LiPA). The CI ε_{410} was calculated after converting all the RS enzyme into CI using 2-10 H_2O_2 equivalents, ensuring there is no auto-reduction to CII. The values obtained were: 53 mM^{-1} cm^{-1} (CI-CaPo), 74 mM^{-1} cm^{-1} (CI-CaD), 68 mM^{-1} cm^{-1} (CI-AVPd), 87 mM^{-1} cm^{-1} (CI-ALiP) and 110 mM^{-1} cm^{-1} (CI- PC-LiPA). The CII ε_{410} was calculated after converting all RS enzyme into CI and then into CII, without further auto-reduction to RS, using 2 equivalents of H_2O_2 and 1 equivalent of ferrocyanide.(Wariishi et al. 1990) The values obtained were: 73 mM^{-1} cm^{-1} (CII-CaPo), 98 mM^{-1} cm^{-1} (CII-CaD), 105 mM^{-1} cm^{-1} (CII-AVPd), 133 mM^{-1} cm^{-1} (CII-ALiP) and 105 mM^{-1} cm^{-1} (CII- PC-LiPA).

Concerning the $E^{o'}$ of CI/CII, we were unable to calculate it experimentally. Despite obtaining enzyme in CI and being able to attain certain equilibrium with CII at specific Tyr concentrations, the CII auto-reduction to RS enzyme made impossible to calculate equilibrium concentrations. However, it was possible to infer the $E^{o'}$ of CI/CII from the experimental values of the CI/RS and CII/RS couples. The standard reaction free energy is connected to $E^{o'}$ according to:

[8] $\Delta G_r'^{o} = -n\, F\, E^{o'}$

with n = 2 electrons for the reduction of CI to RS, $\Delta G_r'^{o}$ equals to -2 x F x [$E^{o'}$(CI/RS)] being the sum of the reaction free energy of one electron reductions (CI to CII, and CII to RS). Therefore, the sum of reaction free energies $\Delta G_r'^{o}$(CI/CII) + $\Delta G_r'^{o}$(CII/RS) and the experimental determination of $E^{o'}$(CII/RS) allows the determination of $E^{o'}$(CI/CII).

^{1}H NMR spectroscopy

The cyanide adducts of the ancestral peroxidases and extant PC-LiPA were obtained by incubating 0.7-1.0 mM of each sample with KCN in 50 mM potassium phosphate, pH 6.5, prepared with ^2H_2O (isotopic purity 99.9%). The ^{1}H-NMR spectra were recorded at 298K using a 600 MHz Bruker spectrometer equipped with cryoprobe. Fast scanning WEFT pulse sequence was used.(Inubushi and Becker 1983) Delays between 13 to 120 ms were tested in order to maximize signals corresponding to fast relaxing protons of the proximal histidine coordinated to Fe of the heme group. Spectra were acquired with up to 40K total scans, 120 ppm spectral width and a

minimum of 1K points. The spectra were processed with TOPSPIN 3.0 (Bruker) with manual base line adjustment and the signal of residual water proton (δ_H 4.701 ppm) was used as internal reference for chemical shifts.

Precise conditions of the presented spectra are: i) CaPo, addition of 5 spectra of 1K scans and 128 dummy scans each, 15 ms delay, 100 ppm width and 1K points; ii) CaD, addition of 5 spectra of 4K scans and 128 dummy scans each, 13 ms delay, 120 ppm width and 1K points; iii) AVPd, addition of 5 spectra of 1K scans and 128 dummy scans each, 13 ms delay, 120 ppm width and 1K points; iv) AliP, addition of 5 spectra of 4K scans and 128 dummy scans each, 23 ms delay, 100 ppm width and 2K points; and v) PC-LiPA, addition of 10 spectra of 4K scans and 128 dummy scans each, 13 ms delay, 120 ppm width and 1K points.

Signal corresponding to Hε1 of the proximal histidine was assigned to the signal with highest upper-field shift by homology with reported data.(Banci et al. 1991a; Banci et al. 1991b; Banci et al. 1992; Banci et al. 1995; Banci et al. 2003) Other signals were assigned by comparison with reported data for PC-LiPA isolated from *P. chrysosporium* cultures, although some differences are observed as the PC-LiPA used in this study was obtained by heterologous expression in *E. coli* and devoid of natural glycosylation. Heme group numbering and protons labeling are indicated in Figure 5B and correspond to those used by Banci et al.(Banci et al. 1995)

Softwood and hardwood lignins

Water solubility is required for the transient kinetic studies, described below. Therefore, to represent natural lignin, two water-soluble sulfonated lignins were used: softwood (*Picea abies*) and hardwood (*Eucalyptus grandis*) lignosulfonates provided by G. E. Fredheim (Borregaard AS, Sapsborg, Norway). Lignosulfonates were prepared for reactions according to Sáez-Jiménez *et al*, (Sáez-Jiménez et al. 2016) including dialysis in 10 mM EDTA, 50 mM Tris (pH 8) to remove Mn^{2+} traces, and finally in Milli-Q water twice.

Enzyme (transient-state) kinetics of lignin oxidation

The reduction of peroxidases CI (to CII) and CII (to RS) was evaluated using hardwood and softwood lignosulfonates or Mn^{2+} as enzyme reducing substrates in a stopped-flow rapid spectrophotometry equipment (Bio-Logic, Claix, France) with a three syringe module (SFM300) synchronized to a diode array detector (J&M, Essingen, Germany), and BioKine software. All reactions were carried out in 0.1 M tartrate (pH 3 for lignosulfonates or pH 5 for Mn^{2+}).

Specifically, the reduction of CI was analyzed by mixing the enzyme (1 μM final concentration) with H_2O_2 (1-8 molar equivalents, depending on the

enzyme) for 3 s, observing total CI formation. Next, different concentrations of either lignosulfonates or Mn^{2+} in 0.1 M tartrate (at the adequate pH) were added, and the reactions were followed at 416 nm, the isosbestic point of CII and the resting state of peroxidases. CII reduction was studied by mixing a solution containing the enzyme and ferrocyanide (both at 1 μM final concentration) with H_2O_2 (1-8 molar equivalents, depending on the enzyme) to ensure the formation of CII after 6 s. After that, different concentrations of lignosulfonates or Mn^{2+} in 0.1 M tartrate (at the adequate pH) were added, following the reaction at 410 nm, the maximum of the resting state for the peroxidases analyzed.

All the traces showed single-exponential character from which pseudo first-order rate constants (k_{2obs} and k_{3obs} for CI and CII reduction, respectively) were calculated. Plots of k_{2obs} and k_{3obs} *vs* substrate concentration fitted to linear, hyperbolic or sigmoid models. From those kinetics that fitted to a linear model, apparent second-order rate constants (k_{2app} and k_{3app} for CI and CII reduction, respectively) were obtained. Plots of k_{obs} *vs* substrate concentration that fitted to a Michaelis-Menten or sigmoid models yielded dissociation constants of the CI-lignin and CII-lignin or CII-Mn^{2+} complexes (K_{D2} and K_{D3} respectively) and first-order rate constants (k_2 and k_3, respectively). The corresponding apparent second-order rate constants, k_{2app} (k_2/K_{D2}) and k3app (k_3/K_{D3}), were calculated with the equation: $k_{obs} = (k/K_D)[S]/(1 + [S]/K_D)$, where [S] indicates substrate concentration. Despite replicated experiments (> 5 replica) measurements of CI/CII reduction resulted in high error values due to the intrinsic instability of CI, together with the polydisperse nature of the lignosulfonate preparations (being higher in the softwood lignosulfonate). For comparison, the lignosulfonate concentrations in these and other experiments were referenced to the basic phenylpropanoid unit corresponding to 260 and 290 Da in the sulfonated softwood and hardwood lignins, respectively (Sáez-Jiménez et al. 2015a).

Lignin treatment under steady-state conditions

To analyze how ancestral peroxidases and extant PC-LiPA modify lignin, hardwood and softwood lignosulfonates (12 g L^{-1}) were treated at 25 ºC over 24 h in 50 mM phosphate (pH 5). The concentration of enzymes were 1 μM, added in two doses, with a H_2O_2 final concentration of 10 mM, added continuously with a syringe pump. Control treatments were performed under the same conditions but without enzyme. It is important to note that although peroxidases exhibit their highest activity at pH 3 (used in stopped-flow measurements), reactions over 24 h were carried out at pH 5 to maintain the enzymes active.

NMR analyses of treated lignins

The treated lignosulfonates and the corresponding controls were freeze-dried for NMR analyses. Spectra were recorded at 25 ºC on a Bruker Avance-III 500 MHz instrument equipped with a cryogenically cooled 5 mm TCI gradient probe with inverse geometry. The lignosulfonate samples (40 mg initial weight, before treatments) were dissolved in 0.75 mL of deuterated dimethylsulfoxide (DMSO-d_6). The central peak of residual non-deuterated DMSO was used as the internal reference (at δ_H/δ_C 2.49/39.5 ppm), and the spectra were normalized to the same intensity of the DMSO signals, since the same DMSO volume and initial amount of lignin was used in all the cases.

The HSQC experiment used Bruker´s "hsqcetgpsisp.2" adiabatic pulse program with spectral widths from 0 to 10 ppm (5000 Hz) and from 0 to 165 ppm (20625 Hz) for the ^{1}H and ^{13}C dimension. The $^1J_{CH}$ used was 145 Hz. Processing used typical matched Gaussian apodization in the ^{1}H dimension and squared cosine-bell apodization in the ^{13}C dimension. Prior to Fourier transformation, the data matrices were zero-filled to 1024 points in the ^{13}C dimension. Signals were assigned by comparison with the literature (Lutnaes et al. 2008; Martínez et al. 2008; Lebo et al. 2008; Ralph et al. 2009; Prasetyo et al. 2010; Magina et al. 2015; Sáez-Jiménez et al. 2015b).

In the aromatic region of the spectrum, the H_2-C_2, H_5-C_5 and H_6-C_6 correlation signals were integrated to estimate the amount of lignins and the S/G ratio. In the aliphatic oxygenated region, the signals of methoxyls, and Hβ-Cβ (or Hα-Cα) correlations in the side chains of sulfonated and non-sulfonated β-*O*-4', phenylcoumaran and resinol substructures were integrated. The intensity corrections introduced by the adiabatic pulse program permits to refer the latter integrals to the previously obtained number of lignin units.

Chapter 1

Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date

General abstract

The main goal of the present thesis is the biochemical characterization of ligninolytic peroxidase ancestors, a process commonly known as enzyme resurrection, as an important aspect in wood-rotting fungi evolution. For that reason, we analyzed the phylogeny of class-II peroxidases from Polyporales, where most wood-rotting fungi are included, and with PAML we obtained the ancestral sequences of all the phylogeny.

We targeted those nodes of the phylogeny leading to extant lignin peroxidases, including the model enzyme lignin peroxidase H8 from *Phanerochaete chrysosporium* (corresponding to gene *lipaA* in the sequenced genome). We targeted this lineage as start of this thesis for several reasons. Extant LiPs oxidize lignin only at the protein surface, so the investigation of their ancestors would elucidate this specialization in lignin degradation. Also, they are among the most-efficient enzymes for lignin degradation, therefore the study of their evolution is interesting *per se* and for biotechnological purposes. Ancestral enzymes have unique properties *a priori* (i.e. higher thermostability) so we may obtain new enzymes for the biorefinery industry. After manually inspection and correction of the reconstructed peroxidase sequences, we resurrected the ancestors in the laboratory using *Escherichia coli* as host.

We characterized four ancestors and one extant enzyme that serve as milestones in the evolution of LiPs. We analyzed their modeled molecular structures, and obtained their kinetic parameters using model compounds to evaluate the change in their activities through time. Additionally, we studied their temperature and pH stability not only for an evolutionary survey, but also due to their possible application in the industry. The results discussed in this chapter represent the first resurrection of ligninolytic peroxidases, and served as the starting point for further questions in evolution that will be addressed in the following chapters.

Article published as:

Ayuso-Fernánez, I., Martínez, AT, Ruiz-Dueñas, FJ. *Experimental Recreation of the Evolution of Lignin Degrading Enzymes from the Jurassic to Date*. Biotechnology for Biofuels, 10:67. **2017**. IF: 5.497.

Author contributions: I.A.-F. performed the experimental work. F.J.R.-D. and A.T.M. designed the research. All authors analyzed the data, wrote the paper, and approved the final manuscript.

Abstract

Background: Floudas et al. (*Science* 336: 1715) established that lignin-degrading fungi appeared at the end of Carboniferous period associated with the production of the first ligninolytic peroxidases. Here, the subsequent evolution of these enzymes in Polyporales, where most wood-rotting fungi are included, is experimentally recreated using genomic information.

Results: With this purpose, we analyzed the evolutionary pathway leading to the most efficient lignin-degrading peroxidases characterizing Polyporales species. After sequence reconstruction from 113 genes of ten sequenced genomes, the main enzyme intermediates were resurrected and characterized. Biochemical changes were analyzed together with predicted sequences and structures, to understand how these enzymes acquired the ability to degrade lignin and how this ability changed with time. The most probable first peroxidase in Polyporales would be a manganese peroxidase (Mn^{3+} oxidizing phenolic lignin) that did not change substantially until the appearance of an exposed tryptophan (oxidizing nonphenolic lignin) originating an ancestral versatile peroxidase. Later, a quick evolution, with loss of the Mn^{2+}-binding site, generated the first lignin peroxidase that evolved to the extant form by improving the catalytic efficiency. Increased stability at acidic pH, which strongly increases the oxidizing power of these enzymes, was observed paralleling the appearance of the exposed catalytic tryptophan.

Conclusions: We show how the change in peroxidase catalytic activities meant an evolutionary exploration for more efficient ways of lignin degradation by fungi, a key step for carbon recycling in land ecosystems. The study provides ancestral enzymes with a potential biotechnological interest for the sustainable production of fuels and chemicals in a biomass-based economy.

Keywords: Fungal genomes, ancestral sequence reconstruction, fungal evolution, lignin biodegradation resurrected enzymes, ligninolytic peroxidases, catalytic properties, pH stability

Background

The large diversity of living organisms that we observe today is associated to the evolution of proteins, making the analysis of how proteins change with time a central issue in molecular evolution. However, the study of ancestral proteins has an important difficulty: they are extinct. Therefore, the use of the tools provided by bioinformatics is mandatory (Liberles 2007). To this point, ancestral sequence reconstruction can give us hints about ancient protein functions, and resurrection of the extinct proteins in the laboratory, using *ad hoc* expression hosts, will allow to directly analyze their properties and confirm evolutionary hypotheses. During the past years,

several examples have shown the power of this technique, from resurrection of the most ancient proteins in the Precambrian (Gaucher et al. 2003; Pérez-Jiménez et al. 2011; Risso et al. 2013) to the explanation of how enzymes evolved and acquired the specific mechanisms and functions that they have today (Thomson et al. 2005; Huang et al. 2012; Wilson et al. 2015; Barriuso and Martínez 2017).

Resurrected proteins are of interest not only because of the essential information about evolution that they can provide, but also due to the biotechnological potential that they have. Ancestral proteins often have higher stability (Risso et al. 2014) and, likely, new activities (Bornscheuer et al. 2012), which can make of them interesting biocatalysts. Also, if we identify the elements that confer this stability or new activity we can improve extant enzymes by rational design. Or even more, the famous James Gould (1990) sentence "replaying the tape of life" will no longer be a metaphor if we evolve the ancient proteins in the laboratory (Alcalde 2015; Alcalde 2016).

A central problem of white biotechnology for establishing a sustainable bioeconomy in the 21st century is processing recalcitrant lignin in vascular plant feedstocks (2014). Biological decay of lignin is essential for carbon recycling in nature, and lignin removal often represents also a key step for the sustainable production of fuels and chemicals in lignocellulose biorefineries (Martínez et al. 2009). Oxidation of the predominantly nonphenolic lignin polymer is a unique ability of extracellular enzymes produced by white-rot basidiomycetes (Ruiz-Dueñas and Martínez 2009). These fungi secrete three families of ligninolytic peroxidases: **i)** Lignin peroxidases (LiPs), which are able to oxidize nonphenolic lignin model compounds (Hammel and Cullen 2008); **ii)** Manganese peroxidases (MnPs), which oxidize Mn^{2+} to Mn^{3+} whose chelates act as diffusing oxidizers of phenolic lignin (Gold et al. 2000); and **iii)** Versatile peroxidases (VPs), which combine the catalytic activities (and oxidation sites) of LiPs, MnPs, and plant peroxidases acting on phenols and some dyes (Ruiz-Dueñas et al. 1999; Camarero et al. 1999b). Often white-rot basidiomycetes also produce generic peroxidases (GPs) with catalytic properties similar to the plant peroxidases (Baunsgaard et al. 1993). These four peroxidase types, which constitute the class II of the peroxidase-catalase superfamily (Zámocký et al. 2015), are well characterized today and they differ in the substrate oxidation sites they have (Ruiz-Dueñas et al. 2009).

In the last years, the evolution of lignin degrading organisms and enzymes has been investigated using genomic information. Molecular clock analyses and reconstruction of ancestral states of catalytic sites showed that the origin of lignin biodegradation occurred in the late Carboniferous with the appearance of the first ligninolytic peroxidase in the common ancestor of

Agaricomycetes (Polyporales included) (Floudas et al. 2012; Nagy et al. 2016). This event provided to the first white-rot fungi the ability to attack the lignocellulosic biomass of vascular plants, enabling carbon recycling in land ecosystems. In the present work we target, by complete sequence reconstruction, the evolution that led to the highly-specialized LiP of *Phanerochaete chrysosporium* (isoenzyme H8) (Kersten and Cullen 2007; Hammel and Cullen 2008), the first sequenced basidiomycete as a model ligninolytic organism (Martinez et al. 2004). With this aim, we resurrected the most relevant enzymes of the evolutionary line from the common ancestor of class II peroxidases in the order Polyporales, where most wood-rotting fungi are included (Binder et al. 2013), and studied their catalytic and stability properties with the aim of recreating in the laboratory the natural evolution of lignin degrading enzymes, which contributed to land colonization by vascular plants (Floudas et al. 2012).

Results

Reconstruction of ancestral sequences from Polyporales genomes

From the information available in ten genomes of Polyporales (phylum Basidiomycota) sequenced at the Joint Genome Institute (JGI), a maximum likelihood (ML) phylogenetic tree of ligninolytic peroxidase and GP sequences (**Fig. 1**) was constructed with RAxML (Stamatakis et al. 2008). The tree, which is consistent with previous results revealing a robust evolutionary history (Floudas et al. 2012; Ruiz-Dueñas et al. 2013), was used to predict ancestral sequences using PAML 4.7 (Yang 2007) and the WAG evolution model (yielding reconstructed sequences with the highest average probabilities), followed by manual curation. In the path from the first peroxidase in Polyporales to LiPH8 (**Fig. 1**, red line), we focused on four proteins (nodes), whose most probable reconstructed sequences are shown in **Fig. 2**, because they are milestones in LiP appearance: CaPo is the Common ancestor of Polyporales peroxidases, CaCD represents the Common ancestor of Cluster D (the largest peroxidase cluster including LiPs), AVP would be the most Ancestral VP in this evolutionary line, and ALiP would be the most Ancestral LiP in Polyporales (and probably in all basidiomycetes).

The posterior probabilities for three of the most probable ancestral sequences (CaCD, AVP and ALiP) are high (mean ≥0.95) while those of the first ancestor (CaPo) were lower (mean 0.82) (Additional file 1: **Fig. S2**). To check the possible functional variability of the four reconstructed nodes, a total of 5000 sequences (including the most probable ones) were obtained for each of them by Monte Carlo sampling (Risso et al. 2013) from the PAML results, and manually inspected for the presence/absence of the two substrate-oxidation sites described below (Glu37/Glu41/Asp183 binding

Mn^{2+}, and Trp172 oxidizing nonphenolic lignin). For the three more recent nodes 100% of the sequences showed the following invariable oxidation site/s: Trp172 in ALiP, both Trp172 and Glu37/Glu41/Asp183 in AVP, and Glu37/Glu41/Asp183 in CaCD sequences. Therefore, only the most probable sequences were resurrected for each of them.

For node CaPo, all sequences lack Trp72 and have at least two of the above acidic residues, a situation associated to Mn^{2+} oxidation ability by normal or atypical MnPs (lacking one of these residues) (Floudas et al. 2012; Ruiz-Dueñas et al. 2013). To verify the activity of the predicted atypical MnPs, the corresponding sequences were classified into two subsets (with either Asp37 or Arg183) that were submitted to a random sampling. In this way, the CaPo-bis and CaPo-tris alternative sequences ("near-ancestors") were selected, which were resurrected together with the most probable CaPo ancestor with a typical Mn^{2+}-oxidation site. Although the resulting sequences were not resurrected, parallel random samplings were also performed on the CaCD, AVP and ALiP sets for *in silico* analysis, showing invariable catalytic sites (their sequences and posterior probability values are shown in Additional file 1: **Figs. S1** and **S2**, respectively, together with those of the CaPo node).

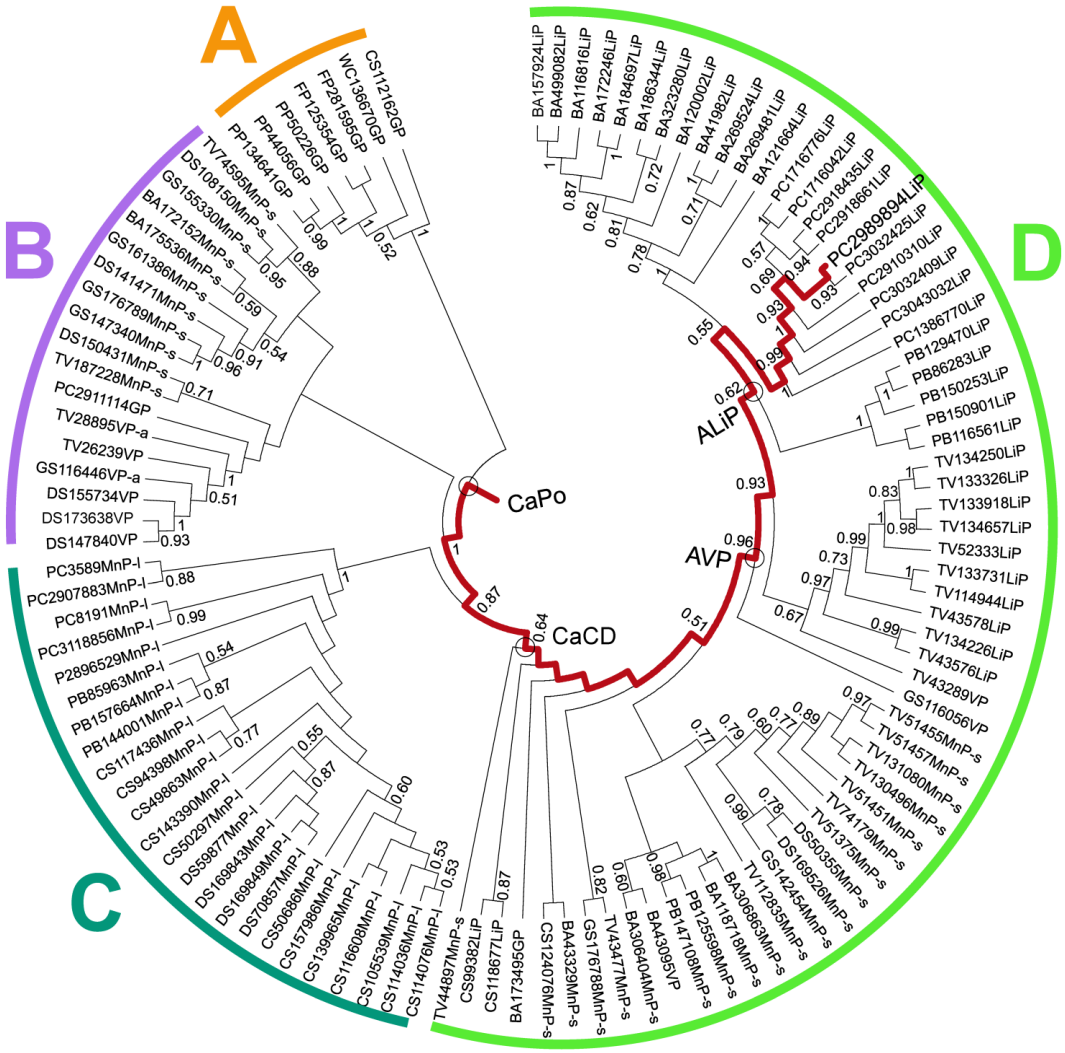


Figure 1. Phylogenetic tree for 113 peroxidases from ten Polyporales genomes (sequences in Additional **file 2**). Clusters A (GPs), B (short MnPs and VPs, plus one GP) C (long MnPs), and D (LiPs, short MnPs and VPs, plus one GP) are shown. The path from the common ancestor to the extant LiPH8 of *P. chrysosporium* (JGI ID# 2989894) is in red. Also, the milestones in this evolutionary line (CaPo, CaCD, AVP and ALiP) are marked (circles). The sequence labels start with the species code (BA, *Bjerkandera adusta*; CS, *Ceriporiopsis subvermispora*; DS, *Dichomitus squalens*; FP, *Fomitopsis pinicola*; GS, *Ganoderma* sp; PB, *Phlebia brevispora*; PC, *P. chrysosporium*; PP, *Postia placenta*; TV, *Trametes versicolor*; and WC, *Wolfiporia cocos*) followed by the JGI ID# and the peroxidase type, including GP, LiP, short MnP (MnP-s), long MnP (MnP-l), VP, and atypical VP (VP-a). Bootstrap values >0.5 are indicated on the different nodes.

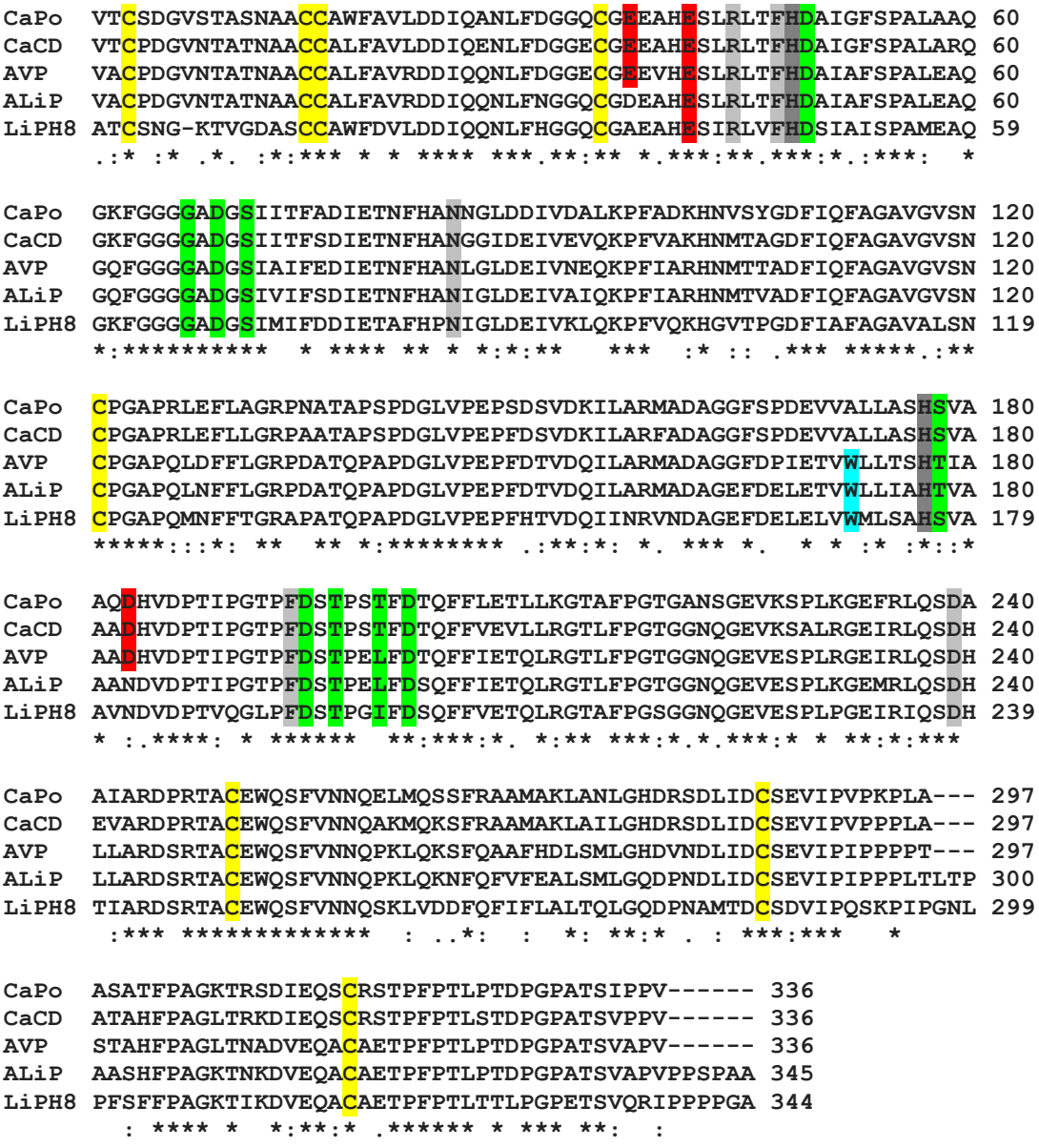


Figure 2. Alignment of the four most probable ancestral sequences (mature proteins) predicted with PAML 4.7 (using the WAG evolution model) and extant LiPH8 (alternative ancestral sequences are in **Fig. S1**). Conserved catalytic and other relevant residues (Ruiz-Dueñas et al. 2009) are indicated including: two active site histidines (*dark gray*); three acidic residues forming the Mn^{2+}-binding site (*red*); other active site conserved residues (*light gray*); one tryptophan involved in lignin direct oxidation (*cyan*); nine ligands of two Ca^{2+} ions (*green*); and eight cysteines forming disulfide bonds (*yellow*). Symbols below indicate full conservation of the same (*asterisk*) or equivalent residues (*colon*) and partial residue conservation (*dot*). The identity between the sequences decreased in the order: 87% (301) for AVP/ALiP, 86% (289) for CaPo/CaCD, 74% (257) for CaCD/ALiP, 74% (248) for CaPo/AVP, 72% (248) for ALiP/LiPH8, 70% (242) for CaPo/ALiP, 66% (226) for AVP/LiPH8, 61% (211) for CaCD/LiPH8 and 61% (209) for CaPo/LiPH8 (with the number of aligned residue pairs in parenthesis).

In silico analysis of ancestral heme pocket and substrate-oxidation sites

Molecular models of the selected ancestral proteins, together with multiple alignments, reveal that their 12 helices, two sites binding structural Ca^{2+} ions, and four disulfide bonds did not change during evolution, despite the differences in sequence (only 61-87% identity) between the ancestors and between them and the extant LiPH8 (see **Fig. 2** legend). This comparison also reveals that most of the essential amino acids for LiPH8 function were already present at the first stages of evolution, including proximal His177 (near the heme iron), Asp239 and Phe194 at one side of heme, and distal His48, Arg44 (both contributing to H_2O_2 reaction in extant peroxidases), Asn85 and Phe47 at the opposite side. Hypothetical His177-Asp239 and Ser/Thr178-Asp202 H-bonds would control the position of the proximal histidine with respect to the heme iron (**Fig. S3**). The proximal Ca^{2+} hypothetical ligands - Ser178 (CaPo and CaCD) or Thr178 (AVP and ALiP), Asp195, Thr197, Thr200 (CaPo, CaCD) or Leu200 (AVP and ALiP) and Asp202 - show some differences with those in LiPH8 (Ser177, Asp194, Thr196, Ile199 and Asp201), while the distal Ca^{2+} ligands would be exactly the same. However, the above differences would not affect proximal Ca^{2+} binding since backbone carbonyls are involved at those positions where differences are predicted.

Then, several substrate oxidation sites were identified in the molecular models and sequences of the ancestral peroxidases. The three acidic residues that define the Mn^{2+}-binding site (red background and asterisks in **Figs. 2** and **3**, respectively) of extant MnP (Kishi et al. 1996) and VP (Ruiz-Dueñas et al. 2007) already appear in CaPo (Glu37, Glu41, Asp183), and continue invariable in evolution until the appearance of ALiP, when Glu37 becomes Asp37 and Asp183 changes to Asn183 (gray arrow in **Fig. 3**). Extant LiPs and VPs oxidize nonphenolic (high redox-potential) lignin model compounds by long-range electron transfer from an exposed tryptophan, being Trp171 in LiPH8 and Trp164 in *Pleurotus eryngii* VP (Doyle et al. 1998; Mester et al. 2001; Pérez-Boada et al. 2005), whose implication in direct oxidation of nonphenolic lignin has been recently demonstrated (Sáez-Jiménez et al. 2016). Analysis of the corresponding residues revealed that CaPo and CaCD have an alanine (Ala172) in that position. The fact that these first two (most probable) ancestors have a Mn^{2+}-binding site, but no catalytic tryptophan, makes of them two putative MnPs (the catalytic activity of the CaPo alternative sequences was established after heterologous expression, described below). Then, we were able to track the point in the evolutionary line where the catalytic tryptophan (blue asterisk in **Fig. 3**) appears: in AVP Ala172 becomes Trp172 (**Fig. 3** blue arrow). Since AVP has also a well structured Mn^{2+}-

binding site, this protein is *a priori* a VP. Finally, in ALiP the Trp172 is conserved but the Mn^{2+}-binding site disappears becoming the first LiP in the evolution of Polyporales. To confirm the above predictions, the most relevant sequences were resurrected.

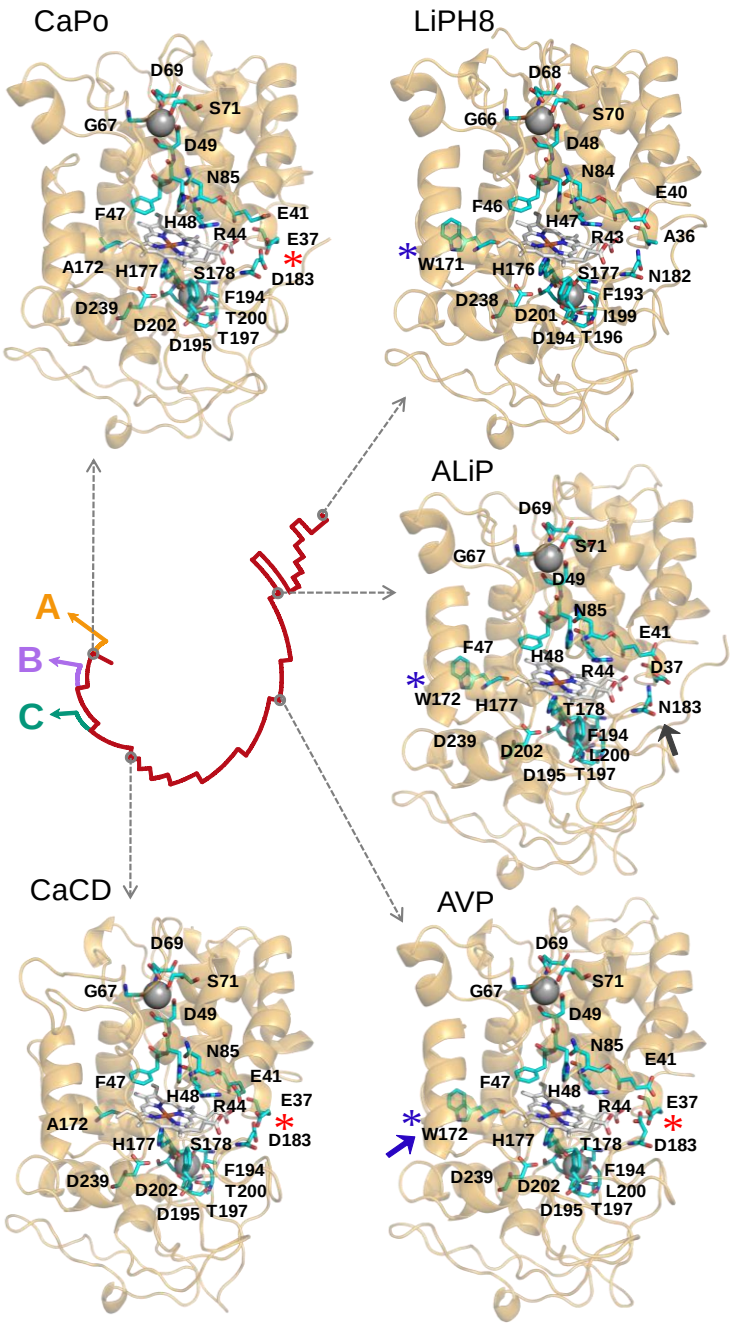


Figure 3. Molecular models of the reconstructed ancestors and extant LiP. The hypothetical location of heme, two Ca^{2+} ions, ancestral Mn^{2+}-binding site formed by two glutamate and one aspartate residues (red asterisks), evolved lignin-oxidizing exposed tryptophan (blue asterisks), and other relevant residues (**Fig. 2**) are shown on the ancestor models, built using three related crystal structures (4BM1, 3FJW and 1QPA) as templates (Biasini et al. 2014), and LiPH8 crystal structure (PDB 1B82). The models are shown on a schematic representation of evolution (see **Fig. 1**). The blue and gray arrows show the amino acids that appeared in the main evolutionary events and defined new activities (gain of catalytic Trp172 and loss of Mn^{2+}-binding site, respectively).

Kinetic properties of resurrected peroxidases

The most probable ancestral and LiPH8 DNA sequences (plus two additional sequences from node CaPo) were synthesized, expressed in *Escherichia coli*, *in vitro* activated, and purified to homogeneity. Then, their steady-state kinetic constants for oxidation of veratryl alcohol (VA)

representing nonphenolic lignin, 2,6-dimethoxyphenol (DMP) representing the minor phenolic moiety in lignin, and Mn^{2+} were estimated, together with those for two dyes with high (Reactive Black 5, RB5) and low (2,2'-azinobis[3-ethylbenzothiazoline-6-sulfonate], ABTS) redox-potentials (**Table 1**).

The above substrates are oxidized at different sites, and the kinetic analysis shows how these sites have evolved giving rise to different peroxidase families. The sites for oxidation of Mn^{2+} and high redox-potential substrates such as VA (and RB5) have been already described above. Phenols, such as DMP (and the generic oxidoreductase substrate ABTS) can be oxidized: **i)** with high efficiency at the same tryptophan oxidizing the high redox-potential substrates; and **ii)** with low efficiency at one of the heme access channels (Morales et al. 2012; Fernández-Fueyo et al. 2014a).

The catalytic efficiency (k_{cat}/K_m) on the different substrates clearly showed the changes produced along the evolution, as illustrated in **Fig. 4**. In this way, the ability to oxidize Mn^{2+} gradually improved from CaPo to AVP (due to 11-fold K_m reduction) and then was completely lost in ALiP. In contrast, CaPo shows the highest activity oxidizing low redox-potential DMP (and ABTS) at the low efficiency site (heme channel), which is reduced from CaPo to AVP (7-15 fold lower catalytic efficiency, due to k_{cat} decrease) and then disappears. The two alternative CaPo ancestors showed similar activity on ABTS and lower or null on DMP and Mn^{2+} (**Table 1** footnote), and were classified as GP (CaPo-bis) and atypical MnP (CaPo-tris). The vanishing of the low efficiency oxidation site comes along with the rise of a high efficiency site in AVP. In this ancestral peroxidase, DMP (and ABTS) begin to be oxidized at a second site in the protein (the catalytic tryptophan discussed below), as revealed by sigmoid kinetic curves.

As regard for the high redox-potential substrates, the turning point where the LiP/VP usual substrate VA (and the recalcitrant dye RB5) begins to be oxidized in this evolutionary line is AVP, with the appearance of the exposed Trp172 at a position that coincides with the catalytic tryptophan in extant VP and LiP (Doyle et al. 1998; Mester et al. 2001; Pérez-Boada et al. 2005). The catalytic efficiency towards VA is maintained from AVP to ALiP and then sharply increases (7-fold) from ALiP to LiPH8. The k_{cat} value did not change, and the higher catalytic efficiency is due to an improvement in VA affinity (10-fold lower K_m value). The appearance of Trp172 also resulted in efficient oxidation of the low redox-potential substrates, fully substituting the low efficiency site in ALiP and extant LiP. The activity of the catalytic tryptophan in the advanced stages of peroxidase evolution will be affected by the electrostatic charge of its surface environment (**Fig. S4**): VA cation radical will be stabilized by the strongly acidic environment in LiPH8 that, in contrast, will prevent oxidation of anionic RB5.

Table 1. Kinetic parameters for oxidation of Mn^{2+}, phenolic (DMP) and nonphenolic (VA) aromatics, and dyes (means and 95% confidence limits) by the resurrected enzymes and extant LiPH8 (a comparison of all catalytic efficiencies is shown in **Fig. 4**).[a]

		CaPo	CaCD	AVP	ALiP	LiPH8
Metal ion:						
	K_m (μM)	700 ± 48	275 ± 41	62 ± 10	-[b]	-
Mn^{2+}	k_{cat} (s^{-1})	185 ± 3	170 ± 6	106 ± 4	-	-
	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	260 ± 15	617 ± 80	1710 ± 240	-	-
Aromatics:						
DMP[c]	K_m (μM)	32900 ± 2700	66800 ± 3800	32500 ± 12100	-	-
(low	k_{cat} (s^{-1})	221 ± 9	109 ± 4	31 ± 5	-	-
efficiency)	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	6.7 ± 0.3	1.6 ± 0.03	1.0 ± 0.2	-	-
DMP	K_m (μM)	-	-	5.3 ± 1.1	34.0 ± 5.4	4.0 ± 0.07
(high	k_{cat} (s^{-1})	-	-	4.5 ± 0.1	18.3 ± 0.7	6.9 ± 0.5
efficiency)	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	-	-	837 ± 162	537 ± 55	600 ± 36
	K_m (μM)	-	-	299 ± 104	773 ± 155	79.3 ± 18
VA	k_{cat} (s^{-1})	-	-	7.1 ± 0.5	21.3 ± 1.0	16.2 ± 0.8
	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	-	-	24 ± 7	28 ± 5	205 ± 4
Dyes:						
ABTS[c]	K_m (μM),	3170 ± 270	1280 ± 350	2150 ± 420	-	-
(low	k_{cat} (s^{-1})	539 ± 24	103 ± 10	25 ± 2	-	-
efficiency)	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	170 ± 8	80 ± 15	12 ± 1	-	-
ABTS	K_m (μM),	-	-	5.4 ± 0.7	13.7 ± 3.5	21.3 ± 2.3
(high	k_{cat} (s^{-1})	-	-	2.1 ± 0.1	12.5 ± 0.7	6.5 ± 0.2
efficiency)	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	-	-	400 ± 44	911 ± 212	300 ± 25
	K_m (μM),	-	-	4.8 ± 0.8	12.6 ± 3.6	-
RB5	k_{cat} (s^{-1})	-	-	2.4 ± 0.2	5.4 ± 0.8	-
	k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	-	-	504 ± 50	428 ± 68	-
***Direction of evolution*:**		→				

[a]Reactions at 25 °C in 0.1 M tartrate at optimal pH 5.0 (CaPo, and its alternative ancestors, and CaCD) or 5.5 (AVP) for Mn^{2+}, pH 3.5 (AVP) or 3.0 (ALiP and LiPH8) for VA, pH 2 (CaPo) or 3.0 (others) for DMP, pH 3.0 (CaPo) and pH 3.5 (others) for ABTS, and pH 3 for RB5; and saturating H_2O_2 concentrations of 0.4 mM for CaPo, CaCD and AVP, 0.2 mM for ALiP, and 0.1 mM for LiPH8. [b]-, absence of activity. [c]Biphasic kinetics for DMP and ABTS oxidation by AVP enabled calculation of two sets of constants assigned to two catalytic sites, as reported for extant VP (Morales et al. 2012), comparable to those present in CaPo/CaCD and ALiP/LiPH8, respectively.

The alternative CaPo ancestors oxidized DMP (low efficiency) with K_m 455000 ± 84000 (CaPo-bis) and 298000 ± 66000 μM (CaPo-tris), k_{cat} 129 ± 22 (CaPo-bis) and 108 ± 21 s^{-1}(CaPo-tris) and k_{cat}/K_m 0.3±0.05 (CaPo-bis) and 0.4 ± 0.01 $s^{-1}\cdot mM^{-1}$ (CaPo-tris); ABTS (low efficiency) with K_m, 1392 ± 345 (CaPo-bis) and 807 ± 104 μM (CaPo-tris), k_{cat}, 241 ± 23 (CaPo-bis) and 128 ± 6 s^{-1} (CaPo-tris), and k_{cat}/K_m, 173 ± 28 (CaPo-bis) and 160 ± 15 $s^{-1}\cdot mM^{-1}$ (CaPo-tris), while Mn^{2+} was only oxidized by CaPo-tris with K_m 5010 ± 274 μM, k_{cat} 47 ± 1 s^{-1}, and k_{cat}/K_m 9.3 ± 0.4 $s^{-1}\cdot mM^{-1}$.

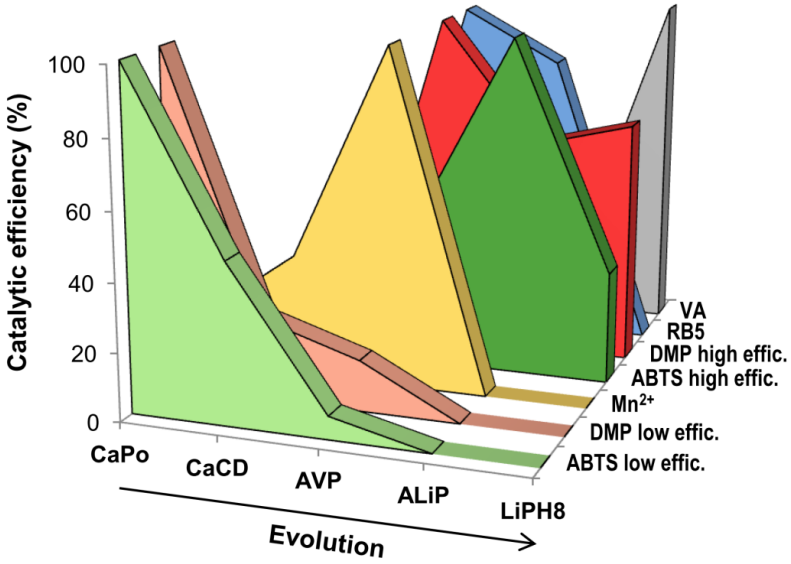


Figure 4. Catalytic efficiency on different substrates during peroxidase evolution. Changes of the relative catalytic efficiencies (k_{cat}/K_m) on ABTS, DMP, Mn^{2+}, RB5 and VA (the two former at high and low efficiency oxidation sites) during evolution (Ruiz-Dueñas et al. 2013), with the maximum for each substrate taken as 100%. The CaPo-bis and CaPo-tris alternative ancestors showed the following relative catalytic efficiencies: **i)** ABTS (low efficiency), 100 and 94%, respectively; **ii)** DMP (low efficiency), 4.5 and 6.0%, respectively; **iii)** Mn^{2+}, 0.5% for CaPo-tris (no activity detected with CaPo-bis); and **iv)** no activity on ABTS (high efficiency), DMP (high efficiency), RB5 and VA.

Stability of ancestral peroxidases

The pH stability of the resurrected peroxidases was analyzed by measuring the residual activity (after 4 h at 25 ºC) in the pH 2-10 range (**Fig. 5A**). All the enzymes retained ≥50% activity at pH 5-7, which strongly decreased at pH 8-10. The most important difference was at pH 3, where AVP, ALiP and LiPH8 retained ≥70% activity, while CaPo (alternative ancestors included) and CaCD were strongly inactivated. As shown in **Fig. 5B**, the increase of acidic pH stability paralleled the introduction of VA oxidation activity (due to the appearance of catalytic Trp172) and was maintained till today.

Thermal stability was analyzed by both residual activity measurements and thermal melting profiles from circular dichroism (**Fig. 6A** and **B**, respectively). From most probable CaPo (and alternative ancestors) to LiPH8 there is a decrease in thermal stability, but AVP leaves that trend. Moreover, Mn^{2+} caused a slight increase of the thermal stability, as revealed by T_{50}/T_m values (**Fig. 6C** and **D**) as reported for other MnPs (Fernández-Fueyo et al. 2014a).

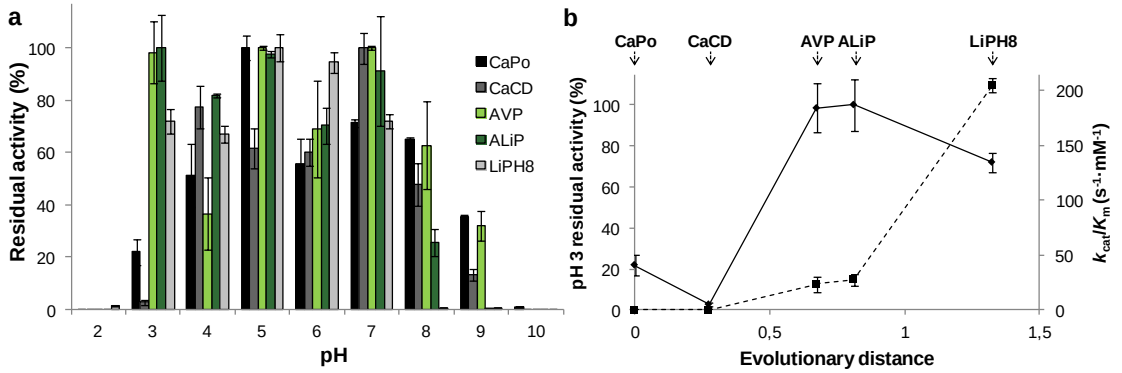


Figure 5. pH stability of the resurrected ancestors and extant LiPH8. **a)** Residual activities were measured with 2.5 mM ABTS at optimal pH (see **Table 1**) after 4 h incubation at different pH values (25 ºC) and referred to activity after 1 min incubation of each of them at pH 5. The CaPo-bis and CaPo-tris alternative ancestors showed the following relative residual activities: **i)** pH 2, 0%; **ii)** pH 3, 10 and 44%, respectively; **iii)** pH 4, 88 and 89%, respectively; **iv)** pH 5, 85 and 90%, respectively; **v)** pH 6, 93 and 97%, respectively; **vi)** pH 7, 93 and 95%, respectively; **vii)** pH 8, 84 and 95%, respectively; **viii)** pH 9, 88 and 28%, respectively; and **ix)** pH 10, 0%. **b**) Comparison of the VA catalytic efficiency (s^{-1}·mM^{-1}, dashed line) and enzyme stability at pH 3 (continuous line) vs evolutionary distance (Ruiz-Dueñas et al. 2013). The pH 3 residual activities were measured as indicated in **Fig 5A**. Means and 95% confidence limits.

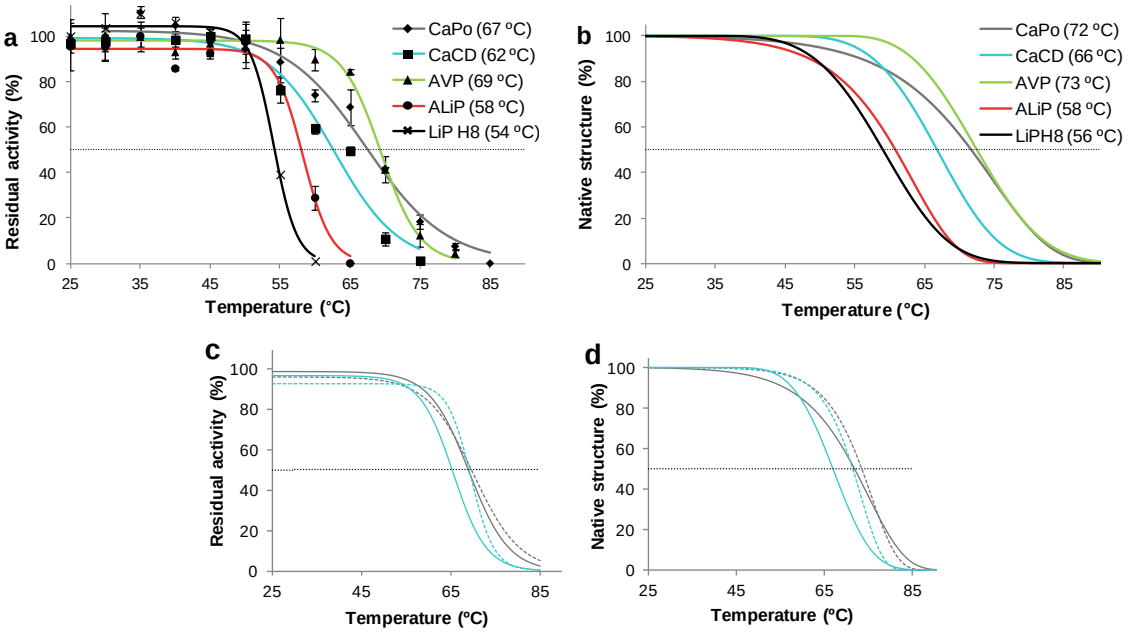


Figure 6. Changes in thermal stability. **a,b)** Thermal stability of the resurrected proteins and LiPH8 estimated from enzyme activity (**a**) and secondary structure loss (**b**). The T_{50} and T_m values are provided in the legends. Alternative CaPo-bis and CaPo-tris showed 67 and 63 ºC T_{50} and 69 and 66 ºC T_m values, respectively. **c,d)** Mn^{2+} (1 mM) addition during incubation (dashed lines) slightly increased the T_{50} (**c**) and T_m (**d**) values of the CaPo (gray) and CaCD (blue) ancestors. Inactivation was measured after 10 min at pH 5.5, using 2.5 mM ABTS (**a**) or 6 mM Mn^{2+} (**c**) as substrate. Secondary structure loss (**b,d**) was estimated by circular dichroism at 222 nm.

Discussion

Over the last years there has been a profound mining of basidiomycete genomes to find new enzymes and obtain evolutionary information (Floudas et al. 2012; Ruiz-Dueñas et al. 2013; 2014b; Floudas et al. 2015). Our analysis of the lineage from the most ancestral peroxidase in Polyporales, a basidiomycete order that appeared near 150 mya at the end of the Jurassic (Floudas et al. 2012; Nagy et al. 2016), to *P. chrysosporium* LiP (Kersten and Cullen 2007; Hammel and Cullen 2008) reveals how these enzymes acquired the ability to degrade lignin and how this ability changed with time. At the same time, the reconstruction of ancestral peroxidases provides proteins of biotechnological interest in lignocellulose biorefineries (Martínez et al. 2009) (e.g. CaPo is a thermostable enzyme and a good candidate for further directed evolution, as it has the potential to become every peroxidase in the phylogenetic tree).

In the above analysis, we used PAML (Yang 2007), a common software for ancestral protein reconstruction, and the published ML phylogenetic tree of class-II peroxidases in Polyporales (Ruiz-Dueñas et al. 2013), whose accuracy is assured by a precise annotation of the peroxidase genes. The program uses the best evolutionary model for the sequences analyzed (WAG model in this case) to predict the most probable amino-acid at every position of each node in the phylogenetic tree (providing a complete matrix with the probabilities of every other amino acids). The reconstruction procedure has inherent limitations for sequence prediction in the most ancestral nodes (such as CaPo) that can be partially overcome by sequence sampling from the PAML matrix and *in silico* and/or experimental evaluation of ancestor variability (reconstruction robustness) with respect to the property/ies most relevant for each phylogenetic analysis (catalytic properties defining the different peroxidase types, in the present study).

Ligninolytic peroxidases are high redox-potential enzymes being this characteristic related to the distance between the heme iron and the proximal histidine (Nε) acting as its fifth ligand (Banci et al. 1991b; Banci 1997). This distance is shorter in peroxidases of prokaryotic origin, such as cytochrome *c* peroxidase (C*c*P), and significantly increased in fungal peroxidases. This is due to an H-bond between a backbone carbonyl (LiPH8 Asp201) and the hydroxyl of a Ser/Thr residue (LiPH8 Ser177, absent from C*c*P), that displaces the contiguous proximal histidine increasing the electron-deficiency of iron in ligninolytic and other eukaryotic peroxidases. The distance between the His177 Nε and the heme iron in the molecular models of the ancestral peroxidases are not significantly different from those found in LiP, due to conserved Ser/Thr178-Asp202 and His177-Asp239 H-bonds, although significantly larger than in C*c*P (Additional File 1: **Fig. S3**). Therefore, no strong changes in peroxidase redox-potential are expected to

be produced during LiP evolution in Polyporales, and the VA oxidation ability of AVP (and ALiP) would be mainly related to the appearance of the surface catalytic tryptophan.

It is important that the heme peroxidase redox-potential, and lignin-degrading ability of basidiomycete peroxidases, drastically increases when pH decreases (Millis et al. 1989). Therefore, the improved stability at acidic pH in the most recent stages of peroxidase evolution in Polyporales (from AVP to extant LiP) represents an important evolutionary adaptation to the acidic conditions where ligninolysis occurs in nature (Martínez 2002). Concerning thermal stability, AVP appears as the most thermostable ancestral peroxidase in Polyporales, especially when compared with its descendants. Experimental evidence is required to know if genetic drift (Wang et al. 2002; Bloom and Arnold 2009) caused the increased stability of AVP related with the appearance and subsequent evolution of the catalytic tryptophan. After that, the loss of the Mn^{2+}-binding site caused a decrease in stability, since the Mn-binding site can (in the presence of Mn^{2+} or other cations) contributes to anchor the heme cofactor (Gold et al. 2000; Ruiz-Dueñas et al. 2007), but proteins were still stable in order to be selected in evolution. In agreement with the present results, increased thermal stability has been reported for both an ancestral plant peroxidase (Loughran et al. 2014) and a LiP containing several ancestral mutations (Semba et al. 2015).

The kinetic constants of the resurrected peroxidases show the high evolvability of these enzymes incorporating new activities. Although some exceptions have been reported (Sayou et al. 2014; Aakre et al. 2015), it is generally assumed that ancestral enzymes have a wide substrate specificity and are specialized after duplication events (Huang et al. 2012; Voordeckers et al. 2012; Risso et al. 2013). However, peroxidases in Polyporales show a different evolutionary history. The first ancestor (CaPo) is able to oxidize only low redox-potential substrates, and Mn^{2+} with low catalytic efficiency. This low Mn^{2+}-oxidizing activity is in agreement with the GP-type ancestor of Dikarya peroxidases (Floudas et al. 2012). Similar substrate specificity has been reported for the extant short MnPs, where CaPo and CaCD would be included according to their short C-terminal tail (Fernández-Fueyo et al. 2014a). This suggests that short MnPs are old enzymes, whose efficiency oxidizing Mn^{2+} was improved later. Then, the efficiency oxidizing Mn^{2+} increased in AVP (the first VP in Polyporales) with activity on this cation similar to extant VP (Ruiz-Dueñas et al. 2007). What is surprising, as mentioned above, is that the AVP appearance, a relatively recent event in the mid-term evolution of Polyporales peroxidases, resulted in the enzyme with the widest specificity: AVP is able to oxidize every substrate assayed. Then, evolution focused in the direction of more efficient oxidation of

nonphenolic lignin at Trp172, as shown using VA, and resulted in the appearance of LiP, the most specialized ligninolytic enzyme only reported in the order Polyporales (Floudas et al. 2012; Ruiz-Dueñas et al. 2013; 2015). However, this "unusual" evolutionary behavior is most probably related to the existence of three different substrate-oxidation sites in these peroxidases, whose appearance/disappearance caused qualitative changes in catalytic activities (compared with enzymes with a progressively specialized unique catalytic site).

From an organismal point of view, the above evolutionary trend can be seen as a search for new and more efficient tools to degrade lignin. From their appearance in the late Jurassic, Polyporales had ligninolytic peroxidases, as shown by the ancestral enzyme reconstruction. The oldest ancestors would use Mn^{2+}, being Mn^{3+} able to oxidize the minor phenolic moiety of lignin (Martínez et al. 2009). Trp172 appeared later giving versatile AVP that oxidizes Mn^{2+} (more efficiently than its MnP-type ancestors), phenols and most probably also nonphenolic lignin, as shown by its ability to oxidize high redox potential substrates. When ALiP appeared, the Mn^{2+}-binding site was lost, as well as the low efficiency site oxidizing phenols. In this way, non-competitive inhibition of lignin oxidation was prevented. Although the ability to oxidize Mn^{2+} was lost in the evolutionary line leading to LiPH8, a family of highly-efficient Mn^{2+}-oxidizing peroxidases (cluster C) evolved from the first peroxidase ancestor CaPo. Finally, after the appearance of the exposed tryptophan in AVP, the peroxidase catalytic efficiency oxidizing VA, was improved, first by increasing the k_{cat} in ALiP and later by improving the K_m in LiPH8, which means a peroxidase specialization towards lignin degradation.

Conclusions

We predict that the evolutionary pathway leading to the most efficient lignin-degrading enzymes (LiP family, only found in Polyporales) included successive incorporation of: **i)** a surface tryptophan to an ancestral short MnP resulting in an ancestral VP; and **ii)** loss of the Mn^{2+}-binding site generating the first LiP. The experimental evaluation of catalytic properties of the resurrected ancestral enzymes was consistent with the bioinformatic analysis and prove the above hypothesis. Interestingly, an increase of stability at acidic pH was found simultaneously with the appearance of the catalytic tryptophan, enabling these enzymes to act under the acidic conditions characterizing lignin decay. The evolutionary and experimental studies also show that some ancestral peroxidases are of biotechnological interest because of their stability and potential evolvability.

Methods

Phylogenetic analysis

113 predicted peroxidase sequences (mature proteins) from the genomes of ten Polyporales (phylum Basidiomycotina) species (namely *B. adusta*, *C. subvermispora*, *D. squalens*, *F. pinicola*, *Ganoderma sp.*, *P. brevispora*, *P. chrysosporium*, *P. placenta*, *T. versicolor* and *W. cocos*) available at the DOE JGI Mycocosm portal (as http://genome.jgi.doe.gov/Bjead1_1/Bjead1_1.home.html, http://genome.jgi.doe.gov/Cersu1/Cersu1.home.html, http://genome.jgi.doe.gov/Dicsq1/Dicsq1.home.html, http://genome.jgi.doe.gov/Fompi3/Fompi3.home.html, http://genome.jgi.doe.gov/Gansp1/Gansp1.home.html, http://genome.jgi.doe.gov/Phlbr1/Phlbr1.home.html, http://genome.jgi.doe.gov/Phchr2/Phchr2.home.html, http://genome.jgi.doe.gov/Pospl1/Pospl1.home.html, http://genome.jgi.doe.gov/Trave1/Trave1.home.html and http://genome.jgi.doe.gov/Wolco1/Wolco1.home.html, respectively), all of them containing one or several peroxidases of the peroxidase-catalase superfamily (Ruiz-Dueñas et al. 2013), have been used in this study (Additional **file 2)**.

The amino-acid sequences were aligned with MUSCLE as implemented in MEGA 7 (Kumar et al. 2016a). ML analysis was then performed using RAxML (Stamatakis et al. 2008) under the GTR model with GAMMA-distributed rate of heterogeneity, using the WAG evolution model, as suggested by ProtTest (Abascal et al. 2005).

Ancestral sequence reconstruction

PAML 4.7 package (Yang 2007) was used to obtain the posterior amino-acid probability per site in each ancestor under the WAG model of evolution, and the most probable whole sequences for each of the nodes, using as inputs the ML phylogeny and the MUSCLE alignment previously obtained (PAML reconstructions using the LG and Dayhoff evolution models were also performed for comparison). Marginal reconstruction was selected for the present work.

Five thousand sequences (including the most probable ancestor predicted by PAML) were selected for each of the nodes by successive Monte Carlo sampling steps (using an *ad hoc* program kindly provided by Dr J.M. Sanchez-Ruiz and 0.2 and 0.5 probability thresholds, referred to the highest probability at every position) (Risso et al. 2013). These sequences were manually corrected for C-terminal and other insertions or deletions (the former often originating from intron to exon transitions) according to the sequences of the ancestor progeny, and inspected for the presence/absence of

the substrate-oxidation sites (Glu37, Glu40 and Asp183 involved in Mn^{2+} binding, and Trp172 responsible for VA oxidation).

The 5000 CaPo sequences were classified into three subsets corresponding to sequences containing: **i)** the three acidic residues forming a typical Mn^{2+}-binding site (including the most probable ancestor); **ii)** an atypical Mn^{2+}-binding site formed by only two of the above acidic residues, without a basic residue in the third position (including CaPo-tris selected later); and **iii)** a site formed by two of the above acidic residues plus a basic residue in the third position (including CaPo-bis selected later), with Trp172 being absent in all the cases. The Monte Carlo sampling provided a significant number of alternative sequences with an arginine (or lysine) residue in position 183 (forming the above subset-iii) whose presence prevented binding and oxidation of the Mn^{2+} cation (as confirmed after CaPo-bis resurrection). However, for the most probable sequence, PAML reconstructed an aspartic acid at this position (resulting in a functional Mn^{2+}-binding site after CaPo resurrection) in agreement with the extant peroxidases from the ten Polyporales genomes analyzed (with only 5% sequences containing a basic residue at this position).

Each of the CaPo subsets was submitted to a random sampling, together with the whole sets for the three other nodes (where no subsets were defined because the 5000 sequences from each of them showed the same catalytic sites), yielding 12 representative sequences, including the four most probable ones, that were analyzed *in silico*.

Then, the DNA sequences encoding the most probable amino-acid sequences for the four reconstructed nodes, together with the two alternative sequences from the CaPo subsets, were synthesized by ATG:biosynthetics (Merzhausen, Germany) after optimizing the codon usage for high expression in *E. coli* using OPTIMIZER (Puigbò et al. 2007).

Protein modeling

Molecular models of the predicted proteins were obtained at the Swiss-Model automated protein homology modeling server (Biasini et al. 2014) using related crystal structures, selected using the GMQE parameter, as templates (PDB entries 4BM1, 3FJW and 1QPA corresponding to *Pleurotus ostreatus* MnP, *P. eryngii* VP and *P. chrysosporium* LiP, respectively). All protein models had great quality taking into account the Swiss-Model parameters (good QMEAN and high GMQE). The electrostatic surfaces were computed with the PyMOL Molecular Graphics System, version 1.8 Schrödinger, LLC (http://pymol.org) using default parameters.

***E. coli* expression**

After gene synthesis, the coding sequences of the most probable CaPo, CaCD, AVP and ALiP ancestral sequences, plus the two additional CaPo sequences described above (CaPo-bis and CaPo-tris), and the extant LiPH8 were cloned into the expression vector pET23b(+) (Novagen). The resulting plasmids were transformed into *E. coli* DH5α for propagation and conservation, and in BL21(DE3)pLysS (ancestors) or W3110 (LiPH8) for expression. With this purpose, cells were grown in Terrific broth until an OD_{600} 0.5-0.6 to reach an adequate expression level, induced with 1 mM isopropyl-ß-D-thiogalactopyranoside, and grown for another 4 h.

The apoenzymes accumulated in inclusion bodies, as reveled by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and were activated *in vitro* (Pérez-Boada et al. 2002; Fernández-Fueyo et al. 2014b). After solubilization in 8 M urea, the refolding conditions for the ancestral proteins included: 0.16 M urea, 5 mM $CaCl_2$, 15 μM hemin, 0.4 mM oxidized glutathione, 0.1 mM dithiothreitol and 0.1 mg/mL of protein in 50 mM Tris-HCl, pH 9.5. The refolding conditions for LiPH8 were: 2.1 M urea, 5 mM $CaCl_2$, 10 μM hemin, 0.7 mM oxidized glutathione, 0.1 mM dithiothreitol and 0.2 mg/mL of protein, in buffer 40 mM Tris-HCl, pH 9.5 (Doyle and Smith 1996). The active enzymes were purified using a Resource-Q column (GE-Healthcare, USA) with 0-400 mM NaCl salt gradient and 2 mL/min flow in 10 mM sodium acetate, pH 5.5, containing 1 mM of $CaCl_2$.

Steady-state kinetics

Five different substrates were selected for the kinetic characterization of the resurrected peroxidases: **i)** Mn^{2+}, which is oxidized to Mn^{3+} (Mn^{3+}-tartrate complex ε_{238} 6500 $M^{-1}\cdot cm^{-1}$); **ii)** VA, whose oxidation product is veratraldehyde (ε_{310} 9300 $M^{-1}\cdot cm^{-1}$); **iii)** DMP, which dimerizes to coerulignone (ε_{469} 55000 $M^{-1}\cdot cm^{-1}$); **iv)** ABTS, which is oxidized yielding the cation radical (ε_{436} 29300 $M^{-1}\cdot cm^{-1}$); and **v)** RB5, whose disappearance/discoloration after oxidation was measured (ε_{598} 30000 $M^{-1}\cdot cm^{-1}$). These reactions were analyzed using a Thermo Scientific Biomate5 spectrophotometer, at 25 ºC and the optimal pH and H_2O_2 concentration for each enzyme, determined using 50 mM Britton-Robinson (B&R) buffer (pH 2-10) and 2.5 mM ABTS as substrate (see **Table 1** footnote). The sigmoid kinetic curves obtained for oxidation of DMP and ABTS enabled calculation of two sets of constants for AVP, corresponding to low and high efficiency oxidation sites, as reported for extant VP (Morales et al. 2012).

pH and temperature stability

To study the effect of pH on enzyme stability, the resurrected peroxidases and extant LiPH8 were incubated in B&R buffer, pH 2-10, at 25 ºC for 4 h.

Then, the residual activity was estimated by the oxidation of ABTS (2.5 mM), under the conditions described above. For every enzyme, the activity after 1 min incubation at 25 ºC in pH 5 buffer was taken as 100%, and the percentages of residual activity at the different pH conditions were referred to this value.

To study their thermal stability, the enzymes were incubated in 10 mM acetate, pH 5.5, for 10 min at 5 ºC intervals in the range 25-85 ºC. Residual activity was measured and calculated as described above. Temperature stability was presented as the 10 min T_{50}, i.e. the temperature at which 50% of the activity was lost after 10 min incubation.

The effect of temperature on circular dichroism spectra was addressed studying the changes at 222 nm from 20 ºC to 95 ºC, 30 ºC/h, using a Jasco J-815 spectropolarimeter equipped with a Peltier temperature controller and a thermostated cell holder on a 0. 1 cm path length quartz cell. A final concentration of 6 μM pure enzyme in 10 mM acetate, pH 5.5 was used. T_m represents the temperature at the midpoint of the unfolding transition in the thermal melting profiles.

The effect of Mn^{2+} in the enzyme thermal inactivation and structure unfolding was analyzed by adding 1 mM SO_4Mn to the incubation mixture, and measuring the residual activity by oxidation of 6 mM Mn^{2+}, as described above.

Funding

This work was supported by the INDOX (KBBE-2013-613549) and EnzOx2 (H2020-BBI-PPP-2015-2-720297) EU projects and the NOESIS (BIO2014-56388-R) project of the Spanish Ministry of Economy and Competitiveness (MINECO). The work conducted by JGI was supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231.

Acknowledgements

The authors thank Valeria Risso and José M. Sánchez-Ruiz (Granada, Spain) for providing the program used for Monte Carlo sampling and their helpful suggestions on ancestral sequence reconstruction, and David S. Hibbett (Clark University, USA) for his comments on phylogenetic analysis. IA-F thanks a MINECO FPI Fellowship. We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI) and the EC OpenAIRE FP7 post-grant Open Access Pilot.

Additional material

Additional file 1: Fig. S1 includes three reconstructed sequences for each of the four nodes in the evolutionary pathway. **Fig. S2** provides the posterior probabilities for each amino acid in the 12 reconstructed sequences included in the previous figure. **Fig. S3** shows the proximal histidine environment in the ancestral

sequences, compared with C*c*P and LiPH8**. Fig. S4** represents the environment of the catalytic tryptophan and homologous region in three ancestral peroxidases and LiPH8 as electrostatic surfaces.

Additional file 2: This file provides 113 sequences of Polyporales peroxidases from ten sequenced genomes in fasta format, aligned in a multiple fas (txt) file.

Additional file 1: Supplemental Figures S1-S4

A

```
CaPo      VTCSDGVSTASNAACCAWFAVLDDIQANLFDGGQCGEEAHESLRLTFHDAIGFSPALAAQGKFGGGGADGSIITFADIETNFHANNGLDDIVDALKPFAD 100
CaPo-bis  VTCSDGVSTATNAACCAWFAVLDDIQANLFDGGQCGEEARESLRLTFHDAIGFSPALAAQGKFTGGGADGSIITFADIETSFHANNGLDDIVEALKHFAD 100
CaPo-tris ATCSDGIRTATNAACCAWFAVLDDIQANLFDGGQCGDQARESLRLTFHDAIGRSPALARQGKFTGGGADGSIITFADIELSYHANNGLEEIVEAQKPFAD 100
          .*****: **:************************::*:*:********** ***** **** *************** .:******::**:* * ***
CaPo      KHNVSYGDFIQFAGAVGVSNCPGAPRLEFLAGRPNATAPSPDGLVPEPSDSVDKILARMADAGGFSPDEVVALLASHSVAAQHHVDPTIPGTPFDSTPST 200
CaPo-bis  KHNVSYGDFIQFAGAVGVSNCPGAPRLEFHAGRPNATAPSPDGLVPEPSDSVDKILARMSDAGGFTPDETVALLASHSVAAQRHVDPTIPGTPFDSTPGT 200
CaPo-tris KHNVSFGDFIQFAGAVGLSNCPGAPRLEFHAGRPNAIAPSPDGLVPEPSDSVDKILARMSDAGGFTPDDTVALLASHSVAAQRTVDPTIPGTPFDSTPGT 200
          *****:**********:********** ****** ********************:*****:**:.********** ***************.*
CaPo      FDTQFFLETLLKGTAFPGTGANSGEVKSPLKGEFRLQSDAAIARDPRTACEWQSFVNNQELMQSSFRAAMAKLANLGHDRSDLIDCSEVIPVPKPLAASA 300
CaPo-bis  FDTQFFVETLLKGTAFPGKGANSGEVKSPSPGEFRLASDAAIARDPRTACEWQSFVNNQEKMQSSFRAAMLKLAIQGHDRSDLIDCSEVIPVPKPLASSA 300
CaPo-tris FDTQFFVETLLKGTAFPGTGANTGEVKSPLKGEFRLASDAAIARHPRTACEWQSFVNNQEKMQASFRAAMAKLANQGHDRSDLVDCSEVIPVPKPLAASA 300
          ******:**********.***:******  ***** ******* *************** **:****** ***  *******:*************:**
CaPo      TFPAGKTRSDIEQSCRSTPFPTLPTDPGPATSIPPV 336
CaPo-bis  TFPAGKTRKDIEQSCPSTPFPTLPTDPGPETSVPPV 336
CaPo-tris TFPAGKSRKDIEQSCPSTPFPTLPSDPGPATSIPPV 336
          ******:*.****** ********:**** **:***
```

B

```
CaCD      VTCPDGVNTATNAACCALFAVLDDIQENLFDGGECGEEAHESLRLTFHDAIGFSPALARQGKFGGGGADGSIITFSDIETNFHANGGIDEIVEVQKPFVA 100
CaCD-bis  VTCPDGVNTATNAACCALFAVVDDIQENLFDGGECGEEVHESLRLTFHDAIGFSPALARQGKFGGGGADGSIITFSDIETNFHANGGIDEIVEVQKPFVA 100
CaCD-tris VTCPDGVNTATNAACCALFAVVDDIQENLFDGGECGEEAHESLRLTFHDAIGFSPALARQGKFGGGGADGSIITFSDIETNFHANGGIDDIVEVQKPFVA 100
          **********************:****************.*************************************************:**********
CaCD      KHNMTAGDFIQFAGAVGVSNCPGAPRLEFLLGRPAATAPSPDGLVPEPFDSVDKILARFADAGGFSPDEVVALLASHSVAAAHHVDPTIPGTPFDSTPST 200
CaCD-bis  KHNMTAGDFIQFAGAVGVSNCPGAPRLEFLLGRPAATAPSPDGLVPEPFDSVDKILARFADAGGFSPDEVVALLASHSVAAAHHVDPTIPGTPFDSTPST 200
CaCD-tris KHNMTAGDFIQFAGAVGVSNCPGAPRLEFLLGRPAATAPSPDGLVPEPFDSVDKILARFADAGGFSPDEVVALLASHSVAAAHHVDPTIPGTPFDSTPST 200
          ************************************************************************************************
CaCD      FDTQFFVEVLLRGTLFPGTGGNQGEVKSALRGEIRLQSDHEVARDPRTACEWQSFVNNQAKMQKSFRAAMAKLAILGHDRSDLIDCSEVIPVPPPLAATA 300
CaCD-bis  FDTQFFVEVLLRGTLFPGTGGNQGEVKSALPGEIRLQSDHEVARDPRTACEWQSFVNNQAKMQKSFRAAMAKLAILGHDRSDLVDCSEVIPVPPPLAATA 300
CaCD-tris FDTQFFVEVLLRGTLFPGTGGNQGEVKSPLRGEIRLQSDNEVARDPRTACEWQSFVNNQAKMQKSFRAAMAKLAILGHDRSDLIDCSEVIPVPPPLAATA 300
          *************************** * ********.****************************************:*****************
CaCD      HFPAGLTRKDIEQSCRSTPFPTLSTDPGPATSVPPV 336
CaCD-bis  HFPAGLTRKDIEQSCRSTPFPTLSTDPGPATSVPPV 336
CaCD-tris HFPAGLTRKDIEQSCRSTPFPTLSTDPGPATSVAPV 336
          ********************************** **
```

C

```
AVP       VACPDGVNTATNAACCALFAVRDDIQQNLFDGGECGEEVHESLRLTFHDAIAFSPALEAQGQFGGGGADGSIAIFEDIETNFHANLGLDEIVNEQKPFIA 100
AVP-bis   VACPDGVNTATNAACCALFAVRDDIQQNLFDGGECGEEVHESLRLTFHDAIAFSPALEAQGQFGGGGADGSIAIFEDIETNFHANLGLDEIVNEQKPFIA 100
AVP-tris  VACPDGVNTATNAACCALFAVRDDIQQNLFDGGECGEEVHESLRLTFHDAIAFSPALEAQGQFGGGGADGSIAIFEDIETNFHANLGLDEIVNEQKPFIA 100
          ************************************************************************************************
AVP       RHNMTTADFIQFAGAVGVSNCPGAPQLDFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGGFDPIETVWLLTSHTIAAAHHVDPTIPGTPFDSTPEL 200
AVP-bis   RHNMTTADFIQFAGAVGVSNCPGAPQLDFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGGFDPIETVWLLTSHTVAAAHHVDPTIPGTPFDSTPEL 200
AVP-tris  RHNMTTADFIQFAGAVGVSNCPGAPQLDFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGGFDPIETVWLLASHTIAAAHHVDPTIPGTPFDSTPEL 200
          ************************************************************:***:********************
AVP       FDTQFFIETQLRGTLFPGTGGNQGEVESPLRGEIRLQSDHLLARDSRTACEWQSFVNNQPKLQKSFQAAFHDLSMLGHDVNDLIDCSEVIPIPPPPTSTA 300
AVP-bis   FDTQFFIETQLRGTLFPGTGGNQGEVESPLRGEIRLQSDHLLARDSRTACEWQSFVNNQPKLQKNFQAAFHDLSVLGHDVNDLIDCSEVIPIPPPPASTA 300
AVP-tris  FDTQFFIETQLRGTLFPGTGGNQGEVESPLRGEIRLQSDHLLARDSRTACEWQSFVNNQPKLQKNFQAAFHDLSMLGHDVNDLIDCSEVIPIPPPPTSTA 300
          ************************************************************.*********:**********************:***
AVP       HFPAGLTNADVEQACAETPFPTLPTDPGPATSVAPV 336
AVP-bis   HFPAGLTNADVEQACAETPFPTLPTDPGPATSVAPV 336
AVP-tris  HFPAGLTNADVEQACAETPFPTLPTDPGPATSVAPV 336
          ************************************
```

D

```
ALiP      VACPDGVNTATNAACCALFAVRDDIQQNLFNGGQCGEEAHESLRLTFHDAIAFSPALEAQGQFGGGGADGSIVIFSDIETNFHANIGLDEIVAIQKPFIA 100
ALiP-bis  VACPDGVNTATNAACCALFAVRDDIQQNLFNGGQCGEEAHESLRLTFHDAIAFSPALEAQGQFGGGGADGSIVIFSDIETNFHANIGLDEIVAIQKPFIA 100
ALiP-tris VACPDGVNTATNAACCALFAVRDDIQQNLFNGGQCGEEAHESLRLTFHDAIAFSPALEAQGQFGGGGADGSIVIFSDIETNFHANIGLDEIVAIQKPFIA 100
          ************************************************************************************************
ALiP      RHNMTVADFIQFAGAVGVSNCPGAPQLNFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGEFDELETVWLLIAHTVAAANDVDPTIPGTPFDSTPEL 200
ALiP-bis  RHNMTVADFIQFAGAVGVSNCPGAPQLNFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGEFDELETVWLLIAHTVAAANDVDPTIPRTPFDSTPEL 200
ALiP-tris RHNMTVADFIQFAGAVGVSNCPGAPQLNFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGEFDELETVWLLIAHTVAAANDVDPTIPRTPFDSTPEL 200
          ************************************************************************************************ *********
ALiP      FDSQFFIETQLRGTLFPGTGGNQGEVESPLKGEMRLQSDHLLARDSRTACEWQSFVNNQPKLQKNFQFVFEALSMLGQDPNDLIDCSEVIPIPPPLTLTP 300
ALiP-bis  FDSQFFIETQLRGTAFPGTGGNQGEVESPLKGEMRLQSDHLLARDSRTACEWQSFVNNQPKLQKNFQFIFEALSMLGQDPNNLIDCSEVIPIPPPLTLTP 300
ALiP-tris FDSQFFIETQLRGTLFPGTGGNQGEVESPLKGEIRLQSDHLLARDSRTACEWQSFVNNQPKLQKNFQFVFEALSMLGQDPNDLIDCSEVIPIPPPLTLTP 300
          ************** ********************:************************************:*************:*****************
ALiP      AASHFPAGKTNKDVEQACAETPFPTLPTDPGPATSVAPVPPSPAA 345
ALiP-bis  AFSHFPAGKTNADVEQACAETPFPTLPTDPGPATSVAPVPPSPAA 345
ALiP-tris GAAHFPAGKTNKDVEQACAETPFPTLPTDPGPATSVAPVPPSPAA 345
          . :******** ********************************
```

Figure S1. Twelve reconstructed sequences representative of the different nodes. Most probable and alternative (bis/tris) sequences (mature proteins) for CaPo (**A**, 88-92% identity), CaCD (**B**, 98-99%), AVP (**C**, 99%) and ALiP (**D**, 98-99%) nodes were obtained by Monte Carlo sampling (yielding 5000 sequences/node using 0.5 threshold) followed by random sampling (see Methods for sampling strategy) yielding 3 sequences/node (red and blue, typical Mn-binding site and catalytic tryptophan, respectively; other colors, see **Fig. 2**).

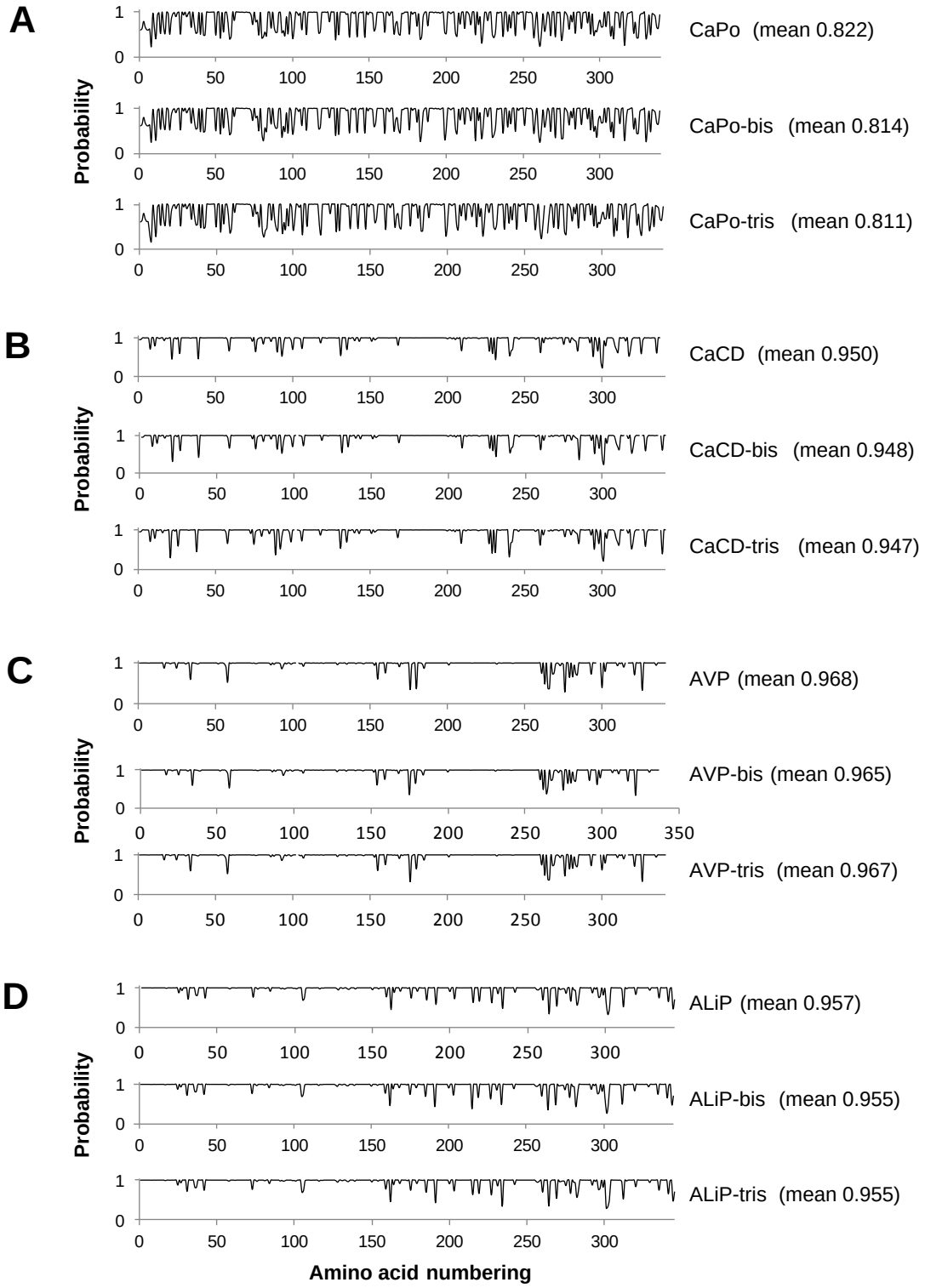


Figure S2. Posterior probability for each amino acid of the 12 reconstructed ancestral sequences shown in **Fig. S1**. The most probable sequence for nodes CaPo (**A**), CaCD (**B**), AVP (**C**) and ALiP (**D**), and two alternative sequences (bis and tris) for each of them (see Methods for sampling strategy) are included. The probabilities of Trp172 in AVP and ALiP sequences are 0.999 and 1.000, respectively; and that of Asn183 in ALiP sequences is 0.993. The mean probabilities for every reconstructed sequence are also shown.

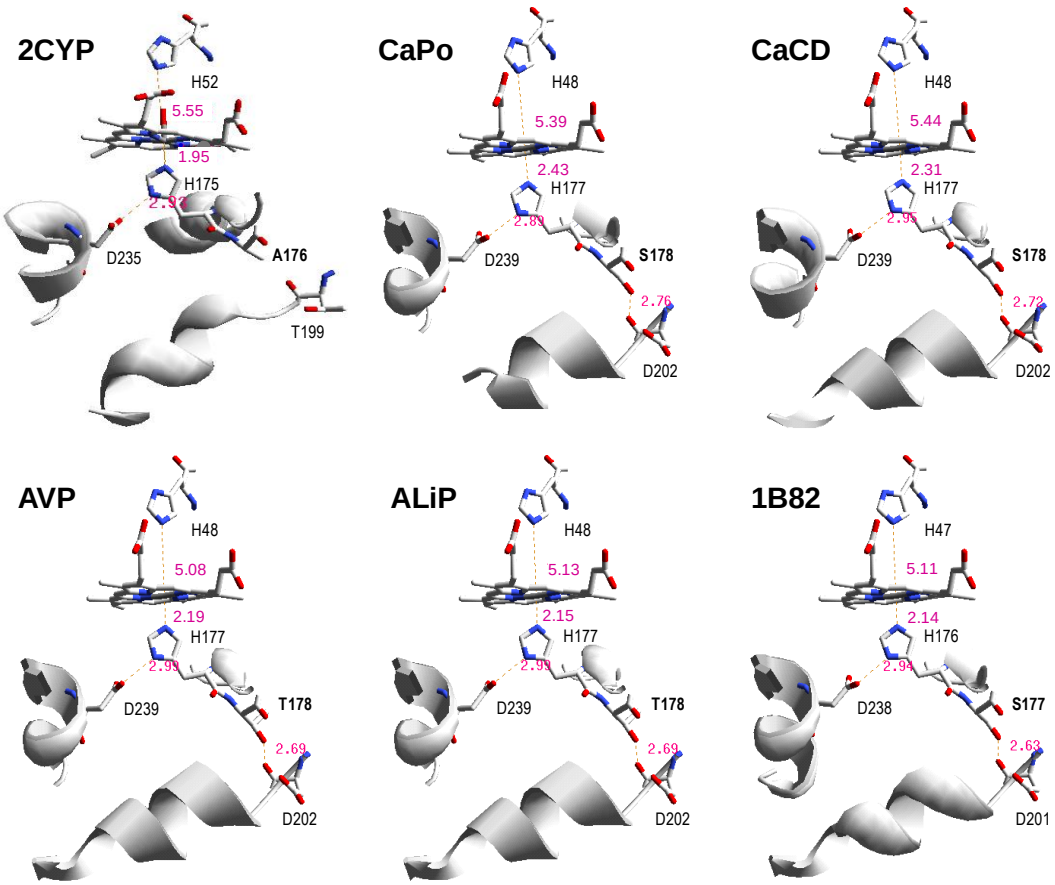


Figure S3. Proximal histidine, neighbor residues and distances (Å) in ancestral peroxidases (His177) and extant C*c*P (2CYP; His175) and *P. chrysosporium* LiPH8 (1B82; His176). His177-Asp239 and Ser/Thr178-Asp202 H-bonds would affect the position of the proximal histidine in the ancestral peroxidases, and consequently the heme iron electron-deficiency (homologous H-bonds are present in LiPH8, but the second one is absent from C*c*P). Distal histidine (His48 in the ancestors, His52 in CcP, and His47 in LiPH8) is also shown.

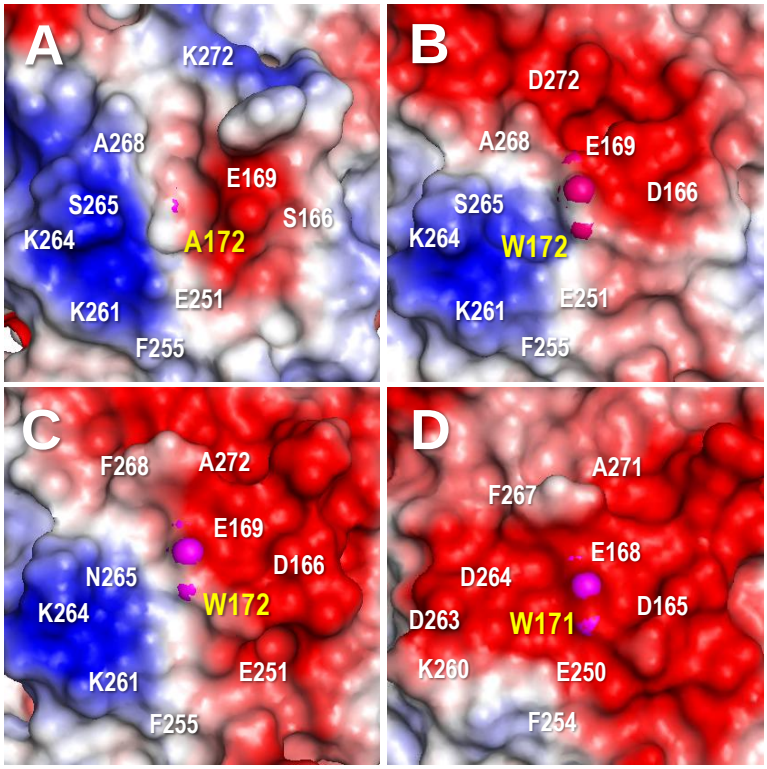


Figure S4. Surface environment of catalytic tryptophan. Electrostatic surfaces computed for ancestral CaCD (**A**), AVP (**B**) and ALiP (**C**) homology models, and extant LiPH8 (**D**) crystal structure (PDB 1B82) showing the environment of the exposed tryptophan (magenta spheres, yellow labels) in **B-D**, and the equivalent alanine residue in **A** (the positions of neighbor residues are also indicated). The presence of the catalytic tryptophan in AVP and ALiP and the absence of a net negative environment, as found in LiPH8 that is unable to oxidize RB5, contribute to the oxidation of this anionic dye by the two ancestral enzymes. The opposite effect is expected for VA, whose cation radical would be stabilized by the negative environment in LiPH8.

Chapter 2

Evolutionary convergence in lignin degrading enzymes

General abstract

After characterization of the ancestral enzymes leading to the most efficient ligninolytic peroxidases (LiPs), we explored other interesting lineages in the phylogeny of class-II peroxidases from Polyporales. This way, we observed that in the enzymes leading to Cluster-B peroxidases (formed by VPs and MnPs mainly) there was another node where the solvent-exposed catalytic tryptophan appeared, meaning that the exploration of new strategies for lignin degradation could have happened several times in evolution. As discussed in the previous chapter, this solvent-exposed tryptophan constitutes an important innovation in evolution: it is a new oxidation site responsible for the direct oxidation of nonphenolic lignin whose bulky nature prevents contact with the heme cofactor.

In this chapter, we considered this new strategy for lignin degradation an important evolutionary trait in wood-rotting fungi, and using ancestral enzyme resurrection we evaluate if it is a convergent process due to the production of the same type of enzymes in two distant lineages. For the first time, convergence in wood-rotting fungi is evaluated here, and we show the results of the biochemical characterization of the two distant lineages where it could have happened.

Article published as:

Ayuso-Fernánez, I., Ruiz-Dueñas, FJ., Martínez, AT. *Evolutionary convergence in lignin degrading-enzymes*. **PNAS**. 2018, 115, 6428-6433. IF: 9.504.

Author contributions: I.A.-F. performed the experimental work. F.J.R.-D. and A.T.M. designed the research. All authors analyzed the data, wrote the paper, and approved the final manuscript.

Abstract

The resurrection of ancestral enzymes of now extinct organisms (paleogenetics) is a developing field that allows the study of evolutionary hypotheses otherwise impossible to be tested. In the present study we target fungal peroxidases that play a key role in lignin degradation, an essential process in the carbon cycle and often a limiting step in biobased industries. Ligninolytic peroxidases are secreted by wood-rotting fungi, whose origin was recently established in the Carboniferous period associated with the appearance of these enzymes. These first peroxidases were not able to degrade lignin directly, and used diffusible metal cations to attack its phenolic moiety. The phylogenetic analysis of the peroxidases of Polyporales, the order where most extant wood-rotting fungi are included, suggests that later in evolution these enzymes would have acquired the ability to degrade nonphenolic lignin using a tryptophanyl radical interacting with the bulky polymer at the surface of the enzyme. Here we track this powerful strategy for lignin degradation as a phenotypic trait in fungi, and show that is not an isolated event in the evolution of Polyporales. Using ancestral enzyme resurrection, we study the molecular changes that led to the appearance of the same surface oxidation site in two distant peroxidase lineages. By characterization of the resurrected enzymes, we demonstrate convergent evolution at the amino-acid level during the evolution of these fungi, and track the different changes leading to phylogenetically-distant ligninolytic peroxidases from ancestors lacking the ability to degrade nonphenolic lignin.

Significance statement

We analyze the molecular mechanisms that led to the rise of a powerful strategy for lignin degradation (i.e. the formation of a solvent exposed tryptophanyl radical capable of oxidizing the bulky lignin polymer) as a convergent trait in different species of fungi (order Polyporales). We use ancestral sequence reconstruction and enzyme resurrection to obtain the ancestors of the two extant types of ligninolytic peroxidases - lignin peroxidase (LiP) and versatile peroxidase (VP) - and compare their predicted molecular structures and catalytic properties after resurrection. The results presented demonstrate convergent evolution in distant LiP and VP lineages, with the exposed tryptophan residue appearing twice, as two independent events, following different molecular changes.

Introduction

Degradation of lignin is essential for carbon recycling in land ecosystems, and often represents a key step for the use of biomass in the industry (Martínez et al. 2009). The main organisms that are able to mineralize lignin are white-rot fungi, using an array of oxidative enzymes (Ruiz-Dueñas and Martínez 2009). Three class-II peroxidases of the peroxidase-catalase superfamily (Zámocký et al. 2015) are involved in the initial attack to lignin: **i)** lignin peroxidases (LiP), which are able to oxidize its major nonphenolic moiety (Hammel and Cullen 2008); **ii)** manganese peroxidases (MnP) including short and long MnPs with slightly different properties (Fernández-Fueyo et al. 2014a) that oxidize Mn^{2+} to Mn^{3+}, whose diffusible chelates oxidize the minor phenolic moiety of lignin (Gold et al. 2000); and **iii)** versatile peroxidases (VP), which combine the catalytic properties of LiP, MnP and plant peroxidases (the latter oxidizing phenolic monolignols) (Ruiz-Dueñas et al. 1999; Camarero et al. 1999b). Also, white-rot fungi produce generic peroxidases (GP), catalytically similar to plant peroxidases. The above four peroxidase types have been characterized, and their structural and kinetic properties are well known (Ruiz-Dueñas and Martínez 2010). Thereby, they can oxidize substrates at three sites: **i)** the main heme access channel, where low redox-potential compounds are oxidized (in all of them); **ii)** a Mn^{2+}-oxidizing site, formed by three acidic residues near one of the heme propionates (Kishi et al. 1996) (in MnP and VP); and **iii)** a surface tryptophan (Pérez-Boada et al. 2005; Smith et al. 2009) that is able to oxidize lignin directly (Sáez-Jiménez et al. 2015; 2016) *via* an aminoacyl radical and long-range electron transfer to heme (Acebes et al. 2017) (in VP and LiP).

In past years, there has been an increasing interest in the evolution of wood-degrading organisms. The origin of lignin degradation by fungi, associated with the appearance of the first ligninolytic peroxidases, has been estimated to have occurred during the Carboniferous period, playing a role in the decline of coal accumulation near the end of the Permo-Carboniferous (Floudas et al. 2012). However, geoclimatic factors would have also significantly contributed to coal formation under everwet tropical conditions, and its decline could also be related to climatic shifts toward drier conditions (Nelsen et al. 2016; Hibbett et al. 2016). Then, the expansion and diversification of genes encoding ligninolytic peroxidases occurred leading to the families existing today, as shown by genomic and evolutionary studies (Floudas et al. 2012; Ruiz-Dueñas et al. 2013; Nagy et al. 2016). The diversity and evolution of these enzymes have been particularly studied in the order Polyporales, where the lignicolous habitat is largely predominant resulting in the most efficient ligninolytic enzymes. Recent studies included first analyses of the appearance and disappearance

of relevant catalytic sites (Ruiz-Dueñas et al. 2013) and later sequence reconstruction, heterologous expression ("resurrection") and experimental characterization of some ancestral enzymes (Ayuso-Fernández et al. 2017). The evolutionary analysis of peroxidases shows phylogenetically distant enzymes (corresponding to the above LiP and VP types) that would be *a priori* able to oxidize nonphenolic lignin at the exposed catalytic tryptophan (Mester et al. 2001; Pérez-Boada et al. 2005; Smith et al. 2009; Sáez-Jiménez et al. 2015; 2016; Acebes et al. 2017). To determine if the LiP/VP distribution in the peroxidase phylogeny is due to duplication of an ancestor and maintenance of function or to a convergent process, we first perform phylogenetic analyses and ancestral sequence reconstruction of Polyporales peroxidases from sequenced genomes. This choice should enable a precise description on the evolution of ligninolytic enzymes within this order, although the reconstruction of enzymes pre-dating Polyporales could be partially biased. Then, we compared in the laboratory the previously described line leading to extant LiP (Ayuso-Fernández et al. 2017) with a new independent evolutionary line leading to extant VP, and established their convergent evolution using empirical analyses of resurrected enzymes.

Results

Ancestral Sequences of Polyporales Peroxidases. For reconstructing the ancestral sequences leading to the extant lignin-degrading peroxidases in Polyporales, we built a phylogenetic tree with RAxML (**Fig. 1**), after manual annotation of the complete set of peroxidase sequences (a total of 113) in the sequenced genomes of *Bjerkandera adusta*, *Ceriporiopsis subvermispora*, *Dichomitus squalens*, *Fomitopsis pinicola*, *Ganoderma* sp., *Phlebia brevispora*, *Phanerochaete chrysosporium*, *Postia placenta*, *Trametes versicolor* and *Wolfiporia cocos*. For the present analysis, the tree was divided into: **i)** clade-A containing GPs; **ii)** clade-B including two sub-clades of short MnPs and VPs; **iii)** clade-C formed exclusively of long MnPs; and **iv)** clade-D containing short MnPs and some VPs, together with the large LiP subclade. B and D are the only clades that contain enzymes that *a priori* would oxidize nonphenolic lignin due to the presence of the exposed catalytic tryptophan. To explain the presence of this residue in two phylogenetically different peroxidase clades, horizontal gene transfer between lines D and B was first ruled out, since no similarities were detected between the flanking regions of genes encoding clade-D LiPs and clade-B VPs. Double gene transfer (from outside Polyporales) was also rejected since BLAST search of both extant genes on the NCBI database only show sequence identities with genes of Polyporales peroxidases. Therefore, duplication of the gene of an ancestral peroxidase containing the tryptophan and its differentiation in lines B and D, or convergent evolution were considered.

To decide between these two alternative hypotheses, we performed ancestral sequence reconstruction with PAML. The average posterior probability, for the predicted amino acids in each reconstructed sequence, always was >0.95, except for the most ancestral reconstructed sequence (SI Appendix, **Fig. S1**). More importantly, no ambiguity was observed for the positions of the catalytic tryptophan, the Mn^{2+}-oxidation site, or other residues relevant for catalysis described below. Then, we selected six ancestral enzymes, and two lignin-degrading extant peroxidases - the highly expressed *T. versicolor* VP2 (Carabajal et al. 2013) and the well known *P. chrysosporium* LiPA (LiPH8) (Hammel and Cullen 2008) - for heterologous expression and comparative characterization. The eight sequences define two well separated evolutionary lines leading to the above extant ligninolytic peroxidases (TV-VP2 and PC-LiPA). Both pathways begin with the Common ancestor of Polyporales peroxidases (CaPo), which is the precursor of gene lineages leading to clade-B VP (blue line in **Fig. 1**) and clade-D LiP (red line in **Fig. 1**). In this way, CaPo gave rise to the Common ancestors of clades D (CaD) and B (CaB). Therefore, the main evolutionary change is the independent and parallel appearance of the surface tryptophan in two ancestral states of both lines: Ancestral VP of line-B (AVP-b) and Ancestral VP of line-D (AVP-d). After this event, there are no significant modifications in the evolution of peroxidases in line-B up to the extant enzyme (TV-VP2). However, in line-D evolution, AVP-d lost its Mn^{2+}-oxidation site giving rise to the first LiP of the phylogeny (Ancestral LiP, ALiP) that will later evolve into extant LiPs of clade-D (including PC-LiPA). The molecular mechanisms that lead to the emergence of the surface tryptophan are different in both lines, as described below (and indicated on the ancestral nodes in **Fig. 1)**.

Structural Comparison of two Peroxidase Lineages. The amino-acid sequences of the ancestors and the two extant peroxidases show 56-87% identity (SI Appendix, **Table S1)**. Their multiple alignment (**Fig. S2**) reveals that there have not been strong modifications in the conserved regions during evolution, beyond the changes in the oxidation sites that we describe below. Moreover, the molecular models (**Fig. 2**) show that the overall structure, with 12 helices, two structural Ca^{2+} ions and four disulfide bonds, is maintained through time. A more detailed analysis shows that all the enzymes have a well defined heme pocket, with a proximal His177 coordinating the $Fe^{3+}$of the heme, Asp239 and Phe194 at one side of the heme (numbers referred to CaPo and line-D); and a distal His48, Arg44 (two residues that participate in reaction with H_2O_2), Asn85 and Phe47 at the opposite side (being His47, Arg43, Asn84 and Phe46 in line-B). The Ca^{2+} ions would be similarly coordinated in all the enzymes with small differences that would not affect the anchorage of the cation.

The main differences are related to the oxidation sites that these peroxidases have. These sites are identified in all the ancestors, and we tracked their changes as they evolved. Mn^{2+}-binding site is defined by three acidic residues that already appear in CaPo (Glu37, Glu41 and Asp183) and are maintained both in line-B until the extant TV-VP2 (Glu36, Glu40 and Asp181) and in line-D until AVP-d. In this line the Mn^{2+}-binding site is lost in ALiP (Asp183 becomes Asn183, **Fig. 2**, red line), and remains absent in all extant LiPs. Note that clade-C, the sister clade of clade-D (**Fig. 1**), evolved maintaining the Mn^{2+} oxidation site.

The site for direct oxidation of lignin is located in a surface tryptophan and electrons are transferred to the heme following a preferred route involving buried Trp251 and Phe205 in PC-LiPA (Acebes et al. 2017). Analysis of the homologous oxidation site in CaPo reveals that the ancestral surface residue was Ala172, while the amino acids of the route are present since the origin (Trp252 and Phe206 in CaPo) and conserved through evolution (**Fig. 2**). Therefore, although the scaffold for electron transfer is present, the absence of the required exposed tryptophan would have impeded for the most ancestral enzymes the direct oxidation of lignin.

The appearance of the catalytic tryptophan in line-D occurs in AVP-d (Trp172, **Figs. 1** and **2** red line), the first enzyme of this line that would be able to modify lignin directly. After AVP-d, both ALiP and PC-LiPA maintain the surface tryptophan. The study of the same oxidation site in line-B shows that in the common ancestor of this clade (CaB) Ala172 became Asp170, which later changed to Trp170 in AVP-b (and is maintained in TV-VP2, **Figs. 1** and **2**, blue line). Thus, the same oxidation site appears twice in evolution and, interestingly, different sequences of changes led to the same catalytic amino acid in parallel processes. Therefore, this event is defined as a convergent trait in lignin degradation by fungal peroxidases. To confirm the ability to oxidize lignin model compounds, and compare the catalytic properties in both convergent lines, we resurrected the described ancestral enzymes as reported below.

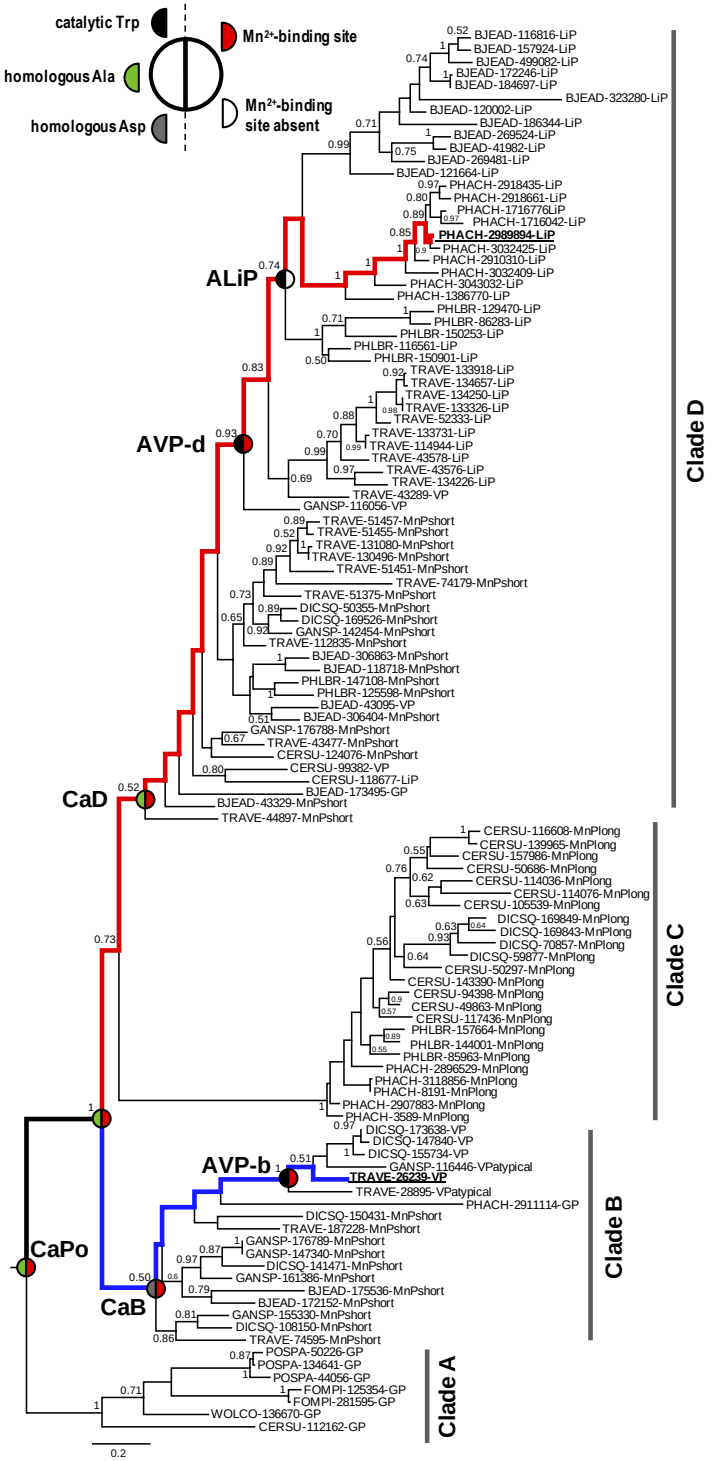


Fig. 1. Phylogenetic tree of 113 peroxidases from 10 Polyporales genomes (sequences in **Dataset S1**; bootstrap values ≥0.5 are indicated). Clades A to D are shown. The paths to the extant LiPA of *P. chrysosporium* (JGI ID# 2989894) and VP2 of *T. versicolor* (JGI ID# 26239) are shown in red and blue, respectively. Also, the milestones in this evolutionary lines (CaPo for both lines; CaD, AVP-d, and ALiP in red line; and CaB and AVP-b in blue line) are marked. The circles show the characteristics of the oxidation sites present in each of these nodes (left: catalytic tryptophan and homologous residues; right: Mn^{2+}-oxidation site). The sequence labels start with the species code (BJEAD, *B. adusta*; CERSU, *C. subvermispora*; DICSQ, *D. squalens*; FOMPI, *F. pinicola*; GANSP, *Ganoderma* sp; PHLBR *P. brevispora*; PHACH, *P. chrysosporium*; POSPL, *P. placenta*; TRAVE, *T. versicolor*; and WOLCO, *W. cocos*) followed by the JGI ID# and the peroxidase type, including GP, LiP, MnP-short, MnP-long, VP, and VP-atypical.

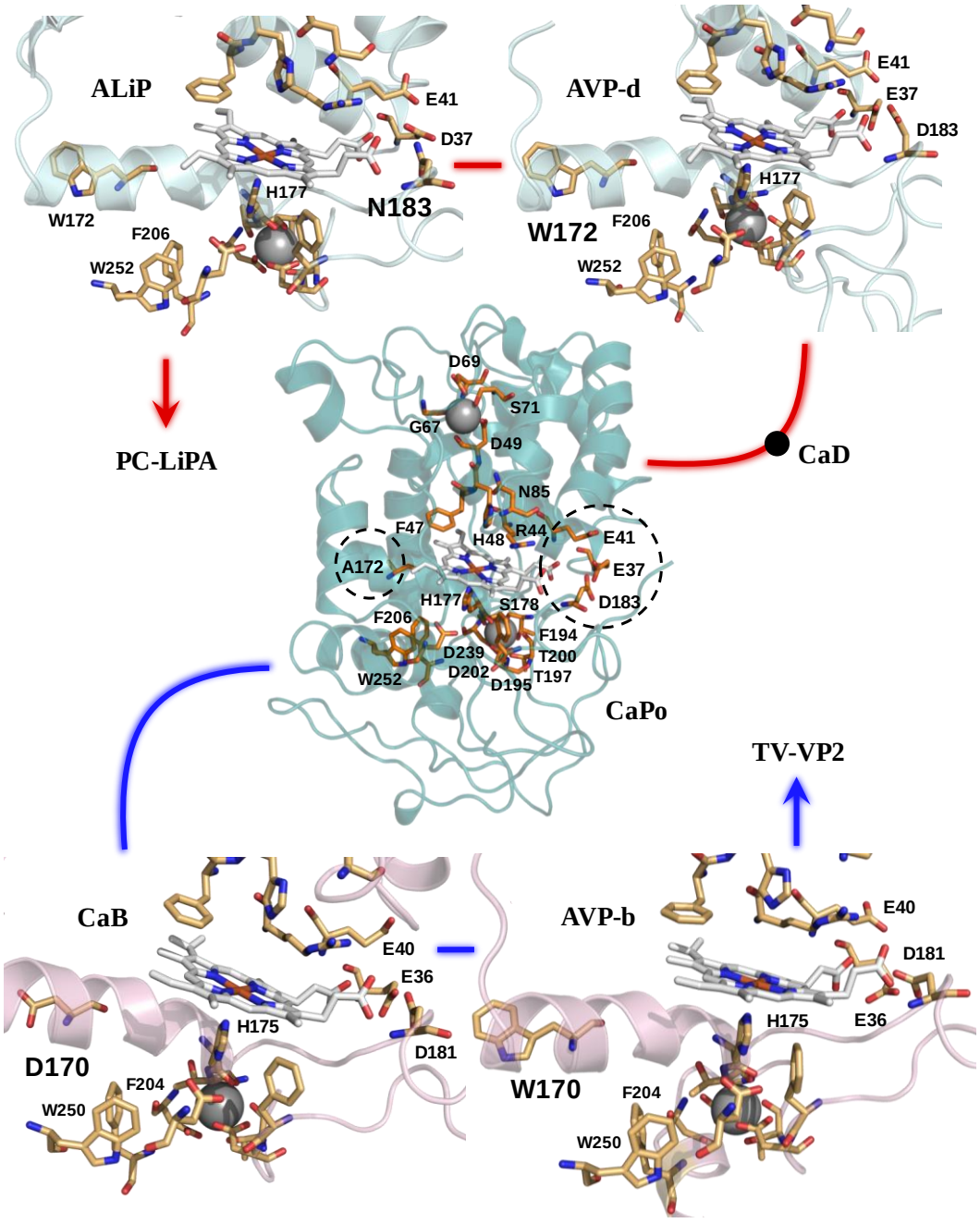


Fig. 2. Molecular model of ancestral CaPo and its main structural changes in evolution. The most relevant amino acids of the common ancestor (CaPo) are labeled, and only the main changes in the oxidation sites are represented for the other peroxidases (two structural Ca^{2+} ions are shown as gray spheres). Mn^{2+}-binding site, formed by three acidic residues, and Ala172, homologous to catalytic tryptophan, are circled in CaPo. Red line: changes in line-D evolution, with appearance of Trp172 in AVP-d and loss of the Mn^{2+}-binding site due to Asn183 appearance in ALiP. Blue line: changes in line-B evolution, with appearance of Asp170 in CaB that later changed to Trp170 for lignin oxidation by AVP-b. Aromatic residues (Phe206 and Trp252 in CaPo) involved in long-range electron transfer from the exposed tryptophan are conserved from the first ancestor.

Reaction Kinetics and Convergent Evolution of Ancestral Peroxidases. The selected enzymes from the two evolutionary lines (**Fig. 1**) were resurrected by *Escherichia coli* expression of the synthesized genes, *in*

vitro activated and purified, and their catalytic properties analyzed using five substrates, which define the different oxidation sites that these peroxidases can have. Veratryl alcohol (VA) was used as a model for nonphenolic lignin, while 2,6-dimethoxyphenol (DMP) was tested representing the minor phenolic moiety of lignin. The oxidation of Mn^{2+} was also analyzed, since Mn^{3+} oxidizes phenolic lignin. Finally, the oxidation of two dyes, often used as high (Reactive Black 5, RB5) and low (2,2'-azinobis[3-ethylbenzothiazoline-6-sulfonate], ABTS) redox-potential peroxidase substrates, was also assayed. As we described above, the sites for the oxidation of Mn^{2+} and high redox-potential substrates (VA and RB5) are well defined in the structure of these enzymes. Moreover, low redox-potential substrates (DMP and ABTS) can be oxidized, with high efficiency, at the same tryptophan responsible for oxidation of high redox-potential substrates and, with low efficiency, at one of the heme-access channels, resulting in biphasic kinetics, as shown for some extant peroxidases (Morales et al. 2012).

Catalytic constant (k_{cat}) and affinity constant (K_m) were calculated for all the resurrected and extant peroxidases (**Table S2**), but what is clear with evolution is the change in the catalytic efficiency (k_{cat}/K_m) with time (**Fig. 3**, upper bars). CaPo, the common ancestor of both lines, is an enzyme that can oxidize low redox-potential substrates in the low-efficiency site, but also Mn^{2+} at the specific binding site. After that, the respective common ancestors of each line (CaB and CaD) are almost identical in their catalytic properties: both are able to oxidize low redox-potential substrates (in their low-efficiency sites) and Mn^{2+}. Interestingly, the same trend is observed in both ancestors: they reduce the catalytic efficiency oxidizing ABTS and DMP while the efficiency oxidizing Mn^{2+} is improved. From this point on, the trends in lines B and D are different taking into account the nature of their catalytic sites and the reactions that they perform. While the efficiency oxidizing Mn^{2+} decreases in AVP-b (blue line in **Fig. 3**), AVP-d attains the highest value among all enzymes analyzed here (red line in **Fig. 3**). Later, in evolution of clade-B there are no changes in the oxidation of the cation (in TV-VP2) while in clade-D the Mn^{2+}-oxidation site/activity is lost.

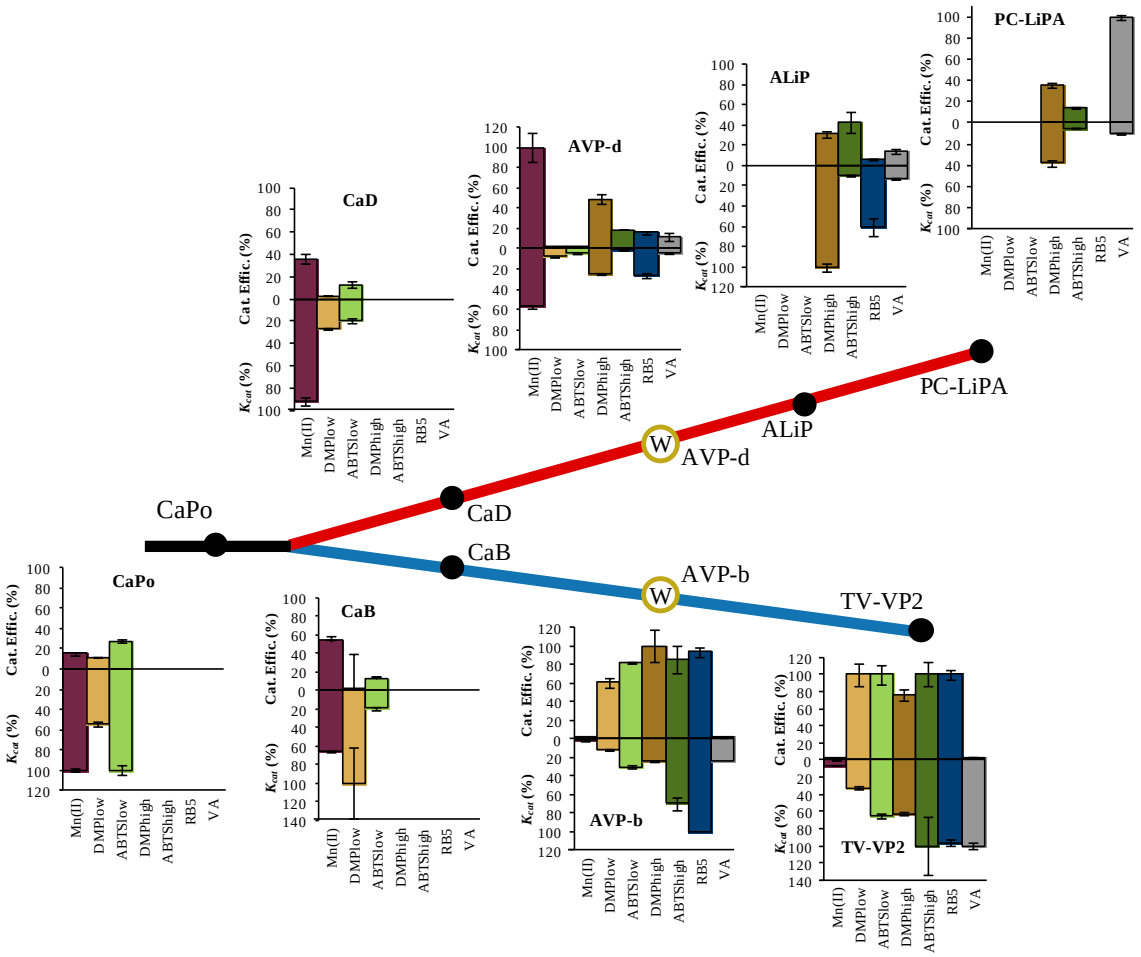


Fig. 3. Evolution of catalytic properties in the D (red line) and B (blue line) evolutionary pathways. Changes of catalytic efficiency (*upper bars*) and k_{cat} (*lower bars)* are shown for oxidation of Mn^{2+} (purple); DMP at low and high efficiency sites (light and dark brown, respectively); ABTS at low and high efficiency sites (light and dark green, respectively); RB5 (blue); and VA (gray) (means and 95% confidence limits). For each substrate, the maximum value was taken as 100% and referred to that for the other enzymes (see **Table S2** for absolute values). The circled W marks the point when the catalytic tryptophan appeared for the first time, and black circles represent the other nodes analyzed.

The nonphenolic lignin model substrate VA begins to be oxidized in parallel in both lines, as a convergent trait in AVP-b and AVP-d, coinciding with the appearance of the catalytic tryptophan. However, the evolution of the catalytic efficiency differs. While in line-D the efficiency oxidizing VA increases in ALiP and is maximal in PC-LiPA (red line in **Fig. 3**), it is maintained at low levels in TV-VP2 of clade-B (blue line in **Fig. 3**). Interestingly, the k_{cat} for VA is always high in line-B, with the TV-VP2 value being 8-fold the observed for ALiP (the highest in line-D) (**Fig. 3,** lower bars). The main reason for a low catalytic efficiency of the two VPs in line-B

is the high K_m they have (**Table S2**), 3-4 magnitude orders greater than observed for the enzymes of line-D. One explanation for differences in the activity of peroxidases in lines B and D is the different charge distribution in the surface environment of the catalytic tryptophan. As shown in **Fig. 4**, the tryptophan (or homologous residue) environment in line-D is progressively more acidic while a similar tendency was not observed in line-B. A more electronegative environment will promote stabilization of positively-charged compounds (such as veratryl alcohol cation radical) and, more importantly, will increase the oxidizing power of the catalytic radical.

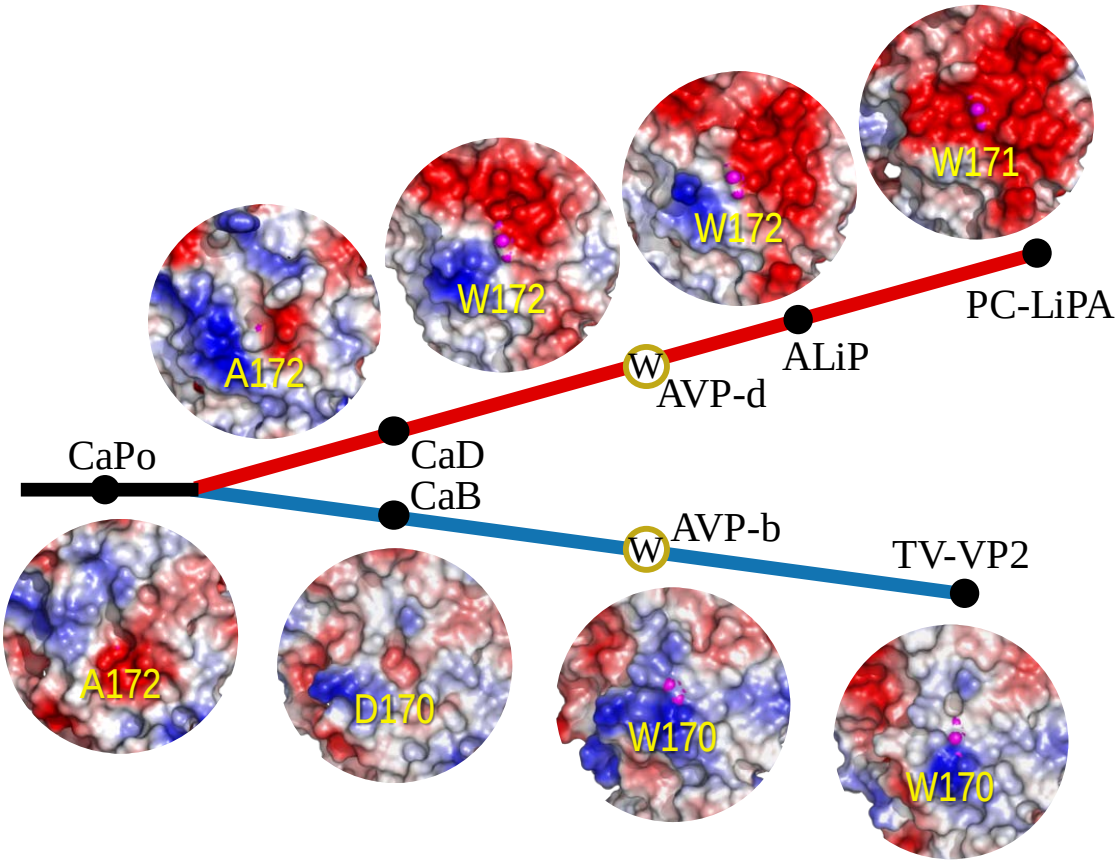


Fig. 4. Changes in the electrostatic surface of the environment of the catalytic tryptophan and homologous residues (pink spheres in the center) in peroxidase evolution. In line-B to TV-VP2 (blue line), the changes are more subtle, but in line-D to PC-LiPA (red line) a clear increase in the negative charge (red) happened with time.

Stability Comparison in the two Peroxidase Lineages. We analyzed the pH stability of the ancestors and extant enzymes in the evolutionary lines B (**Fig. S3**) and D (**Fig. S4**) by measuring the residual activity after 4-h incubation at 25 ºC, in the pH 2-10 range. Overall, the stability at pH 4-6 is higher in the ancestors of clade-B while the ancestors of clade-D are more stable at pH > 6 (where CaB, AVP-b and TV-VP2 are inactivated). More interestingly, the stability at pH 3 (where ligninolysis takes place in nature) strongly increases during the last evolution steps (**Fig. 5A**), either in

parallel with the appearance of the catalytic tryptophan (line-D) or after its appearance (line-B).

Thermal denaturation was studied by circular dichroism (T_m values) and residual activity measurement (T_{50} values) of peroxidases in lines B (**Fig. S5**) and D (**Fig. S6**). The melting profiles parallel the changes observed in activity in all cases, and tended to decrease during evolution (line-D), although all the T_{50} values were in the 55-65 ºC range (**Fig. 5B**). The main change was observed when the Mn^{2+}-binding site disappears in line-D, diminishing then the thermal stability (in ALiP and PC-LIPA). The higher stability in line-B, and in more ancestral line-D enzymes, is in agreement with Mn^{2+} contribution to cofactor binding.

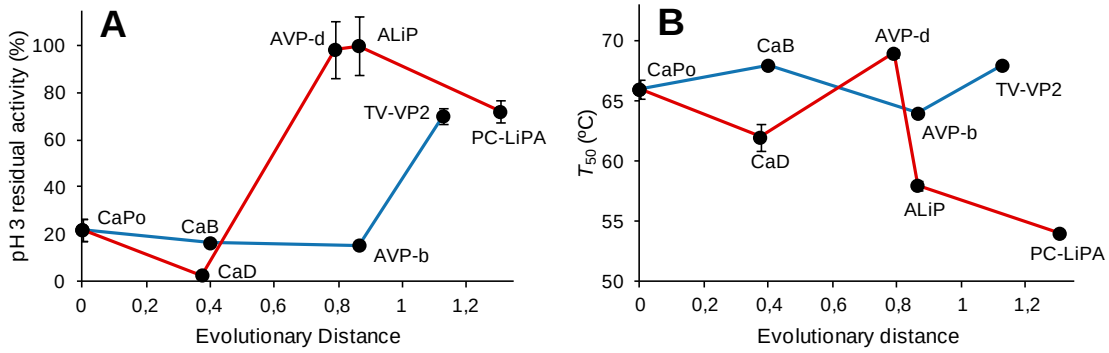


Fig. 5. pH and thermal stabilities in the D (red line) and B (blue line) evolutionary pathways. **A)** Changes in residual activity after incubation at pH 3. **B)** Changes in T_{50}. Means and 95% confidence limits are shown.

Discussion

The main evidence about the basidiomycete enzymes involved in lignin degradation comes from the genomic information available in the last years. Every study shows the presence of LiP, MnP or VP genes in the sequenced genomes of all typical white-rot (lignin-degrading) fungi, and their absence from all typical brown-rot (cellulose-degrading) fungi and some poor wood rotters (Floudas et al. 2012; Ruiz-Dueñas et al. 2013; Riley et al. 2014; Nagy et al. 2016).

Here, we analyzed the appearance and subsequent evolution of phylogenetically-distant LiP and VP type genes within the evolution of Polyporales, resulting in the most efficient ligninolytic enzymes. Although the first class-II fungal peroxidase(s) most probably appeared by horizontal transfer of a prokaryotic peroxidase (such as cytochrome *c* peroxidase) gene from an ancestral organelle (Passardi et al. 2007), no evidence for subsequent horizontal transfer in Polyporales was obtained by BLAST searches (Zhaxybayeva 2009) in agreement with the very rare nature of such events in basidiomycetes (Fitzpatrick 2012). However, evolutionary

clues on the presence of the same lignin-degradation mechanism in distant Polyporales peroxidases could be obtained by ancestral sequence reconstruction and characterization of the resurrected enzymes.

For sequence reconstruction, we used maximum likelihood (ML) methods that have some advantages over other approaches (Omland 1999). To deal with the inherent limitations and uncertainties in ancestral reconstructions (Royer-Carenzi et al. 2013), *(i)* we used the best data available, i.e. all the class-II peroxidase genes in ten Polyporales genomes after their careful revision and manual annotation (Ruiz-Dueñas et al. 2013), and *(ii)* we verified that no ambiguity exists in the amino acids forming the different oxidation sites (Liberles 2007). The reconstructed sequences revealed that the appearance of the surface tryptophan abstracting electrons from lignin (Sáez-Jiménez et al. 2015; 2016) was not an isolated event in the evolution in Polyporales. Moreover, the biochemical characterization of the resurrected ancestral peroxidases enabled us to confirm their predicted new catalytic properties and revealed how they progressively changed in the two evolutionary lines analyzed, as discussed below.

First of all, the experimental characterization of the resurrected enzymes showed the evolvability (Colegrave and Collins 2008) of fungal peroxidases in the exploration of new mechanisms to modify lignin, at different points of their phylogeny. The common ancestor of Polyporales peroxidases (CaPo), was most probably a short MnP that used Mn^{3+} to attack the phenolic moiety of lignin and other phenolic molecules, acting synergistically with secreted oxidases and intracellular oxidoreductases. By comparison to analyses with a broader sampling of genomes (although including less Polyporales species) (Floudas et al. 2012), CaPo would correspond with the common ancestor of all Agaricomycetes class-II ligninolytic peroxidases, not just those of Polyporales, and is estimated to have existed roughly 400 mya (results from molecular clock analyses with fossil calibration). Note that the common ancestor of Polyporales fungi, appearing at the early Cretaceous, would already had several (3-13) peroxidase genes (ligninolytic and GP included) (Floudas et al. 2012). Therefore, the higher expansion and specialization of peroxidases would post-date the Carboniferous, although the fungal capacity to degrade lignin would occur earlier using ancestral peroxidases (like CaPo) oxidizing the minor phenolic moiety of lignin and phenolic ancestral polymers. After CaPo, the common ancestors of clades D (CaD) and B (CaB) would have almost identical properties (both in activity and stability). This includes an increase in the efficiency of oxidizing Mn^{2+} that reveals a similar initial degradative strategy in the two branches, using Mn^{3+} chelates. Reconstruction of these old peroxidases would be more uncertain than for the three more recent ancestors, which already appeared within Polyporales. The oxidation site for high redox-potential substrates

appeared in both lines, and VA (the laboratory model substrate for lignin degradation studies) was oxidized by the resurrected enzymes. MnPs would be efficient degrading ancestral phenolic polymers and phenolic lignin in plants, but MnP's action on nonphenolic lignin would require the concerted action of LiPs/VPs. In this way, ancient fungi would incorporate a powerful tool to its degradative machinery.

After the catalytic tryptophan appearance, we observed an increase in the peroxidase efficiency oxidizing VA in both branches, but the kinetic parameters and the evolution of other oxidation sites were different. Evolution in line-D focused on a better degradation of lignin by removing other oxidation sites (at expenses of the stability conferred by the Mn^{2+} binding site) and maintaining the surface tryptophan, with a progressively more acidic environment (that increases the tryptophanyl radical reactivity). However, VPs in clade-B maintain both oxidation sites. Although VPs could be seen as mere evolutionary intermediates, as found in line-D, a significant improvement in the oxidation of VA was observed in both branches after the appearance of the catalytic tryptophan. In addition to the changes in architecture and activity of the oxidation sites, we also observed an increase with evolution in the peroxidase stability under the acidic conditions that characterize the hyphal microenvironment where ligninolysis takes place in nature (Martínez 2002). In clade-D, the appearance of the new oxidation site comes along with a huge increase in acidic stability, but in clade-B this stability is acquired later. Either way, there is a clear improvement towards the stabilization at pH 3, where the oxidizing power of these enzymes is the highest (Millis et al. 1989; Oyadomari et al. 2003).

The above evolutionary trend, which results in more efficient oxidation of lignin, was most probably related to changes in plant cell wall and tissue anatomy. Despite the evolutionary history of lignins remaining unclear, there have been significant changes in their composition and structure including convergent evolution between different vascular plants (Weng and Chapple 2010; Novo-Uzal et al. 2012). Angiosperms are the plants with the more complex lignin (Ralph et al. 2004; Martínez et al. 2008) including a higher relative abundance of syringyl units with the C_3 and C_5 positions of the aromatic ring blocked by methoxyls, compared with most gymnosperms, that results in a predominance of nonphenolic (C_4-etherified) units (Camarero et al. 1999a). The angiosperm appearance (140-250 mya) (Magallón et al. 2015) roughly corresponded with the age of the two most recent ancestors of major clades B and D of ligninolytic peroxidases in Polyporales (~200 mya) that subsequently incorporated the exposed catalytic tryptophan almost at the same time (Floudas et al. 2012). This evolutionary event resulted in the most efficient peroxidases that oxidize

nonphenolic lignin by long-range electron transfer from the protein surface, as shown using methylated lignin (Sáez-Jiménez et al. 2016). Interestingly, a similar electron transfer mechanism has been developed by plant peroxidases involved in lignin polymerization, with the appearance of an enzyme being able to oxidize the bulkier (dimethoxylated) sinapyl alcohol monolignol characterizing angiosperm lignin, at a surface aromatic residue (Shigeto et al. 2012).

Materials and methods

The 113 sequences of class-II peroxidases in the genomes of *B. adusta*, *C. subvermispora*, *D. squalens*, *F. pinicola*, *Ganoderma* sp., *P. brevispora*, *P. chrysosporium*, *P. placenta*, *T. versicolor* and *W. cocos* (see **Dataset S1**) available at the DOE JGI were used in this study. ML phylogeny was constructed with RAxML (Stamatakis et al. 2008) and PAML 4.7 (Yang 2007) was used to obtain the most probable ancestral sequences that were manually corrected for insertions or deletions, and synthesized for *E. coli* expression. Molecular models were obtained at the Swiss-Model server (Biasini et al. 2014). The coding DNA sequences of ancestral and extant peroxidases were cloned and used to transform *E. coli*. The apoenzymes were recovered from inclusion bodies, *in vitro* activated and purified. Mn^{2+}, VA, ABTS, DMP and RB5 were used for kinetic characterization at the optimal pH and H_2O_2 concentrations (see **Table S2** footnote *a*). For pH stability, the peroxidases were incubated at pH 2-10 for 4 h, and activity was estimated with ABTS. For thermal stability, the enzymes were incubated for 10 min at 25-85 ºC (pH 5.5) to obtain T_{50} values. Circular dichroism was used to obtain T_m values. See **SI Appendix** for details.

ACKNOWLEDGEMENTS. This work was supported by the IndOx (KBBE-2013-613549, www.indoxproject.eu) and EnzOx2 (H2020-BBI-PPP-2015-2-720297, www.enzox2.eu) EU projects and the BIO2014-56388-R and BIO2017-86559-R Spanish projects. The work conducted by JGI was supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231. I.A.-F. thanks a Spanish FPI Fellowship.

Suporting Information

Evolutionary convergence in lignin degrading enzymes

The SI Appendix includes Supporting Materials and Methods, Supporting Figs. S1-S6 and Tables S1 and S2 and Supporting References.

Supporting Materials and Methods

Phylogenetic analysis, ancestral enzyme reconstruction and molecular modeling. Every class-II peroxidase (113 sequences) annotated in the genomes of ten Polyporales (phylum Basidiomycota) species (namely *B. adusta*, *C. subvermispora*, *D. squalens*, *F. pinicola*, *Ganoderma* sp., *P. brevispora*, *P. chrysosporium*, *P. placenta*, *T. versicolor* and *W. cocos*, see file S1) were used in this study, being available at the DOE JGI Mycocosm portal as http://genome.jgi.doe.gov/Bjead1_1/Bjead1_1.home.html (Binder et al. 2013), http://genome.jgi.doe.gov/Cersu1/Cersu1.home.html (Fernández-Fueyo et al. 2012), http://genome.jgi.doe.gov/Dicsq1/Dicsq1.home.html (Floudas et al. 2012), http://genome.jgi.doe.gov/Fompi3/Fompi3.home.html (Floudas et al. 2012), http://genome.jgi.doe.gov/Gansp1/Gansp1.home.html (Binder et al. 2013), http://genome.jgi.doe.gov/Phlbr1/Phlbr1.home.html (Binder et al. 2013), http://genome.jgi.doe.gov/Phchr2/Phchr2.home.html (Ohm et al. 2014), http://genome.jgi.doe.gov/Pospl1/Pospl1.home.html (Martinez et al. 2009), http://genome.jgi.doe.gov/Trave1/Trave1.home.html (Floudas et al. 2012) and http://genome.jgi.doe.gov/Wolco1/Wolco1.home.html (Floudas et al. 2012), respectively. BLAST (Basic Local Alignment Search Tool) was used for the search of sequences similar to line-D LiPs and line-B VPs at the NCBI databases (with Polyporales genes included/excluded).

The amino-acid sequences were aligned using MUSCLE as implemented in MEGA 7 (Kumar et al. 2016b). The sequences were tested using ProtTest (Abascal et al. 2005) to determine the evolutionary model that best fits the data for the ML analysis among 64 empirical models of evolution. The ML phylogeny was then constructed using RAxML (Stamatakis et al. 2008), under the Whelan and Goldman (2001) model of evolution using gamma-distributed rate of heterogeneity (gamma shape with 4 rates of categories = 1.33) with empirical amino-acid frequencies and invariant sites (proportion of invariant sites = 0.073) (WAG+I+G+F).

PAML 4.7 package (Yang 2007) was used to obtain the most probable sequence at each node of the phylogeny and the posterior amino-acid

probability per site in each ancestor. We used the WAG model of evolution, and the previously obtained ML phylogeny and the MUSCLE alignment as inputs for the software. The most probable sequences of the marginal reconstruction were selected and manually corrected for C-terminal and other insertions or deletions according to the sequences of the ancestor progeny. The genes were synthesized by ATG:biosynthetics (Merzhausen, Germany) using the most frequent codons for high expression in *E. coli*.

Molecular models of the predicted proteins were obtained at the Swiss-Model protein homology modeling server (Guex et al. 2009; Biasini et al. 2014) using related crystal structures, selected using the GMQE parameter, as templates (PDB entries 4BM1, 3FJW, 1QPA and 1B85 corresponding to *Pleurotus ostreatus* MnP4, *Pleurotus eryngii* VPL and *P. chrysosporium* LiPD and LiPA, respectively). All protein models had great quality taking into account the Swiss-Model parameters (good QMEAN and high GMQE). The electrostatic surfaces were computed with PyMOL v1.8 (Schrödinger LLC, http://pymol.org) using default parameters.

***E. coli* expression and characterization of the resurrected enzymes.** After gene synthesis, the coding DNA sequences of the selected ancestors (CaPo, CaB, CaD, AVP-d, AVP-b and ALiP) and the extant PC-LiPA (JGI ID# 2989894) and TV-VP2 (JGI ID# 29236) were cloned into pET23b(+) (Novagen). The resulting plasmids were transformed into BL21(DE3)pLysS (ancestors and extant TV-VP2) or W3110 (PC-LiPA) as reported by Pérez-Boada et al (2002). The apoenzymes accumulated in inclusion bodies and, after solubilization in 8 M urea, the *in vitro* activation conditions of the ancestral proteins and extant TV-VP2 were optimized - including 0.16 M urea, 5 mM $CaCl_2$, 15 μM hemin, 0.4 mM oxidized glutathione, 0.1 mM dithiothreitol and 0.1 mg/mL of protein in 50 mM Tris-HCl, pH 9.5 - while those reported by Doyle and Smith (1996) where used for PC-LiPA. The active enzymes were purified with a Resource-Q column (GE-Healthcare) using a 0-400 mM NaCl salt gradient, 2 mL/min flow, in 10 mM sodium tartrate, pH 5.5, containing 1 mM of $CaCl_2$.

Five substrates were selected for the kinetic characterization of the resurrected peroxidases: Mn^{2+} (Mn^{3+}-tartrate complex ε_{238} 6500 $M^{-1}\cdot cm^{-1}$), VA (veratraldehyde ε_{310} 9300 $M^{-1}\cdot cm^{-1}$), ABTS (cation radical ε_{436} 29300 $M^{-1}\cdot cm^{-1}$), DMP (dimeric coerulignone ε_{469} 55000 $M^{-1}\cdot cm^{-1}$), and RB5 (ε_{598} 30000 $M^{-1}\cdot cm^{-1}$). These reactions were analyzed at 25 ºC and the optimal pH (determined using 50 mM Britton-Robinson, B&R, buffer, pH 2-10) and H_2O_2 concentration (determined using 2.5 mM ABTS as substrate) for each enzyme (see **Table S2** footnote).

To study the pH stability, the resurrected peroxidases were incubated in B&R buffer, pH 2-10, at 25 ºC for 4 h. The residual activity was estimated

by the oxidation of ABTS (2.5 mM), with the activity after 1 min at pH 5.5 (25 ºC) taken as 100%. To study thermal stability, the enzymes were incubated in 10 mM acetate, pH 5.5, for 10 min in the range 25-85 ºC, to obtain the T_{50} values (i.e. the temperature at which 50% of the activity was lost after 10 min incubation). The effect of temperature on enzyme (6 μM) circular dichroism spectra was addressed studying the changes at 222 nm from 20ºC to 95ºC, 30ºC/h, using a Jasco J-815 spectropolarimeter with a 0.01 cm path length quartz cell. T_m represents the temperature at the midpoint of the unfolding transition in the thermal melting profiles.

Supporting Figures and Tables

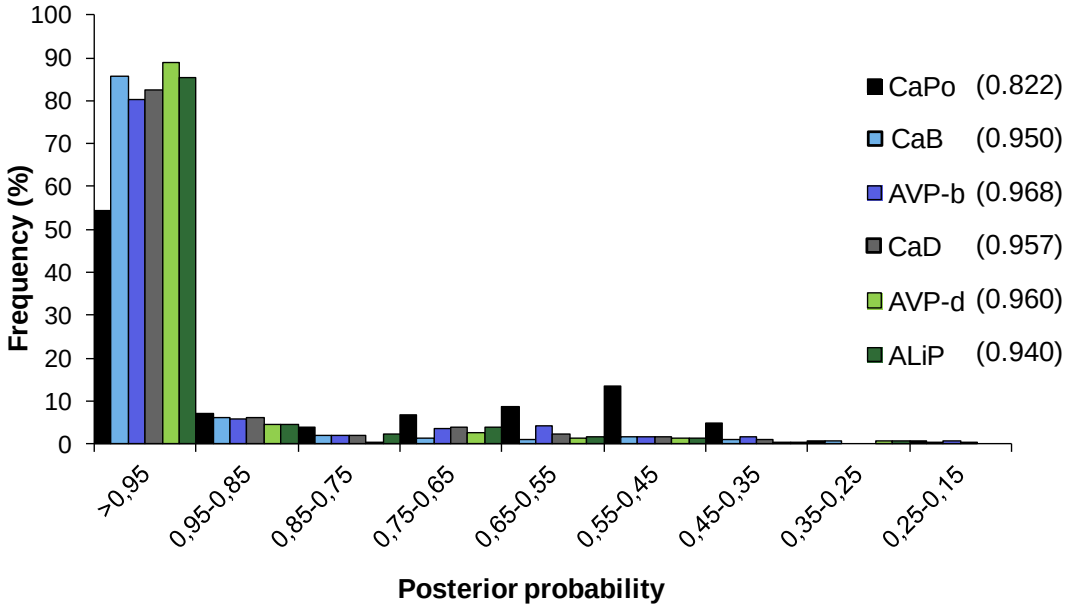


Fig. S1. Frequency distribution of amino acids with different posterior probabilities (i.e. the probability that a hypothesis is true after the data have been analyzed) for the six reconstructed ancestral sequences (**Fig. S2**) from the evolutionary lines B and D (**Fig. 1**). Most of the amino acids for each sequence have a posterior probability >0.95, showing the robustness of the sequence reconstruction. CaPo has a wider distribution, while for the other ancestors 80-90% of the predicted amino acids have a probability >0.95. The values in parenthesis represent the average probability for the complete sequences.

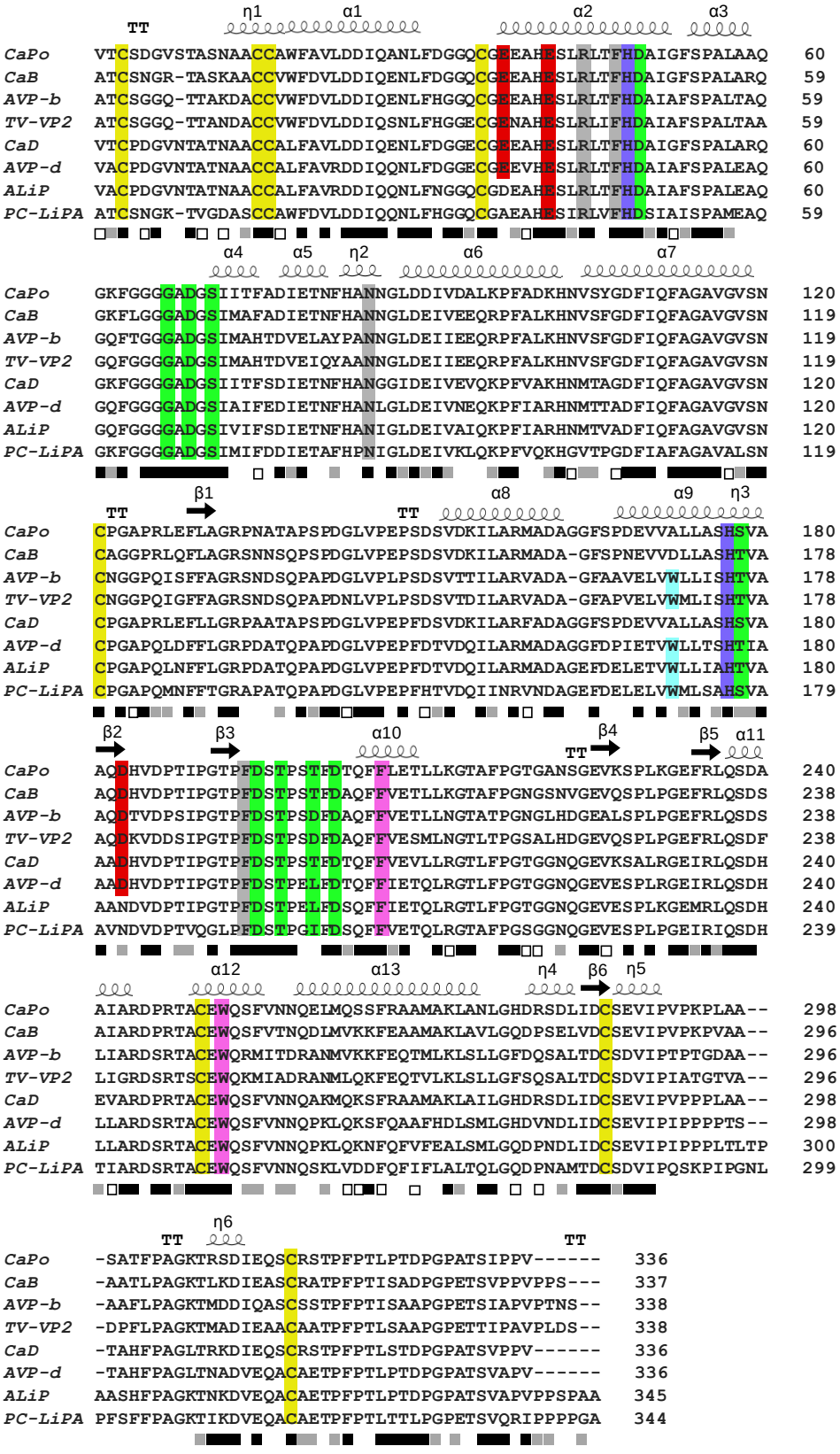


Fig. S2. Alignment of ancestral and extant (PC-LiPA and TV-VP2) sequences, with indication of: catalytic tryptophan (cyan), electron-transfer residues (pink), Mn^{2+}-binding site (red), active-site histidines (purple), other active-site residues (gray), Ca^{2+} ligands (green) and disulfide bonds (yellow) (Ruiz-Dueñas et al. 2009). Boxes (bottom) indicate full conservation of the same (black) or equivalent (gray) residues, and partial conservation (white). Secondary structure (top) includes: α, α-helices; η, 3_{10} helices; ß, ß-strands; and TT, strict ß –turn.

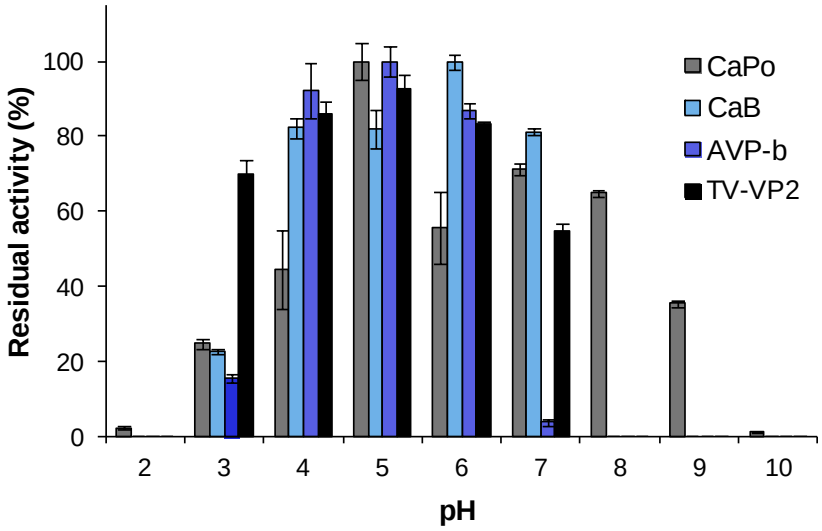


FIGURE S3. pH stability of line-B ancestral enzymes. Stability in the range of pH 2-10 was compared after 4-h incubation at 25ºC. Means and 95% confidence limits are shown.

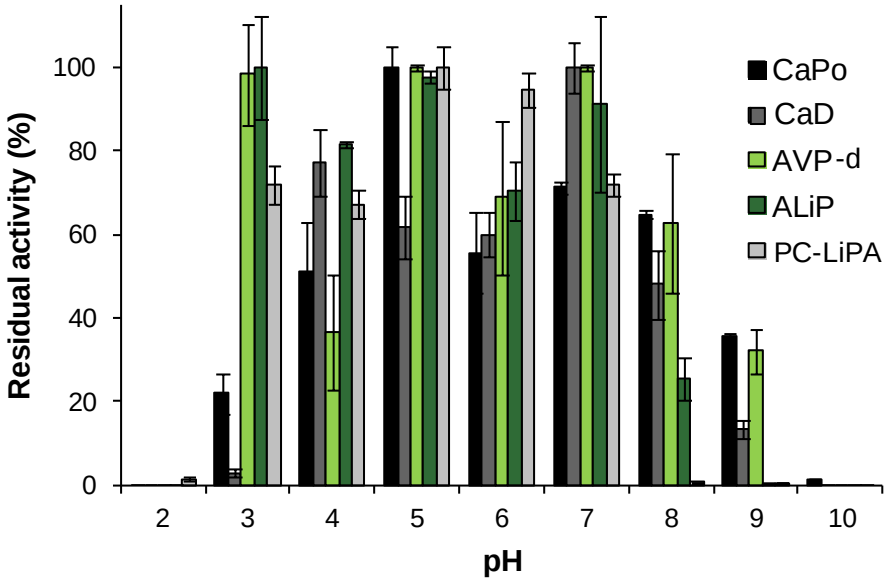


FIGURE S4. pH stability of line-D ancestral enzymes. Stability in the range of pH 2-10 was compared after 4-h incubation at 25ºC. Means and 95% confidence limits are shown. Adapted from Ayuso-Fernández et al. (2017).

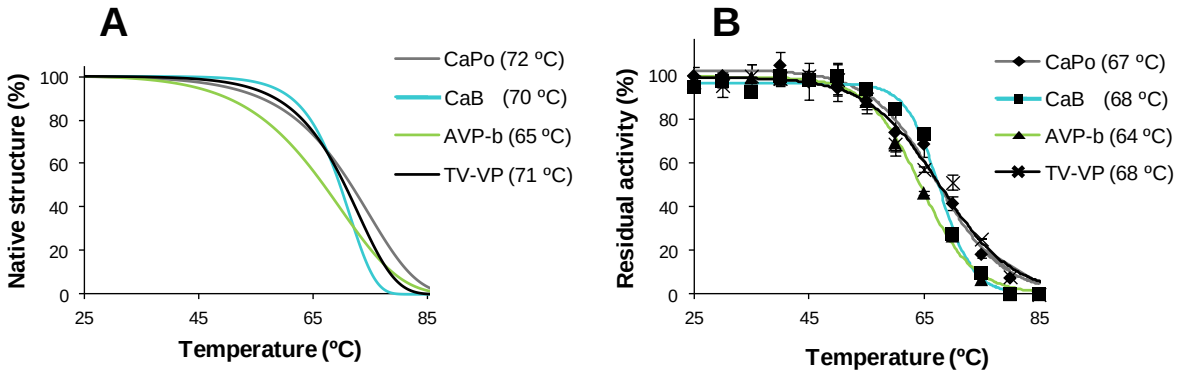


Fig. S5. Thermal stability of line-B ancestral peroxidases. **A)** Loss of secondary structure. **B)** Loss of activity. The T_m (**A**) and T_{50} (**B**) values are shown in the legends. Means and 95% confidence limits are shown in **B**.

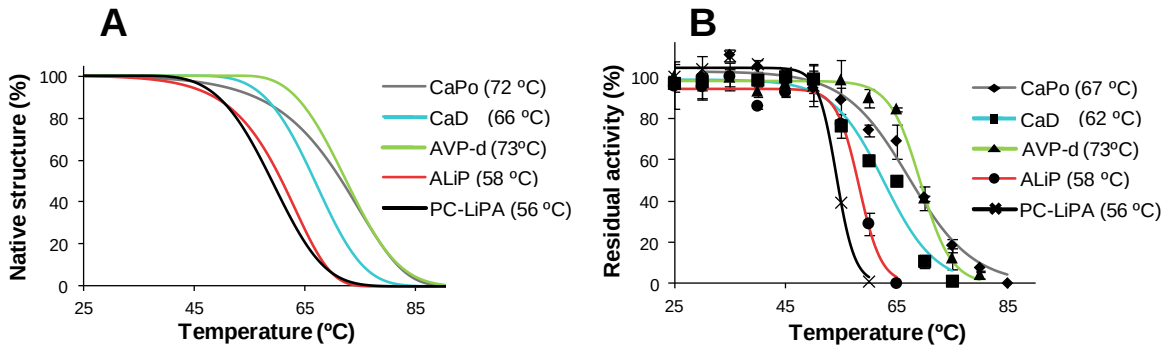


Fig. S6. Thermal stability of line-D ancestral peroxidases. **A)** Loss of secondary structure. **B)** Loss of activity. The T_m (**A**) and T_{50} (**B**) values are shown in the legends. Means and 95% confidence limits are shown in **B**. Adapted from Ayuso-Fernández et al. (2017).

TABLE S1. Pairwise identity percentages between ancestral and extant (TV-VP2 and PC-LiPA) peroxidase sequences (the number of aligned residues are in parentheses)

	CaPo	CaB	AVP-b	TV-VP2	CaD	AVP-d	ALiP	PC-LiPA
CaPo	100 (336)	81 (275)	66 (223)	61 (208)	86 (289)	74 (248)	70 (242)	61 (209)
CaB		100 (337)	76 (257)	69 (233)	76 (259)	67 (226)	66 (227)	61 (211)
AVP-b			100 (338)	85 (288)	62 (211)	60 (205)	60 (207)	58 (199)
TV-VP2				100 (338)	59 (199)	59 (200)	57 (198)	56 (193)
CaD					100 (336)	80 (269)	75 (257)	61 (211)
AVP-d						100 (336)	87 (301)	66 (227)
ALiP							100 (345)	72 (249)
PC-LiPA								100 (344)

TABLE S2. Kinetic parameters - K_m (μM), k_{cat} (s^{-1}) and k_{cat}/K_m ($s^{-1} \cdot mM^{-1}$) - for oxidation of Mn^{2+}, DMP, ABTS, RB5 and VA by resurrected peroxidases from evolutionary lines B and D (**fig. 1**) and extant TV-VP2 and PC-LiPA.[a] See **Fig. 3** for graphical comparison of the peroxidase kinetic constants.

		Ancestral	Line-B (to VP)			Line-D (to LiP)[b]			
		CaPo[b]	CaB	AVP-b	TV-VP2	CaD	AVP-d	ALiP	PC-LiPA
Mn^{2+}	K_m	700 ± 48	132 ± 12	396 ± 82	3150 ± 550	275 ± 41	62 ± 10	-[c]	-
	k_{cat}	185 ± 3	121 ± 2	4 ± 0	11 ± 1	170 ± 6	106 ± 4	-	-
	k_{cat}/K_m	260±15	919 ± 73	9 ± 1	4 ± 1	617 ± 80	1710 ± 240	-	-
DMP[d]	K_m	32900 ± 2700	635000 ± 27000	1400 ± 170	2300 ± 260	66800 ± 3800	32500 ± 12100	-	-
low	k_{cat}	221 ± 9	404 ± 162	48 ± 1.7	130 ± 6	109 ± 4	31 ± 5	-	-
efficiency	k_{cat}/K_m	6.7 ± 0.3	0.6 ± 0.2	34.0 ± 3.0	56.0 ± 7.0	1.6 ± 0.0	1.0 ± 0.2	-	-
ABTS[d]	K_m	3170 ± 270	1280 ± 350	314 ± 68	580 ± 80	1280 ± 350	2150 ± 420	-	-
low	k_{cat}	539 ± 24	103 ± 10	163 ± 11	352 ± 13	103 ± 10	25 ± 2	-	-
efficiency	k_{cat}/K_m	170 ± 8	80 ± 10	500 ± 9	610 ± 70	80 ± 15	12 ± 1	-	-
DMP	K_m	-	-	2.6 ± 0.6	8.9 ± 1.1	-	5.3 ± 1.1	34.0 ± 5.4	4.0 ± 0.1
high	k_{cat}	-	-	4.4 ± 0.2	11.4 ± 0.4	-	4.5 ± 0.1	18.3 ± 0.7	6.9 ± 0.5
efficiency	k_{cat}/K_m	-	-	1700 ± 300	1300 ± 100	-	837 ± 162	537 ± 55	600 ± 36
ABTS	K_m	-	-	48 ± 11	103 ± 36	-	5 ± 1	14 ± 4	21 ± 2
high	k_{cat}	-	-	89 ± 8.7	128 ± 44	-	2 ± 0	13 ± 1	7 ± 0
efficiency	k_{cat}/K_m	-	-	1800 ± 300	2100 ± 300	-	400 ± 44	911 ± 212	300 ± 25
RB5	K_m	-	-	3.0 ± 0.4	2.8 ± 0.2	-	4.8 ± 0.8	12.6 ± 3.6	-
	k_{cat}	-	-	8.9 ± 0.5	8.6 ± 0.3	-	2.4 ± 0.2	5.4 ± 0.8	-
	k_{cat}/K_m	-	-	2900 ± 160	3100 ± 180	-	504 ± 50	428 ± 68	-
VA	K_m	-	-	10000 ± 100	24600 ± 200	-	299 ± 104	773 ± 155	79.3 ± 18
	k_{cat}	-	-	41 ± 1	172 ± 6.4	-	7 ± 1	21 ± 1	16 ± 1
	k_{cat}/K_m	-	-	4 ± 0	7 ± 0	-	24 ± 7	28 ± 5	205 ± 4

[a]Reactions at 25 °C in 0.1 M tartrate at optimal pH 5.0 (CaPo and CaD) or 5.5 (CaB, AVP-b, VP-d and TV-VP2) for Mn^{2+}, pH 3.5 (AVP-d), 3.0 (AVP-b, ALiP and PC-LiPA) or 2.0 (TV-VP2) for VA, pH 2 (CaPo and TV-VP2) or 3.0 (others) for DMP, pH 2 (TV-VP2), 3.0 (CaPo, CaB and AVP-b) or 3.5 (others) for ABTS, and pH 3 for RB5; and saturating H_2O_2 concentrations of 0.4 mM (CaPo, CaD, CaB, AVP-b, AVP-d and TV-VP2), 0.2 mM (ALiP) or 0.1 mM (PC-LiPA). [b]From Ayuso-Fernández et al. (2017). [c]–, Absence of activityc. [d]Biphasic kinetics for DMP and ABTS oxidation by AVP-b, AVP-d and TV-VP2 enabled calculation of two sets of constants assigned to two catalytic sites, as reported for *P. eryngii* VP (Morales et al. 2012). Means and 95% confidence limits are provided.

Chapter 3

Redox potential increased during the evolution of enzymes degrading recalcitrant lignin

General abstract

One of the most intriguing properties of ligninolytic peroxidases is their uniquely high reduction potential. This feature allows them to oxidize high redox-potential substrates, including the complex and recalcitrant lignin polymer. To study the structural features that tune the reduction potential of peroxidases, the literature has focused in its majority in the measurement of the Fe^{2+}/Fe^{3+} pair due to its simplicity. However, since Fe^{2+} is not involved in the catalytic cycle, this information has to be taken carefully.

In this chapter we analyze the changes in reduction potential in ligninolytic peroxidase ancestors to investigate how they tuned this property during evolution. We used stopped-flow spectrophotometry to measure the reduction potentials of all the catalytic species in the enzymes leading to extant LiPs. Together with protein NMR, we explore possible structural changes in the heme environment responsible of the values obtained.

The results presented in this chapter show how ancestral peroxidases evolved from ancestors with lower reduction potentials and, during evolution, a boost in their reduction potentials correlated with structural changes in the proximal side led to the high redox-potential enzymes that we observe today.

Article published as:

Ayuso-Fernández, I., López De Lacey, A., Cañada, FJ., Ruiz-Dueñas, FJ., Martínez, AT. *Increase of Redox Potential during the Evolution of Enzymes Degrading Recalcitrant Lignin*. **Chemistry - A European Journal**. 2019. 25:11, 2708-2712. IF: 5.161. doi:10.1002/chem.201805679.

F.J.R.-D. and A.T.M. designed research. I.A.-F. performed research. F.J.C measured and analyzed the NMR spectra. A.L. de L. helped with Fe^{3+}/Fe^{2+} measurements and gave thoughtful advice through the work. I.A.-F., F.J.R.-D., and A.T.M. wrote the paper. All authors contributed to data analysis and discussion of the results, and read, revised and approved the manuscript.

Abstract
To investigate how ligninolytic peroxidases acquired the uniquely high redox-potential they show today, their ancestors were resurrected and characterized. Unfortunately, the transient Compounds I (CI) and II (CII) from peroxide activation of the enzyme resting state (RS) are unstable. Therefore, the reduction potentials ($E^{o'}$) of the three redox-couples (CI/RS, CI/CII and CII/RS) were estimated (for the first time in a ligninolytic peroxidase) from equilibria concentrations analyzed by stopped-flow UV/vis spectroscopy. Interestingly, the $E^{o'}$ of rate-limiting CII reduction to RS increased 70 mV from the common peroxidase ancestor to extant lignin peroxidase (LiP), and the same boost was observed for CI/RS and CI/CII, although with higher $E^{o'}$ values. A straightforward correlation was found between the $E^{o'}$ and the progressive displacement of the proximal histidine Hε1 shift in NMR spectra, due to higher paramagnetic effect of the heme Fe3+. More interestingly, the $E^{o'}$ and NMR data also correlate with the evolutionary time, revealing that ancestral peroxidases increased their reduction potential in the evolution to LiP thanks to molecular rearrangements in their heme pocket during the last 400 my.

Results & Discussion
The use of paleogenetics to investigate protein properties in evolution is a developing field that has been used to prove different evolutionary hypotheses (Gumulya and Gillam 2017). Using a robust phylogeny and the ancestral sequence reconstruction tools, the time travel investigation of any family of proteins is virtually possible, even to the early stages of Earth.(Risso et al. 2013) Lignin peroxidases (LiP), versatile peroxidases (VP) and manganese peroxidases (MnP) (Ruiz-Dueñas et al. 2009) belong to the peroxidase-catalase superfamily (Zámocký et al. 2015). These fungal heme-proteins are characterized by a high reduction potential enabling oxidation of the recalcitrant (nonphenolic) lignin polymer formed by the dehydrogenative polymerization of phenolic monolignols by low redox-potential plant peroxidases (Battistuzzi et al. 2010a). Lignin biodegradation, which was defined as an enzymatic combustion (Kirk and Farrell 1987), is a key step for carbon recycling in land ecosystems and a process of biotechnological interest for the use of plant biomass in a bio-based economy (Martínez et al. 2009; Ragauskas et al. 2014).

The evolutionary changes in ligninolytic peroxidases have been recently investigated (Ayuso-Fernández et al. 2017; Ayuso-Fernández et al. 2018) by resurrection of ancestral enzymes from sequenced genomes of Polyporales (Ruiz-Dueñas et al. 2013), where most lignin-degrading fungi are included. In this way, the transition from an ancestor that oxidized lignin poorly using diffusible Mn^{3+} chelates, around 400 million years ago (Floudas et al.

2012), into more efficient enzymes that oxidized lignin directly was demonstrated (Ayuso-Fernández et al. 2017) . That powerful degradative strategy was acquired when a solvent-exposed tryptophanyl radical appeared in an ancestral VP, and became more efficient in ancestral and extant LiPs. The latter include LiP-H8 from the model fungus *Phanerochaete chrysosporium* (Hammel and Cullen 2008), corresponding to LiPA (PC-LiPA) in the sequenced genome (Martinez et al. 2004), whose molecular structure is shown in **Figure 1**. Direct degradation of lignin emerged independently several times in peroxidase evolution (Ayuso-Fernández et al. 2018), showing the importance of lignin decay during land colonization by vascular plants (Floudas et al. 2012). To elucidate if ligninolytic peroxidases (i) acquired their oxidizing power with evolution or (ii) they were already high redox-potential enzymes in ancestral Polyporales peroxidases and later shaped their oxidation sites, we analyze here the evolutionary lineage leading to PC-LiPA (supplementary Figure S1) by ancestral sequence reconstruction from the genomes of ten Polyporales species (Ruiz-Dueñas et al. 2013), using the PAML software (Yang 2007). The most relevant ancestral sequences in the PC-LiPA lineage were "resurrected", corresponding to: i) Common ancestor of Polyporales peroxidases (CaPo); ii) Common ancestor of clade D peroxidases (CaD); iii) First ancestral VP in clade D (AVPd); and iv) First ancestral LiP (ALiP) (multiple alignment in supplementary Figure S2). The four ancestral peroxidases and the extant LiPA were characterized, including reduction-potential estimation for the different redox pairs in the catalytic cycle.

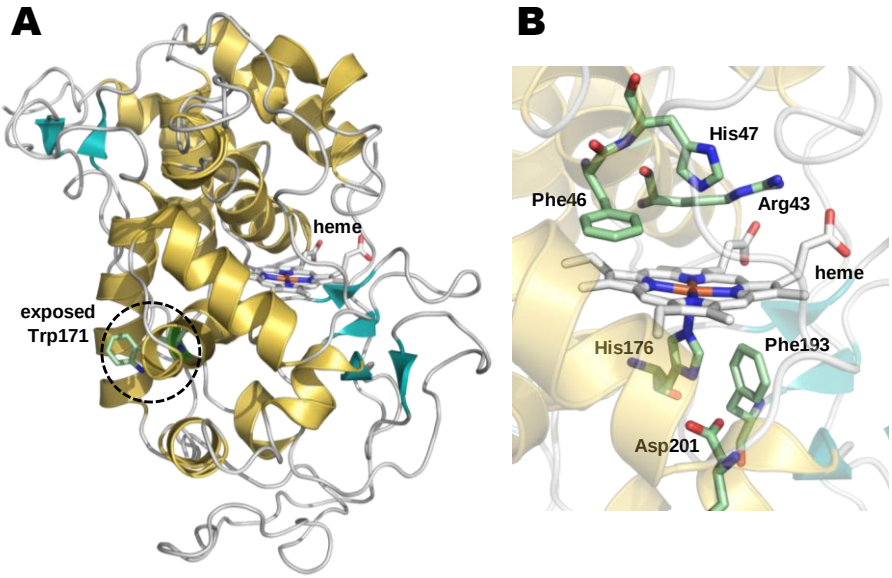


Figure 1. General model of *P. chrysosporium* LiP showing the secondary structure of the protein, the buried heme cofactor and the catalytic tryptophan exposed to the solvent (**A**), and detail of the heme pocket with two axial histidines and other residues (**B**) (from PDB 1LGA).

To understand the evolution of the peroxidase redox properties, the three steps of their catalytic cycle (Poulos 2014) are analyzed (**Figure 2**). The initial reaction is the oxidation of the resting state (RS) enzyme by H_2O_2, extracting two electrons from the cofactor. This generates Compound I (CI) with two oxidizing equivalents (as Fe^{4+}=O and porphyrin cation radical complex). Alternatively, CI of LiP/VP transfers one of these equivalents forming the above mentioned solvent-exposed tryptophanyl radical directly oxidizing the bulky lignin polymer (Sáez-Jiménez et al. 2015; Sáez-Jiménez et al. 2016), which is unable to access the buried cofactor (Ruiz-Dueñas and Martínez 2009). CI is reduced to Compound II (CII) during oxidation of a substrate molecule. CII has one oxidation equivalent (Fe^{4+}=O, in equilibrium with the tryptophanyl radical in LiP/VP), and is further reduced to the RS by oxidation of another molecule of substrate.

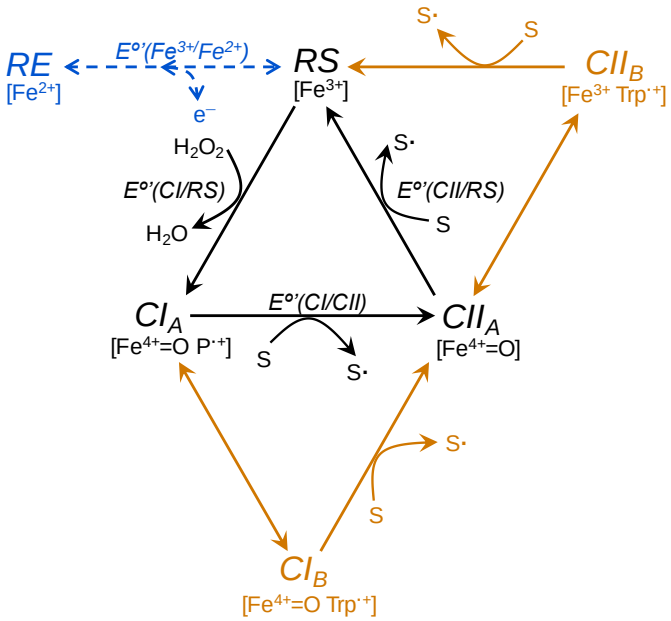


Figure 2. Ligninolytic peroxidase catalytic cycle. In black is the common cycle initiated with oxidation of the RS enzyme by peroxide, followed by two substrate (S) one-electron oxidations (by CI and CII). In orange is the expanded cycle of VPs and LiPs, with formation of a tryptophanyl radical. The blue dashed line to ferrous reduced enzyme (RE-Fe^{2+}) does not form part of the cycle, but has been studied here. The formal reduction potentials ($E^{o'}$) analyzed are indicated. Adapted from Sáez-Jiménez et al (2015).

The most common redox measurement in heme proteins, basidiomycete peroxidases included (Millis et al. 1989; Whitwam et al. 1999; Santucci et al. 2000; Ciaccio et al. 2003; Oyadomari et al. 2003; Battistuzzi et al. 2006), is the midpoint potential of the ferric/ferrous transition (supplementary Figures S3 and S4), even though it is not part of the catalytic cycle (Figure 2). To explore the mechanistic implications of reduction potential in the peroxidase catalytic cycle, we used stopped-flow spectrophotometry (Sørlie

et al. 2000; Arnhold et al. 2001; Furtmüller et al. 2003; Furtmüller et al. 2005) to measure the concentration of the oxidized and reduced forms of the different enzymes in their RS/CI and CII/RS transitions (supplementary Figures S5 and S6, respectively) as illustrated in **Figure 3** for the most recent ancestor. Stopped-flow is used here to assign the equilibrium concentrations of the two redox states of the protein and the two redox states of the substrates used providing an equilibrium constant that, with the use of the Nerst equation, will allow the determination of the midpoint potential. In this way, the reduction potentials of the catalytic iron couples (those of CI/RS and CII/RS directly, and the CI/CII E$^{o'}$ by difference) were calculated for the first time in ligninolytic peroxidases (see supplementary Material and Methods and Tables for details). Analysis of the three half-reactions revealed a general boost of the reduction potentials with evolution (**Figure 4**). For all the redox pairs (CI/RS, CI/CII and CII/RS) we obtained differences of approximately 70 mV between the oldest ancestor (CaPo) and ALiP, which had the highest reduction potentials in evolution (close to the values of extant PC-LiPA).

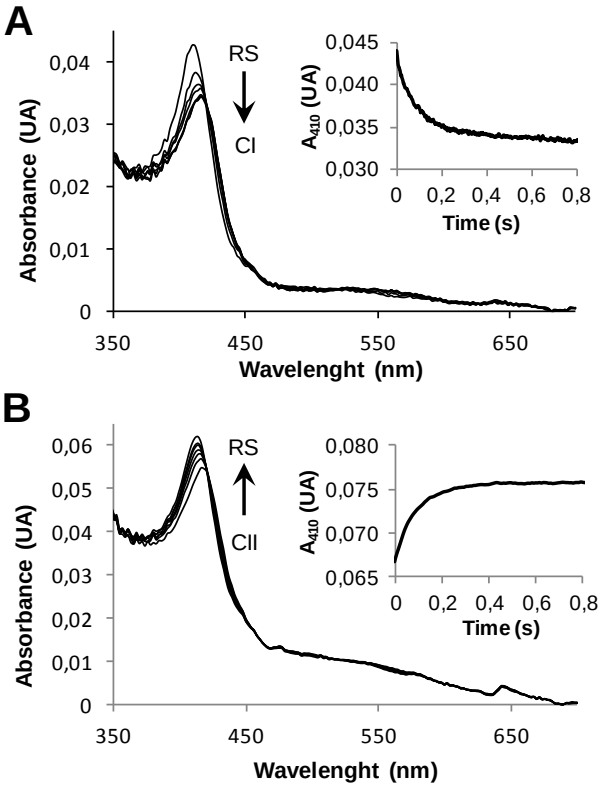


Figure 3. **(A)** Spectral changes upon rapid mixing of ALiP RS with H_2O_2 to follow CI formation (from 1.6 to 800 ms after peroxide addition). **(B)** Spectral changes during Tyr reduction of CII, formed by enzyme mixing with H_2O_2 and ferrocyanide (from 1.6 to 800 ms after Tyr addition). The insets show time traces at 410 nm (near Soret maximum) to attain equilibrium conditions. All reactions were at optimal pH 3, and 25 ºC.

Starting with the activation by H_2O_2, the E°'(CI/RS) values increase progressively from CaPo (1.34 V) to ALiP (1.41 V), being that value stabilized in extant LiP (**Figure 4, *middle***). Although the increasing difficulty to oxidize the enzyme through evolution seems a disadvantage, the reduction potential of this two-electron oxidation is connected to the subsequent one-electron oxidations of substrates (Furtmüller et al. 2005) (equation 8 in supplementary Materials and Methods): the free energy of the activation reaction is the sum of the free energies of the two one-electron oxidations of substrate. Therefore, the higher the E°'(CI/RS), the higher will be the free energy in the subsequent steps, being the oxidation of high redox-potential substrates favored in evolution.

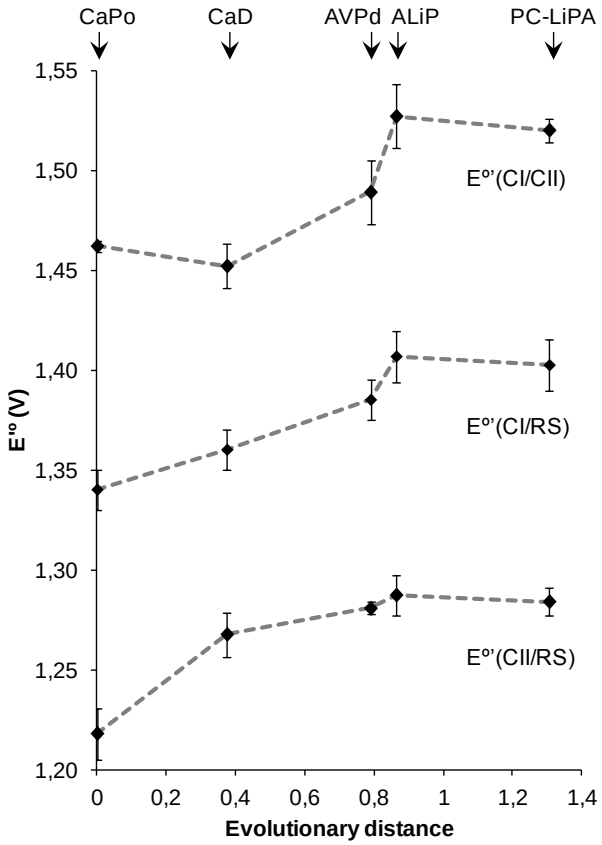


Figure 4. E°' values for the three redox couples in the catalytic cycle (**Figure 2**) *vs* evolutionary distance.(Ayuso-Fernández et al. 2018) The E°'(CI/RS) and E°'(CII/RS) values were from stopped-flow, while the E°'(CI/CII) values were from equation $\Delta G_r^{o\prime} = -n\,F\,E^{o\prime}$ (see supplementary Materials and Methods and Tables).

The two one-electron oxidations showed strong differences between the E°'(CI/CII) and E°'(CII/RS) values (**Figure 4, *top* and *bottom***, respectively). Variations of ~0,2 V were observed in all cases, with the E°'(CI/CII) values being the highest, as reported for other peroxidases (Arnhold et al. 2001; Furtmüller et al. 2005; Efimov et al. 2007). As a result for this high reduction potential, it is expected for CI to be rapidly reduced to CII in presence of the electron donor ferrocyanide, as observed in every enzyme

analyzed. Moreover, it has been estimated from quantum mechanics/ molecular mechanics calculations that at low pH the reduction potential will increase due to protonation of negative charges that stabilize CI (Castro et al. 2016), with an estimation of E°'>1.2 V for the PC-LiPA CI/CII couple at pH 3 (Kersten et al. 1990). Here we have estimated for the first time this reduction potential of PC-LiPA at pH 3, the eco-physiological pH of lignin biodegradation in nature (Martínez 2002), with a value (1.52 V) that widely exceeds those estimations. More importantly, the E°'(CII/RS) is also improved in evolution, and the potentials obtained for extant and ancestral peroxidases are the lowest in the catalytic cycle, explaining why this half-reaction is the limiting step in the oxidation of lignin (Sáez-Jiménez et al. 2015).

The reduction potentials of the three redox couples in extant/ancestral fungal ligninolytic peroxidases are higher than those reported for most animal (Arnhold et al. 2001), plant (Farhangrazi et al. 1995) or prokaryotic (Efimov et al. 2007) peroxidases, usually estimated at pH 7. However, the differences with the E°'(CI/RS) values of some animal peroxidases (Arnhold et al. 2001), which were estimated with a similar stopped-flow method, would be small after considering the variation (around 0.2 V) existing between the pH 7 and pH 3 estimations (Conroy et al. 1978; Millis et al. 1989; Wang et al. 2013), due to the presence of one or several protonable residues contributing to peroxidase catalysis (Yamazaki 1996; Santucci et al. 2000). The values shown here are also slightly higher than the E°' of 1.3 V estimated at pH 3, with the stopped-flow method, for the CI/RS couple of fungal chloroperoxidase and peroxygenase (Wang et al. 2013), two heme-thiolate proteins related to cytochrome P450 monooxygenases.

Additionally, some clues on the structural changes that led to the above boost of reduction potential in the PC-LiPA lineage were provided by ^{1}H-NMR of the cyanide adducts of the extant and ancestral enzymes (Figure 5A), where the cyanide carbon occupies the position of oxygen in CI/CII, acquired at pH 6.5 to promote enzyme stability and solubility. Analyses under similar conditions had shown that the geometry and strength of the bond between the proximal histidine (His176) and the heme Fe^{3+} (Figure 5B) would influence the redox properties of heme-peroxidases, from ancient peroxidases of prokaryotic origin to the more recent peroxidases of basidiomycetes (Banci et al. 1991a; Banci et al. 1991b; Banci et al. 1992; Banci et al. 1995; Banci et al. 2003), by controlling the imidazolate character of the histidine and the electron deficiency of the iron. The above correlates with the characteristic hyperfine chemical shifts of imidazole protons. These strong hyperfine shifts, induced by the iron paramagnetism, have been extensively used to assign the ^{1}H-NMR signals on the amino acids surrounding the heme group in peroxidases (Bertini et al. 1993; La Mar

2007). Interestingly, the ring protons of proximal histidine appear as very broad signals being the Hε1 isolated and located in the upfield area, whereas the Hδ2 is located in the more crowded downfield area (La Mar et al. 1982; Satterlee and Erman 1991).

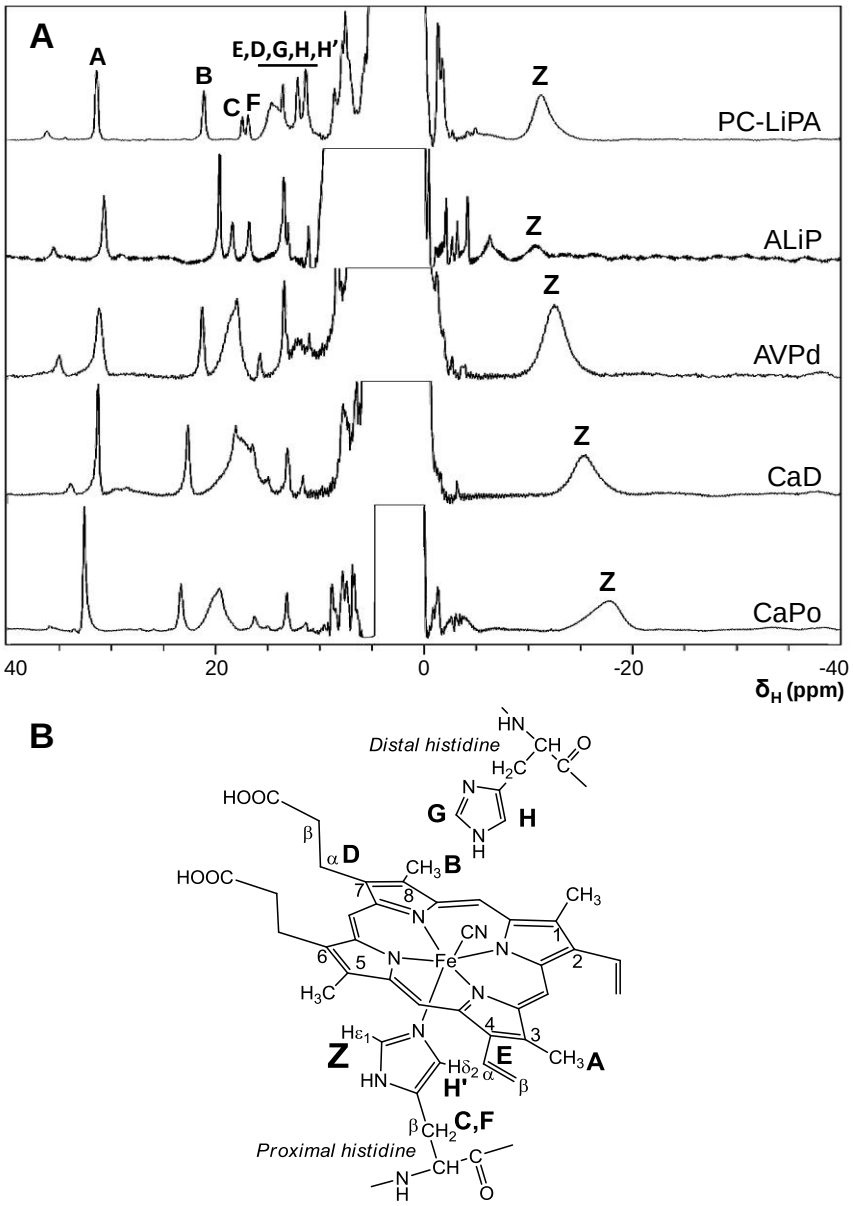


Figure 5. **(A)** ^{1}H-NMR spectra of CN-adducts of the ancestral and extant enzymes. Some signals were assigned by comparison with literature. By homology, in all enzymes the signal Z with the largest high field shift was assigned to the Hε1 of the proximal histidine directly coordinated to the heme Fe^{3+}, while the other low-field shifted labeled signals (>10 ppm), correspond to the heme $3CH_3$ (A), $8CH_3$ (B), 7α (D) and 4α (E), the proximal histidine Hβ (C), Hβ' (F) and Hδ2 (H') and the distal histidine Hε1(G) and Hδ2(H). Spectra (in 50 mM phosphate, pH 6.5, prepared with $^{2}H_2O$) were acquired with high scanning rate using WEFT pulse sequence,(Inubushi and Becker 1983) with short delay to maximize fast relaxing signals. **(B)** Representation of heme-CN complex and proximal and distal histidines, with numbering and labels as in Banci *et al* (1995).

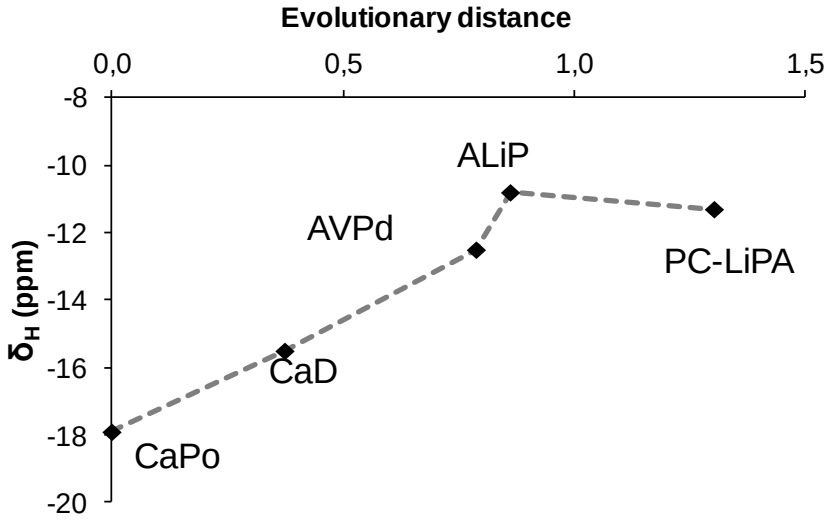


Figure 6. Representation of the evolutionary distance of the ancestral and extant enzymes (Ayuso-Fernández et al. 2018) with the chemical shift (δ_H) of signal Z in ^{1}H-NMR spectra of their CN-adducts shown in Figure 5.

Here we show a significant change in evolution of the Hε1 chemical shift of the proximal histidine (Figure 6), and the same trend was found at pH 3 (not shown) despite the lower peroxidase stability than at pH 6.5. The gradual δ_H displacement from CaPo (-18 ppm) to ALiP/PC-LiPA (-11 ppm) correlates with a reduction of the imidazolate character of the proximal histidine (Banci et al. 1991b) (due to changes in side-chain distance and/or orientation caused by variable heme/histidine environments) meaning a weaker axial bond between the histidine side-chain and the heme Fe^{3+} that, in turns, destabilizes the higher oxidation state of the heme iron (Banci et al. 1995). Interestingly, this change correlates with the changes observed in the reduction potentials: progressively higher $E^{o'}$ for the three steps in the peroxidase cycle. This is uniquely interesting, since it is the first time that structural data from NMR are correlated with reduction-potential measurements of the catalytic cycle couples. Moreover, NMR and stopped-flow changes correlate also with evolutionary time from the common ancestor of Polyporales peroxidases ~400 mya (Floudas et al. 2012). This leads to the main conclusion that, during evolution, ligninolytic peroxidases increased their already high redox-potential by molecular rearrangement in the proximal side of the heme pocket.

The general boost of reduction potentials sums to other adaptations reported in ligninolytic peroxidase evolution. It has been demonstrated that these enzymes became more stable at acidic pH, at which they act in nature (Martínez 2002), when they acquired the solvent-exposed tryptophan that oxidize lignin directly (Ayuso-Fernández et al. 2017). The environment surrounding this tryptophan was also shaped in evolution increasing the oxidation power of the tryptophanyl radical and stabilizing the substrate

radical-cation intermediates. In particular, the more than 5-fold higher efficiency of PC-LiPA oxidizing the simple lignin model compound veratryl alcohol can be connected with the above environment being more negative than in its ancestors (Ayuso-Fernández et al. 2018). Therefore, it seems that, once a high enough reduction potential was attained (in ALiP) the peroxidase final evolutionary step focused on increasing the acidic stability and selecting a more acidic tryptophanyl radical environment in extant LiP. As a whole, ancestral enzyme resurrection illustrates how LiPs became the most efficient enzymes degrading lignin.

The present study provides new information on the uniquely-high reduction potential of enzymes of biotechnological interest due to their ability to degrade lignin, a key issue in lignocellulose biorefineries for the production of bio-based chemicals, fuels and materials (Martínez et al. 2009).

Experimental Section

The 113 sequences of class-II peroxidases(Zámocký et al. 2015) in the genomes of *Bjerkandera adusta*, *Ceriporiopsis subvermispora*, *Dichomitus squalens*, *Fomitopsis pinicola*, *Ganoderma* sp., *Phlebia brevispora*, *P. chrysosporium*, *Postia placenta*, *Trametes versicolor* and *Wolfiporia cocos*,(Ruiz-Dueñas et al. 2013) available at the DOE JGI (https://genome.jgi.doe.gov/programs/fungi/index.jsf) were used in this study. After sequence alignment with MUSCLE as implemented in Mega X (Kumar et al. 2018), ML phylogeny was constructed with RAxML (Stamatakis et al. 2008) and PAML 4.7 (Yang 2007) was used to obtain the most probable ancestral sequences that were manually corrected for insertions or deletions, and synthesized for *E. coli* expression. The coding DNA sequences of ancestral and extant peroxidases were cloned and used to transform *E. coli*, and expressed and purified as previously described (Ayuso-Fernández et al. 2017). E'° values of the catalytic cycle were calculated using stopped-flow spectrophotometry (Sørlie et al. 2000; Arnhold et al. 2001; Furtmüller et al. 2003; Furtmüller et al. 2005), and the E'° (Fe^{3+}/Fe^{2+}) was obtained with spectrophotometric titration (Millis et al. 1989). ^{1}H-NMR spectra of CN-adducts of all enzymes were collected, as previously reported for other heme peroxidases (Banci et al. 1995), to analyze the displacement with evolution of the signal corresponding to the proximal histidine. For more details see supplementary Methods.

Acknowledgements

This work has been funded by the EnzOx2 (BBI-PPP-2015-2-720297; www.enzox2.eu) EU project, and the CTQ2015-71290-R, BFU2016-77835-R and BIO2017-86559-R projects of the Spanish Ministry of Economy, Industry and Competitiveness, co-financed by FEDER funds. The NMR service of CIB and M.

Carmen Fernández are acknowledged for acquiring the NMR spectra. The authors thank Paul G. Furtmüller and Roland Ludwig (BOKU, Vienna) for advice in stopped-flow determination of reduction potentials, and Werner Mäntele (University of Frankfurt) for kindly providing the spectro-electrochemical cell.

Suporting Information

Redox potential increased during the evolution of enzymes degrading recalcitrant lignin

Supporting Information includes supplementary Materials and Methods, Tables on stopped-flow calculation of E°'(CI/RS) of CaPo, CaD, AVPd, ALiP and PC-LiPA (**Tables S1-S5**, respectively) and E°'(CII/RS) of CaPo, CaD, AVPd, ALiP and PC-LiPA (**Tables S6-S10**, respectively**)**, inferred E°'(CI/CII) of CaPo, CaD, AVPd, ALiP and PC-LiPA (**Table S11**), Figures on evolution of ligninolytic peroxidases in Polyporales (**Figure S1**), multiple alignment of ancestral peroxidase sequences (**Figure S2**), spectro-electrochemical titration of the Fe^{3+}/Fe^{2+} couple (**Figure S3**), evolutionary distance *vs* E°'(Fe^{3+}/Fe^{2+}) (**Figure S4**), spectral changes during CI formation (**Figure S5**), and CII reduction (**Figure S6**), and Supporting Information References.

SUPPLEMENTARY MATERIALS AND METHODS

Ancestral sequence reconstruction and enzyme production (resurrection). The ancestral sequences were reconstructed as previously described.(Ayuso-Fernández et al. 2017) Briefly, every class-II peroxidase (113 sequences) annotated in the genomes of ten Polyporales (phylum Basidiomycota)(Ruiz-Dueñas et al. 2013) were aligned with MUSCLE to obtain a phylogeny using RAxML(Stamatakis et al. 2008) (with gamma distribution of parameters under the Whelan and Goldman model of evolution and empirical frequencies). Then, PAML 4.7(Yang 2007) was used to obtain the most probable sequence of the nodes of interest, and the marginal reconstruction sequences were manually corrected for C-terminal and other insertions/deletions according to the ancestor progeny. The genes were synthesized by ATG:biosynthetics (Merzhausen, Germany) using the most frequent codons for high expression in *Escherichia coli*.

The coding DNA sequences of the selected ancestors (CaPo, CaD, AVPd and ALiP in the simplified phylogenetic tree of Figure 1) and the extant PC-LiPA from the *P. chrysosporium* genome (JGI ID# 2989894)(Martinez et al. 2004) were cloned into pET23b(+) (Novagen). The resulting plasmids were transformed into BL21(DE3)pLysS cells, which were grown at 37ºC in Terrific broth till 0.6 OD_{600}, and four additional hours after addition of 1 mM isopropyl ß-D-1-thiogalactopyranoside. The apoenzymes accumulated in inclusion bodies and, after solubilization in 8 M urea, an *in vitro* activation protocol was optimized for the ancestral proteins. The activation conditions for the ancestral enzymes include 0.16 M urea, 5 mM $CaCl_2$, 15 μM hemin, 0.4 mM oxidized glutathione, 0.1 mM dithiothreitol and 0.1 mg/mL of protein in 50 mM Tris-HCl (pH 9.5) while those reported by Doyle and Smith(Doyle and Smith 1996) were used for PC-LiPA. The active enzymes were purified in a Resource-Q column (GE-Healthcare) using a 0-400 mM NaCl gradient in 35 mL, 2 mL/min flow, of 10 mM sodium tartrate (pH 5.5) containing 1 mM $CaCl_2$.

Spectro-electrochemical measurement of Fe^{3+}/Fe^{2+} reduction potential. Electronic absorption spectra showing the characteristic Soret band were measured with a Shimadzu UV-2401PC spectrophotometer. Redox titrations were controlled with a BAS-CV27 potentiostat and a FLUKE 77 Series II voltmeter. The potentials were calculated *vs* the standard hydrogen electrode at pH 7 (Tris/HCl 20 mm with 0.2 M of KCl) and 25 ºC for comparison with the literature. For each measurement, 10 μL of enzyme (0.6-1.0 mM concentration) were placed in an *ad hoc* cell,(Moss et al. 1990) provided by Prof. Mäntele, with a 6 μm gold mesh (Buckbee-Mears) as working electrode, a platinum plate as auxiliary electrode and a silver/silver chloride reference electrode whose potential was checked prior and after each experiment. The potentiometric titrations were carried out in the presence of 50 μM of the following redox mediators: methylene blue, 2-hydroxy-1,4-naphtoquinone, anthraquinone-1,5-disulfonate, anthraquinone-2-sulfonate, neutral red and benzyl viologen. To quantify the oxidized and

reduced enzyme, we took the absorbance at the 410 nm (Fe^{3+}) and 438 nm (Fe^{2+}) maxima, and adjusted the values to the Nernst equation:

[1] $$A_{410} = A_{\max 410} \frac{e^{\frac{(E-E^{o'})nF}{RT}}}{1+e^{\frac{(E-E^{o'})nF}{RT}}}$$

[2]. $$A_{438} = A_{\max 438} \frac{1}{1+e^{\frac{(E-E^{o'})nF}{RT}}}$$

Transient state stopped-flow experiments. The enzyme in RS, CI or CII states was followed in a stopped-flow rapid spectrophotometry equipment (Bio-Logic) synchronized with a diode array detector (J&M), and the BioKine software. All experiments were made in 100 mM tartrate, pH 3, at 25 ºC. For CI formation, a typical experiment consisted of mixing 4 μM enzyme with different concentrations of H_2O_2 (1-12 molar equivalents till equilibrium) for 3 seconds. To analyze the reduction of CII to RS, a typical experiment started by mixing 8 μM enzyme with enough equivalents (usually 2) of H_2O_2 and 1 equivalent of potassium hexacyanoferrate(II) (ferrocyanide) during 1 second, to ensure total conversion to CII. After that, adding different concentrations of tyrosine, which is oxidized to tyrosinate radical (Tyr·) by the peroxidases, allowed to quantify the reduction of CII to RS, measuring up to 1 min. All experiments were at least triplicates.

Standard reduction potentials of CI/RS, CI/CII and CII/RS couples. The formal reduction potentials ($E^{o\prime}$) of CI/RS and CII/RS were calculated using the above-described stopped-flow spectrophotometry,(Sørlie et al. 2000; Furtmüller et al. 2003) and the $E^{o\prime}$ of CI/CII was inferred. In the two former cases, the reduction potential was determined using the Nernst equation at equilibrium:

[3] $\Delta E^{o\prime} = (RT / nF)\, lnK'$

that correlates the difference of reduction potentials between enzyme and substrate with the equilibrium constant *K'*. *R* is equal to 8.31 J K^{-1} mol^{-1}, *T* is set to 298 K, *n* represents the number of electrons transferred in a single reaction step of the redox couple, and *F* (the Faraday constant) is 96,485 J V^{-1} mol^{-1}. *K'* represents the equilibrium constant, and is calculated as follows for the couples CI/RS (equation 4) and CII/RS (equation 5):

[4] $K' = ([H_2O_2][RS])/[CI]$

[5] $K' = ([Tyr^{\cdot}][RS])/([Tyr][CII])$

The different redox species at equilibrium were quantified as follows: i) for the pair CI/RS, the amounts of CI and RS were estimated with their extinction coefficients at 410 nm (see below) at the equilibrium (i.e. when the spectral changes ended, during H_2O_2 addition; see Figure 3A) and the reduction potential for the pair H_2O_2/H_2O at pH 3 ($E^{o'}$ = 1.56 V);(Koppenol 1987) and ii) for the CII/RS pair, the equilibrium concentrations of CII and RS (see Figure 3B) were calculated using their extinction coefficients at 410 nm, after incubation with different amounts of tyrosine, and the calculated $E^{o'}$ = 1.18 V for the pair $Tyr^{\cdot}$/Tyr.(DeFelippis et al. 1989) The selection of Tyr as substrate was made taking into account that: i) all enzymes of the PC-LiPA lineage were able to oxidize it (data not shown); and ii) it has been previously used to calculate the $E^{o'}$ of mammalian peroxidases using a similar approach.(Arnhold et al. 2001; Furtmüller et al. 2003; Furtmüller et al. 2005; Battistuzzi et al. 2010b)

At a specific wavelength, the absorbance is an additive measurement of those of the individual components of a mixture. Therefore, using the 410 nm (Soret band) extinction coefficients for RS, CI and CII (given below) the quantification of the different redox pairs at equilibrium is possible using the equations 6 (for equilibrium of CI and RS) and 7 (for equilibrium of CII and RS):

[6] $A_{410} = \varepsilon_{410\text{-}RS}[RS]\,l + \varepsilon_{410\text{-}CI}[CI]\,l$

[7] $A_{410} = \varepsilon_{410\text{-}CI}[CI]\, l + \varepsilon_{410\text{-}CII}[CII]\, l$

where l is the path length of the stopped-flow cuvette.

The RS ε_{410} values were: 138 mM^{-1} cm^{-1} (RS-CaPo), 177 mM^{-1} cm^{-1} (RS-CaD), 149 mM^{-1} cm^{-1} (RS-AVPd), 171 mM^{-1} cm^{-1} (RS-ALiP) and 168 mM^{-1} cm^{-1} (RS-PC- LiPA). The CI ε_{410} was calculated after converting all the RS enzyme into CI using 2-10 H_2O_2 equivalents, ensuring there is no auto-reduction to CII. The values obtained were: 53 mM^{-1} cm^{-1} (CI-CaPo), 74 mM^{-1} cm^{-1} (CI-CaD), 68 mM^{-1} cm^{-1} (CI-AVPd), 87 mM^{-1} cm^{-1} (CI-ALiP) and 110 mM^{-1} cm^{-1} (CI- PC-LiPA). The CII ε_{410} was calculated after converting all RS enzyme into CI and then into CII, without further auto-reduction to RS, using 2 equivalents of H_2O_2 and 1 equivalent of ferrocyanide.(Wariishi et al. 1990) The values obtained were: 73 mM^{-1} cm^{-1} (CII-CaPo), 98 mM^{-1} cm^{-1} (CII-CaD), 105 mM^{-1} cm^{-1} (CII-AVPd), 133 mM^{-1} cm^{-1} (CII-ALiP) and 105 mM^{-1} cm^{-1} (CII- PC-LiPA).

Concerning the $E^{o'}$ of CI/CII, we were unable to calculate it experimentally. Despite obtaining enzyme in CI and being able to attain certain equilibrium with CII at specific Tyr concentrations, the CII auto-reduction to RS enzyme made impossible to calculate equilibrium concentrations. However, it was possible to infer the $E^{o'}$ of CI/CII from the experimental values of the CI/RS and CII/RS couples. The standard reaction free energy is connected to $E^{o'}$ according to:

[8] $\Delta G_r'^{o} = -n\, F\, E^{o'}$

with n = 2 electrons for the reduction of CI to RS, $\Delta G_r'^{o}$ equals to -2 x F x [$E^{o'}$(CI/RS)] being the sum of the reaction free energy of one electron reductions (CI to CII, and CII to RS). Therefore, the sum of reaction free energies $\Delta G_r'^{o}$(CI/CII) + $\Delta G_r'^{o}$(CII/RS) and the experimental determination of $E^{o'}$(CII/RS) allows the determination of $E^{o'}$(CI/CII).

^{1}H NMR Spectroscopy. The cyanide adducts of the ancestral peroxidases and extant PC-LiPA were obtained by incubating 0.7-1.0 mM of each sample with KCN in 50 mM potassium phosphate, pH 6.5, prepared with $^{2}H_2O$ (isotopic purity 99.9%). The ^{1}H-NMR spectra were recorded at 298K using a 600 MHz Bruker spectrometer equipped with cryoprobe. Fast scanning WEFT pulse sequence was used.(Inubushi and Becker 1983) Delays between 13 to 120 ms were tested in order to maximize signals corresponding to fast relaxing protons of the proximal histidine coordinated to Fe of the heme group. Spectra were acquired with up to 40K total scans, 120 ppm spectral width and a minimum of 1K points. The spectra were processed with TOPSPIN 3.0 (Bruker) with manual base line adjustment and the signal of residual water proton (δ_H 4.701 ppm) was used as internal reference for chemical shifts.

Precise conditions of the presented spectra are: i) CaPo, addition of 5 spectra of 1K scans and 128 dummy scans each, 15 ms delay, 100 ppm width and 1K points; ii) CaD, addition of 5 spectra of 4K scans and 128 dummy scans each, 13 ms delay, 120 ppm width and 1K points; iii) AVPd, addition of 5 spectra of 1K scans and 128 dummy scans each, 13 ms delay, 120 ppm width and 1K points; iv) AliP, addition of 5 spectra of 4K scans and 128 dummy scans each, 23 ms delay, 100 ppm width and 2K points; and v) PC-LiPA, addition of 10 spectra of 4K scans and 128 dummy scans each, 13 ms delay, 120 ppm width and 1K points.

Signal corresponding to Hε1 of the proximal histidine was assigned to the signal with highest upper-field shift by homology with reported data.(Banci et al. 1991a; Banci et al. 1991b; Banci et al. 1992; Banci et al. 1995; Banci et al. 2003) Other signals were assigned by comparison with reported data for PC-LiPA isolated from *P. chrysosporium* cultures, although some differences are observed as the PC-LiPA used in this study was obtained by heterologous expression in *E. coli* and devoid of natural glycosylation. Heme group numbering and protons labeling are indicated in Figure 5B and correspond to those used by Banci et al.(Banci et al. 1995)

SUPPLEMENTARY TABLES

Table S1. Parameters of the redox equilibrium and calculated $E^{o'}$ of the CI/RS redox couple of CaPo as a function of the initial concentration of H_2O_2 (all reactions at optimal pH 3).

Initial H_2O_2	Equilibrium concentrations (μM)			$E^{o'}$ (V)
(μM)	CI-CaPo	RS-CaPo	H_2O_2	
2.00	1.73	0.16	0.27	1.334
3.00	1.85	0.05	1.15	1.335
4.00	1.85	0.05	2.16	1.345
8.00	1.87	0.02	6.13	1.339
Mean and 95% confidence interval:				1.338 ± 0.004

Table S2. Parameters of the redox equilibrium and calculated $E^{o'}$ of the CI/RS redox couple of CaD as a function of the initial concentration of H_2O_2 (all reactions at optimal pH 3).

Initial H_2O_2	Equilibrium concentrations (μM)			$E^{o'}$ (V)
(μM)	CI-CaD	RS-CaD	H_2O_2	
1.00	0.77	1.15	0.23	1.369
2.00	1.09	0.83	0.91	1.378
3.00	1.55	0.3	1.45	1.367
4.00	1.83	0.09	2.17	1.353
Mean and 95% confidence interval:				1.367 ± 0.011

Table S3. Parameters of the redox equilibrium and calculated $E^{o'}$ of the CI/RS redox couple of AVPd as a function of the initial concentration of H_2O_2 (all reactions at optimal pH 3).

Initial H_2O_2	Equilibrium concentrations (μM)			$E^{o'}$ (V)
(μM)	CI-AVPd	RS-AVPd	H_2O_2	
1.00	0.28	1.08	0.72	1.392
2.00	0.55	0.81	1.45	1.392
3.00	0.94	0.42	2.06	1.382
4.00	1.18	0.18	2.82	1.372
Mean and 95% confidence interval:				1.385 ± 0.011

Table S4. Parameters of the redox equilibrium and calculated $E^{o'}$ of the CI/RS redox couple of ALiP as a function of the initial concentration of H_2O_2 (all reactions at optimal pH 3).

Initial H_2O_2	Equilibrium concentrations (μM)			$E^{o'}$ (V)
(μM)	CI-ALiP	RS-ALiP	H_2O_2	
1.00	0.06	1.75	0.094	1.429
2.00	0.50	1.31	1.50	1.400
3.00	0.67	1.14	2.33	1.400
4.00	0.77	1.04	3.23	1.402
8.00	1.10	0.71	6.90	1.402
Mean and 95% confidence interval:				1.407 ± 0.011

Table S5. Parameters of the redox equilibrium and calculated E°' of the CI/RS redox couple of PC-LiPA as a function of the initial concentration of H_2O_2 (all reactions at optimal pH 3).

Initial H_2O_2	Equilibrium concentrations (μM)			E°' (V)
(μM)	CI-PCLiPA	RS-PCLiPA	H_2O_2	
1.50	0.27	1.97	1.23	1.411
2.00	0.41	1.85	1.59	1.408
3.00	0.98	1.28	2.02	1.395
4.00	1.14	1.13	2.86	1.396
8.00	1.53	0.73	6.47	1.399
Mean and 95% confidence interval:				1.402 ± 0.006

Table S6. Parameters of the redox equilibrium and calculated E°' of the CII/RS redox couple of CaPo as a function of the initial concentration of Tyr (all reactions at optimal pH 3).

Initial Tyr	Equilibrium concentrations (μM)				E°' (V)
(μM)	RS-CaPo	CII-CaPo	Tyr	Tyr·	
5	0.79	1.35	4.21	0.79	1.233
10	1.22	0.92	8.78	1.22	1.219
25	1.66	0.49	23.35	1.66	1.212
50	1.88	0.26	48.12	1.88	1.207
Mean and 95% confidence interval:					1.217 ± 0.011

Table S7. Parameters of the redox equilibrium and calculated $E^{o'}$ of the CII/RS redox couple of CaD as a function of the initial concentration of Tyr (all reactions at optimal pH 3).

Initial Tyr	Equilibrium concentrations (μM)				$E^{o'}$ (V)
(μM)	RS-CaD	CII-CaD	Tyr	Tyr·	
2.5	0.35	1.47	2.15	0.35	1.258
5	0.48	1.34	4.52	0.48	1.260
10	0.60	1.22	9.4	0.6	1.264
20	0.60	1.22	19.4	0.6	1.283
Mean and 95% confidence interval:					1.266 ± 0.011

Table S8. Parameters of the redox equilibrium and calculated $E^{o'}$ of the CII/RS redox couple of AVPd as a function of the initial concentration of Tyr (all reactions at optimal pH 3).

Initial Tyr	Equilibrium concentrations (μM)				$E^{o'}$ (V)
(μM)	RS-AVPd	CII-AVPd	Tyr	Tyr·	
10	0.50	1.85	9.51	0.50	1.286
20	0.79	1.56	19.21	0.79	1.275
50	1.17	1.18	48.83	1.17	1.270
100	1.35	1.00	98.65	1.35	1.278
Mean and 95% confidence interval:					1.278 ± 0.006

Table S9. Parameters of the redox equilibrium and calculated E°' of the CII/RS redox couple of ALiP as a function of the initial concentration of Tyr (all reactions at optimal pH 3).

Initial Tyr	Equilibrium concentrations (μM)				E°' (V)
(μM)	RS-ALiP	CII-ALiP	Tyr	Tyr·	
10	0.46	2.02	9.54	0.46	1.294
20	0.64	1.84	19.36	0.64	1.290
50	1.04	1.44	48.96	1.04	1.283
100	1.35	1.13	98.65	1.35	1.281
Mean and 95% confidence interval:					1.287 ± 0.006

Table S10. Parameters of the redox equilibrium and calculated E°' of the CII/RS redox couple of PC-LiPA as a function of the initial concentration of Tyr (all reactions at optimal pH 3).

Initial Tyr	Equilibrium concentrations (μM)				E°' (V)
(μM)	RS-PCLiPA	CII-PCLiPA	Tyr	Tyr·	
10	0.42	1.50	9.58	0.42	1.289
20	0.62	1.30	19.38	0.62	1.283
50	0.95	0.97	49.05	0.95	1.279
100	1.19	0.73	98.81	1.19	1.277
Mean and 95% confidence interval:					1.284 ± 0.006

Table S11. E°'(CI/CII) inferred for each enzyme using the equation for the reaction free energy ($\Delta G_r^{o\prime} = -n\ F\ E^{o\prime}$). E°'(CI/RS) and E°'(CII/RS) were obtained experimentally at optimal pH 3 (**Tables S1-S5** and **S6-S10**, respectively). For RS/CI n = 2, and its $\Delta G_r^{o\prime}$ equals to the sum of the $\Delta G_r^{o\prime}$(CI/CII) and $\Delta G_r^{o\prime}$(CII/R), both with n = 1.

	$\Delta G_r^{o\prime}$(R/CI) (kJ)	$\Delta G_r^{o\prime}$(CII/R) (kJ)	$\Delta G_r^{o\prime}$(CI/CII) (kJ)	$E^{o\prime}$(CI/CII) (V)
CaPo	-258.6	-117.5	-141.1	1,462
CaD	-262.4	-122.4	-140.1	1,452
AVPd	-267.3	-123.6	-143.7	1,489
ALiP	-271.5	-124.2	-147.3	1,527
PCLiPA	-270.5	-123.9	-146.7	1,520

SUPPLEMENTARY FIGURES

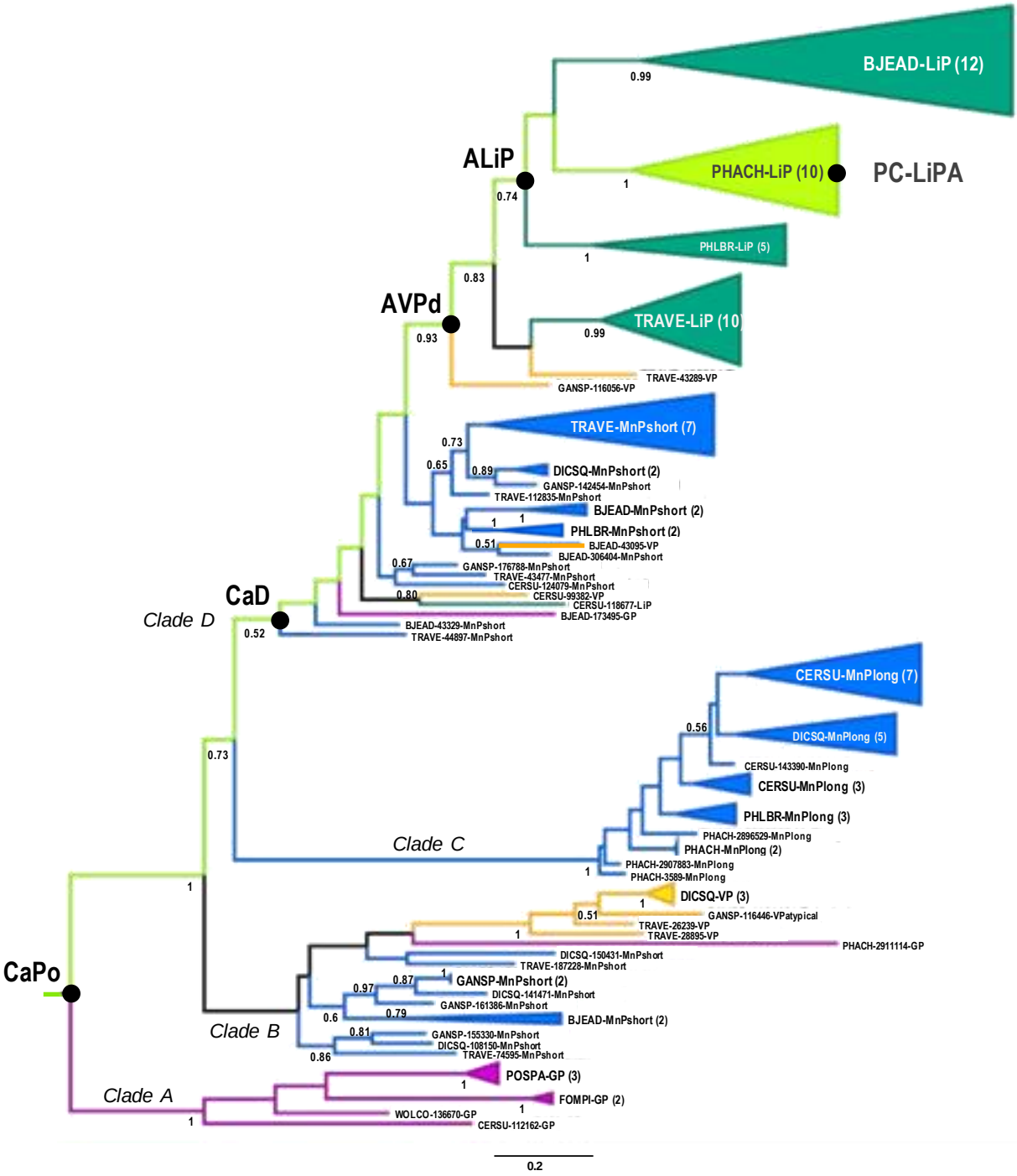


Figure S1. Evolution of fungal ligninolytic peroxidases in Polyporales (bootstrap values > 0.5 are indicated) adapted from Ayuso-Fernández *et al.*(2018) LiPs are shown in green, VPs in yellow, MnPs in blue and GPs in purple. The path to the extant LiPA of *P. chrysosporium* (JGI ID# 2989894) is marked in light green. Also, the enzymes analyzed in the present work (common ancestor of Polyporales peroxidases, CaPo; common ancestor of clade D, CaD; ancestral versatile peroxidase in clade D, AVPd; and ancestral LiP, ALiP) are marked.

```
CaPo        VTCSDGVSTASNAACCAWFAVLDDIQANLFDGGQCGEEAHESLRLTFHDAIGFSPALAAQ    60
CaD         VTCPDGVNTATNAACCALFAVLDDIQENLFDGGECGEEAHESLRLTFHDAIGFSPALARQ    60
AVPd        VACPDGVNTATNAACCALFAVRDDIQQNLFDGGECGEEVHESLRLTFHDAIAFSPALEAQ    60
ALiP        VACPDGVNTATNAACCALFAVRDDIQQNLFNGGQCGDEAHESLRLTFHDAIAFSPALEAQ    60
PC-LiPA     ATCSNG-KTVGDASCCAWFDVLDDIQQNLFHGGQCGAEAHESIRLVFHDSIAISPAMEAQ    59
            .:* :* .*. :*:*** * * **** ***.**:** *.***:**.***:*.:***:  *

CaPo        GKFGGGGADGSIITFADIETNFHANNGLDDIVDALKPFADKHNVSYGDFIQFAGAVGVSN    120
CaD         GKFGGGGADGSIITFSDIETNFHANGGIDEIVEVQKPFVAKHNMTAGDFIQFAGAVGVSN    120
AVPd        GQFGGGGADGSIAIFEDIETNFHANLGLDEIVNEQKPFIARHNMTTADFIQFAGAVGVSN    120
ALiP        GQFGGGGADGSIVIFSDIETNFHANIGLDEIVAIQKPFIARHNMTVADFIQFAGAVGVSN    120
PC-LiPA     GKFGGGGADGSIMIFDDIETAFHPNIGLDEIVKLQKPFVQKHGVTPGDFIAFAGAVALSN    119
            *:**********  * **** ** * *:*:**   ***  :* :: .*** *****.:**

CaPo        CPGAPRLEFLAGRPNATAPSPDGLVPEPSDSVDKILARMADAGGFSPDEVVALLASHSVA    180
CaD         CPGAPRLEFLLGRPAATAPSPDGLVPEPFDSVDKILARFADAGGFSPDEVVALLASHSVA    180
AVPd        CPGAPQLDFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGGFDPIETVWLLTSHTIA    180
ALiP        CPGAPQLNFFLGRPDATQPAPDGLVPEPFDTVDQILARMADAGEFDELETVWLLIAHTVA    180
PC-LiPA     CPGAPQMNFFTGRAPATQPAPDGLVPEPFHTVDQIINRVNDAGEFDELELVWMLSAHSVA    179
            *****:::*: **  ** *:******** .:**:*: *. *** *.  * * :* :*::*

CaPo        AQDHVDPTIPGTPFDSTPSTFDTQFFLETLLKGTAFPGTGANSGEVKSPLKGEFRLQSDA    240
CaD         AADHVDPTIPGTPFDSTPSTFDTQFFVEVLLRGTLFPGTGGNQGEVKSALRGEIRLQSDH    240
AVPd        AADHVDPTIPGTPFDSTPELFDTQFFIETQLRGTLFPGTGGNQGEVESPLRGEIRLQSDH    240
ALiP        AANDVDPTIPGTPFDSTPELFDSQFFIETQLRGTLFPGTGGNQGEVESPLKGEMRLQSDH    240
PC-LiPA     AVNDVDPTVQGLPFDSTPGIFDSQFFVETQLRGTAFPGSGGNQGEVESPLPGEIRIQSDH    239
            * :.****: * ******  **:***:*. *:** ***:*.*.***:* * **:*:***

CaPo        AIARDPRTACEWQSFVNNQELMQSSFRAAMAKLANLGHDRSDLIDCSEVIPVPKPLA---    297
CaD         EVARDPRTACEWQSFVNNQAKMQKSFRAAMAKLAILGHDRSDLIDCSEVIPVPPPLA---    297
AVPd        LLARDSRTACEWQSFVNNQPKLQKSFQAAFHDLSMLGHDVNDLIDCSEVIPIPPPPT---    297
ALiP        LLARDSRTACEWQSFVNNQPKLQKNFQFVFEALSMLGQDPNDLIDCSEVIPIPPPLTLTP    300
PC-LiPA     TIARDSRTACEWQSFVNNQSKLVDDFQFIFLALTQLGQDPNAMTDCSDVIPQSKPIPGNL    299
             :*** *************  : ..*:  :  *: **:* . : ***:***   *

CaPo        ASATFPAGKTRSDIEQSCRSTPFPTLPTDPGPATSIPPV------    336
CaD         ATAHFPAGLTRKDIEQSCRSTPFPTLSTDPGPATSVPPV------    336
AVPd        STAHFPAGLTNADVEQACAETPFPTLPTDPGPATSVAPV------    336
ALiP        AASHFPAGKTNKDVEQACAETPFPTLPTDPGPATSVAPVPPSPAA    345
PC-LiPA     PFSFFPAGKTIKDVEQACAETPFPTLTTLPGPETSVQRIPPPPGA    344
           : **** *  *:**:* .****** * *** **:  :
```

Figure S2. Alignment of the four ancestral sequences (mature proteins) predicted with PAML 4.7 and extant LiPA. Conserved catalytic and other relevant residues(Ruiz-Dueñas et al. 2009) are indicated including: two active site histidines (*dark gray*); three acidic residues forming the Mn^{2+}-binding site (*red*); other active site conserved residues (*light gray*); one tryptophan involved in lignin direct oxidation (*cyan*); nine ligands of two Ca^{2+} ions (*green*); and eight cysteines forming disulfide bonds (*yellow*). Symbols below indicate full conservation of the same (*asterisk*) or equivalent residues (*colon*) and partial residue conservation (*dot*). Adapted from Ayuso-Fernández et al.(Ayuso-Fernández et al. 2017)

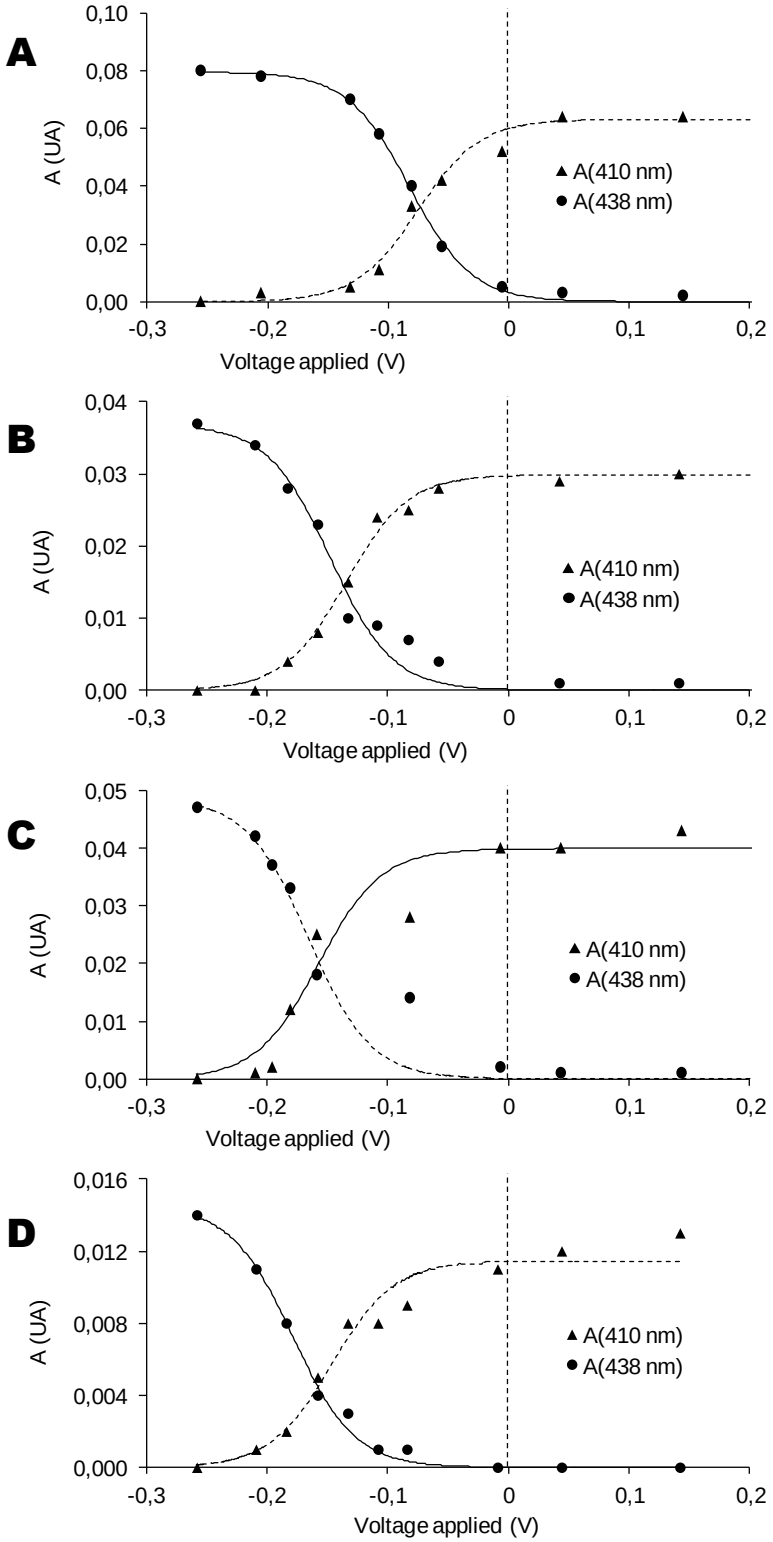


Figure S3 (next page)**.** Spectro-electrochemical titration of the Fe^{3+}/Fe^{2+} couple in ancestral CaPo (**A**), CaD (**B**), AVPd (**C**) and ALiP (**D**) peroxidases. As described in supplementary Materials and Methods, the use of different redox mediators allowed the conversion of RS-Fe^{3+} to RE-Fe^{2+} after applying several potential differences. All the potentials are referred to the standard hydrogen electrode. The data obtained were fitted to the Nernst equation for a reversible one electron transition. Continuous line: changes in the $A_{410\text{-nm}}$ (Fe^{3+}); dashed line: changes in $A_{438\text{-nm}}$ (Fe^{2+}).

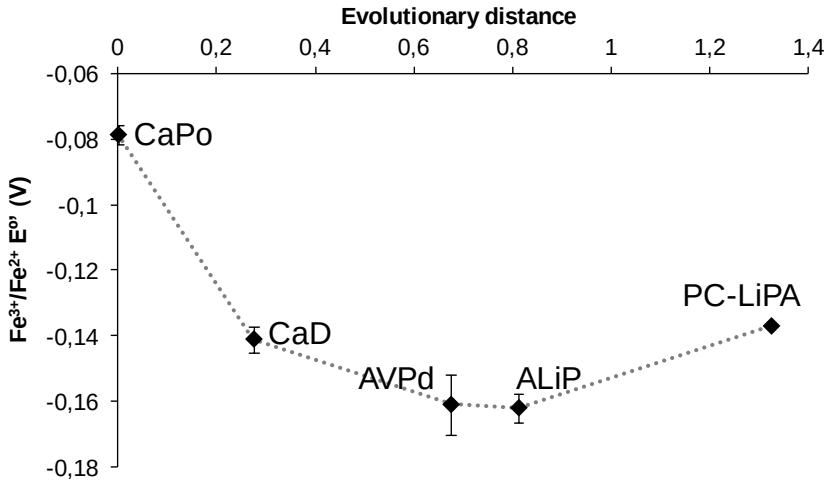


Figure S4. $E^{o'}(Fe^{3+}/Fe^{2+})$ values *vs* evolutionary distance of ancestral peroxidases (CaPo, CaD, AVPd and ALiP) and PC-LiPA. See **Figure S3** for the spectro-electrochemical titration of the ancestral enzymes. The values for PC-LiP have been reported several times in the bibliography, and here we take the most accepted value of -0.137 V.(Millis et al. 1989) The negative values ensure that the peroxidases have a heme with Fe^{3+}, the active form of iron, being in the range of those reported for other basidiomycete peroxidases.(Millis et al. 1989; Whitwam et al. 1999; Santucci et al. 2000; Ciaccio et al. 2003; Oyadomari et al. 2003; Battistuzzi et al. 2006) Although it is considered that a more positive $E^{o'}$ would mean a higher oxidation power,(Battistuzzi et al. 2010b) PC-LiPA, along with AVPd and ALiP, are able to oxidize high redox-potential substrates in spite of the more negative $E^{o'}(Fe^{3+}/Fe^{2+})$ than the evolutionarily older CaPo and CaD, which only oxidize low redox-potential substrates and Mn^{2+},(Ayuso-Fernández et al. 2017) estimated here.

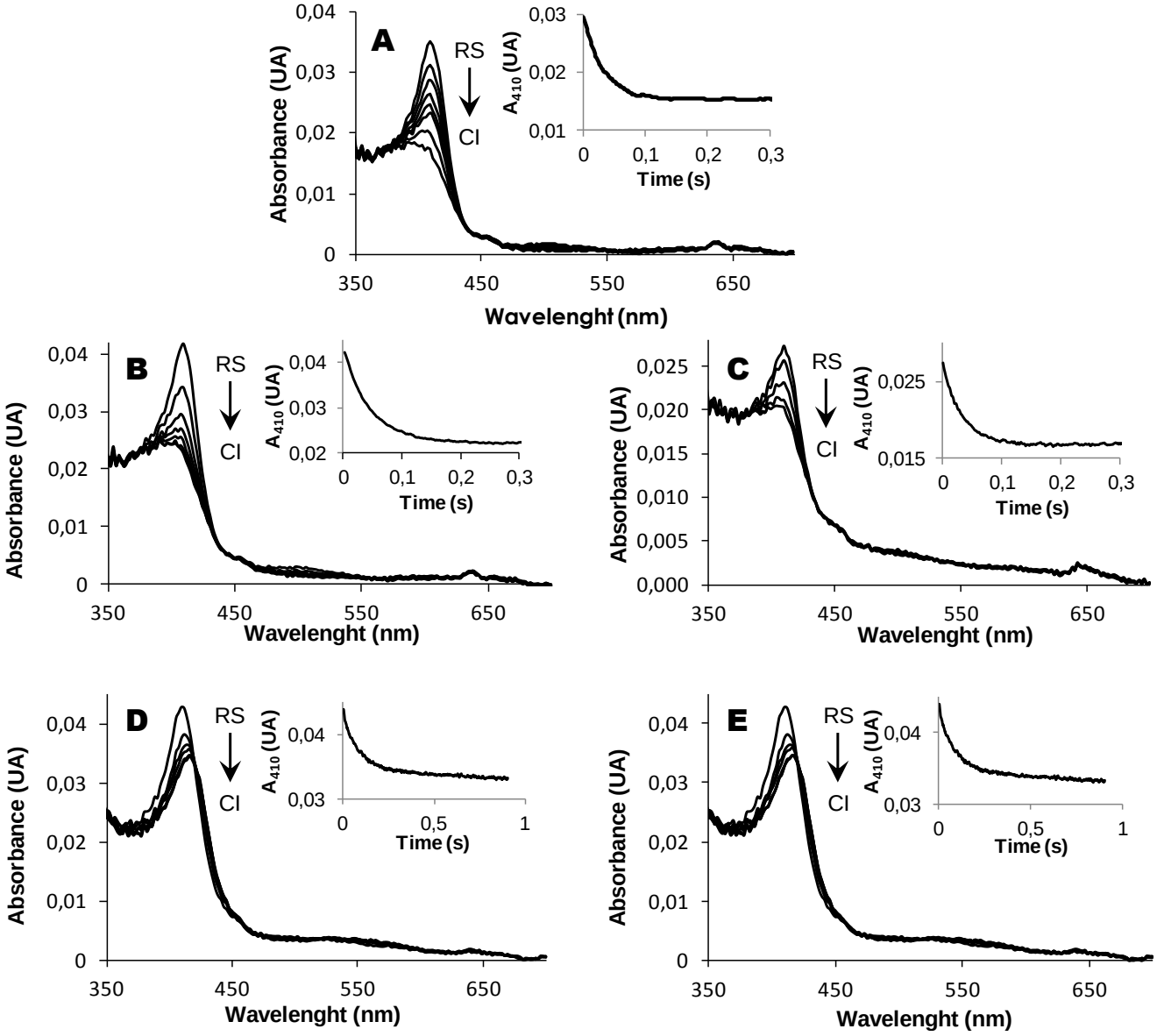


Figure S5. Spectral changes upon mixing of CaPo (**A**), CaD (**B**), AVPd (**C**), ALiP (**D**) and PC-LiPA (**E**) with H_2O_2 to follow CI formation (from 1.6 to 225 ms after peroxide addition). The insets show time traces at 410 nm (near Soret maximum) to attain equilibrium conditions. All reactions were at optimal pH 3, and 25 ºC.

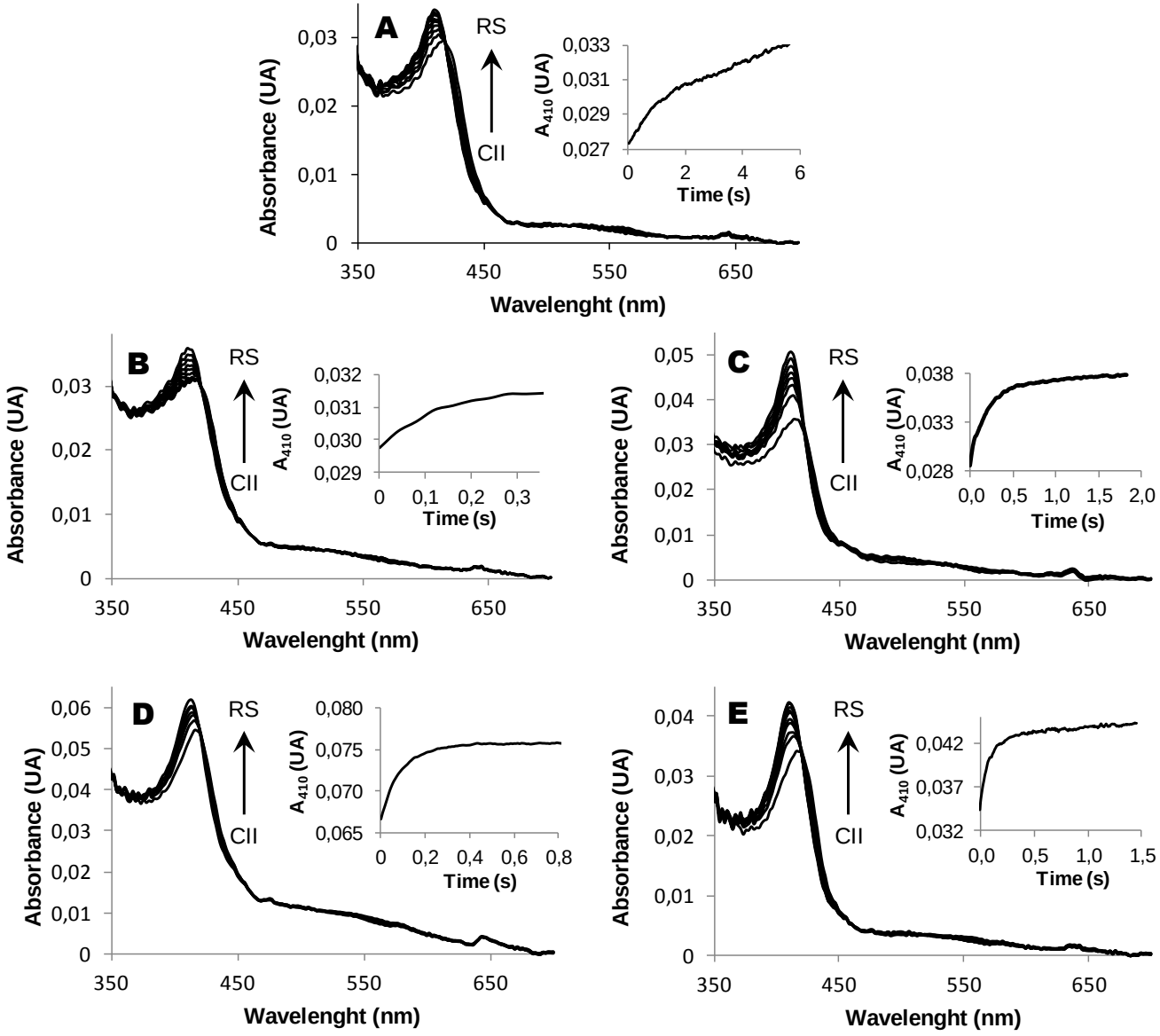


Figure S6. Spectral changes during Tyr reduction of CII, formed by CaPo (**A**), CaD (**B**), AVPd (**C**), ALiP (**D**) and PC-LiPA (**E**) mixing with H_2O_2 and ferrocyanide (from 1.6 to 650 ms after Tyr addition).The insets show time traces at 410 nm (near Soret maximum) to attain equilibrium conditions. All reactions were at optimal pH 3, and 25 ºC.

Chapter 4

Peroxidase evolution in white-rot fungi follows wood lignin evolution in plants

General abstract

The results presented so far have focused in the analysis of evolution of ligninolytic peroxidases using model compounds. However, for a better biological interpretation, more *realistic* substrates have to be used. For this reason we employ here lignosulfonates for the characterization of the lineage leading to extant lignin peroxidases. Lignosulfonates are excellent substrates to represent the real lignin polymer: they are water soluble macromolecules extracted from lignin with a phenolic content similar to the natural polymer and their sulfonation does not affect the reactivity of peroxidases.

In this final chapter the results of the characterization of LiP ancestors using lignosulfonates are presented. We used lignosulfonates from gymnosperms and angiosperms, representing more basal and more recent lignins, respectively. We characterized the transient state kinetics for enzyme reduction extracting electrons from the lignosulfonates with rapid stopped-flow spectrophotometry. We also analyzed steady-state reactions of prolonged treatments, to evaluate the modification of the macromolecule using 2D-NMR.

Together with the time-calibration of peroxidases phylogeny and the plant-host distribution for Polyporales, the results obtained are discussed in the light of plant evolution.

Manuscript in preparation (tentative title and authors):

Ayuso-Fernández, I, Rencoret, J., Gutiérrez, A., Ruiz-Dueñas, F.J., and Martínez, A.T. Peroxidase evolution in white-rot fungi follows wood lignin evolution in plants.

Author contributions: I.A.-F. and J.R. performed the experimental work. I.A.-F. performed the bioinformatic analyses. F.J.R.-D. and A.T.M. designed the research. A.G. contributed to the chemical analyses. All authors analyzed the data, contributed to writing, and approved the final manuscript.

Abstract

Comparison of sequenced Agaricomycotina genomes suggests that the ability to degrade wood lignin was associated to the appearance of secreted peroxidases with a solvent-exposed catalytic tryptophan. This hypothesis is experimentally demonstrated here by resurrecting ancestral fungal peroxidases, after sequence reconstruction from genomes of extant white-rot Polyporales, and evaluating their oxidative attack on the lignin polymer by state-of-the-art analytical techniques. Rapid stopped-flow estimation of the transient-state constants of the two successive one-electron transfers from lignin to the peroxide-activated enzyme (k_{2app} and k_{3app}) first showed a progressive increase during peroxidase evolution (up to 50-fold higher values for the rate-limiting k_{3app}). The above agreed with 2D-NMR analyses during steady-state treatments of hardwood lignin showing that the degradation of lignin (estimated from the normalized aromatic signals of lignin units, compared with a control) and its S/G ratio increased with the enzyme evolutionary distance from the first peroxidase ancestor. More interestingly, the stopped-flow estimations of electron-transfer rates also showed how the most recent peroxidase ancestors that already incorporated the exposed tryptophan to their molecular structure (as well as the extant lignin peroxidase) were comparatively more efficient oxidizing hardwood (angiosperm) lignin, while the most ancestral "tryptophan-less" enzymes were more efficient abstracting electrons from softwood (gymnosperm) lignin. A time-calibration of the ancestry of Polyporales peroxidases localized the appearance of the first peroxidase with a solvent-exposed catalytic tryptophan 194 ± 70 mya, coincident with the diversification of angiosperm plants characterized by the appearance of dimethoxylated syringyl lignin units.

Significance statement

We analyze the evolution of ligninolytic peroxidases from wood-rotting fungi using lignin from gymnosperms and angiosperms, as representative for two steps of lignin evolution in plants. By enzyme resurrection, we show that, during fungal evolution, these enzymes improved their activity and switched their degradative preferences with the rise of a solvent-exposed tryptophan conferring them the ability to oxidize nonphenolic lignin. We calibrated the peroxidases phylogeny and determined that this residue appeared coincident with angiosperm diversification, characterized by the synthesis of a more complex and less phenolic lignin due to incorporation of a new unit in its structure. This way, we show that fungal evolution followed that of lignin synthesis, pointing to a co-evolution between fungal saprotrophs and their plant hosts.

Introduction

Lignin (Ralph et al. 2004) is an aromatic polymer formed by dehydrogenative polymerization of *p*-hydroxycinnamyl alcohols and other monomers by peroxidases/laccases (Tobimatsu and Schuetz 2019) that confers resistance to plant tissues (Fengel and Wegener 1984). Due to lignin recalcitrance, its degradation is essential for carbon recycling in nature and often represents a limiting step in the biomass industry (Martínez et al. 2009). The biological degradation of lignin is described as an enzymatic combustion made by white-rot fungi (Kirk and Farrell 1987). It involves the so-called manganese peroxidases (MnPs), versatile peroxidases (VPs) and lignin peroxidases (LiPs) together with auxiliary enzymes (Martínez et al. 2018). MnPs would oxidize the minor phenolic moiety of lignin *via* Mn^{3+} chelates (Glenn et al. 1986) or directly in contact with the heme cofactor (Fernández-Fueyo et al. 2014). In contrast, LiPs would attack nonphenolic lignin directly using a solvent-exposed catalytic tryptophan forming a reactive radical, as shown using model compounds (Smith et al. 2009). Finally, VPs combine the structural and catalytic properties of MnPs and LiPs (Ruiz-Dueñas et al. 2009) including the surface tryptophan whose role in direct oxidation of lignin has been recently demonstrated for this peroxidase (Sáez-Jiménez et al. 2015b; Sáez-Jiménez et al. 2016).

Fungal degradation of lignin started in the Carboniferous period associated to the production of the first ligninolytic peroxidases (Floudas et al. 2012) that, together with geochemical factors, contributed to the end of coal accumulation from undecayed plant biomass (Hibbett et al. 2016). Recent works have resurrected ancestral ligninolytic peroxidases and analyzed their evolution in Polyporales (Ayuso-Fernández et al. 2017), where most wood-rotting fungi are included, using high and low redox-potential simple model compounds. These studies suggested that the most ancestral ligninolytic enzymes acted on phenolic lignin, and later shaped several times their oxidations sites to become first VPs and later LiPs, the most-efficient lignin-degrading enzymes, as a convergent trait in fungal evolution (Ayuso-Fernández et al. 2018b). Also, ligninolytic peroxidases progressively increased the redox potential of their reactive species during evolution, in parallel with changes of catalytic sites. Finally, their stability at acidic pH, where lignin decay takes place in nature, increased and, at the latest stages of evolution, their catalytic tryptophan environment became more negative to stabilize lignin cation radicals (Ayuso-Fernández et al. 2017; Ayuso-Fernández et al. 2018a).

Recent phylogenomic studies using 1157 Agaricomycetes species show that, during evolution, wood-rotting fungi became host-specific depending on whether they were brown-rot, often acting on gymnosperms, or white-rot fungi, specializing on angiosperms (Krah et al. 2018). The evolution of host-

specialization is notably interesting since gymnosperm lignin basically consists of guaiacyl (G) units (derived from coniferyl alcohol) while angiosperms introduce also syringyl (S) units (derived from sinapyl alcohol) (Ralph et al. 2004). Lignin in gymnosperms has water conduction as primary function in evolution, without detectable changes in its composition with time (Sarkanen and Hergert 1971). This way, the inclusion of S-units during angiosperms evolution, together with the specialization of vessels fibers for conduction and xylem fibers for support, provided significant mechanical and physiological advantages (Li et al. 2001). Also the inclusion of S-units represents a better protection against pathogens, as demonstrated by the rise of S-lignin upon fungal penetration on plant tissues (Menden et al. 2007). Importantly, from a structural point of view, the appearance of syringyl units in angiosperm lignin increases the polymer complexity (Ralph et al. 2004) and results in lower average phenolic content (Lai and Guo 1991; Camarero et al. 1999).

So far, the evolution of ancestral peroxidase properties has not been analyzed on the lignin polymer. The use of lignosulfonates represents an excellent approach to this issue, since they are water-soluble lignins, whose sulfonation scarcely affects peroxidase reactivity, and have been recently used to probe the catalytic role of the surface tryptophan of ligninolytic peroxidases (Sáez-Jiménez et al. 2015b; Sáez-Jiménez et al. 2016). Here we study how ancestral peroxidases modify lignin through their evolution, and estimate the electron-transfer rates with softwood (gymnosperm) and hardwood (angiosperm) lignosulfonates as substrates. With this purpose, we use stopped-flow rapid spectrophotometry to calculate transient-state kinetic parameters, and 2D-NMR spectroscopy to evaluate the polymer modification in prolonged steady-state treatments.

Results

Evolution of Polyporales lignin-degrading peroxidases. To investigate how the ability to oxidize the lignin polymer evolved in the ancestry of Polyporales peroxidases, we reconstructed the lineage leading to the well-known isoenzyme-H8 of *P. chrysosporium* LiP (Cullen and Kersten 2004), corresponding to gene-A in its sequenced genome (PC-LiPA) (Martinez et al. 2004). Then, we resurrected by *Escherichia coli* expression (Ayuso-Fernández et al. 2017) the main nodes in the evolutionary line (**Fig. 1**) and characterized the degradative abilities of these ancestral enzymes using polymeric lignin as substrate, as described in the next sections.

Previously, to calibrate the age of the ancestral peroxidases in **Fig. 1**, we built an ultrametric tree with PATHd8 (Britton et al. 2007), which uses mean path lengths from the leaves of a tree with a correction of the

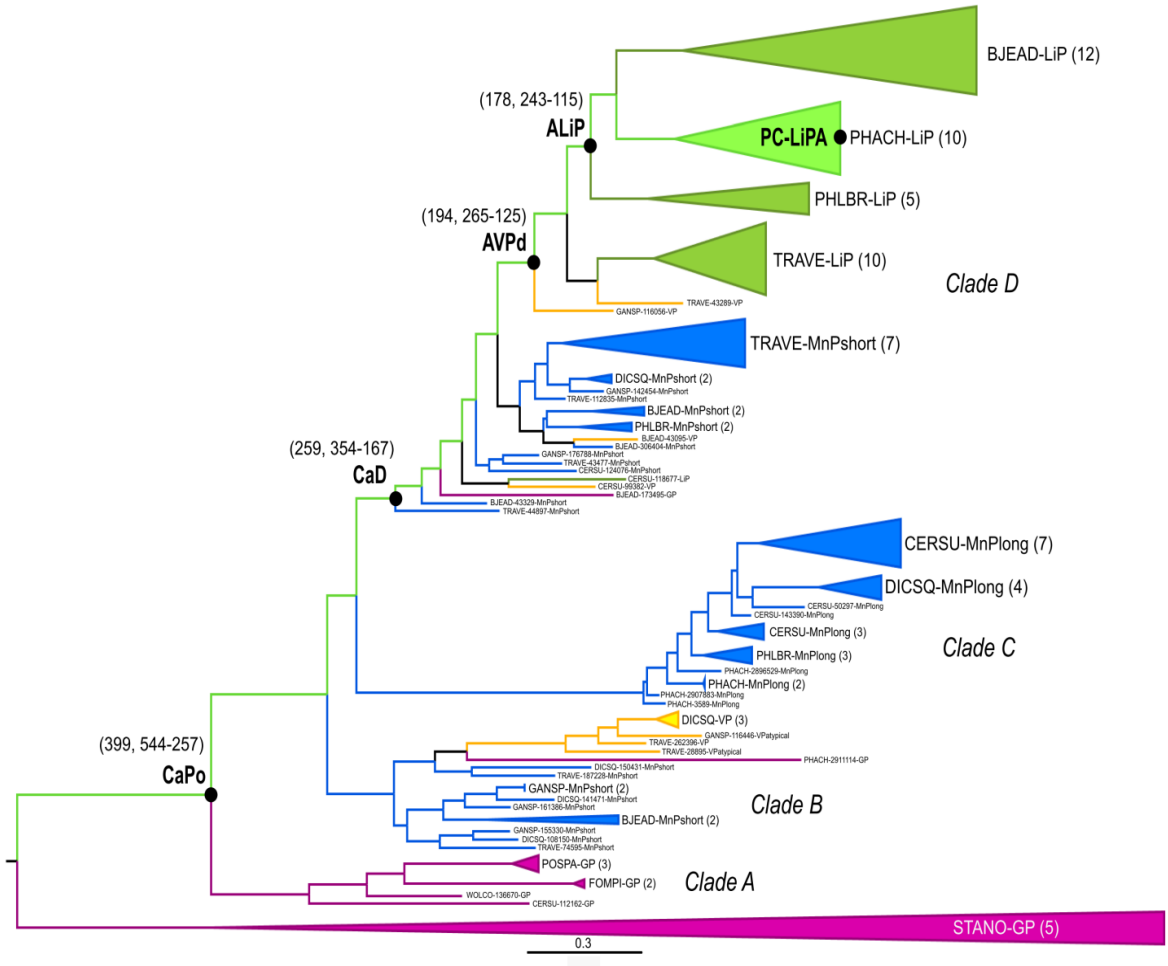


Fig. 1. Evolution of Polyporales ligninolytic peroxidases from ten selected sequenced genomes (2018b), including the *Stagonospora nodorum* generic peroxidases (GPs) for tree rooting and time calibration. Collapsed clusters are shown in dark green (LiPs), yellow (VPs), blue (MnPs) and purple (GPs) with indication of the number of sequences in each of them. The path to the extant LiPA of *P. chrysosporium* is marked in green, with indication of the resurrected enzymes analyzed in this work (common ancestor of Polyporales peroxidases, CaPo; common ancestor of clade D, CaD; ancestral versatile peroxidase in clade D, AVPd; and ancestral LiP, ALiP) and their ages (average and limits in mya from **Fig. S1**). The sequence labels start with the species code (BJEAD, *Bjerkandera adusta*; CERSU, *Ceriporiopsis subvermispora*; DICSQ, *Dichomitus squalens*; FOMPI, *Fomitopsis pinicola*; GANSP, *Ganoderma* sp; PHLBR *Phlebia brevispora*; PHACH, *Phanerochaete chrysosporium*; POSPL, *Postia placenta*; STANO, *S. nodorum*; TRAVE, *Trametes versicolor*; and WOLCO, *Wolfiporia cocos*) followed by the JGI ID# and the peroxidase type (including GP, LiP, MnP-short, MnP-long, VP, and VP-atypical).

deviation from a molecular clock by smoothing substitution rates locally (Supplementary **Fig. S1**). As time constraints for PATHd8 we used the node calibration obtained by Floudas et al (Floudas et al. 2012) that could be

applied to our phylogeny. The ages obtained show that the evolution of ligninolytic peroxidases started 399 ± 140 mya with CaPo, which most probably was the ancestor of all Agaricomycetes class-II peroxidases (Floudas et al. 2012; Ayuso-Fernández et al. 2018b). CaPo had a Mn^{2+}-binding site for oxidation of this metal cation, but also oxidized substrates in direct contact with the heme. From CaPo, the lineage to extant LiPs did not experiment significant changes during 200 my, with CaD (259 ± 94 mya) still being a MnP-type enzyme, until the appearance of the catalytic tryptophan in AVPd 194 ± 70 mya. AVPd retained the Mn^{2+} oxidation site together with the new solvent-exposed tryptophan putatively responsible for nonphenolic lignin oxidation. Then, a quick change, losing the Mn^{2+}-binding site, occurred in a few my originating the first LiP of the phylogeny (ALiP, 178 ± 64 mya) that further led to the extant LiPs of Polyporales maintaining the oxidation sites architecture. This last evolutionary step resulted in a highly-specialized enzyme, whose action on nonphenolic lignin is not non-competively inhibited by Mn^{2+}, as in the case of VPs.

Kinetics of lignin oxidation through peroxidase evolution. For better biological interpretation of peroxidase evolution, we analyzed their catalytic cycle with lignosulfonates as enzyme reducing substrates. This cycle (Supplementary **Fig. S2**) includes two-electron activation of the resting state (RS) enzyme by H_2O_2 forming compound I (CI), which is reduced back to RS *via* a one-electron oxidized compound II (CII) yielding two product molecules/ions.

The different UV-visible spectra of these intermediates make possible to follow the cycle reactions by stopped-flow rapid spectrophotometry. The CI/CII and CII/RS reduction kinetic curves were obtained and the corresponding transient-state constants for hardwood and softwood lignins were estimated (**Table 1**) with dissociation (K_D) and first-order rate (k) constants being calculated when k_{obs} exhibit saturation behavior (as illustrated in **Fig. 2** and Supplementary **Fig. S3**).

For all enzymes, the CII/RS reduction is the rate-limiting step of the catalytic cycle, with CaPo and CaD (the most ancestral enzymes lacking a solvent-exposed tryptophan) having significantly slower rates that the enzymes with the catalytic tryptophan (AVPd, ALiP and PC-LiPA) (**Fig. 3**).

For softwood lignin there is a progressive increase of k_{3app} through evolution, but the differences between enzymes do not exceed 13-fold increase. However, for hardwood lignin there is a huge boost in k_{3app}, increasing near 50-fold from CaPo to AVPd, when the solvent-exposed tryptophan appeared for the first time. These increments can be interpreted in terms of the dissociation constant (K_{D3}) indicating a greater affinity for the polymer in the enzymes with the solvent exposed tryptophan (as

illustrated in **Fig. 2B**). Therefore, a change in lignin preference was produced through peroxidase evolution: the ratio between hardwood and softwood lignosulfonate k_{3app} being lower than 1 for the enzymes lacking the catalytic tryptophan and greater (up to 7) for those incorporating that residue.

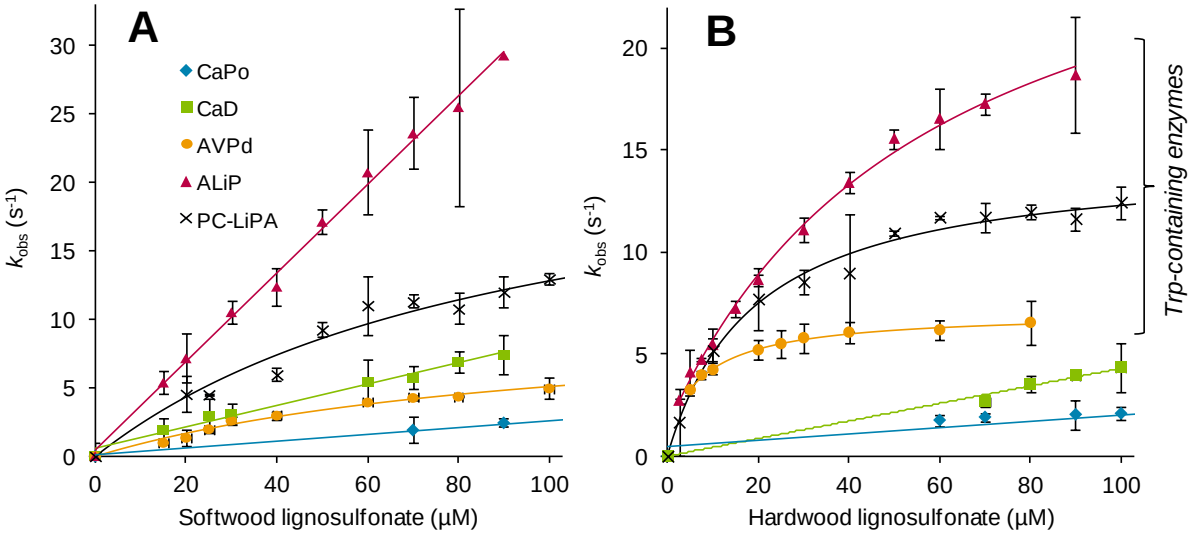


Fig. 2. Kinetics of reduction of CII of ancestral (CaPo, CaD, AVPd and ALiP) and extant (PC-LiPA) peroxidases by softwood (**A**) and hardwood (**B**) lignosulfonates. The lignosulfonate concentrations refers to the basic phenylpropanoid unit. Means and 95% confidence limits. More efficient enzymes containing a catalytic tryptophan residue are indicated in **B**.

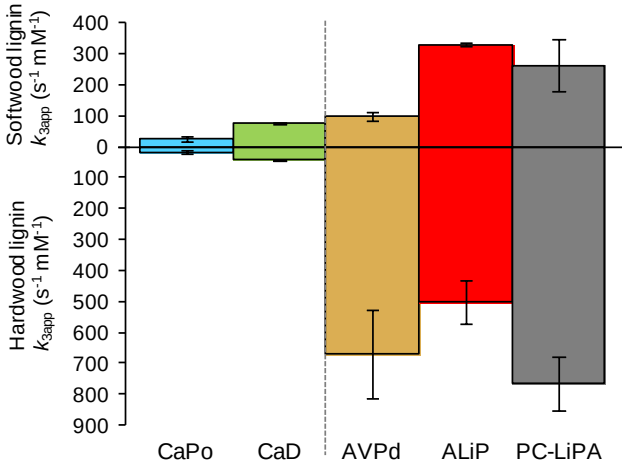


Fig. 3. Comparison of rate-limiting step constant (k_{3app}) in lignin stopped-flow transient-state kinetic experiments with ancestral and extant peroxidases showing a shift from softwood to hardwood lignin preference with the appearance of the catalytic tryptophan in AVPd (dashed line).

Additionally, the transient-state kinetics for Mn^{2+} oxidation (**Table 1**) showed an increase of the Mn^{2+}-oxidation activity till AVPd, being lost in ALiP, in agreement with the evolution of catalytic sites mentioned above.

Table 1. Kinetic constants for CI/CII and CII/RS reduction in ancestral (CaPo, CaD, AVPd and ALiP) and extant (PC-LiPA) peroxidases by softwood and hardwood lignosulfonates (LS) and Mn^{2+}.[a]

			CaPo	CaD	AVPd	ALiP	PC-LiPA
CI/CII	Soft-wood LS	K_{D2} (μM)	46±3	57±3.9	46±3.3	44±15	23±1
		k_2 (s^{-1})	9±0	20±1	19±1	42±9	33±2
		k_{2app} ($s^{-1}\cdot mM^{-1}$)	197±14	351±34	413±33	955±385	1440±120
	Hard-wood LS	K_{D2} (μM)	-	-	8±1	40±5	8±2
		k_2 (s^{-1})	-	-	18±0.6	27±1.4	25±0.9
		k_{2app} ($s^{-1}\cdot mM^{-1}$)	86±18	202±7	2280±280	674±97	2940±740
	Mn^{2+}	K_{D2} (μM)	-	-	-	na[c]	na
		k_2 (s^{-1})	-	-	-	0	0
		k_{2app} ($s^{-1}\cdot mM^{-1}$)	~11800	~19000	~34600	0	0
CII/RS	Soft-wood LS	K_{D3} (μM)	-[b]	-	102±12	-	95±326
		k_3 (s^{-1})	-	-	10±1	-	25±4
		k_{3app} ($s^{-1}\cdot mM^{-1}$)	25±8	78±4	99±14	331±4	263±83
	Hard-wood LS	K_{D3} (μM)	-	-	9±2	58±8	19±2
		k_3 (s^{-1})	-	-	6±0	29±1	14±0
		k_{3app} ($s^{-1}\cdot mM^{-1}$)	16±3	43±1	670±145	500±70	764±86
	Mn^{2+}	K_{D3} (μM)	-	-	62±19	na	na
		k_3 (s^{-1})	-	-	44±9	0	0
		k_{3app} ($s^{-1}\cdot mM^{-1}$)	355±7	530±5	700±250	0	0

[a] Means and 95% confidence limits from reactions at 25 °C, in 0.1 M tartrate, pH 3 for lignosulfonates and pH 5 for Mn^{2+}. [b]-, not determined because saturation was not reached. [c] na, no activity.

Steady-state treatment of lignins. To further explore the evolution of lignin degradation, the structural modification of guaiacyl and syringyl-guaiacyl lignins from softwood and hardwood, respectively, were analyzed by heteronuclear single quantum correlation (HSQC) 2D-NMR spectroscopy. The HSQC signals of the main aromatic units and side-chain interunit linkages in hardwood (**Fig. 4**) and softwood (Supplementary Fig. S4) lignosulfonates included: sulfonated guaiacyl units (**G**), and Cα-sulfonated (**S**), non-sulfonated (***S***) and Cα-oxidized (**S'**) syringyl units, together with Cα-sulfonated (**A**) and non-sulfonated (***A***) ß-*O*-4' substructures and less abundant phenylcoumaran (***B***) and resinol (***C***) substructures. A semi-quantitative analysis of these signals was performed to obtain: i) the lignin composition referred to 100 aromatic units, and S/G ratio; and ii) the lignin

(aromatic signal) decrease referred to the control treatment without enzymes (Supplementary Table S1).

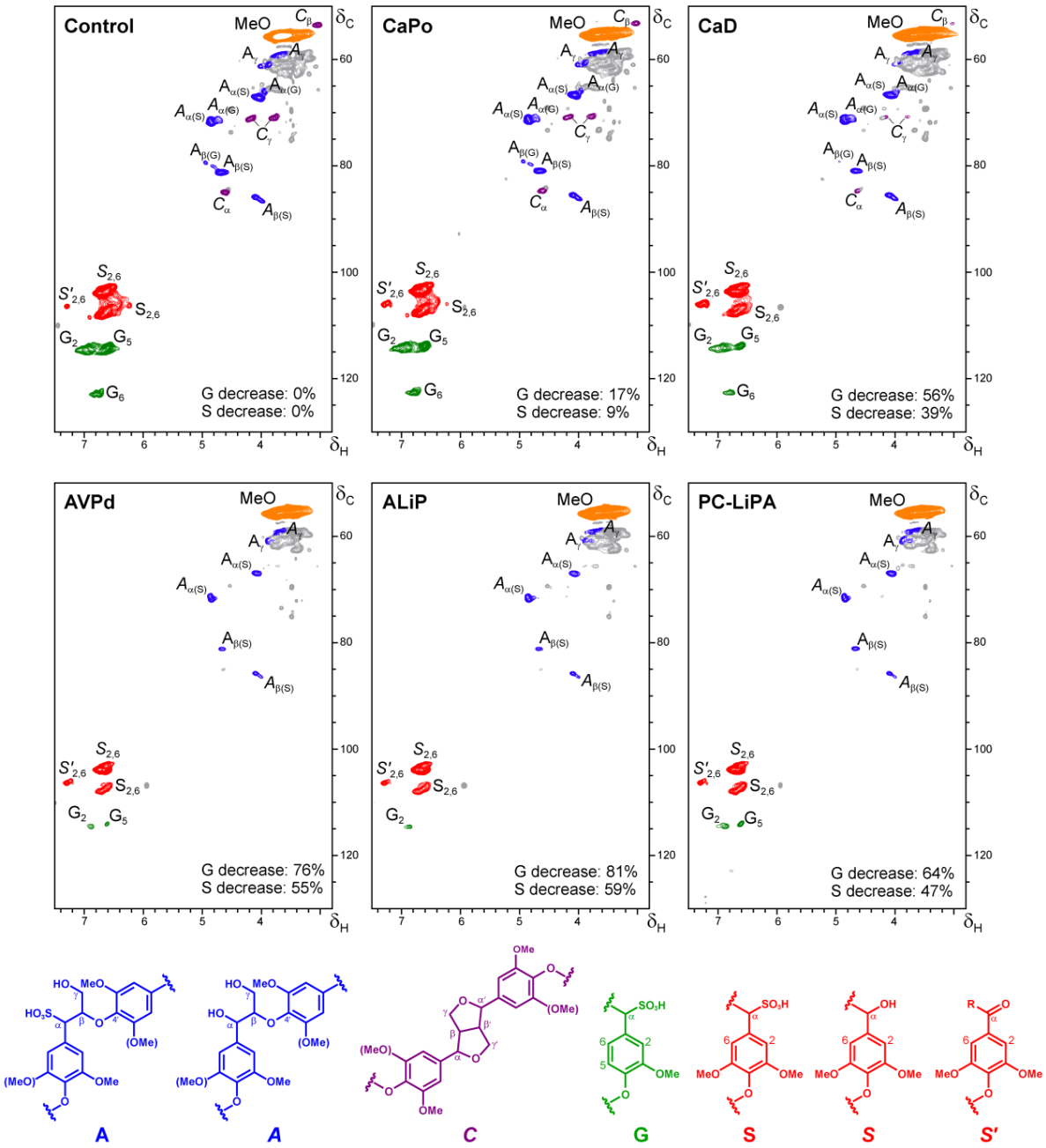


Fig. 4. HSQC NMR spectra of hardwood lignosulfonate treated for 24 h with ancestral (CaPo, CaD, AVPd and ALiP) and extant (PC-LiPA) peroxidases, and control without enzyme. The formulae of the main structures identified are included below, and the decrease of G and S units (with respect to the control) are indicated in each panel. Signals correspond to ^{1}H-^{13}C correlations at the different positions of Cα-sulfonated guaiacyl units (**G**), Cα-sulfonated, non-sulfonated and Cα-oxidized syringyl units (**S**, ***S*** and **S'**, respectively), Cα-sulfonated and non-sulfonated ß-*O*-4′ substructures (**A** and ***A***, respectively) and resinol (***C***) substructures, and methoxyls (**MeO**).

24-h treatments of softwood lignosulfonate with the ancestral enzymes and PC-LiPA did not reveal strong differences between them and compared to the enzyme-free control (**Fig. S4**). However, the same treatments on hardwood lignosulfonate resulted in significant changes in the 2D-NMR spectra (**Fig. 4** and **Table S1**), a fact related to the peroxidase evolution to act on angiosperm lignin, discussed below. Integrating the aromatic signals in the hardwood lignosulfonate spectra (as an estimation of the amount of lignin) we observe a decrease through evolution, being the maximal change produced by ALiP, with 59% decrease of syringyl units, and 81% decrease of guaiacyl units. This is accompanied by an overall increase (up to 2.1-fold) of the S/G ratio. Concerning side-chain signals, no strong changes in their relative abundance (per 100 aromatic units) were observed.

The progressive decrease of lignin signals in the steady-state treatments through peroxidase evolution supports the trend observed in stopped-flow measurements. Interestingly, in both cases, the strongest change coincided with the evolutionary appearance of the surface catalytic tryptophan.

Polyporales specialization for plant hosts. Using the USDA Fungus-Host Distribution Database (https://nt.ars-grin.gov/fungaldatabases/fungushost/fungushost.cfm) we obtained a list of plant hosts for the ten Polyporales species analyzed. This information was compared with the number and type of peroxidase genes in the genomes of the same fungi, extracted from the JGI Mycocosm database (https://genome.jgi.doe.gov/mycocosm/home) (Supplementary Table S2). With the only exception of *P. brevispora*, with only three citations, a correlation could be found between the specialization towards angiosperm wood decay and the number of peroxidases with the solvent exposed catalytic tryptophan (VPs or LiPs).

Discussion

Lignin-degrading fungi are fundamental pieces of the land biogeochemical cycles, recycling the carbon stored in lignin, and opening the way for biodegradation of plant polysaccharides protected by recalcitrant lignin (Martínez et al. 2018). Moreover, these organisms and their enzymes are being studied in the bio-based industries, to implement greener technologies for modification of lignin and use of cellulosic sugars/fibers and simple aromatics for value-added products and materials in the biorefinery context (Martínez et al. 2009; Ragauskas et al. 2014).

Genomic studies from 2004 (Martinez et al. 2004) have demonstrated the presence of ligninolytic peroxidase (LiP, VP and/or MnP) genes in all typical white-rot fungi sequenced (Floudas et al. 2012; Ruiz-Dueñas et al. 2013; Nagy et al. 2016). More recently, some aspects of their evolution have been

addressed by ancestral sequence reconstruction and characterization of the resurrected ancestors using high and low redox-potential substrates, redox potential estimation and other techniques (Ayuso-Fernández et al. 2017; Ayuso-Fernández et al. 2018a; Ayuso-Fernández et al. 2018b). Here, we use for the first time softwood and hardwood lignin as substrate to analyze the peroxidase evolution leading to the most efficient ligninolytic enzymes, Polyporales LiPs (Cullen and Kersten 2004).

Initial reaction rates of lignin oxidation cannot be obtained under steady-state conditions, but electron transfer was accurately estimated from the reduction of peroxide-activated enzymes under stopped-flow conditions, as recently reported for extant VP (Sáez-Jiménez et al. 2015a; Sáez-Jiménez et al. 2016). Considering the rate-limiting step (CII reduction to RS) the constants for the enzymes containing the catalytic tryptophan (AVPd, ALiP and PC-LiPA) were higher for both types of lignin. This is in agreement with the decreased electron-transfer rate observed for tryptophan-less VP variants (Sáez-Jiménez et al. 2015a). As shown by stopped-flow spectrophotometry, uniquely interesting is the change of specificity when the solvent-exposed tryptophan appeared in AVPd, revealing a switch from better direct oxidation of softwood lignin by the oldest CaPo and CaD ancestors to hardwood lignin since then.

Long-term (steady-state) treatments of lignosulfonates through peroxidase evolution were analyzed by state-of-the-art 2D NMR (Ralph et al. 1999; Ralph and Landucci 2010; Mansfield et al. 2012). Using this technique, enzymatic modification of lignin was observed, as previously reported for extant VP (Sáez-Jiménez et al. 2016). Especially informative are the treatments of hardwood lignosulfonates with the resurrected peroxidase lineage, revealing that lignin degradation increases through evolution being higher in the enzymes containing the solvent exposed catalytic tryptophan, in agreement with the stopped-flow measurements.

During evolution of vascular plants, the route to syringyl units appeared before the diversification of angiosperms, which started 250-140 mya (Morris et al. 2018), although convergent synthesis of syringyl lignin has been demonstrated in basal plants lineages (Weng and Chapple 2010). Lignin containing S-units has mechanical and physiological evolutionary advantages (Li et al. 2001; Menden et al. 2007), together with higher complexity (Ralph et al. 2004) and lower average phenolic content (Lai and Guo 1991; Camarero et al. 1999) reducing biodegradability. Due to the apparent implication of the above surface tryptophan in the oxidation of nonphenolic lignin, we calibrated its evolutionary appearance. Our time calibration suggests that AVPd, the first peroxidase with this exposed tryptophan, appeared roughly 200 mya, concurrent with angiosperm diversification.

Further studies are required to establish a detailed co-evolution between fungi and plants, but our results show how saprotrophic fungi followed lignin evolution in plants incorporating a catalytic tryptophan (Ayuso-Fernández et al. 2018b) responsible for oxidation of the more complex and less phenolic lignin developed by angiosperms. This is coherent with the correlation existing between the number of Polyporales ligninolytic peroxidases with a solvent-exposed catalytic tryptophan (VPs and LiPs oxidizing nonphenolic lignin) (Sáez-Jiménez et al. 2016) and the frequency of angiosperms as their host. Recent character reconstruction based on 5284 mushroom-forming species determines that the ancestor of Polyporales had mainly gymnosperms as hosts (Varga et al. 2019), and during fungal evolution, white-rot fungi specialized in angiosperms while the brown-rot fungi, without ligninolytic peroxidases, became more generalist (Krah et al. 2018). Our results analyzing ancestral ligninolytic peroxidases, and their behavior using real lignin substrates, support these conclusions, and point to a fascinating co-evolution between fungi and plants.

Acknowledgements
This work was supported by the EnzOx2 (H2020-BBI-PPP-2015-2-720297, https://www.enzox2.eu) European project and the BIO2017-86559-R and AGL2017-83036-R Spanish projects. The work conducted by JGI was supported by the Office of Science of the US Department of Energy under Contract DE-AC02-05CH11231. The authors thank Guro E. Fredheim (Borregaard AS, Sarpsborg, Norway) for providing the lignosulfonate preparations, and Manuel Angulo (CITIUS, University of Seville) for the NMR analyses.

Materials and methods
118 peroxidase sequences were obtained from the genomes of ten Polyporales (Ruiz-Dueñas et al. 2013) and the ascomycete *S. nodorum* sequenced at DOE JGI and available at JGI Mycocosm (https://genome.jgi.doe.gov/mycocosm). PAML 4.7 (Yang 2007) was used for ancestral sequence reconstruction, using the phylogeny from RAxML (Stamatakis et al. 2008). PATHd8 (Britton et al. 2007) was used for time calibration of an ultrametric phylogenetic tree, using the diversification of Pezizomycotina (631 mya, 533-762 mya), the internal diversification of *Stagonospora nodorium* GPs (403 mya, 276-533 mya) and the diversification of the Antrodia clade peroxidases (399 mya, 446-207) (Floudas et al. 2012) as time constrainst.

Four reconstructed ancestral peroxidase sequences and extant PC-LIPA were expressed in *E. coli* and activated *in vitro* (Ayuso-Fernández et al. 2017). Then, transient-state kinetic constants were estimated by stopped-flow rapid spectrophotometry (at 25ºC in 0.1 M tartrate, pH 3 for lignosulfonate and pH 5 for Mn^{2+}) using softwood and hardwood lignosulfonates, whose concentrations were referred to the basic phenylpropanoid unit (260 and 290 Da, respectively). Lignosulfonate modification in steady-state treatments (24 h, at 25ºC in 50 mM

phosphate, pH 5) were analyzed by HSQC 2D-NMR, after spectra normalization with respect to the signal at δ_H/δ_C 2.49/39.5 ppm (not shown) corresponding to the non-deuterated fraction of the dimethylsulfoxide used as NMR solvent. See SI Appendix for details.

SI Appendix

Peroxidase evolution in white-rot fungi follows wood lignin evolution in plants

Iván Ayuso-Fernández‡, Jorge Rencoret§, Ana Gutiérrez§, Francisco Javier Ruiz-Dueñas‡ and Angel T. Martínez‡

‡Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040 Madrid, Spain; and §Instituto de Recursos Naturales y Agrobiología de Sevilla, CSIC, PO Box 1052, E-41080 Seville, Spain

The SI Appendix includes Supporting Materials and Methods, Tables S1 and S2, Figures S1-S4, and References.

Supporting Materials and Methods

Phylogenetic analysis and ancestral enzyme reconstruction. Every class-II peroxidase (113 sequences) annotated in the genomes of the following Polyporales, available at the DOE JGI Mycocosm portal, were used in the present study:
- *B. adusta* (http://genome.jgi.doe.gov/Bjead1_1/Bjead1_1.home.html) (Binder et al. 2013)
- *C. subvermispora* (http://genome.jgi.doe.gov/Cersu1/Cersu1.home.html (Fernández-Fueyo et al. 2012)
- *D. squalens* (http://genome.jgi.doe.gov/Dicsq1/Dicsq1.home.html) (Floudas et al. 2012)
- *F. pinicola* (http://genome.jgi.doe.gov/Fompi3/Fompi3.home.html) (Floudas et al. 2012)
- *Ganoderma* sp. (http://genome.jgi.doe.gov/Gansp1/Gansp1.home.html) (Binder et al. 2013)
- *P. brevispora* (http://genome.jgi.doe.gov/Phlbr1/Phlbr1.home.html) (Binder et al. 2013)
- *P. chrysosporium* (http://genome.jgi.doe.gov/Phchr2/Phchr2.home.html) (Ohm et al. 2014)
- *P. placenta* (http://genome.jgi.doe.gov/Pospl1/Pospl1.home.html) (Martinez et al. 2009)
- *T. versicolor* (http://genome.jgi.doe.gov/Trave1/Trave1.home.html) (Floudas et al. 2012) and

- *W. cocos* (http://genome.jgi.doe.gov/Wolco1/Wolco1.home.html) (Floudas et al. 2012).

For tree rooting and time calibration, we used the GP sequences of the ascomycete *S. nodorum* available at JGI Mycocosm (http://genome.jgi.doe.gov/Stano2/Stano2.home.html) (Hane et al. 2007; Floudas et al. 2012).

The amino-acid sequences were aligned with MUSCLE as implemented in MEGA 7 (Kumar et al. 2016). The sequences were tested with ProtTest (Abascal et al. 2005) to determine the evolutionary model that best fits the data for maximum likelihood (ML) analysis, among 64 empirical models of evolution. A ML phylogeny was then constructed using RAxML (Stamatakis et al. 2008), under the Whelan and Goldman (2001) model of evolution using gamma-distributed rate of heterogeneity (gamma shape with 4 rates of categories = 1.33) with empirical amino-acid frequencies and invariant sites (proportion of invariant sites = 0.073) (WAG+I+G+F).

The PAML 4.7 package (Yang 2007) was used to obtain the most probable sequence at each node of the phylogeny and the posterior amino-acid probability per site in each ancestor. We used the WAG model of evolution, and the previously obtained ML phylogeny and MUSCLE alignment as inputs for the software. Both joint and marginal reconstructions were performed and the most probable sequences of the marginal reconstruction were selected and manually corrected for C-terminal and other insertions or deletions according to the sequences of the ancestor progeny. The genes were synthesized by ATG:biosynthetics (Merzhausen, Germany) using the most frequent codons for high expression in *E. coli*.

Time calibration of ancestral nodes in Polyporales phylogeny. To determine the ages of the ancestral nodes, the peroxidase phylogenetic tree was converted in an ultrametric tree using the recognized non-parametric method implemented in PATHd8 (Britton et al. 2007). The algorithm in PATHd8 only considers the mean path lengths from a node to its descendants and addresses deviations from the molecular clock locally. We used as time constraints those obtained by Floudas *et al.* (Floudas et al. 2012) using Bayesian methods corresponding to: i) the diversification of Pezizomycotina from the rest of the phylogeny (631 mya with 533-762 mya maximum deviations); ii) the internal diversification of *S. nodorium* GPs (403 mya with 276-533 mya maximum deviations); and iii) the diversification of the Antrodia clade peroxidases (formed by *F. pinicola, P. placenta* and *W. cocos*) from the rest of Polyporales peroxidases (399 mya with 446-207 mya maximum deviations).

***E. coli* expression and isolation of the resurrected enzymes.** After gene synthesis, the coding DNA sequences of the selected ancestors (CaPo, CaD, AVPd and ALiP) and the extant PC-LiPA (JGI ID# 2989894) were cloned into pET23b(+) (Novagen). The resulting plasmids were transformed into BL21(DE3)pLysS (ancestors) or W3110 (PC-LiPA) as reported by Pérez-Boada et al (2002). The apoenzymes accumulated in inclusion bodies and, after solubilization in 8 M urea, the *in vitro* activation conditions of the ancestral proteins were optimized - including 0.16 M urea, 5 mM $CaCl_2$, 15 μM hemin, 0.4 mM oxidized glutathione, 0.1 mM dithiothreitol and 0.1 mg/mL of protein in 50 mM Tris-HCl, pH 9.5 - while those reported by Doyle and Smith (1996) where used for PC-LiPA. The active enzymes were purified with a Resource-Q column (GE-Healthcare) using a 0-400 mM NaCl salt gradient, 2 mL/min flow, in 10 mM sodium tartrate, pH 5.5, containing 1 mM of $CaCl_2$.

Softwood and hardwood lignins. Water solubility is required for the transient kinetic studies, described below. Therefore, to represent natural lignin, two water-soluble sulfonated lignins were used: softwood (*Picea abies*) and hardwood (*Eucalyptus grandis*) lignosulfonates provided by G. E. Fredheim (Borregaard AS, Sapsborg, Norway). Lignosulfonates were prepared for reactions according to Sáez-Jiménez *et al*, (Sáez-Jiménez et al. 2016) including dialysis in 10 mM EDTA, 50 mM Tris (pH 8) to remove Mn^{2+} traces, and finally in Milli-Q water twice.

Enzyme (transient-state) kinetics. The reduction of peroxidases CI (to CII) and CII (to RS) was evaluated using hardwood and softwood lignosulfonates or Mn^{2+} as enzyme reducing substrates in a stopped-flow rapid spectrophotometry equipment (Bio-Logic, Claix, France) with a three syringe module (SFM300) synchronized to a diode array detector (J&M, Essingen, Germany), and BioKine software. All reactions were carried out in 0.1 M tartrate (pH 3 for lignosulfonates or pH 5 for Mn^{2+}).

Specifically, the reduction of CI was analyzed by mixing the enzyme (1 μM final concentration) with H_2O_2 (1-8 molar equivalents, depending on the enzyme) for 3 s, observing total CI formation. Next, different concentrations of either lignosulfonates or Mn^{2+} in 0.1 M tartrate (at the adequate pH) were added, and the reactions were followed at 416 nm, the isosbestic point of CII and the resting state of peroxidases. CII reduction was studied by mixing a solution containing the enzyme and ferrocyanide (both at 1 μM final concentration) with H_2O_2 (1-8 molar equivalents, depending on the enzyme) to ensure the formation of CII after 6 s. After that, different concentrations of lignosulfonates or Mn^{2+} in 0.1 M tartrate (at the adequate pH) were

added, following the reaction at 410 nm, the maximum of the resting state for the peroxidases analyzed.

All the traces showed single-exponential character from which pseudo first-order rate constants (k_{2obs} and k_{3obs} for CI and CII reduction, respectively) were calculated. Plots of k_{2obs} and k_{3obs} *vs* substrate concentration fitted to linear, hyperbolic or sigmoid models. From those kinetics that fitted to a linear model, apparent second-order rate constants (k_{2app} and k_{3app} for CI and CII reduction, respectively) were obtained. Plots of k_{obs} *vs* substrate concentration that fitted to a Michaelis-Menten or sigmoid models yielded dissociation constants of the CI-lignin and CII-lignin or CII-Mn^{2+} complexes (K_{D2} and K_{D3} respectively) and first-order rate constants (k_2 and k_3, respectively). The corresponding apparent second-order rate constants, k_{2app} (k_2/K_{D2}) and k3app (k_3/K_{D3}), were calculated with the equation: $k_{obs} = (k/K_D)[S]/(1 + [S]/K_D)$, where [S] indicates substrate concentration. Despite replicated experiments (> 5 replica) measurements of CI/CII reduction resulted in high error values due to the intrinsic instability of CI, together with the polydisperse nature of the lignosulfonate preparations (being higher in the softwood lignosulfonate). For comparison, the lignosulfonate concentrations in these and other experiments were referenced to the basic phenylpropanoid unit corresponding to 260 and 290 Da in the sulfonated softwood and hardwood lignins, respectively (Sáez-Jiménez et al. 2015a).

Lignin treatment under steady-state conditions. To analyze how ancestral peroxidases and extant PC-LiPA modify lignin, hardwood and softwood lignosulfonates (12 g L^{-1}) were treated at 25 ºC over 24 h in 50 mM phosphate (pH 5). The concentration of enzymes were 1 μM, added in two doses, with a H_2O_2 final concentration of 10 mM, added continuously with a syringe pump. Control treatments were performed under the same conditions but without enzyme. It is important to note that although peroxidases exhibit their highest activity at pH 3 (used in stopped-flow measurements), reactions over 24 h were carried out at pH 5 to maintain the enzymes active.

NMR analyses. The treated lignosulfonates and the corresponding controls were freeze-dried for NMR analyses. Spectra were recorded at 25 ºC on a Bruker Avance-III 500 MHz instrument equipped with a cryogenically cooled 5 mm TCI gradient probe with inverse geometry. The lignosulfonate samples (40 mg initial weight, before treatments) were dissolved in 0.75 mL of deuterated dimethylsulfoxide (DMSO-d_6). The central peak of residual non-deuterated DMSO was used as the internal reference (at δ_H/δ_C 2.49/39.5 ppm), and the spectra were normalized to the same intensity of the DMSO

signals, since the same DMSO volume and initial amount of lignin was used in all the cases.

The HSQC experiment used Bruker´s "hsqcetgpsisp.2" adiabatic pulse program with spectral widths from 0 to 10 ppm (5000 Hz) and from 0 to 165 ppm (20625 Hz) for the ^{1}H and ^{13}C dimension. The $^{1}J_{CH}$ used was 145 Hz. Processing used typical matched Gaussian apodization in the ^{1}H dimension and squared cosine-bell apodization in the ^{13}C dimension. Prior to Fourier transformation, the data matrices were zero-filled to 1024 points in the ^{13}C dimension. Signals were assigned by comparison with the literature (Lutnaes et al. 2008; Martínez et al. 2008; Lebo et al. 2008; Ralph et al. 2009; Prasetyo et al. 2010; Magina et al. 2015; Sáez-Jiménez et al. 2015b).

In the aromatic region of the spectrum, the H_2-C_2, H_5-C_5 and H_6-C_6 correlation signals were integrated to estimate the amount of lignins and the S/G ratio. In the aliphatic oxygenated region, the signals of methoxyls, and Hß-Cß (or Hα-Cα) correlations in the side chains of sulfonated and non-sulfonated ß-*O*-4', phenylcoumaran and resinol substructures were integrated. The intensity corrections introduced by the adiabatic pulse program permits to refer the latter integrals to the previously obtained number of lignin units.

Supporting Tables and Figures

Table S1. Softwood and hardwood lignin decrease by ancestral (CaPo, CaD, AVPd and ALiP) and extant (PC-LiPA) peroxidases referred to enzyme-free control treatment, composition including sulfonated (**A**) and nonsulfonated ß-*O*-4' (***A***), phenylcoumaran (***B***) and resinol (***C***) substructures and oxidized syringyl units (**S'**), and S/G ratio (only in hardwood lignin). From integration of 2D-NMR signals in **Figs 4** and **S4**.

	Control	CaPo	CaD	AVPd	ALiP	PC-LiPA
Softwood lignin						
Decrease [a]	-	8%	13%	14%	25%	19%
Composition: [b]						
- ß-*O*-4' (**A**)	32	30	30	28	29	29
- ß-*O*-4' (***A***)	3	3	3	3	3	3
- Phenylcoumaran (***B***)	4	4	4	4	4	4
- Resinol (***C***)	4	4	4	5	5	4
Hardwood lignin						
Decrease: [a]						
- G decrease	-	17%	56%	76%	81%	64%
- S decrease	-	9%	39%	55%	59%	47%
- Total	-	11%	43%	61%	65%	51%
Composition: [b]						
- ß-*O*-4' (**A**)	20	20	20	18	19	19
- ß-*O*-4' (***A***)	12	13	15	17	16	14
- Resinol (***C***)	10	11	10	9	9	8
- S'	3	6	10	11	10	10
S/G ratio	2.9	3.1	4.0	5.4	6.0	4.2

[a]From aromatic signals referred to the control (%). [b] Referred to 100 aromatic units in the sample.

Table S2. Distribution of plant host for the Polyporales analyzed (*left*) and number and types of peroxidases annotated in their genomes, including percentage of peroxidases with the surface catalytic tryptophan (*right*).[a]

	Host (citations and %)		Peroxidase type				
	Gymnosperms	Angiosperms	GP	MnP	VP	LiP	Surface tryptophan
P. placenta	7 (78%)	2 (22%)	3	0	0	0	0
W. cocos	27 (59%)	19 (41%)	1	0	0	0	0
F. pinicola	126 (61%)	81 (39%)	2	0	0	0	0
C. subvermispora	8 (62%)	5 (38%)	1	13	1	1	13%
D. squalens	46 (96%)	2 (4%)	0	9	3	0	25%
Ganoderma sp.	2 (4%)	45 (96%)	0	6	2	0	25%
T. versicolor	78 (17%)	376 (83%)	0	12	3	10	52%
B. adusta	33 (16%)	176 (84%)	1	6	1	12	65%
P. chrysosporium	0	5 (100%)	1	5	0	10	63%
P. brevispora	3 (100%)	0	0	5	0	5	50%

[a]From USDA Fungus-Host Distribution Database (https://nt.ars-grin.gov/fungaldatabases/fungushost/fungushost.cfm) and JGI Mycocosm (https://genome.jgi.doe.gov/mycocosm), respectively

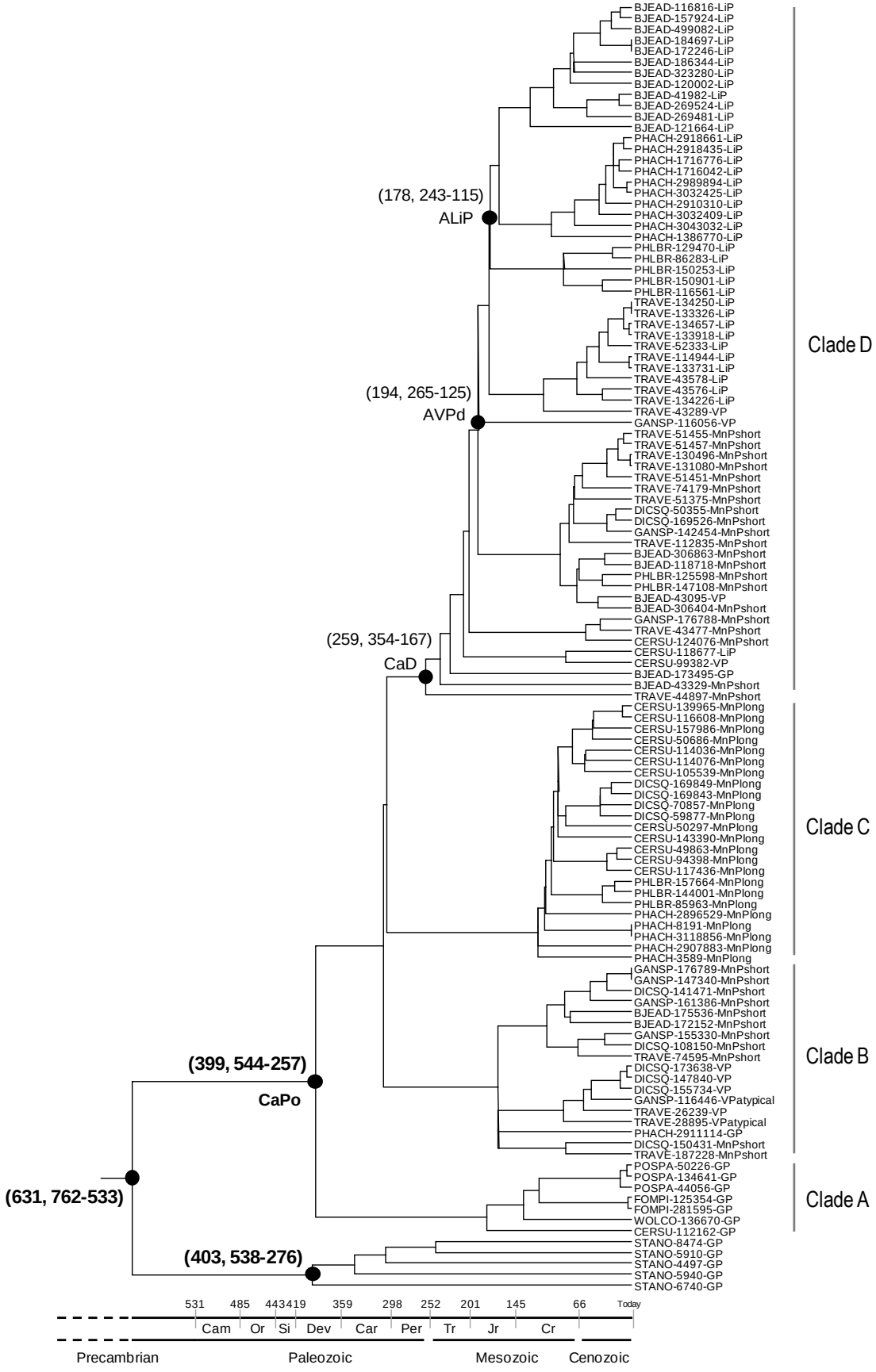


Fig. S1. Ultrametric tree obtained for calibration of the ancestral nodes in the present study. We used the phylogeny of Polyporales peroxidases rooted with *S. nodorum* GPs and the software PATHd8 with the time constraints (average and limits in mya) marked in bold.

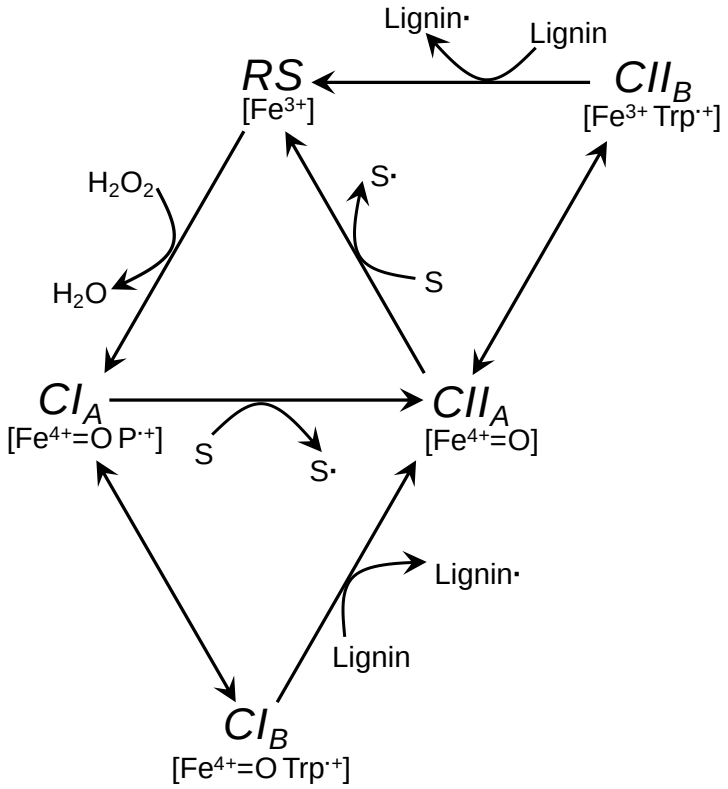


Fig. S2. Catalytic cycle of ligninolytic peroxidases. The common cycle is initiated with the oxidation of the RS enzyme by H_2O_2 to form CI. Two consecutive one-electron oxidations of substrate (S) lead to CII and RS, successively. The expanded cycle for VPs and LiPs is also shown, with the displacement of one oxidation equivalent to a solvent-exposed tryptophan able to attack nonphenolic lignin.

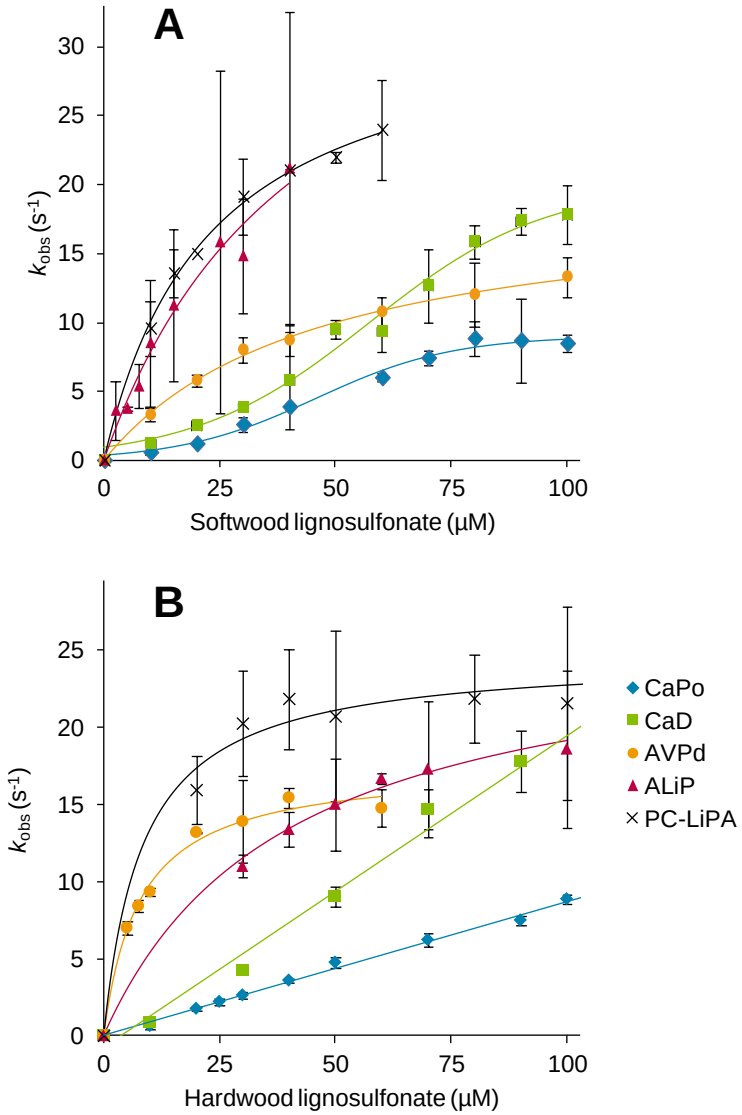


Fig. S3. Kinetics of reduction of CI of ancestral (CaPo, CaD, AVPd and ALiP) and extant (PC-LiPA) peroxidases by softwood (**A**) and hardwood (**B**) lignosulfonates. The lignosulfonates concentrations refer to the basis phenylpropanoid unit (260 and 290 Da for sulfonated softwood and hardwood lignins, respectively). Means and 95% confidence limits.

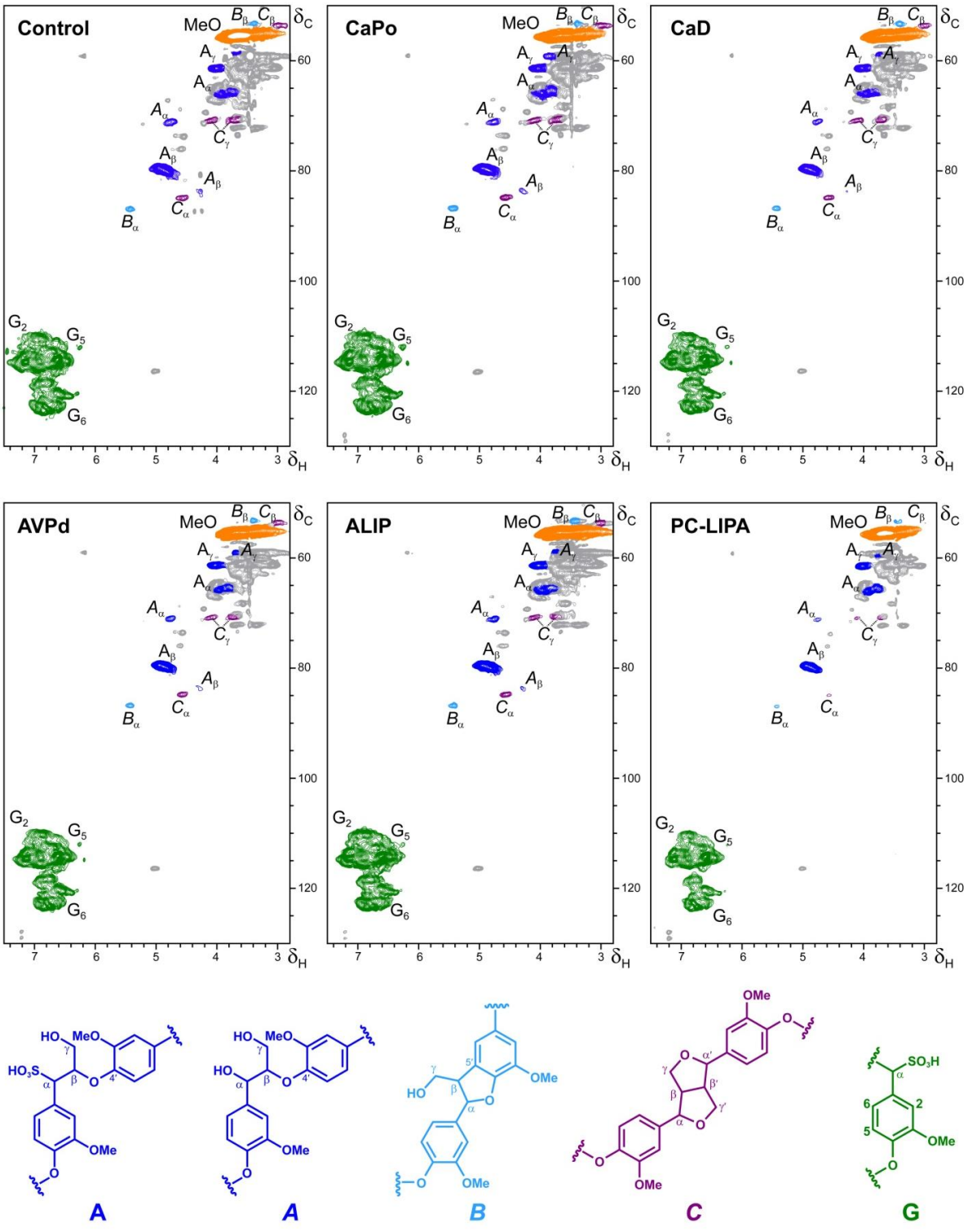


Fig. S4. HSQC 2D-NMR spectra of softwood lignosulfonate treated for 24 h with ancestral (CaPo, CaD, AVPd and ALiP) and extant (PC-LiPA) peroxidases, and control without enzyme. Formulae of the main structures identified are included. Signals correspond to ^{1}H-^{13}C correlations at different positions of Cα-sulfonated guaiacyl units (**G**), and Cα-sulfonated and non-sulfonated side chains in ß-*O*-4′ (**A** and ***A***) phenylcoumaran (***B***) and resinol (***C***) substructures, and methoxyls (**MeO**).

General discussion

The starting point of this thesis is the establishment of the origin of fungal degradation of lignin in the Carboniferous period (Floudas et al. 2012), which, together with other geochemical factors, contributed to the end of coal accumulation signature of this period (Hibbett et al. 2016). These ancestral fungi would have secreted the first ligninolytic peroxidases, starting then the enzymatic oxidation of the recalcitrant polymer of lignin.

The main goal of this thesis is the resurrection of ancestral ligninolytic peroxidases to analyze the subsequent evolution that led to the enzymes existing today. Wood-rotting fungi are essential organisms of land ecosystems enabling recycling the carbon stored in lignocellulosic biomass, whose degradation is hampered by the presence of recalcitrant lignin (Martínez et al. 2018). Moreover, they are being extensively studied to implement greener technologies for lignin modification in the bio-based industries (Martínez et al. 2009; Ragauskas et al. 2014). The study of peroxidases evolution will unveil essential aspects of their mechanisms related to the biology of wood-rotting fungi, and will *a priori* provide interesting enzymes for the industry. The results obtained in this thesis are discussed below.

In this thesis we use the main tools of paleogenetics, i.e. the resurrection of ancestral enzymes of now extinct organisms, which is a developing field that allows the time travel investigation of virtually any family of proteins (chapters 1 and 2). With a proper sequence alignment, a robust phylogeny and an evolutionary model adequate to the sequences used, it is possible to obtain the ancestors of the proteins studied (Liberles 2007). We have targeted ancestral ligninolytic peroxidases from Polyporales, where most wood-rotting fungi are included. We chose ML methods as implemented in PAML (Yang 2007) for ancestral sequence reconstruction, which have some advantages over other reconstruction procedures (Omland 1999). To cover the inherent uncertainty in ancestral inference (Royer-Carenzi et al. 2013) we have used the best data available at the moment (all class-II peroxidases annotated in ten Polyporales genomes) and checked that no ambiguity exists in the oxidation sites. Moreover, for the oldest targeted node (CaPo, 399 ± 140 my), which has the lowest posterior probability, we reconstructed several alternative ancestors. We have also manually inspected and curated all the selected sequences, so the confidence in our reconstruction robustness is high.

After sequence reconstruction, we selected *E. coli* as host for the heterologous expression of the ancestral genes designed. Ancestral proteins were obtained as inclusion bodies and, after *in vitro* enzyme activation, all the ancestors were successfully purified, showing the typical spectral properties of peroxidases, meaning a successful resurrection of the ancestral enzymes.

We focused our analyses in lignin peroxidases, which are high redox-potential and specialized enzymes that can directly extract electrons from lignin (Sáez-Jiménez et al. 2015a). Among LiPs, the isoenzime H8 (corresponding to the gene *lip*A) from the model ligninolytic fungus *P. chrysosporium* is one of the most studied (Kersten and Cullen 2007; Hammel and Cullen 2008). To understand how these enzymes appeared and evolved, we resurrected the most relevant ligninolytic peroxidases in the lineage of Polyporales LiPs (chapters 1, 3 and 4).

Their kinetic characterization using simple model compounds (chapter 1) revealed an exploration of new strategies for lignin degradation during evolution. CaPo, the most ancestral ligninolytic enzyme resurrected, was most probably a MnP that poorly oxidized Mn^{2+} (for indirect oxidation of phenolic lignin) and phenols. In the evolution to LiPs, first increased the ability to oxidize Mn^{2+}, and later appeared the first VP in Polyporales evolution, with the rise of a solvent-exposed tryptophan for direct lignin degradation 194 ± 70 mya, according to our time-calibration using PATHd8 (Britton et al. 2007) (chapter 4). From that point on, the Mn^{2+} binding site was lost, and the direct oxidation of lignin was improved, leading to the most efficient LiPs existing today. Additionally, we show that the generally accepted idea of promiscuous ancestors (Siddiq et al. 2017) is not true for ligninolytic peroxidases, since the ancestor with wider substrate specificity appeared at a midpoint in evolution.

The temperature stability of ancestral enzymes is usually higher than in their descendants (Gaucher et al. 2008), but for ligninolytic peroxidases we did not observe the expected decrease of T_m or T_{50} with evolution (chapters 1 and 2). However, with the loss of the Mn^{2+}-binding site there was a drop in temperature stability, associated to an structural effect of this site anchoring the heme cofactor (Gold et al. 2000; Ruiz-Dueñas et al. 2007).

Concerning pH, we observed an improved stability towards acidic pH when the solvent exposed catalytic tryptophan appeared in AVPd. This boost in stability in recent stages of evolution is an important adaptation to the acidic pH at which ligninolysis takes place in nature (Martínez 2002) and at which the reduction potential of ligninolytic enzymes is the highest (Millis et al. 1989; Oyadomari et al. 2003).

An elusive aspect of ligninolytic peroxidases is how they tune their uniquely high reduction potentials, among the highest observed for any oxidoreductase (Battistuzzi et al. 2010). For that purpose we used our model evolutionary lineage leading to high redox-potential LiPs to evaluate how this property changed with time (chapter 3).

We used stopped-flow spectrophotometry to accurately calculate the equilibrium concentrations of each catalytic pair in the peroxidase cycle.

Applying the Nerst equation, we were able to obtain the reduction potentials of every step in their cycle through evolution. This way we observed a general boost of redox power through evolution, with a difference of +70 mV between CaPo and the enzymes with the highest reduction potential (ALiP and PC-LiPA). Moreover, with protein NMR we observed the change in evolution of the signal corresponding to the proximal histidine, meaning a lower imidazolate character of the amino acid, as previously postulated (Banci et al. 1991a; 1991b; 1992; 1995; 2003). This led to a weaker histidine-heme Fe^{3+} axial bond that, in turns, destabilizes the higher oxidation state of the heme iron (Banci et al. 1995). More importantly is that stopped-flow and protein NMR data are correlated here for the first time, showing that rearrangements in the heme pocket led to enzymes with higher reduction potentials in evolution. The boost in reduction potentials through evolution is coherent with the better catalytic properties of the most recent enzymes (chapter 1), although with the appearance of the catalytic tryptophan the main changes in activity are apparently due to the negative charge increase in its environment (chapter 2).

The results discussed so far involve simple model compounds for the characterization of the evolution of oxidation sites in ligninolytic peroxidases. To further analyze the evolution of LiPs we employed the real lignin polymer as substrate (chapter 4). We selected the same water-soluble lignosulfonates, from angiosperms and gymnosperms, recently used to characterize the role of the surface tryptophan of ligninolytic peroxidases (Sáez-Jiménez et al. 2015b; 2016) representing two different compositions of lignin in nature.

Angiosperm lignin represents a later step in plant evolution (Weng and Chapple 2010), incorporating to their chemical structure S-units, which are dimethoxylated. While gymnosperms lignin have been considered unchanged through evolution (Sarkanen and Hergert 1971), angiosperm lignin would represent an innovation with mechanical and physiological advantages (Li et al. 2001). The inclusion of S-units in their structure makes angiosperm lignin more complex (Ralph et al. 2004) and, in average, less phenolic than gymnosperm lignin (Lai and Guo 1991; Camarero et al. 1999).

For characterization of LiP evolution with lignosulfonates we used: i) stopped-flow spectrophotometry to estimate the transient-state kinetic constants for the oxidation of both substrates; and ii) 2D-NMR as the state-of-the-art analytical technique for the analysis of lignin modification in prolonged treatments. Our results show, that during evolution, there was a progressive increase in the oxidation of lignin from both origins, but there was a clear switch in peroxidase preference when the solvent exposed tryptophan appeared. This way, from a preferential oxidation of gymnosperm lignin in the most ancestral states, the specificity changed to

angiosperm lignin when the oxidation site for direct lignin degradation appeared 194 ± 70 mya. The kinetic results are coherent with the 2D-NMR characterization, showing an increase in the decay of the polymer through evolution. Interestingly, in both experiments, the strongest modifications coincided with the rise of the surface catalytic tryptophan. The results obtained here correlate with the characterization using simple substrates (chapter 1), using both phenolic and non-phenolic compounds, and with the general increase of reduction potentials observed through evolution (chapter 3).

The explosive angiosperm diversification took place 250-140 mya (Morris et al. 2018), leading to the huge number of species (>350000) observed today. During plant evolution, the appearance of lignin containing S-units (typical of angiosperms) involved evolutionary advantages, as said above. Our time calibration (chapter 4) suggests that AVPd, the first peroxidase in Polyporales evolution with a site for direct lignin oxidation, appeared roughly 200 mya, concurrent with angiosperm diversification. The results discussed here show how wood-rotting fungi followed the progressive complexity of lignin in plant evolution with the incorporation of an efficient oxidation site for direct lignin oxidation. This is also coherent with recent studies showing a change from an ancestral gymnosperm host in Polyporales to an especialization in angiosperm host coincident with their diversification (Krah et al. 2018; Varga et al. 2019), and with our time-calibration of LiPs evolution, pointing to a fascinating coevolution between fungi and plants.

Finally, the analysis of Polyporales peroxidases phylogeny reveals that enzymes with a solvent exposed homologous tryptophan are present at least in two distant lineages, clade-D LiPs and clade-B VPs. As said above this is an essential trait in wood-rotting fungi, and we explored the origin of this important strategy for non-phenolic lignin degradation.

First we ruled out the possibility of horizontal gene transfer since BLAST searches did not reveal any signal of that (Zhaxybayeva 2009). Then, to evaluate the appearance of the same lignin-degradation mechanism in distant Polyporales peroxidases we resurrected the ancestral enzymes of both lineages (chapter 2).

Our results demonstrate that direct lignin oxidation appeared as a convergent trait in distant peroxidase lineages. We show that from the common ancestor of Polyporales peroxidases (and most probably of all Agaricomycetes peroxidases; CaPo, 399 ± 140 mya) (Floudas et al. 2012) the evolution in both lineages led first to an increase in Mn^{2+} oxidation and then to ancestral VPs due to the appearance of an equivalent solvent-exposed residue. Later in the LiP lineage the Mn^{2+} binding-site was lost for a better

lignin oxidation, while in the lineage leading to VPs the Mn^{2+}-site is kept but the oxidation of lignin was also improved.

The different kinetic parameters observed are most probably related to the evolution of the tryptophan environment: while in LiP lineage it becomes much more negative stabilizing the cation radicals and increasing the catalytic efficiency, in VP lineage the change is not so obvious. Either way, we show a convergent switch from the oxidation of the minor phenolic moiety of lignin (direct or indirectly) to a direct oxidation of nonphenolic lignin in two distant lineages.

We also observed similar stabilization towards the acidic conditions where ligninolysis takes place in nature (Martínez 2002), as said above. While in clade-D this stabilization coincided with the rise of the catalytic tryptophan, in clade-B it was acquired immediately after. Anyway there was a huge stabilization towards pH 3 in both lineages for a better performance of these enzymes, including their increased reduction potentials with pH.

In conclusion, our results resurrecting ancestral ligninolytic peroxidases have demonstrated both the convergent exploration of new strategies for lignin degradation and the improvement of their catalytic and stability properties for a better oxidation of lignin. We also show that ligninolytic peroxidases tuned their high-reduction potentials by rearrangements in the heme pocket. Additionally, using lignosulfonates from gymnosperms and angiosperms, some insights into the coevolution of plants and saprotrophic fungi were obtained, showing the change in lignin substrate preference with the rise of the solvent-exposed catalytic tryptophan.

Conclusiones / Conclusions

Conclusiones

1. El ancestro común de las peroxidasas ligninolíticas de Polyporales, y probablemente de todos los Agarycomicetes, apareció hace 399 ± 140 ma. Esta enzima fue una MnP con capacidad para oxidar pobremente manganeso y fenoles. La evolución posterior hacia las LiPs se enfocó primero en mejorar la eficiencia de oxidación del Mn^{2+} y más tarde, con la aparición de un triptófano catalítico expuesto en la superficie hace 194 ± 70 ma, en la oxidación directa de la lignina.

2. La aparición del triptófano catalítico expuesto en la superficie se produjo de forma independiente en varios linajes de peroxidasas ligninolíticas, suponiendo un ejemplo de convergencia evolutiva en la oxidación directa de lignina en hongos degradadores de la madera.

3. El aumento de la estabilidad a pH ácido, al cual actúan las peroxidasas ligninolíticas en la naturaleza, se produjo simultáneamente con la aparición del triptófano catalítico en los dos linajes independientes estudiados.

4. Las diferencias en reactividad de las enzimas que contienen el triptófano expuesto se debe en parte a la evolución diferente de la superficie que rodea dicho sitio de oxidación. Mientras que en las LiPs el entorno del triptófano se vuelve claramente más negativo con la evolución, pudiendo estabilizar intermedios catiónicos en la oxidación de lignina, en la evolución de las VPs no se produce un cambio tan claro.

5. Las peroxidasas ligninolíticas incrementaron su potencial redox durante la evolución, como se ha demostrado usando espectrofotometría de *stopped-flow* en equilibrio redox, debido a reordenamientos en el bolsillo del hemo. Estos se debieron principalmente a cambios en la geometría del enlace histidina proximal - hierro del hemo, como se ha visto utilizando NMR de proteínas.

6. Con la aparición del triptófano catalítico, se produjo un cambio en la preferencia de oxidación de lignina de gimnospermas, más antigua, a lignina de angiospermas, más reciente y compleja, como se muestra usando cinéticas de *stopped-flow* y 2D-NMR en reacciones con lignosulfonatos. Dicho cambio coincide con el origen de las angiospermas en la evolución y, junto al análisis de las especies hospedadoras de los hongos saprótrofos, apunta a una coevolución entre plantas y hongos.

1. The common ancestor of Polyporales ligninolytic peroxidases, and probably of all Agaricomycetes, appeared 399 ± 140 mya. This enzyme was a MnP able to poorly oxidize Mn^{2+} and phenols. The posterior evolution to LiPs focused on an improvement in the Mn^{2+} oxidation efficiency first, and later, with the appearance of a solvent-exposed catalytic tryptophan 194 ± 70 mya, in the direct oxidation of lignin.

2. The solvent-exposed catalytic tryptophan appeared independently in several ligninolytic peroxidases lineages, being an example of evolutionary convergence in lignin direct oxidation that took place in wood-rotting fungi.

3. The rise in acidic pH stability, at which ligninolytic peroxidases act in nature, was simultaneous to the appearance of the catalytic tryptophan in both independent lineages studied.

4. The differences in enzyme reactivity of the enzymes with solvent-exposed tryptophan were in part due to the different evolution in the surface surrounding the oxidation site. While in LiPs the tryptophan environment becomes clearly more negative through evolution, stabilizing the cationic intermediates in lignin oxidation, in VPs evolution there is no such a clear change.

5. Ligninolytic peroxidases incremented their redox-potential during evolution, as demonstrated using redox equilibrium stopped-flow, due to rearrangements in the heme pocket. These changes were mainly produced by modifications in the geometry of the proximal histidine – heme iron bond, as seen using protein NMR.

6. With the rise of the catalytic tryptophan there was a change in the preferential oxidation of lignin from gymnosperms, older, to lignin from angiosperms, more recent and complex, as demonstrated using stopped-flow kinetics and 2D-NMR in reactions with lignosulfonates. This change is coincident with the origin of angiosperms in evolution and, together with the analysis of hosts of saprotrophic fungi, point to a coevolution between plants and fungi

Bibliography

Introduction references

Abramson M, Shoseyov O, Shani Z (2010) Plant cell wall reconstruction toward improved lignocellulosic production and processability. Plant Sci 178:61-72

Akanuma S, Nakajima Y, Yokobori S, Kimura M, Nemoto N, Mase T, Miyazono K, Tanokura M, Yamagishi A (2013) Experimental evidence for the thermophilicity of ancestral life. Proc Natl Acad Sci USA 110:11067-11072

Alcalde M (2017) When directed evolution met ancestral enzyme resurrection. Microbial Biotechnol 10:22-24

Andrawis A, Johnson KA, Tien M (1988) Studies on compound I formation of the lignin peroxidase from *Phanerochaete chrysosporium*. J Biol Chem 263:1195-1198

Arnhold J, Furtmüller PG, Regelsberger G, Obinger C (2001) Redox properties of the couple compound I/native enzyme of myeloperoxidase and eosinophil peroxidase. Eur J Biochem 268:5142-5148

Baldauf LS (2008) An overview of the phylogeny and diversity of eukaryotes. J Syst Evol 46:263-273

Balestrini R, Lumini E (2018) Focus on mycorrhizal symbioses. Applied Soil Ecology 123:299-304

Banci L, Bertini I, Pease EA, Tien M, Turano P (1992) ^{1}H NMR investigation of manganese peroxidase from *Phanerochaete chrysosporium*. A comparison with other peroxidases. Biochemistry 31:10009-10017

Banci L, Bertini I, Pierattelli R, Tien M, Vila AJ (1995) Factoring of the hyperfine shifts in the cyanide adduct of lignin peroxidase from *P. chrysosporium*. J Am Chem Soc 117:8659-8667

Banci L, Bertini I, Turano P, Tien M, Kirk TK (1991) Proton NMR investigation into the basis for the relatively high redox potential of lignin peroxidase. Proc Natl Acad Sci USA 88:6956-6960

Battistuzzi G, Bellei M, Bortolotti CA, Sola M (2010) Redox properties of heme peroxidases. Arch Biochem Biophys 500:21-36

Beckham GT, Johnson CW, Karp EM, Salvachúa D, Vardon DR (2016) Opportunities and challenges in biological lignin valorization. Curr Opin Biotechnol 42:40-53

Bengtson S, Rasmussen B, Ivarsson M, Muhling J, Broman C, Marone F, Stampanoni M, Bekker A (2017) Fungus-like mycelial fossils in 2.4-billion-year-old vesicular basalt. Nat Ecol Evol 1:0141

Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F, Sjokvist E, Copeland A, Foster B, Sun H, Larsson E and others (2013) Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105:1350-1373

Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. Proc Natl Acad Sci USA 103:5869-5874

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519-546

Bonnarme P, Jeffries TW (1990) Mn(II) regulation of lignin peroxidases and manganese-dependent peroxidases from lignin-degrading white rot fungi. Appl Environ Microbiol 56:210-217

Brown ME, Barros T, Chang MCY (2012) Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium. ACS Chem Biol 7:2074-2081

Butterfield NJ (2005) Probable proterozoic fungi. Paleobiology 31:165-182

Camarero S, Barrasa JM, Pelayo M, Martínez AT (1998) Evaluation of *Pleurotus* species for wheat-straw biopulping. J Pulp Paper Sci 24:197-203

Cañas AI, Camarero S (2010) Laccases and their natural mediators: Biotechnological tools for sustainable eco-friendly processes. Biotechnol Adv 28:694-705

Carrigan MA, Uryasev O, Frye CB, Eckman BL, Myers CR, Hurley ThD, Benner SA (2015) Hominids adapted to metabolize ethanol long before human-directed fermentation. Proc Natl Acad Sci USA 112:458

Chang BSW, Jonsson K, Kazmi MA, Donoghue MJ, Sakmar TP (2002) Recreating a functional ancestral archosaur visual pigment. Mol Biol Evol 19:1483-1489

Cohen R, Persky L, Hadar Y (2002) Biotechnological applications and potential of wood-degrading mushrooms of the genus *Pleurotus*. Appl Microbiol Biotechnol 58:582-594

Cox VE, Gaucher EA (2014) Engineering Proteins by Reconstructing Evolutionary Adaptive Paths. In: Gillam EMJ, Copp JN, Ackerley D (eds) Directed Evolution Library Creation: Methods and Protocols. Springer New York, New York, NY, pp 353-363

Dighton J,White J (2017) The Fungal Community: Its Organization and Role in the Ecosystem, Fourth Edition. CRC Press, Boca Raton, USA

Erman JE, Vitello LB, Miller MA, Shaw A, Brown KA, Kraut J (1993) Histidine 52 is a critical residue for rapid formation of cytochrome *c* peroxidase compound I. Biochemistry 32:9798-9806

Fengel D,Wegener G (1984) Wood: Chemistry, ultrastructure, reactions. De Gruyter, Berlin

Fernández-Fueyo E, Acebes S, Ruiz-Dueñas FJ, Martínez MJ, Romero A, Medrano FJ, Guallar V, Martínez AT (2014) Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi. Acta Crystallogr D Biol Crystallogr 70:3253-3265

Fernández-Fueyo E, Linde D, Almendral D, López-Lucendo MF, Ruiz-Dueñas FJ, Martínez AT (2015) Description of the first fungal dye-decolorizing peroxidase oxidizing manganese(II). Appl Microbiol Biotechnol 99:8927-8942

Fernández-Fueyo E, Ruiz-Dueñas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, Larrondo L, James TY, Seelenfreund D, Lobos S and others (2012) Comparative genomics of *Ceriporiopisis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc Natl Acad Sci USA 109:5458-5463

Ferreira P, Carro J, Serrano A, Martínez AT (2015) A survey of genes encoding H2O2-producing GMC oxidoreductases in 10 Polyporales genomes. Mycologia 107:1105-1119

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Gaucher EA, Govindarajan S, Ganesh OK (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature 451:704

Gaucher EA, Thomson JM, Burgan MF, Benner SA (2003) Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature 425:285-288

Gibling MR, Davies NS (2012) Palaeozoic landscapes shaped by plant evolution. Nat Geosci 5:99

Glembo TJ, Farrell DW, Gerek ZN, Thorpe MF, Ozkan SB (2012) Collective Dynamics Differentiates Functional Divergence in Protein Evolution. PLoS Comput Biol 8:e1002428

Gold MH, Youngs HL, Gelpke MD (2000) Manganese peroxidase. Met Ions Biol Syst 37:559-586

Gumulya Y, Gillam EMJ (2017) Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering. Biochem J 474:1

Hammel KE, Cullen D (2008) Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol 11:349-355

Hanson-Smith V, Kolaczkowski B, Thornton JW (2010) Robustness of Ancestral Sequence Reconstruction to Phylogenetic Uncertainty. Mol Biol Evol 27:1988-1999

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001) Molecular Evidence for the Early Colonization of Land by Fungi and Plants. Science 293:1129

Hedges SB, Tao Q, Walker M, Kumar S (2018) Accurate timetrees require accurate calibrations. Proc Natl Acad Sci USA 115:E9510

Hibbett D, Blanchette R, Kenrick P, Mills B (2016) Climate, decay, and the death of the coal forests. Curr Biol 26:R563-R567

Higuchi T (1997) Biochemistry and molecular biology of wood. Springer Verlag, London

Hildén K, Makela MR, Steffen KT, Hofrichter M, Hatakka A, Archer DB, Lundell TK (2014) Biochemical and molecular characterization of an atypical manganese peroxidase of the litter-decomposing fungus *Agrocybe praecox*. Fungal Genet Biol 72:131-136

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass recalcitrance: Engineering plants and enzymes for biofuels production. Science 315:804-807

Hiner ANP, Raven EL, Thorneley RNF, García-Canovas F, Rodríguez-López JN (2002) Mechanisms of compound I formation in heme peroxidases. J Inorg Biochem 91:27-34

Holah JC, Wilson MV, Hansen EM (1993) Effects of a native forest pathogen, Phellinusweirii, on Douglas-fir forest composition in western Oregon. Can J For Res 23:2473-2480

Hoysted GA, Kowal J, Jacob A, Rimington WR, Duckett JG, Pressel S, Orchard S, Ryan MH, Field KJ, Bidartondo MI (2018) A mycorrhizal revolution. Curr Opin Plant Biol 44:1-6

Huang RQ, Hippauf F, Rohrbeck D, Haustein M, Wenke K, Feike J, Sorrelle N, Piechulla B, Barkman TJ (2012) Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. Proc Natl Acad Sci USA 109:2966-2971

Ikeda-Saito M, Prince RC (1985) The effect of chloride on the redox and EPR properties of myeloperoxidase. J Biol Chem 260:8301-8305

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J and others (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443:818

Jarosz AM, Davelos AL (1995) Effects of disease in wild plant populations and the evolution of pathogen aggressiveness. New Phytol 129:371-387

Justo A, Miettinen O, Floudas D, Ortiz-Santana B, Sjökvist E, Lindner D, Nakasone K, Niemelä T, Larsson KH, Ryvarden L and others (2017) A revised family-level classification of the Polyporales (Basidiomycota). Fungal Biol 121:798-824

Kawaoka A, Matsunaga E, Endo S, Kondo S, Yoshida K, Shinmyo A, Ebinuma H (2003) Ectopic expression of a horseradish peroxidase enhances growth rate and increases oxidative stress resistance in hybrid aspen. Plant Physiol 132:1177-1185

Kenrick P, Wellman CH, Schneider H, Edgecombe GD (2012) A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. Philos Trans R Soc Lond B Biol Sci 367:519-536

Khersonsky O, Tawfik DS (2010) Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. Annu Rev Biochem 79:471-505

Kirk TK, Farrell RL (1987) Enzymatic "combustion": The microbial degradation of lignin. Annu Rev Microbiol 41:465-505

Kishi K, Hildebrand DP, Kusters-van Someren M, Gettemy J, Mauk AG, Gold MH (1997) Site-directed mutations at phenylalanine-190 of manganese peroxidase: Effect on stability, functions and coordination. Biochemistry 36:4268-4277

Kratzer JT, Lanaspa MA, Murphy MN, Cicerchi C, Graves CL, Tipton PA, Ortlund EA, Johnson RJ, Gaucher EA (2014) Evolutionary history and metabolic insights of ancient mammalian uricases. Proc Natl Acad Sci USA 111:3763

Krishnan NM, Seligmann H, Stewart CB, de Koning APJ, Pollock DD (2004) Ancestral Sequence Reconstruction in Primate Mitochondrial DNA: Compositional Bias and Effect on Functional Inference. Mol Biol Evol 21:1871-1883

Lenton TM, Dahl TW, Daines SJ, Mills BJW, Ozaki K, Saltzman MR, Porada P (2016) Earliest land plants created modern levels of atmospheric oxygen. Proc Natl Acad Sci USA 113:9704

Li LG, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL (2001) The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. Plant Cell 13:1567-1585

Liberles DA (2007) Ancestral sequence reconstruction. Oxford University Press, Oxford

Lowry B, Lee D, Hèbant C (1980) The Origin of Land Plants: A New Look at an Old Problem. Taxon 29:183-197

Malcolm BA, Wilson KP, Matthews BW, Kirsch JF, Wilson AC (1990) Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing. Nature 345:86-89

Martínez AT (2002) Molecular biology and structure-function of lignin-degrading heme peroxidases. Enzyme Microb Technol 30:425-444

Martínez AT, Camarero S, Ruiz-Dueñas FJ,Martínez MJ (2018) Biological lignin degradation. In: Beckham GT (ed) Lignin valorization: Emerging approaches, Chapter 8. RSC, Cambridge, pp 199-225

Martínez AT, Rencoret J, Marques G, Gutiérrez A, Ibarra D, Jiménez-Barbero J, del Río JC (2008) Monolignol acylation and lignin structure in some nonwoody plants: A 2D NMR study. Phytochemistry 69:2831-2843

Martínez AT, Rencoret J, Nieto L, Jiménez-Barbero J, Gutiérrez A, del Río JC (2011) Selective lignin and polysaccharide removal in natural fungal decay of wood as evidenced by *in situ* structural analyses. Environ Microbiol 13:96-107

Martínez AT, Ruiz-Dueñas FJ, Martínez MJ, del Río JC, Gutiérrez A (2009) Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol 20:348-357

Martínez AT, Speranza M, Ruiz-Dueñas FJ, Ferreira P, Camarero S, Guillén F, Martínez MJ, Gutiérrez A, del Río JC (2005) Biodegradation of lignocellulosics: Microbiological, chemical and enzymatic aspects of fungal attack to lignin. Int Microbiol 8:195-204

Martinez D, Challacombe J, Morgenstern I, Hibbett DS, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Dueñas FJ, Martínez AT, Kersten P and others (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA 106:1954-1959

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F and others (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22:695-700

McCarl BA, Metting FB, Rice C (2007) Soil carbon sequestration. Clim Change 80:1-3

Menden B, Kohlhoff M, Moerschbacher BM (2007) Wheat cells accumulate a syringyl-rich lignin during the hypersensitive resistance response. Phytochemistry 68:513-520

Moerschbacher BM, Noll U, Gorrichon L, Reisener HJ (1990) Specific Inhibition of Lignification Breaks Hypersensitive Resistance of Wheat to Stem Rust. Plant Physiol 93:465-470

Morales M, Mate MJ, Romero A, Martínez MJ, Martínez AT, Ruiz-Dueñas FJ (2012) Two oxidation sites for low redox-potential substrates: A directed mutagenesis, kinetic and crystallographic study on *Pleurotus eryngii* versatile peroxidase. J Biol Chem 287:41053-41067

Morris JL, Leake JR, Stein WE, Berry CM, Marshall JEA, Wellman CH, Milton JA, Hillier Sn, Mannolini F, Quirk J and others (2015) Investigating Devonian trees as geo-engineers of past climates: linking palaeosols to palaeobotany and experimental geobiology. Palaeontology 58:787-801

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ (2018a) Reply to Hedges et al.: Accurate timetrees do indeed require accurate calibrations. Proc Natl Acad Sci USA 115:E9512

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ (2018b) The timescale of early land plant evolution. Proc Natl Acad Sci USA 115:E2274

Nagy LG, Kovacs GM, Krizsan K (2018) Complex multicellularity in fungi: evolutionary convergence, single origin, or both? Biol Rev 93:1778-1794

Nagy LG, Riley R, Bergmann PJ, Krizsán K, Martin FM, Grigoriev IV, Cullen D, Hibbett DS (2016a) Genetic bases of fungal white rot wood decay predicted by phylogenomic analysis of correlated gene-phenotype evolution. Mol Biol Evol 34:35-44

Nagy LG, Riley R, Tritt A, Adam C, Daum C, Floudas D, Sun H, Yadav JS, Pangilinan J, Larsson KH and others (2016b) Comparative genomics of early-diverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. Mol Biol Evol 33:959-970

Nei M (2005) Selectionism and Neutralism in Molecular Evolution. Mol Biol Evol 22:2318-2342

Nelsen MP, DiMichele WA, Peters SE, Boyce C (2016) Delayed fungal evolution did not cause the Paleozoic peak in coal production. Proc Natl Acad Sci USA 113:2442-2447

Nguyen V, Wilson C, Hoemberger M, Stiller JB, Agafonov RV, Kutter S, English J, Theobald DL, Kern D (2017) Evolutionary drivers of thermoadaptation in enzyme catalysis. Science 355:289

Novo-Uzal E, Pomar F, Ros LVG, Espineira JM, Barcelo AR (2012) Evolutionary history of lignins. Adv Bot Res (Lignins: Biosynthesis, Biodegradation and Bioengineering) 61:311-350

Oyadomari M, Shinohara H, Johjima T, Wariishi H, Tanaka H (2003) Electrochemical characterization of lignin peroxidase from the white-rot basidiomycete *Phanerochaete chrysosporium*. J Mol Catal B-Enzym 21:291-297

Pandya C, Farelli JD, Dunaway-Mariano D, Allen KN (2014) Enzyme Promiscuity: Engine of Evolutionary Innovation. J Biol Chem 289:30229-30236

Parfrey LW, Lahr DJG, Knoll AH, Katz LA (2011) Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc Natl Acad Sci USA 108:13624

Pauling L, Zuckerkandl E (1963) Chemical Paleogenetics. Molecular "Restoration Studies" of Extinct Forms of Life. Acta Chem Scand 17 supl.:9-16

Pettersen RC (1984) The chemical composition of wood. In: Powell RM (ed) The Chemistry of Solid Wood. ACS, pp 57-126

Piontek K, Glumoff T, Winterhalter K (1993) Low pH crystal structure of glycosylated lignin peroxidase from *Phanerochaete chrysosporium* at 2.5 Å resolution. FEBS Lett 315:119-124

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M and others (2014) Lignin valorization: improving lignin processing in the biorefinery. Science 344:1246843

Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ, Hallett JP, Leak DJ, Liotta CL and others (2006) The path forward for biofuels and biomaterials. Science 311:484-489

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH and others (2004) Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev 3:29-60

Raven EL,Dunford HB (2016) Heme peroxidases. RSC, London

Raven JA (1984) Physiological correlates of the morphology of early vascular plants. Bot J Linn Soc 88:105-126

Rencoret J (2008) Estudio de lignina y lípidos en madera de eucalipto: Caracterización química en distintas especies y su evolución durante la fabricación y blanqueo químico y enzimático de la pasta de papel. PhD thesis, Universidad de Sevilla,

Risso VA, Gavira JA, Mejia-Carmona DF, Gaucher EA, Sanchez-Ruiz JM (2013) Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. J Am Chem Soc 135:2899-2902

Risso VA, Martinez-Rodriguez S, Candel AM, Krüger DM, Pantoja-Uceda D, Ortega-Muñoz M, Santoyo-Gonzalez F, Gaucher EA, Kamerlin SCL, Bruix M and others (2017) De novo active sites for resurrected Precambrian enzymes. Nat Commun 8:16113

Risso VA, Sanchez-Ruiz JM, Ozkan SB (2018) Biotechnological and protein-engineering implications of ancestral protein resurrection. Curr Opin Struct Biol 51:106-115

Ruiz-Dueñas FJ, Guillén F, Camarero S, Pérez-Boada M, Martínez MJ, Martínez AT (1999) Regulation of peroxidase transcript levels in liquid cultures of the ligninolytic fungus *Pleurotus eryngii*. Appl Environ Microbiol 65:4458-4463

Ruiz-Dueñas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martínez AT (2013) Lignin-degrading peroxidases in Polyporales: An evolutionary survey based on ten sequenced genomes. Mycologia 105:1428-1444

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441-452

Ruiz-Dueñas FJ, Morales M, Mate MJ, Romero A, Martínez MJ, Smith AT, Martínez AT (2008) Site-directed mutagenesis of the catalytic tryptophan environment in *Pleurotus eryngii* versatile peroxidase. Biochemistry 47:1685-1695

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015a) Demonstration of lignin-to-peroxidase direct electron transfer. A transient-state kinetics, directed mutagenesis, EPR and NMR study (vol 290, pag 23201, 2015). J Biol Chem 290:30268

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015b) Demonstration of lignin-to-peroxidase direct electron transfer: A transient-state kinetics, directed mutagenesis, EPR and NMR study. J Biol Chem 290:23201-23213

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Salvachúa D, Martínez AT, Tien M, López-Lucendo MF, García F, de los Ríos V, Martínez MJ, Prieto A (2013) Differential proteomic analysis of the secretome of *Irpex lacteus* and other white-rot fungi during wheat straw pretreatment. Biotechnol Biofuels 6:115

Sanderson MJ, Wojciechowski MF, Hu JM, Sher Khan T, Brady SG (2000) Error, Bias, and Long-Branch Attraction in Data for Two Chloroplast Photosystem Genes in Seed Plants. Mol Biol Evol 17:782-797

Siddiq MA, Hochberg GKA, Thornton JW (2017) Evolution of protein specificity: insights from ancestral protein reconstruction. Curr Opin Struct Biol 47:113-122

Sigoillot JC, Berrin JG, Bey M, Lesage-Meessen L, Levasseur A, Lomascolo A, Record E, Uzan-Boukhris E (2012) Fungal strategies for lignin degradation. Adv Bot Res 61:263-308

Sørlie M, Seefeldt LC, Parker VD (2000) Use of stopped-flow spectrophotometry to establish midpoint potentials for redox proteins. Anal Biochem 287:118-125

Spribille T, Tuovinen V, Resl P, Vanderpool D, Wolinski H, Aime MC, Schneider K, Stabentheiner E, Toome-Heller M, Thor G and others (2016) Basidiomycete yeasts in the cortex of ascomycete macrolichens. Science 353:488

Stackhouse J, Presnell SR, McGeehan GM, Nambiar KP, Benner SA (1990) The ribonuclease from an extinct bovid ruminant. FEBS Lett 262:104-106

Thornton JW (2004) Resurrecting ancient genes: experimental analysis of extinct molecules. Nat Rev Genet 5:366-375

Tien M, Kirk TK (1984) Lignin-degrading enzyme from *Phanerochaete chrysosporium*: purification, characterization, and catalytic properties of a unique H_2O_2-requiring oxygenase. Proc Natl Acad Sci USA 81:2280-2284

Tokuriki N, Tawfik DS (2009) Protein Dynamism and Evolvability. Science 324:203

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010) Lignin Biosynthesis and Structure. Plant Physiol 153:895-905

Vanholme R, Morreel K, Ralph J, Boerjan W (2008) Lignin engineering. Curr Opin Plant Biol 11:278-285

Vitello LB, Erman JE, Miller MA, Wang J, Kraut J (1993) Effect of Arginine-48 replacement on the reaction between cytochrome *c* peroxidase and hydrogen peroxide. Biochemistry 32:9807-9818

Wariishi H, Huang J, Dunford HB, Gold MH (1991) Reactions of lignin peroxidase compounds I and II with veratryl alcohol. Transient-state kinetic characterization. J Biol Chem 266:20694-20699

Weng JK, Chapple C (2010) The origin and evolution of lignin biosynthesis. New Phytol 187:273-285

Weng JK, Li X, Stout J, Chapple C (2008) Independent origins of syringyl lignin in vascular plants. Proc Natl Acad Sci USA 105:7887-7892

Wijma HJ, Floor RJ, Janssen DB (2013) Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability. Curr Opin Struct Biol 23:588-594

Wilson C, Agafonov RV, Hoemberger M, Kutter S, Zorba A, Halpin J, Buosi V, Otten R, Waterman D, Theobald DL and others (2015) Using ancient protein kinases to unravel a modern cancer drug's mechanism. Science 347:882-886

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Zámocký M, Hofbauer S, Schaffner I, Gasselhuber B, Nicolussi A, Soudi M, Pirker KF, Furtmüller PG, Obinger C (2015) Independent evolution of four heme peroxidase superfamilies. Arch Biochem Biophys 574:108-119

Zhao RL, Li GJ, Sánchez-Ramírez S, Stata M, Yang ZL, Wu G, Dai YC, He SH, Cui BK, Zhou JL and others (2017) A six-gene phylogenetic overview of Basidiomycota and allied phyla with estimated divergence times of higher taxa and a phyloproteomics perspective. Fungal Divers 84:43-74

Materials & Methods references

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104-2105

Arnhold J, Furtmüller PG, Regelsberger G, Obinger C (2001) Redox properties of the couple compound I/native enzyme of myeloperoxidase and eosinophil peroxidase. Eur J Biochem 268:5142-5148

Banci L, Bertini I, Pease EA, Tien M, Turano P (1992) ^{1}H NMR investigation of manganese peroxidase from *Phanerochaete chrysosporium*. A comparison with other peroxidases. Biochemistry 31:10009-10017

Banci L, Bertini I, Pierattelli R, Tien M, Vila AJ (1995) Factoring of the hyperfine shifts in the cyanide adduct of lignin peroxidase from *P. chrysosporium*. J Am Chem Soc 117:8659-8667

Banci L, Bertini I, Turano P, Ferrer JC, Mauk AG (1991a) Comparative ^{1}H NMR study of ferric low-spin cytochrome c peroxidase and horseradish peroxidase. Inorg Chem 30:4510-4516

Banci L, Bertini I, Turano P, Tien M, Kirk TK (1991b) Proton NMR investigation into the basis for the relatively high redox potential of lignin peroxidase. Proc Natl Acad Sci USA 88:6956-6960

Banci L, Camarero S, Martínez AT, Martínez MJ, Pérez-Boada M, Pierattelli R, Ruiz-Dueñas FJ (2003) NMR study of Mn(II) binding by the new versatile peroxidase from the white-rot fungus *Pleurotus eryngii*. J Biol Inorg Chem 8:751-760

Battistuzzi G, Bellei M, Vlasits J, Banerjee S, Furtmüller PG, Sola M, Obinger C (2010) Redox thermodynamics of lactoperoxidase and eosinophil peroxidase. Arch Biochem Biophys 494:72-77

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L and others (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252-W258

Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F, Sjokvist E, Copeland A, Foster B, Sun H, Larsson E and others (2013) Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105:1350-1373

Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating Divergence Times in Large Phylogenetic Trees. Syst Biol 56:741-752

DeFelippis MR, Murthy CP, Faraggi M, Klapper MH (1989) Pulse radiolytic measurement of redox potentials: the tyrosine and tryptophan radicals. Biochemistry 28:4847-4853

Doyle WA, Smith AT (1996) Expression of lignin peroxidase H8 in *Escherichia coli*: Folding and activation of the recombinant enzyme with Ca^{2+} and haem. Biochem J 315:15-19

Fernández-Fueyo E, Ruiz-Dueñas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, Larrondo L, James TY, Seelenfreund D, Lobos S and others (2012) Comparative genomics of *Ceriporiopisis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc Natl Acad Sci USA 109:5458-5463

Fernández-Fueyo E, Ruiz-Dueñas FJ, Martínez MJ, Romero A, Hammel KE, Medrano FJ, Martínez AT (2014) Ligninolytic peroxidase genes in the oyster mushroom genome: Heterologous expression, molecular structure, catalytic and stability properties and lignin-degrading ability. Biotechnol Biofuels 7:2

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Furtmüller PG, Arnhold J, Jantschko W, Pichler H, Obinger C (2003) Redox properties of the couples compound I/compound II and compound II/native enzyme of human myeloperoxidase. Biochem Biophys Res Commun 301:551-557

Furtmüller PG, Arnhold J, Jantschko W, Zederbauer M, Jakopitsch C, Obinger C (2005) Standard reduction potentials of all couples of the peroxidase cycle of lactoperoxidase. J Inorg Biochem 99:1220-1229

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30:S162-S173

Inubushi T, Becker ED (1983) Efficient detection of paramagnetically shifted NMR resonances by optimizing the WEFT pulse sequence. J Magnet Reson 51:128-133

Koppenol WH (1987) Thermodynamics of reactions involving oxyradicals and hydrogen peroxide. Bioelectrochem Bioenerg 18:3-11

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870-1874

Lebo SE, Braaten SM, Fredheim GE, Lutnaes BF, Lauten RA, Myrvold BO,McNally TJ (2008) Recent advances in the characterization of lignosulfonates. In: Hu T (ed) Characterization of lignocellulosic materials. Blackwell Pub., New York, pp 189-205

Lutnaes BF, Myrvold BO, Lauten RA, Endeshaw MM (2008) ^{1}H and ^{13}C NMR data of benzylsulfonic acids - model compounds for lignosulfonate. Magn Reson Chem 46:299-305

Magina S, Marques AP, Evtuguin DV (2015) Study on the residual lignin in *Eucalyptus globulus* sulphite pulp. Holzforschung 69:513-522

Martínez AT, Rencoret J, Marques G, Gutiérrez A, Ibarra D, Jiménez-Barbero J, del Río JC (2008) Monolignol acylation and lignin structure in some nonwoody plants: A 2D NMR study. Phytochemistry 69:2831-2843

Martinez D, Challacombe J, Morgenstern I, Hibbett DS, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Dueñas FJ, Martínez AT, Kersten P and others (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA 106:1954-1959

Morales M, Mate MJ, Romero A, Martínez MJ, Martínez AT, Ruiz-Dueñas FJ (2012) Two oxidation sites for low redox-potential substrates: A directed mutagenesis, kinetic and crystallographic study on *Pleurotus eryngii* versatile peroxidase. J Biol Chem 287:41053-41067

Moss DS, Naberdryk E, Breton J, Mäntele W (1990) Redox-linked conformational changes in proteins detected by a combination of infrared spectroscopy and protein electrochemistry. Eur J Biochem 187:565-572

Ohm RA, Riley R, Salamov A, Min B, Choi IG, Grigoriev IV (2014) Genomics of wood-degrading fungi. Fungal Genet Biol 72:82-90

Pérez-Boada M, Doyle WA, Ruiz-Dueñas FJ, Martínez MJ, Martínez AT, Smith AT (2002) Expression of *Pleurotus eryngii* versatile peroxidase in *Escherichia coli* and optimisation of *in vitro* folding. Enzyme Microb Technol 30:518-524

Prasetyo EN, Kudanga T, Ostergaard L, Rencoret J, Gutiérrez A, del Río JC, Santos JI, Nieto L, Jimenez-Barbero J, Martínez AT and others (2010) Polymerization of lignosulfonates by the laccase-HBT (1-hydroxybenzotriazole) system improves dispersibility. Bioresource Technol 101:5054-5062

Ralph S, Ralph J, Landucci L (2009) NMR database of lignin and cell wall model compounds. available at URL www.glbrc.org/databases_and_software/nmrdatabase),

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015a) Demonstration of lignin-to-peroxidase direct electron transfer. A transient-state kinetics, directed mutagenesis, EPR and NMR study (vol 290, pag 23201, 2015). J Biol Chem 290:30268

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015b) Demonstration of lignin-to-peroxidase direct electron transfer: A transient-state kinetics, directed mutagenesis, EPR and NMR study. J Biol Chem 290:23201-23213

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Sørlie M, Seefeldt LC, Parker VD (2000) Use of stopped-flow spectrophotometry to establish midpoint potentials for redox proteins. Anal Biochem 287:118-125

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Wariishi H, Márquez L, Dunford HB, Gold MH (1990) Lignin peroxidase compounds II and III. Spectral and kinetic characterization of reactions with peroxides. J Biol Chem 265:11137-11142

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691-699

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Chapter 1 references

Aakre CD, Herrou J, Phung TN, Perchuk BS, Crosson S, Laub MT (2015) Evolving new protein-protein interaction specificity through promiscuous intermediates. Cell 163:594-606

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104-2105

Alcalde M (2015) Engineering the ligninolytic enzyme consortium. Trends Biotechnol 33:155-162

Alcalde M (2016) When directed evolution met ancestral enzyme resurrection. Microbial Biotechnol 10:22-24

Banci L (1997) Structural properties of peroxidases. J Biotechnol 53:253-263

Banci L, Bertini I, Turano P, Tien M, Kirk TK (1991) Proton NMR investigation into the basis for the relatively high redox potential of lignin peroxidase. Proc Natl Acad Sci USA 88:6956-6960

Barriuso J, Martínez MJ (2017) Evolutionary history of versatile-lipases from Agaricales through reconstruction of ancestral structures. BMC Genomics 18:12

Baunsgaard L, Dalboge H, Houen G, Rasmussen EM, Welinder KG (1993) Amino acid sequence of *Coprinus macrorhizus* peroxidase and cDNA sequence encoding *Coprinus cinereus* peroxidase - A new family of fungal peroxidases. Eur J Biochem 213:605-611

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L and others (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252-W258

Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F, Sjokvist E, Copeland A, Foster B, Sun H, Larsson E and others (2013) Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105:1350-1373

Bloom JD, Arnold FH (2009) In the light of directed evolution: Pathways of adaptive protein evolution. Proc Natl Acad Sci USA 106:9995-10000

Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K (2012) Engineering the third wave of biocatalysis. Nature 485:185-194

Camarero S, Sarkar S, Ruiz-Dueñas FJ, Martínez MJ, Martínez AT (1999) Description of a versatile peroxidase involved in natural degradation of lignin that has both Mn-peroxidase and lignin-peroxidase substrate binding sites. J Biol Chem 274:10324-10330

Doyle WA, Blodig W, Veitch NC, Piontek K, Smith AT (1998) Two substrate interaction sites in lignin peroxidase revealed by site-directed mutagenesis. Biochemistry 37:15097-15105

Doyle WA, Smith AT (1996) Expression of lignin peroxidase H8 in *Escherichia coli*: Folding and activation of the recombinant enzyme with Ca^{2+} and haem. Biochem J 315:15-19

Fernández-Fueyo E, Acebes S, Ruiz-Dueñas FJ, Martínez MJ, Romero A, Medrano FJ, Guallar V, Martínez AT (2014a) Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi. Acta Crystallogr D Biol Crystallogr 70:3253-3265

Fernández-Fueyo E, Ruiz-Dueñas FJ, Martínez MJ, Romero A, Hammel KE, Medrano FJ, Martínez AT (2014b) Ligninolytic peroxidase genes in the oyster mushroom genome: Heterologous expression, molecular structure, catalytic and stability properties and lignin-degrading ability. Biotechnol Biofuels 7:2

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Floudas D, Held BW, Riley R, Nagy LG, Koehler G, Ransdell AS, Younus H, Chow J, Chiniqui J, Lipzen A and others (2015) Evolution of novel wood decay mechanisms in Agaricales revealed by the genome sequences of *Fistulina hepatica* and *Cylindrobasidium torrendii*. Fungal Genet Biol 76:78-92

Gaucher EA, Thomson JM, Burgan MF, Benner SA (2003) Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature 425:285-288

Gold MH, Youngs HL, Gelpke MD (2000) Manganese peroxidase. Met Ions Biol Syst 37:559-586

Gould JS (1990) Wonderful life: The burgess shale and the nature of history. Norton & Company, NY

Hammel KE, Cullen D (2008) Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol 11:349-355

Huang RQ, Hippauf F, Rohrbeck D, Haustein M, Wenke K, Feike J, Sorrelle N, Piechulla B, Barkman TJ (2012) Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. Proc Natl Acad Sci USA 109:2966-2971

Kersten P, Cullen D (2007) Extracellular oxidative systems of the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Fungal Genet Biol 44:77-87

Kishi K, Kusters-van Someren M, Mayfield MB, Sun J, Loehr TM, Gold MH (1996) Characterization of manganese(II) binding site mutants of manganese peroxidase. Biochemistry 35:8986-8994

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol

Liberles DA (2007) Ancestral sequence reconstruction. Oxford University Press, Oxford

Loughran NB, O'Connell MJ, O'Connor B, O'Fagain C (2014) Stability properties of an ancient plant peroxidase. Biochimie 104C:156-159

Martínez AT (2002) Molecular biology and structure-function of lignin-degrading heme peroxidases. Enzyme Microb Technol 30:425-444

Martínez AT, Ruiz-Dueñas FJ, Martínez MJ, del Río JC, Gutiérrez A (2009) Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol 20:348-357

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F and others (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22:695-700

Mester T, Ambert-Balay K, Ciofi-Baffoni S, Banci L, Jones AD, Tien M (2001) Oxidation of a tetrameric nonphenolic lignin model compound by lignin peroxidase. J Biol Chem 276:22985-22990

Millis CD, Cai D, Stankovich MT, Tien M (1989) Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*. Biochemistry 28:8484-8489

Morales M, Mate MJ, Romero A, Martínez MJ, Martínez AT, Ruiz-Dueñas FJ (2012) Two oxidation sites for low redox-potential substrates: A directed mutagenesis, kinetic and crystallographic study on *Pleurotus eryngii* versatile peroxidase. J Biol Chem 287:41053-41067

Nagy LG, Riley R, Tritt A, Adam C, Daum C, Floudas D, Sun H, Yadav JS, Pangilinan J, Larsson KH and others (2016) Comparative genomics of early-diverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. Mol Biol Evol 33:959-970

Pérez-Boada M, Doyle WA, Ruiz-Dueñas FJ, Martínez MJ, Martínez AT, Smith AT (2002) Expression of *Pleurotus eryngii* versatile peroxidase in *Escherichia coli* and optimisation of *in vitro* folding. Enzyme Microb Technol 30:518-524

Pérez-Boada M, Ruiz-Dueñas FJ, Pogni R, Basosi R, Choinowski T, Martínez MJ, Piontek K, Martínez AT (2005) Versatile peroxidase oxidation of high redox potential aromatic compounds: Site-directed mutagenesis, spectroscopic and crystallographic investigations of three long-range electron transfer pathways. J Mol Biol 354:385-402

Pérez-Jiménez R, Ingles-Prieto A, Zhao ZM, Sánchez-Romero I, Alegre-Cebollada J, Kosuri P, García-Manyes S, Kappock TJ, Tanokura M, Holmgren A and others (2011) Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. Nature Structural & Molecular Biology 18:592-U99

Puigbò P, Guzmán E, Romeu A, Garcia-Vallvé S (2007) OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res 35:W126-W131

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M and others (2014) Lignin valorization: improving lignin processing in the biorefinery. Science 344:1246843

Risso VA, Gavira JA, Mejia-Carmona DF, Gaucher EA, Sanchez-Ruiz JM (2013) Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. J Am Chem Soc 135:2899-2902

Risso VA, Gavira JA, Sanchez-Ruiz JM (2014) Thermostable and promiscuous Precambrian proteins. Environ Microbiol 16:1485-1489

Ruiz-Dueñas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martínez AT (2013) Lignin-degrading peroxidases in Polyporales: An evolutionary survey based on ten sequenced genomes. Mycologia 105:1428-1444

Ruiz-Dueñas FJ, Martínez AT (2009) Microbial degradation of lignin: How a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. Microbial Biotechnol 2:164-177

Ruiz-Dueñas FJ, Martínez MJ, Martínez AT (1999) Molecular characterization of a novel peroxidase isolated from the ligninolytic fungus *Pleurotus eryngii*. Mol Microbiol 31:223-236

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441-452

Ruiz-Dueñas FJ, Morales M, Pérez-Boada M, Choinowski T, Martínez MJ, Piontek K, Martínez AT (2007) Manganese oxidation site in *Pleurotus eryngii* versatile peroxidase: A site-directed mutagenesis, kinetic and crystallographic study. Biochemistry 46:66-77

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Sayou C, Monniaux M, Nanao MH, Moyroud E, Brockington SF, Thévenon E, Chahtane H, Warthmann N, Melkonian M, Zhang Y and others (2014) A promiscuous intermediate underlies the evolution of LEAFY DNA binding specificity. Science 343:645

Semba Y, Ishida M, Yokobori Si, Yamagishi A (2015) Ancestral amino acid substitution improves the thermal stability of recombinant lignin-peroxidase from white-rot fungi, *Phanerochaete chrysosporium* strain UAMH 3641. Protein Engineering Design and Selection 28:221-230

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Thomson JM, Gaucher EA, Burgan MF, De Kee DW, Li T, Aris JP, Benner SA (2005) Resurrecting ancestral alcohol dehydrogenases from yeast. Nature Genetics 37:630-635

Voordeckers K, Brown CA, Vanneste K, van der Zande E, Voet A, Maere S, Verstrepen KJ (2012) Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. PLoS Biology 10:

Wang XJ, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. J Mol Biol 320:85-95

Wilson C, Agafonov RV, Hoemberger M, Kutter S, Zorba A, Halpin J, Buosi V, Otten R, Waterman D, Theobald DL and others (2015) Using ancient protein kinases to unravel a modern cancer drug's mechanism. Science 347:882-886

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Zámocký M, Hofbauer S, Schaffner I, Gasselhuber B, Nicolussi A, Soudi M, Pirker KF, Furtmüller PG, Obinger C (2015) Independent evolution of four heme peroxidase superfamilies. Arch Biochem Biophys 574:108-119

Chapter 2 references

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104-2105

Acebes S, Ruiz-Dueñas FJ, Toubes M, Sáez-Jiménez V, Pérez-Boada M, Lucas F, Martínez AT, Guallar V (2017) Mapping the long-range electron transfer route in ligninolytic peroxidases . J Phys Chem B 121:3946-3954

Ayuso-Fernández I, Martínez AT, Ruiz-Dueñas FJ (2017) Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date. Biotechnol Biofuels 10:67

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L and others (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252-W258

Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F, Sjokvist E, Copeland A, Foster B, Sun H, Larsson E and others (2013) Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105:1350-1373

Camarero S, Bocchini P, Galletti GC, Martínez AT (1999a) Pyrolysis-gas chromatography/mass spectrometry analysis of phenolic and etherified units in natural and industrial lignins. Rapid Commun Mass Spectrom 13:630-636

Camarero S, Sarkar S, Ruiz-Dueñas FJ, Martínez MJ, Martínez AT (1999b) Description of a versatile peroxidase involved in natural degradation of lignin that has both Mn-peroxidase and lignin-peroxidase substrate binding sites. J Biol Chem 274:10324-10330

Carabajal M, Kellner H, Levin L, Jehmlich N, Hofrichter M, Ullrich R (2013) The secretome of *Trametes versicolor* grown on tomato juice medium and purification of the secreted oxidoreductases including a versatile peroxidase. J Biotechnol 168:15-23

Colegrave N, Collins S (2008) Experimental evolution: experimental evolution and evolvability. Heredity 100:464-470

Doyle WA, Smith AT (1996) Expression of lignin peroxidase H8 in *Escherichia coli*: Folding and activation of the recombinant enzyme with Ca^{2+} and haem. Biochem J 315:15-19

Fernández-Fueyo E, Acebes S, Ruiz-Dueñas FJ, Martínez MJ, Romero A, Medrano FJ, Guallar V, Martínez AT (2014) Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi. Acta Crystallogr D Biol Crystallogr 70:3253-3265

Fernández-Fueyo E, Ruiz-Dueñas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, Larrondo L, James TY, Seelenfreund D, Lobos S and others (2012) Comparative genomics of *Ceriporiopisis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc Natl Acad Sci USA 109:5458-5463

Fitzpatrick DA (2012) Horizontal gene transfer in fungi. FEMS Microbiol Lett 329:1-8

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Gold MH, Youngs HL, Gelpke MD (2000) Manganese peroxidase. Met Ions Biol Syst 37:559-586

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30:S162-S173

Hammel KE, Cullen D (2008) Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol 11:349-355

Hibbett D, Blanchette R, Kenrick P, Mills B (2016) Climate, decay, and the death of the coal forests. Curr Biol 26:R563-R567

Kishi K, Kusters-van Someren M, Mayfield MB, Sun J, Loehr TM, Gold MH (1996) Characterization of manganese(II) binding site mutants of manganese peroxidase. Biochemistry 35:8986-8994

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870-1874

Liberles DA (2007) Ancestral sequence reconstruction. Oxford University Press, Oxford

Magallón S, Gómez-Acevedo S, Sánchez-Reyes LL, Hernández-Hernández T (2015) A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol 207:437-453

Martínez AT (2002) Molecular biology and structure-function of lignin-degrading heme peroxidases. Enzyme Microb Technol 30:425-444

Martínez AT, Rencoret J, Marques G, Gutiérrez A, Ibarra D, Jiménez-Barbero J, del Río JC (2008) Monolignol acylation and lignin structure in some nonwoody plants: A 2D NMR study. Phytochemistry 69:2831-2843

Martínez AT, Ruiz-Dueñas FJ, Martínez MJ, del Río JC, Gutiérrez A (2009) Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol 20:348-357

Martinez D, Challacombe J, Morgenstern I, Hibbett DS, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Dueñas FJ, Martínez AT, Kersten P and others (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA 106:1954-1959

Mester T, Ambert-Balay K, Ciofi-Baffoni S, Banci L, Jones AD, Tien M (2001) Oxidation of a tetrameric nonphenolic lignin model compound by lignin peroxidase. J Biol Chem 276:22985-22990

Millis CD, Cai D, Stankovich MT, Tien M (1989) Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*. Biochemistry 28:8484-8489

Morales M, Mate MJ, Romero A, Martínez MJ, Martínez AT, Ruiz-Dueñas FJ (2012) Two oxidation sites for low redox-potential substrates: A directed mutagenesis, kinetic and crystallographic study on *Pleurotus eryngii* versatile peroxidase. J Biol Chem 287:41053-41067

Nagy LG, Riley R, Tritt A, Adam C, Daum C, Floudas D, Sun H, Yadav JS, Pangilinan J, Larsson KH and others (2016) Comparative genomics of early-diverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. Mol Biol Evol 33:959-970

Nelsen MP, DiMichele WA, Peters SE, Boyce C (2016) Delayed fungal evolution did not cause the Paleozoic peak in coal production. Proc Natl Acad Sci USA 113:2442-2447

Novo-Uzal E, Pomar F, Ros LVG, Espineira JM, Barcelo AR (2012) Evolutionary history of lignins. Adv Bot Res (Lignins: Biosynthesis, Biodegradation and Bioengineering) 61:311-350

Ohm RA, Riley R, Salamov A, Min B, Choi IG, Grigoriev IV (2014) Genomics of wood-degrading fungi. Fungal Genet Biol 72:82-90

Omland KE (1999) The assumptions and challenges of ancestral state reconstructions. Syst Biol 48:604-611

Oyadomari M, Shinohara H, Johjima T, Wariishi H, Tanaka H (2003) Electrochemical characterization of lignin peroxidase from the white-rot basidiomycete *Phanerochaete chrysosporium*. J Mol Catal B-Enzym 21:291-297

Passardi F, Bakalovic N, Teixeira FK, Margis-Pinheiro M, Penel C, Dunand C (2007) Prokaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes. Genomics 89:567-579

Pérez-Boada M, Doyle WA, Ruiz-Dueñas FJ, Martínez MJ, Martínez AT, Smith AT (2002) Expression of *Pleurotus eryngii* versatile peroxidase in *Escherichia coli* and optimisation of *in vitro* folding. Enzyme Microb Technol 30:518-524

Pérez-Boada M, Ruiz-Dueñas FJ, Pogni R, Basosi R, Choinowski T, Martínez MJ, Piontek K, Martínez AT (2005) Versatile peroxidase oxidation of high redox potential aromatic compounds: Site-directed mutagenesis, spectroscopic and crystallographic investigations of three long-range electron transfer pathways. J Mol Biol 354:385-402

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH and others (2004) Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev 3:29-60

Riley R, Salamov AA, Brown DW, Nagy LG, Floudas D, Held BW, Levasseur A, Lombard V, Morin E, Otillar R and others (2014) Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc Natl Acad Sci USA 111:9923-9928

Royer-Carenzi M, Pontarotti P, Didier G (2013) Choosing the best ancestral character state reconstruction method. Mathematical Biosciences 242:95-109

Ruiz-Dueñas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martínez AT (2013) Lignin-degrading peroxidases in Polyporales: An evolutionary survey based on ten sequenced genomes. Mycologia 105:1428-1444

Ruiz-Dueñas FJ, Martínez AT (2009) Microbial degradation of lignin: How a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. Microbial Biotechnol 2:164-177

Ruiz-Dueñas FJ, Martínez AT (2010) Structural and functional features of peroxidases with a potential as industrial biocatalysts. In: Torres E, Ayala M (eds) Biocatalysts based on heme peroxidases. Springer-Verlag, Berlin, pp 37-59

Ruiz-Dueñas FJ, Martínez MJ, Martínez AT (1999) Molecular characterization of a novel peroxidase isolated from the ligninolytic fungus *Pleurotus eryngii*. Mol Microbiol 31:223-236

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441-452

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015) Demonstration of lignin-to-peroxidase direct electron transfer: A transient-state kinetics, directed mutagenesis, EPR and NMR study. J Biol Chem 290:23201-23213

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Shigeto J, Itoh Y, Tsutsumi Y, Kondo R (2012) Identification of Tyr74 and Tyr177 as substrate oxidation sites in cationic cell wall-bound peroxidase from *Populus alba* L. FEBS J 279:348-357

Smith AT, Doyle WA, Dorlet P, Ivancich A (2009) Spectroscopic evidence for an engineered, catalytically active Trp radical that creates the unique reactivity of lignin peroxidase. Proc Natl Acad Sci USA 106:16084-16089

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Weng JK, Chapple C (2010) The origin and evolution of lignin biosynthesis. New Phytol 187:273-285

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691-699

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Zámocký M, Hofbauer S, Schaffner I, Gasselhuber B, Nicolussi A, Soudi M, Pirker KF, Furtmüller PG, Obinger C (2015) Independent evolution of four heme peroxidase superfamilies. Arch Biochem Biophys 574:108-119

Zhaxybayeva O (2009) Detection and quantitative assessment of horizontal gene transfer. In: Gogarten MB, Gogarten JP, Olendzenski LC (eds) Horizontal Gene Transfer: Genomes in Flux. Humana Press, Totowa, NJ, pp 195-213

Chapter 2 Supporting Information references

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104-2105

Ayuso-Fernández I, Martínez AT, Ruiz-Dueñas FJ (2017) Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date. Biotechnol Biofuels 10:67

Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L and others (2014) SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252-W258

Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F, Sjokvist E, Copeland A, Foster B, Sun H, Larsson E and others (2013) Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105:1350-1373

Doyle WA, Smith AT (1996) Expression of lignin peroxidase H8 in *Escherichia coli*: Folding and activation of the recombinant enzyme with Ca^{2+} and haem. Biochem J 315:15-19

Fernández-Fueyo E, Ruiz-Dueñas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, Larrondo L, James TY, Seelenfreund D, Lobos S and others (2012) Comparative genomics of *Ceriporiopisis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc Natl Acad Sci USA 109:5458-5463

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis 30:S162-S173

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870-1874

Martinez D, Challacombe J, Morgenstern I, Hibbett DS, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Dueñas FJ, Martínez AT, Kersten P and others (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA 106:1954-1959

Morales M, Mate MJ, Romero A, Martínez MJ, Martínez AT, Ruiz-Dueñas FJ (2012) Two oxidation sites for low redox-potential substrates: A directed mutagenesis, kinetic and crystallographic study on *Pleurotus eryngii* versatile peroxidase. J Biol Chem 287:41053-41067

Ohm RA, Riley R, Salamov A, Min B, Choi IG, Grigoriev IV (2014) Genomics of wood-degrading fungi. Fungal Genet Biol 72:82-90

Pérez-Boada M, Doyle WA, Ruiz-Dueñas FJ, Martínez MJ, Martínez AT, Smith AT (2002) Expression of *Pleurotus eryngii* versatile peroxidase in *Escherichia coli* and optimisation of *in vitro* folding. Enzyme Microb Technol 30:518-524

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441-452

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691-699

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Chapter 3 references

Arnhold J, Furtmüller PG, Regelsberger G, Obinger C (2001) Redox properties of the couple compound I/native enzyme of myeloperoxidase and eosinophil peroxidase. Eur J Biochem 268:5142-5148

Ayuso-Fernández I, Martínez AT, Ruiz-Dueñas FJ (2017) Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date. Biotechnol Biofuels 10:67

Ayuso-Fernández I, Ruiz-Dueñas FJ, Martínez AT (2018) Evolutionary convergence in lignin degrading enzymes. Proc Natl Acad Sci USA 115:6428-6433

Banci L, Bertini I, Pease EA, Tien M, Turano P (1992) ^{1}H NMR investigation of manganese peroxidase from *Phanerochaete chrysosporium*. A comparison with other peroxidases. Biochemistry 31:10009-10017

Banci L, Bertini I, Pierattelli R, Tien M, Vila AJ (1995) Factoring of the hyperfine shifts in the cyanide adduct of lignin peroxidase from *P. chrysosporium*. J Am Chem Soc 117:8659-8667

Banci L, Bertini I, Turano P, Ferrer JC, Mauk AG (1991a) Comparative ^{1}H NMR study of ferric low-spin cytochrome c peroxidase and horseradish peroxidase. Inorg Chem 30:4510-4516

Banci L, Bertini I, Turano P, Tien M, Kirk TK (1991b) Proton NMR investigation into the basis for the relatively high redox potential of lignin peroxidase. Proc Natl Acad Sci USA 88:6956-6960

Banci L, Camarero S, Martínez AT, Martínez MJ, Pérez-Boada M, Pierattelli R, Ruiz-Dueñas FJ (2003) NMR study of Mn(II) binding by the new versatile peroxidase from the white-rot fungus *Pleurotus eryngii*. J Biol Inorg Chem 8:751-760

Battistuzzi G, Bellei M, Bortolotti CA, Sola M (2010) Redox properties of heme peroxidases. Arch Biochem Biophys 500:21-36

Battistuzzi G, Bellei M, De Rienzo F, Sola M (2006) Redox properties of the Fe3+/Fe2+ couple in *Arthromyces ramosus* class II peroxidase and its cyanide adduct. J Biol Inorg Chem 11:586-592

Bertini I, Turano P, Vila AJ (1993) Nuclear-Magnetic-Resonance of Paramagnetic Metalloproteins. Chem Rev 93:2833-2932

Castro L, Crawford L, Mutengwa A, Goetze JP, Buehl M (2016) Insights into structure and redox potential of lignin peroxidase from QM/MM calculations. Org Biomol Chem 14:2385-2389

Ciaccio C, Rosati A, De Sanctis G, Marini S, Santucci R, Ascenzi P, Welinder KG, Coletta M (2003) Relationships of ligand binding, redox properties and protonation in *Coprinus cinereus* peroxidase. J Biol Chem 278:18730-18737

Conroy CW, Tyma P, Daum PH, Erman JE (1978) Oxidation-reduction potential measurements of cytochrome c peroxidase and pH dependent spectral transitions in the ferrous enzyme. Biochimica et Biophysica Acta (BBA) - Protein Structure 537:62-69

Efimov I, Papadopoulou ND, McLean KJ, Badyal SK, Macdonald IK, Munro AW, Moody PCE, Raven EL (2007) The redox properties of ascorbate peroxidase. Biochemistry 46:8017-8023

Farhangrazi ZS, Fossett ME, Powers LS, Ellis WR (1995) Variable-temperature spectroelectrochemical study of horseradish peroxidase. Biochemistry 34:2866-2871

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Furtmüller PG, Arnhold J, Jantschko W, Pichler H, Obinger C (2003) Redox properties of the couples compound I/compound II and compound II/native enzyme of human myeloperoxidase. Biochem Biophys Res Commun 301:551-557

Furtmüller PG, Arnhold J, Jantschko W, Zederbauer M, Jakopitsch C, Obinger C (2005) Standard reduction potentials of all couples of the peroxidase cycle of lactoperoxidase. J Inorg Biochem 99:1220-1229

Gumulya Y, Gillam EMJ (2017) Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering. Biochem J 474:1

Hammel KE, Cullen D (2008) Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol 11:349-355

Inubushi T, Becker ED (1983) Efficient detection of paramagnetically shifted NMR resonances by optimizing the WEFT pulse sequence. J Magnet Reson 51:128-133

Kersten PJ, Kalyanaraman B, Hammel KE, Reinhammar B, Kirk TK (1990) Comparison of lignin peroxidase, horseradish peroxidase and laccase in the oxidation of methoxybenzenes. Biochem J 268:475-480

Kirk TK, Farrell RL (1987) Enzymatic "combustion": The microbial degradation of lignin. Annu Rev Microbiol 41:465-505

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547-1549

La Mar GN (2007) Application of the paramagnetic dipole field for solution NMR active site structure determination in low-spin, cyanide-inhibited ferric hemoproteins. IUBMB Life 59:513-527

La Mar GN, Deropp JS, Chacko VP, Satterlee JD, Erman JE (1982) Axial Histidyl Imidazole Non-Exchangeable Proton Resonances As Indicators of Imidazole Hydrogen-Bonding in Ferric Cyanide Complexes of Heme Peroxidases. Biochim Biophys Acta 708:317-325

Martínez AT (2002) Molecular biology and structure-function of lignin-degrading heme peroxidases. Enzyme Microb Technol 30:425-444

Martínez AT, Ruiz-Dueñas FJ, Martínez MJ, del Río JC, Gutiérrez A (2009) Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol 20:348-357

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F and others (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22:695-700

Millis CD, Cai D, Stankovich MT, Tien M (1989) Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*. Biochemistry 28:8484-8489

Oyadomari M, Shinohara H, Johjima T, Wariishi H, Tanaka H (2003) Electrochemical characterization of lignin peroxidase from the white-rot basidiomycete *Phanerochaete chrysosporium*. J Mol Catal B-Enzym 21:291-297

Poulos TL (2014) Heme enzyme structure and function. Chem Rev 114:3919-3962

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M and others (2014) Lignin valorization: improving lignin processing in the biorefinery. Science 344:1246843

Risso VA, Gavira JA, Mejia-Carmona DF, Gaucher EA, Sanchez-Ruiz JM (2013) Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. J Am Chem Soc 135:2899-2902

Ruiz-Dueñas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martínez AT (2013) Lignin-degrading peroxidases in Polyporales: An evolutionary survey based on ten sequenced genomes. Mycologia 105:1428-1444

Ruiz-Dueñas FJ, Martínez AT (2009) Microbial degradation of lignin: How a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. Microbial Biotechnol 2:164-177

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441-452

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015) Demonstration of lignin-to-peroxidase direct electron transfer: A transient-state kinetics, directed mutagenesis, EPR and NMR study. J Biol Chem 290:23201-23213

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Santucci R, Bongiovanni C, Marini S, del Conte R, Tien M, Banci L, Coletta M (2000) Redox equilibria of manganese peroxidase from *Phanerochaetes chrysosporium*: functional role of residues on the proximal side of the haem pocket. Biochem J 349:85-90

Satterlee JD, Erman JE (1991) Proton NMR assignments of heme contacts and catalytically implicated amino acids in cyanide-ligated cytochrome *c* peroxidase determined from one- and two-dimensional nuclear Overhauser effects. Biochemistry 30:4398-4405

Sørlie M, Seefeldt LC, Parker VD (2000) Use of stopped-flow spectrophotometry to establish midpoint potentials for redox proteins. Anal Biochem 287:118-125

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Wang XS, Peter S, Ullrich R, Hofrichter M, Groves JT (2013) Driving force for oxygen-atom transfer by heme-thiolate enzymes. Angew Chem Int Ed 52:9238-9241

Whitwam RE, Koduri RS, Natan M, Tien M (1999) Role of axial ligands in the reactivity of Mn peroxidase from *Phanerochaete chrysosporium*. Biochemistry 38:9608-9616

Yamazaki I (1996) Factors that control one-electron reduction potentials of ferric, compound II and compound I of peroxidase. In: Obinger C, Burner U, Ebermann R, Penel C, Greppin H (eds) Plant peroxidases: Biochemistry and physiology. University of Geneva, Geneva, pp 34-41

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Zámocký M, Hofbauer S, Schaffner I, Gasselhuber B, Nicolussi A, Soudi M, Pirker KF, Furtmüller PG, Obinger C (2015) Independent evolution of four heme peroxidase superfamilies. Arch Biochem Biophys 574:108-119

Chapter 3 Supporting Information references

Arnhold J, Furtmüller PG, Regelsberger G, Obinger C (2001) Redox properties of the couple compound I/native enzyme of myeloperoxidase and eosinophil peroxidase. Eur J Biochem 268:5142-5148

Ayuso-Fernández I, Martínez AT, Ruiz-Dueñas FJ (2017) Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date. Biotechnol Biofuels 10:67

Ayuso-Fernández I, Ruiz-Dueñas FJ, Martínez AT (2018) Evolutionary convergence in lignin degrading enzymes. Proc Natl Acad Sci USA 115:6428-6433

Banci L, Bertini I, Pease EA, Tien M, Turano P (1992) ^{1}H NMR investigation of manganese peroxidase from *Phanerochaete chrysosporium*. A comparison with other peroxidases. Biochemistry 31:10009-10017

Banci L, Bertini I, Pierattelli R, Tien M, Vila AJ (1995) Factoring of the hyperfine shifts in the cyanide adduct of lignin peroxidase from *P. chrysosporium*. J Am Chem Soc 117:8659-8667

Banci L, Bertini I, Turano P, Ferrer JC, Mauk AG (1991a) Comparative ^{1}H NMR study of ferric low-spin cytochrome c peroxidase and horseradish peroxidase. Inorg Chem 30:4510-4516

Banci L, Bertini I, Turano P, Tien M, Kirk TK (1991b) Proton NMR investigation into the basis for the relatively high redox potential of lignin peroxidase. Proc Natl Acad Sci USA 88:6956-6960

Banci L, Camarero S, Martínez AT, Martínez MJ, Pérez-Boada M, Pierattelli R, Ruiz-Dueñas FJ (2003) NMR study of Mn(II) binding by the new versatile peroxidase from the white-rot fungus *Pleurotus eryngii*. J Biol Inorg Chem 8:751-760

Battistuzzi G, Bellei M, De Rienzo F, Sola M (2006) Redox properties of the Fe3+/Fe2+ couple in *Arthromyces ramosus* class II peroxidase and its cyanide adduct. J Biol Inorg Chem 11:586-592

Battistuzzi G, Bellei M, Vlasits J, Banerjee S, Furtmüller PG, Sola M, Obinger C (2010) Redox thermodynamics of lactoperoxidase and eosinophil peroxidase. Arch Biochem Biophys 494:72-77

Ciaccio C, Rosati A, De Sanctis G, Marini S, Santucci R, Ascenzi P, Welinder KG, Coletta M (2003) Relationships of ligand binding, redox properties and protonation in *Coprinus cinereus* peroxidase. J Biol Chem 278:18730-18737

DeFelippis MR, Murthy CP, Faraggi M, Klapper MH (1989) Pulse radiolytic measurement of redox potentials: the tyrosine and tryptophan radicals. Biochemistry 28:4847-4853

Doyle WA, Smith AT (1996) Expression of lignin peroxidase H8 in *Escherichia coli*: Folding and activation of the recombinant enzyme with Ca^{2+} and haem. Biochem J 315:15-19

Furtmüller PG, Arnhold J, Jantschko W, Pichler H, Obinger C (2003) Redox properties of the couples compound I/compound II and compound II/native enzyme of human myeloperoxidase. Biochem Biophys Res Commun 301:551-557

Furtmüller PG, Arnhold J, Jantschko W, Zederbauer M, Jakopitsch C, Obinger C (2005) Standard reduction potentials of all couples of the peroxidase cycle of lactoperoxidase. J Inorg Biochem 99:1220-1229

Inubushi T, Becker ED (1983) Efficient detection of paramagnetically shifted NMR resonances by optimizing the WEFT pulse sequence. J Magnet Reson 51:128-133

Koppenol WH (1987) Thermodynamics of reactions involving oxyradicals and hydrogen peroxide. Bioelectrochem Bioenerg 18:3-11

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F and others (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22:695-700

Millis CD, Cai D, Stankovich MT, Tien M (1989) Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*. Biochemistry 28:8484-8489

Moss DS, Naberdryk E, Breton J, Mäntele W (1990) Redox-linked conformational changes in proteins detected by a combination of infrared spectroscopy and protein electrochemistry. Eur J Biochem 187:565-572

Oyadomari M, Shinohara H, Johjima T, Wariishi H, Tanaka H (2003) Electrochemical characterization of lignin peroxidase from the white-rot basidiomycete *Phanerochaete chrysosporium*. J Mol Catal B-Enzym 21:291-297

Ruiz-Dueñas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martínez AT (2013) Lignin-degrading peroxidases in Polyporales: An evolutionary survey based on ten sequenced genomes. Mycologia 105:1428-1444

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441-452

Santucci R, Bongiovanni C, Marini S, del Conte R, Tien M, Banci L, Coletta M (2000) Redox equilibria of manganese peroxidase from *Phanerochaetes chrysosporium*: functional role of residues on the proximal side of the haem pocket. Biochem J 349:85-90

Sørlie M, Seefeldt LC, Parker VD (2000) Use of stopped-flow spectrophotometry to establish midpoint potentials for redox proteins. Anal Biochem 287:118-125

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Wariishi H, Márquez L, Dunford HB, Gold MH (1990) Lignin peroxidase compounds II and III. Spectral and kinetic characterization of reactions with peroxides. J Biol Chem 265:11137-11142

Whitwam RE, Koduri RS, Natan M, Tien M (1999) Role of axial ligands in the reactivity of Mn peroxidase from *Phanerochaete chrysosporium*. Biochemistry 38:9608-9616

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Chapter 4 references

Ayuso-Fernández I, De Lacey AL, Cañada FJ, Ruiz-Dueñas FJ, Martínez AT (2018a) Redox potential increased during the evolution of enzymes degrading recalcitrant lignin. Chem -Eur J 25:2708-2712

Ayuso-Fernández I, Martínez AT, Ruiz-Dueñas FJ (2017) Experimental recreation of the evolution of lignin degrading enzymes from the Jurassic to date. Biotechnol Biofuels 10:67

Ayuso-Fernández I, Ruiz-Dueñas FJ, Martínez AT (2018b) Evolutionary convergence in lignin degrading enzymes. Proc Natl Acad Sci USA 115:6428-6433

Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating Divergence Times in Large Phylogenetic Trees. Syst Biol 56:741-752

Camarero S, Bocchini P, Galletti GC, Martínez AT (1999) Pyrolysis-gas chromatography/mass spectrometry analysis of phenolic and etherified units in natural and industrial lignins. Rapid Commun Mass Spectrom 13:630-636

Cullen D, Kersten PJ (2004) Enzymology and molecular biology of lignin degradation. In: Brambl R, Marzluf GA (eds) Mycota III: Biochemistry and Molecular Biology. Springer, Berlin, pp 249-273

Fengel D,Wegener G (1984) Wood: Chemistry, ultrastructure, reactions. De Gruyter, Berlin

Fernández-Fueyo E, Acebes S, Ruiz-Dueñas FJ, Martínez MJ, Romero A, Medrano FJ, Guallar V, Martínez AT (2014) Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi. Acta Crystallogr D Biol Crystallogr 70:3253-3265

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Glenn JK, Akileswaran L, Gold MH (1986) Mn(II) oxidation is the principal function of the extracellular Mn-peroxidase from *Phanerochaete chrysosporium*. Arch Biochem Biophys 251:688-696

Hibbett D, Blanchette R, Kenrick P, Mills B (2016) Climate, decay, and the death of the coal forests. Curr Biol 26:R563-R567

Kirk TK, Farrell RL (1987) Enzymatic "combustion": The microbial degradation of lignin. Annu Rev Microbiol 41:465-505

Krah FS, Bässler C, Heibl C, Soghigian J, Schaefer H, Hibbett DS (2018) Evolutionary dynamics of host specialization in wood-decay fungi. BMC Evolutionary Biology 18:119

Lai YZ, Guo XP (1991) Variation of the phenolic hydroxyl group content in wood lignins. Wood Sci Technol 25:467-472

Li LG, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL (2001) The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. Plant Cell 13:1567-1585

Mansfield SD, Kim H, Lu FC, Ralph J (2012) Whole plant cell wall characterization using solution-state 2D NMR. Nature Protocols 7:1579-1589

Martínez AT, Camarero S, Ruiz-Dueñas FJ,Martínez MJ (2018) Biological lignin degradation. In: Beckham GT (ed) Lignin valorization: Emerging approaches, Chapter 8. RSC, Cambridge, pp 199-225

Martínez AT, Ruiz-Dueñas FJ, Martínez MJ, del Río JC, Gutiérrez A (2009) Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol 20:348-357

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F and others (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat Biotechnol 22:695-700

Menden B, Kohlhoff M, Moerschbacher BM (2007) Wheat cells accumulate a syringyl-rich lignin during the hypersensitive resistance response. Phytochemistry 68:513-520

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ (2018) The timescale of early land plant evolution. Proc Natl Acad Sci USA 115:E2274

Nagy LG, Riley R, Tritt A, Adam C, Daum C, Floudas D, Sun H, Yadav JS, Pangilinan J, Larsson KH and others (2016) Comparative genomics of early-diverging mushroom-forming fungi provides insights into the origins of lignocellulose decay capabilities. Mol Biol Evol 33:959-970

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M and others (2014) Lignin valorization: improving lignin processing in the biorefinery. Science 344:1246843

Ralph J, Landucci LL (2010) NMR of lignin. In: Heitner C, Dimmel D, Schmidt J (eds) Lignin and lignans: Advances in chemistry. CRC Press, Boca Raton, pp 137-243

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH and others (2004) Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev 3:29-60

Ralph J, Marita JM, Ralph SA, Hatfield RD, Lu F, Ede RM, Peng J, Quideau S, Helm RF, Grabber JH and others (1999) Solution-state NMR of lignin. In: Argyropoulos DS (ed) Advances in lignocellulosics characterization. Tappi Press, Atlanta, pp 55-108

Ruiz-Dueñas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martínez AT (2013) Lignin-degrading peroxidases in Polyporales: An evolutionary survey based on ten sequenced genomes. Mycologia 105:1428-1444

Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot 60:441-452

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015a) Demonstration of lignin-to-peroxidase direct electron transfer. A transient-state kinetics, directed mutagenesis, EPR and NMR study (vol 290, pag 23201, 2015). J Biol Chem 290:30268

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015b) Demonstration of lignin-to-peroxidase direct electron transfer: A transient-state kinetics, directed mutagenesis, EPR and NMR study. J Biol Chem 290:23201-23213

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Sarkanen KV, Hergert HL (1971) Classification and distribution. In: Sarkanen KV, Ludwig CH (eds) Lignins - Occurrence, Formation, Structure and Reactions. Wiley-Interscience, New York, pp 43-94

Smith AT, Doyle WA, Dorlet P, Ivancich A (2009) Spectroscopic evidence for an engineered, catalytically active Trp radical that creates the unique reactivity of lignin peroxidase. Proc Natl Acad Sci USA 106:16084-16089

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Tobimatsu Y, Schuetz M (2019) Lignin polymerization: how do plants manage the chemistry so well? Curr Opin Biotechnol 56:75-81

Varga T, Krizsán K, Földi C, Dima B, Sánchez-García M, Sánchez-Ramírez S, Szöllösi GJ, Szarkándi JG, Papp V, Albert L and others (2019) Megaphylogeny resolves global patterns of mushroom evolution. Nat Ecol Evol

Weng JK, Chapple C (2010) The origin and evolution of lignin biosynthesis. New Phytol 187:273-285

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Chapter 4 Supporting references

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104-2105

Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F, Sjokvist E, Copeland A, Foster B, Sun H, Larsson E and others (2013) Phylogenetic and phylogenomic overview of the Polyporales. Mycologia 105:1350-1373

Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating Divergence Times in Large Phylogenetic Trees. Syst Biol 56:741-752

Doyle WA, Smith AT (1996) Expression of lignin peroxidase H8 in *Escherichia coli*: Folding and activation of the recombinant enzyme with Ca^{2+} and haem. Biochem J 315:15-19

Fernández-Fueyo E, Ruiz-Dueñas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, Larrondo L, James TY, Seelenfreund D, Lobos S and others (2012) Comparative genomics of *Ceriporiopisis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc Natl Acad Sci USA 109:5458-5463

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Hane JK, Lowe RGT, Solomon PS, Tan KC, Schoch CL, Spatafora JW, Crous PW, Kodira C, Birren BW, Galagan JE and others (2007) DothideomyceteΓÇôPlant Interactions Illuminated by Genome Sequencing and EST Analysis of the Wheat Pathogen <em>Stagonospora nodorum</em&gt. The Plant Cell 19:3347

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870-1874

Lebo SE, Braaten SM, Fredheim GE, Lutnaes BF, Lauten RA, Myrvold BO,McNally TJ (2008) Recent advances in the characterization of lignosulfonates. In: Hu T (ed) Characterization of lignocellulosic materials. Blackwell Pub., New York, pp 189-205

Lutnaes BF, Myrvold BO, Lauten RA, Endeshaw MM (2008) ^{1}H and ^{13}C NMR data of benzylsulfonic acids - model compounds for lignosulfonate. Magn Reson Chem 46:299-305

Magina S, Marques AP, Evtuguin DV (2015) Study on the residual lignin in *Eucalyptus globulus* sulphite pulp. Holzforschung 69:513-522

Martínez AT, Rencoret J, Marques G, Gutiérrez A, Ibarra D, Jiménez-Barbero J, del Río JC (2008) Monolignol acylation and lignin structure in some nonwoody plants: A 2D NMR study. Phytochemistry 69:2831-2843

Martinez D, Challacombe J, Morgenstern I, Hibbett DS, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Dueñas FJ, Martínez AT, Kersten P and others (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA 106:1954-1959

Ohm RA, Riley R, Salamov A, Min B, Choi IG, Grigoriev IV (2014) Genomics of wood-degrading fungi. Fungal Genet Biol 72:82-90

Pérez-Boada M, Doyle WA, Ruiz-Dueñas FJ, Martínez MJ, Martínez AT, Smith AT (2002) Expression of *Pleurotus eryngii* versatile peroxidase in *Escherichia coli* and optimisation of *in vitro* folding. Enzyme Microb Technol 30:518-524

Prasetyo EN, Kudanga T, Ostergaard L, Rencoret J, Gutiérrez A, del Río JC, Santos JI, Nieto L, Jimenez-Barbero J, Martínez AT and others (2010) Polymerization of lignosulfonates by the laccase-HBT (1-hydroxybenzotriazole) system improves dispersibility. Bioresource Technol 101:5054-5062

Ralph S, Ralph J, Landucci L (2009) NMR database of lignin and cell wall model compounds. available at URL www.glbrc.org/databases_and_software/nmrdatabase),

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015a) Demonstration of lignin-to-peroxidase direct electron transfer. A transient-state kinetics, directed mutagenesis, EPR and NMR study (vol 290, pag 23201, 2015). J Biol Chem 290:30268

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015b) Demonstration of lignin-to-peroxidase direct electron transfer: A transient-state kinetics, directed mutagenesis, EPR and NMR study. J Biol Chem 290:23201-23213

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57:758-771

Whelan S, Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691-699

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

General discussion references

Banci L, Bertini I, Pease EA, Tien M, Turano P (1992) ^{1}H NMR investigation of manganese peroxidase from *Phanerochaete chrysosporium*. A comparison with other peroxidases. Biochemistry 31:10009-10017

Banci L, Bertini I, Pierattelli R, Tien M, Vila AJ (1995) Factoring of the hyperfine shifts in the cyanide adduct of lignin peroxidase from *P. chrysosporium*. J Am Chem Soc 117:8659-8667

Banci L, Bertini I, Turano P, Ferrer JC, Mauk AG (1991a) Comparative ^{1}H NMR study of ferric low-spin cytochrome c peroxidase and horseradish peroxidase. Inorg Chem 30:4510-4516

Banci L, Bertini I, Turano P, Tien M, Kirk TK (1991b) Proton NMR investigation into the basis for the relatively high redox potential of lignin peroxidase. Proc Natl Acad Sci USA 88:6956-6960

Banci L, Camarero S, Martínez AT, Martínez MJ, Pérez-Boada M, Pierattelli R, Ruiz-Dueñas FJ (2003) NMR study of Mn(II) binding by the new versatile peroxidase from the white-rot fungus *Pleurotus eryngii*. J Biol Inorg Chem 8:751-760

Battistuzzi G, Bellei M, Bortolotti CA, Sola M (2010) Redox properties of heme peroxidases. Arch Biochem Biophys 500:21-36

Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating Divergence Times in Large Phylogenetic Trees. Syst Biol 56:741-752

Camarero S, Bocchini P, Galletti GC, Martínez AT (1999) Pyrolysis-gas chromatography/mass spectrometry analysis of phenolic and etherified units in natural and industrial lignins. Rapid Commun Mass Spectrom 13:630-636

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS and others (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715-1719

Gaucher EA, Govindarajan S, Ganesh OK (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins. Nature 451:704

Gold MH, Youngs HL, Gelpke MD (2000) Manganese peroxidase. Met Ions Biol Syst 37:559-586

Hammel KE, Cullen D (2008) Role of fungal peroxidases in biological ligninolysis. Curr Opin Plant Biol 11:349-355

Hibbett D, Blanchette R, Kenrick P, Mills B (2016) Climate, decay, and the death of the coal forests. Curr Biol 26:R563-R567

Kersten P, Cullen D (2007) Extracellular oxidative systems of the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. Fungal Genet Biol 44:77-87

Krah FS, Bässler C, Heibl C, Soghigian J, Schaefer H, Hibbett DS (2018) Evolutionary dynamics of host specialization in wood-decay fungi. BMC Evolutionary Biology 18:119

Lai YZ, Guo XP (1991) Variation of the phenolic hydroxyl group content in wood lignins. Wood Sci Technol 25:467-472

Li LG, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL (2001) The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. Plant Cell 13:1567-1585

Liberles DA (2007) Ancestral sequence reconstruction. Oxford University Press, Oxford

Martínez AT (2002) Molecular biology and structure-function of lignin-degrading heme peroxidases. Enzyme Microb Technol 30:425-444

Martínez AT, Camarero S, Ruiz-Dueñas FJ,Martínez MJ (2018) Biological lignin degradation. In: Beckham GT (ed) Lignin valorization: Emerging approaches, Chapter 8. RSC, Cambridge, pp 199-225

Martínez AT, Ruiz-Dueñas FJ, Martínez MJ, del Río JC, Gutiérrez A (2009) Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol 20:348-357

Millis CD, Cai D, Stankovich MT, Tien M (1989) Oxidation-reduction potentials and ionization states of extracellular peroxidases from the lignin-degrading fungus *Phanerochaete chrysosporium*. Biochemistry 28:8484-8489

Morris JL, Puttick MN, Clark JW, Edwards D, Kenrick P, Pressel S, Wellman CH, Yang Z, Schneider H, Donoghue PCJ (2018) The timescale of early land plant evolution. Proc Natl Acad Sci USA 115:E2274

Omland KE (1999) The assumptions and challenges of ancestral state reconstructions. Syst Biol 48:604-611

Oyadomari M, Shinohara H, Johjima T, Wariishi H, Tanaka H (2003) Electrochemical characterization of lignin peroxidase from the white-rot basidiomycete *Phanerochaete chrysosporium*. J Mol Catal B-Enzym 21:291-297

Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Davis MF, Davison BH, Dixon RA, Gilna P, Keller M and others (2014) Lignin valorization: improving lignin processing in the biorefinery. Science 344:1246843

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH and others (2004) Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev 3:29-60

Royer-Carenzi M, Pontarotti P, Didier G (2013) Choosing the best ancestral character state reconstruction method. Mathematical Biosciences 242:95-109

Ruiz-Dueñas FJ, Morales M, Pérez-Boada M, Choinowski T, Martínez MJ, Piontek K, Martínez AT (2007) Manganese oxidation site in *Pleurotus eryngii* versatile peroxidase: A site-directed mutagenesis, kinetic and crystallographic study. Biochemistry 46:66-77

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015a) Demonstration of lignin-to-peroxidase direct electron transfer. A transient-state kinetics, directed mutagenesis, EPR and NMR study (vol 290, pag 23201, 2015). J Biol Chem 290:30268

Sáez-Jiménez V, Baratto MC, Pogni R, Rencoret J, Gutiérrez A, Santos JI, Martínez AT, Ruiz-Dueñas FJ (2015b) Demonstration of lignin-to-peroxidase direct electron transfer: A transient-state kinetics, directed mutagenesis, EPR and NMR study. J Biol Chem 290:23201-23213

Sáez-Jiménez V, Rencoret J, Rodríguez-Carvajal MA, Gutiérrez A, Ruiz-Dueñas FJ, Martínez AT (2016) Role of surface tryptophan for peroxidase oxidation of nonphenolic lignin. Biotechnol Biofuels 9:198:

Sarkanen KV, Hergert HL (1971) Classification and distribution. In: Sarkanen KV, Ludwig CH (eds) Lignins - Occurrence, Formation, Structure and Reactions. Wiley-Interscience, New York, pp 43-94

Siddiq MA, Hochberg GKA, Thornton JW (2017) Evolution of protein specificity: insights from ancestral protein reconstruction. Current Opinion in Structural Biology 47:113-122

Varga T, Krizsán K, Földi C, Dima B, Sánchez-García M, Sánchez-Ramírez S, Szöllösi GJ, Szarkándi JG, Papp V, Albert L and others (2019) Megaphylogeny resolves global patterns of mushroom evolution. Nat Ecol Evol

Weng JK, Chapple C (2010) The origin and evolution of lignin biosynthesis. New Phytol 187:273-285

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586-1591

Zhaxybayeva O (2009) Detection and quantitative assessment of horizontal gene transfer. In: Gogarten MB, Gogarten JP, Olendzenski LC (eds) Horizontal Gene Transfer: Genomes in Flux. Humana Press, Totowa, NJ, pp 195-213