



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA QUÍMICA Y BIOTECNOLOGÍA

**BÚSQUEDA DE ENZIMAS LPMO DE HONGOS PARA LA PRODUCCIÓN DE
BIOETANOL A PARTIR DE MATERIAL LIGNOCELULÓSICO.**

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN QUÍMICA
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL EN BIOTECNOLOGÍA

GONZALO NICOLÁS CARVAJAL LOREN

PROFESORA GUÍA:

ORIANA SALAZAR AGUIRRE

MIEMBROS DE LA COMISIÓN:

MARÍA ELENA LIENQUEO CONTRERAS
ÁLVARO OLIVERA NAPPA
JUAN PABLO ACEVEDO COX

SANTIAGO DE CHILE

ABRIL 2014

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE:
Ingeniero Civil en Biotecnología y al grado de Magíster en Ciencias
de la Ingeniería, mención Química
POR: Gonzalo Nicolás Carvajal Loren
FECHA: Abril 2014
PROFESORA GUÍA: Oriana Salazar Aguirre

BÚSQUEDA DE ENZIMAS LPMO DE HONGOS PARA LA PRODUCCIÓN DE BIOETANOL A PARTIR DE MATERIAL LIGNOCELULÓSICO.

La sociedad se enfrenta de forma cada vez más fehaciente y cercana a una crisis del modelo energético actual. El uso indiscriminado de combustibles fósiles, y los problemas medioambientales que su uso acarrea han llevado lentamente hacia un cambio de paradigma. Una de las alternativas renovables al uso de estos es la producción de bioetanol a partir de desechos lignocelulósicos. Este proceso, aunque prometedor, presenta problemas en la eficiencia de las etapas de pretratamiento y de hidrólisis, que no le permiten afianzarse en el mercado energético. En la búsqueda de soluciones, se ha apuntado en los últimos años a las proteínas de hongos capaces de degradar la madera para su consumo. Entre estas, las monooxigenasas líticas de polisacáridos (LPMO) han cobrado gran importancia en el último tiempo. Estas son parte importante del proceso de despolimerización de la celulosa, y su capacidad de aumentar la eficiencia de éste las han puesto en el centro de atención.

Este trabajo de tesis tiene como objetivo la búsqueda de nuevas LPMOs desde dos hongos: *Fusarium oxysporum* y *Gloeophyllum trabeum*. Más específicamente, se buscó identificar enzimas de interés en el secretoma del hongo y secuenciarlas para su posterior análisis *in silico*.

Estos objetivos se desarrollaron en una primera instancia a través de la búsqueda de medios de cultivo que indujeran la producción de celulasas por parte de los hongos. A partir de la biomasa producida en los cultivos se realizó una extracción de RNA, para luego transcribir el mRNA a cDNA. El cDNA fue utilizado en reacciones de PCR con partidores degenerados. Esto tuvo como fin identificar LPMOs de entre las proteínas transcritas por el hongo. Los genes identificados fueron amplificados usando partidores específicos, y transformados en *E. coli* unidos a vectores de clonación. Finalmente los genes fueron secuenciados y esta información se utilizó para analizar *in silico* las propiedades de las LPMOs codificadas por los genes clonados.

Se pudo determinar durante el trabajo medios idóneos para la inducción de la producción de celulasas, y se tuvo la oportunidad de mejorar el proceso de extracción de RNA desde hongos. Del proceso de secuenciación se obtuvo las secuencias de un gen perteneciente a *G. trabeum* y de dos genes pertenecientes a *F. oxysporum*. Se observó que las proteínas que estos codifican poseen los componentes estructurales descritos como indispensables para ser enzimas activas, vale decir las dos histidinas y la tirosina que forman el sitio activo; se estableció el mapa de restricción de los genes, y se realizó un análisis filogenético de estos en relación a proteínas de referencia. Del análisis de las secuencias obtenidas se determinó que, según la clasificación vigente, serían parte de la subfamilia 3 de las LPMOs. Finalmente, se pudo aproximar mediante herramientas computacionales la estructura tridimensional de estas proteínas. Se observó que la estructura predicha sigue los patrones que presentan las LPMOs ya caracterizadas, otro indicador de que se podría tratar de enzimas funcionales.

En conclusión, se puede decir que se cumplieron los objetivos propuestos en este trabajo. Se logró identificar y secuenciar dos LPMOs expresadas por *F. oxysporum* y una por *G. trabeum*. Asimismo, se logró realizar un análisis en profundidad de las secuencias utilizadas. Este permite suponer que se trata efectivamente de proteínas con el potencial de mejorar el proceso de producción de biocombustibles, y que se debiese seguir al siguiente paso lógico, que correspondería al análisis de la funcionalidad de la proteína nativa o expresada en forma heteróloga, y su rol en la degradación de lignocelulosa.

Este trabajo fue financiado por el proyecto Fondecyt Regular N° 1121088.

A mis padres.

Agradecimientos

Me gustaría partir agradeciendo a mi familia. Ellos han sido mi principal apoyo durante todos mis años de desarrollo y estudio. Agradecimientos especiales a mi madre y mi padre, que han trabajado duro por entregarme todo lo que he necesitado, y han dado lo mejor de sí para criarme como una persona íntegra. Por todo esto les estaré por siempre agradecido.

Quisiera agradecer además a mi profesora guía, Oriana Salazar. Pocas veces me he encontrado con una profesora tan preocupada por la docencia y el trabajo de sus alumnos. Su apoyo y alta exigencia son características que he llegado a valorar, y sin las cuales este trabajo no habría alcanzado el nivel que tiene.

Agradecer también a Nancy Carrasco, con su infinita buena voluntad y paciencia para impartir todo su conocimiento. El laboratorio no sería lo mismo sin ella.

Fue además un agrado y privilegio poder compartir con los otros estudiantes del centro. Por nombrar sólo algunos, Javier M., Lisset M., Javier D., Daniela S., entre muchos otros. Todos ustedes y el ambiente que generan hacen del trabajo en el laboratorio una experiencia enriquecedora y entretenida.

Agradecer finalmente a Sandy Contreras. Su amor, comprensión y paciencia me han acompañado en todo este proceso. Gracias por ser mi compañera y apoyarme en los momentos buenos y sobre todo en los malos.

Tabla de contenido

Resumen	II
Agradecimientos.....	IV
Tabla de contenido	V
Índice de tablas.....	IX
Índice de ilustraciones	X
Lista de abreviaciones	XI
I Introducción	1
I.1 MARCO TEÓRICO.....	1
I.1.1 Antecedentes generales.....	1
I.1.1.1 Problema energético mundial	1
I.1.2 Biocombustibles.....	2
I.1.2.1 Biocombustibles de primera y tercera generación	3
I.1.2.2 Biocombustibles de segunda generación	3
I.1.3 Monooxigenasas Líticas de Polisacáridos (LPMO).....	7
I.1.3.1 Historia.....	7
I.1.3.2 Estructura y actividad.....	7
I.1.4 Hongos productores de celulasas.....	10
I.2 DESCRIPCIÓN Y JUSTIFICACIÓN DEL PROYECTO	12
I.3 OBJETIVOS.....	13
I.3.1 Objetivo general.....	13
I.3.2 Objetivos específicos	13
II Materiales y métodos	14
II.1 MATERIALES.....	14
II.1.1 Lista de reactivos y kits utilizados.....	14
II.1.2 Equipos.....	14
II.1.3 Herramientas computacionales	15
II.1.4 Vector de clonación.....	15
II.2 MÉTODOS.....	17
II.2.1 Medios de cultivo.....	17

II.2.1.1	Medios de cultivo sólido	17
II.2.1.2	Medios de cultivo líquidos	17
II.2.1.3	Ensayo de actividad en medios líquidos.....	18
II.2.2	Extracción de RNA	18
II.2.2.1	Verificación de la integridad del RNA extraído	19
II.2.3	Síntesis de cDNA	19
II.2.3.1	Retrotranscripción de RNA a cDNA.....	19
II.2.3.2	Purificación del cDNA	19
II.2.4	Extracción de DNA genómico	19
II.2.5	Diseño de partidores	20
II.2.5.1	Diseño de partidores degenerados	20
II.2.5.2	Diseño de partidores específicos	20
II.2.6	Amplificación mediante PCR.....	20
II.2.6.1	Amplificación mediante PCR con partidores degenerados.....	20
II.2.6.2	Amplificación mediante PCR con polimerasa de alta fidelidad	21
II.2.6.3	Amplificación de fragmento 18S.....	21
II.2.7	Clonación de genes.....	22
II.2.7.1	Ligación	22
II.2.7.2	A-Tailing.....	22
II.2.7.3	Electroporación	22
II.2.7.4	Verificación de la transformación	22
II.2.7.5	Secuenciación.....	23
II.2.8	Análisis bioinformático.....	23
II.2.8.1	Mapa de restricción	23
II.2.8.2	Alineamiento de secuencias con otras de la familia AA9.....	23
II.2.8.3	Determinación de la subfamilia de las LPMOs	23
II.2.8.4	Modelamiento de la estructura de las LPMOs en base a homología.	23
II.2.8.1	Análisis filogenético.....	24
III	Resultados y Discusión	25
III.1	IDENTIFICACIÓN DE MEDIOS DE CULTIVO INDUCTORES DE LA PRODUCCIÓN DE CELULASAS.....	25

III.1.1	Medios de cultivos para <i>Fusarium oxysporum</i>	25
III.1.2	Medios de cultivo para <i>Gloeophyllum trabeum</i>	28
III.2	IDENTIFICACIÓN DE LOS HONGOS MEDIANTE SECUENCIACIÓN DEL rRNA 18S	31
III.3	EVALUACIÓN DE MÉTODO DE EXTRACCIÓN DE RNA DESDE HONGOS	33
III.4	DISEÑO DE PARTIDORES	35
III.4.1	Partidores degenerados para la identificación de LPMOs	35
III.4.2	Partidores específicos para la amplificación de LPMOs	37
III.5	IDENTIFICACIÓN DE GENES USANDO PARTIDORES DEGENERADOS	37
III.5.1	Amplificación mediante partidores degenerados.....	37
III.5.2	Clonamiento en vector pGEM-T Easy y secuenciación	38
III.6	CLONACIÓN DE GENES COMPLETOS USANDO PARTIDORES ESPECÍFICOS	39
III.6.1	Amplificación de genes completos de LPMO.....	39
III.6.2	Obtención de secuencias de LPMOs completas.....	40
III.7	ANÁLISIS BIOINFORMÁTICO DE LAS SECUENCIAS DE LPMOS IDENTIFICADAS.....	42
III.7.1	Mapa de restricción de las LPMOs	42
III.7.2	Comparación con otras secuencias de LPMO de la familia AA9.....	44
III.7.3	Determinación de la subfamilia de las LPMOs	45
III.7.4	Estructura tridimensional de las LPMOs identificadas.....	49
III.7.5	Análisis filogenético de las LPMOs identificadas	58
IV	Conclusiones	62
V	Recomendaciones	64
VI	Bibliografía	65
VII	Anexos	71
VII.1	PREPARACIÓN DE BACTERIAS E. COLI DH5A ELECTROCOMPETENTES.....	71
VII.2	MEDIOS DE CULTIVO SÓLIDOS	71
VII.3	TRATAMIENTO DE IMPLEMENTOS MEDIANTE DEPC.....	71
VII.4	NOMENCLATURA IUPAC PARA PARTIDORES DEGENERADOS	71
VII.5	ALINEAMIENTOS PARA EL DISEÑO DE PARTIDORES DEGENERADOS	72
VII.5.1	Alineamiento para el diseño de los partidores GH61_G1	72

VII.5.2	<i>Alineamiento para el diseño de los partidores GH61_G2</i>	73
VII.5.3	<i>Alineamiento para el diseño de los partidores GH61_G4</i>	74
VII.6	MAPAS DE LOS VECTORES DE EXPRESIÓN SELECCIONADOS	75
VII.7	SECUENCIAS DEL RRNA 18S DE LOS HONGOS.....	77
VII.8	SECUENCIAS PARCIALES DE LPMOS	78
VII.9	SECUENCIAS COMPLETAS DE LAS LPMOS.....	79
VII.9.1	<i>Secuencias nucleotídicas de las LPMOs identificadas</i>	79
VII.9.2	<i>Secuencias proteicas de las LPMOs identificadas</i>	80
VII.10	RESULTADOS DE LA MODELACIÓN TRIDIMENSIONAL.....	81
VII.10.1	<i>Herramientas de evaluación de modelos tridimensionales de proteínas</i>	81
VII.10.2	<i>Resultados para el mejor modelo obtenido para FoPMO9A</i>	82
VII.10.3	<i>Resultados para el mejor modelo obtenido para FoPMO9B</i>	83
VII.10.4	<i>Resultados para el mejor modelo obtenido para GtPMO9A</i>	84

Índice de tablas

TABLA 1: LISTA DE REACTIVOS Y KITS UTILIZADOS.	14
TABLA 2: LISTA DE EQUIPOS UTILIZADOS.....	14
TABLA 3: LISTA DE HERRAMIENTAS COMPUTACIONALES UTILIZADAS.....	15
TABLA 4: SECUENCIA DE PARTIDORES UTILIZADOS PARA LA AMPLIFICACIÓN DEL FRAGMENTO 18S.	21
TABLA 5: COMPOSICIÓN DE MEDIO DE CULTIVO PARA <i>FUSARIUM OXYSPORUM</i>	25
TABLA 6: SOLUCIÓN DE ELEMENTOS TRAZA PARA MEDIOS DE CULTIVO.	25
TABLA 7: DESCRIPCIÓN DE LOS MEDIOS DE CULTIVO ENSAYADOS PARA <i>FUSARIUM OXYSPORUM</i>	26
TABLA 8: COMPOSICIÓN DE LOS MEDIOS DE CULTIVO ENSAYADOS PARA <i>GLOEOPHYLLUM TRABEUM</i>	29
TABLA 9: RESULTADOS DE LA BÚSQUEDA DE LAS SECUENCIAS DEL 18S DEL <i>GLOEOPHYLLUM TRABEUM</i> EN LA BASE DE DATOS BLAST	32
TABLA 10: RESULTADOS DE LA BÚSQUEDA DE LAS SECUENCIAS DEL 18S DEL <i>FUSARIUM OXYSPORUM</i> EN LA BASE DE DATOS BLAST	33
TABLA 11: COMPARACIÓN DEL RENDIMIENTO DE LA EXTRACCIÓN DE RNA	34
TABLA 12: LISTA DE SECUENCIAS NUCLEOTÍDICAS UTILIZADAS EN LOS ALINEAMIENTO PARA EL DISEÑO DE PARTIDORES DEGENERADOS	35
TABLA 13: LISTA DE PARTIDORES DEGENERADOS Y ESPECÍFICOS DISEÑADOS.	37
TABLA 14: RESULTADOS DE LA IDENTIFICACIÓN Y SECUENCIACIÓN PARCIAL DE LAS LPMO USANDO PARTIDORES DEGENERADOS.....	39
TABLA 15: COMPARACIÓN DE LAS SECUENCIAS OBTENIDAS CON LAS UTILIZADAS COMO HEBRA MOLDE PARA EL DISEÑO DE PARTIDORES....	40
TABLA 16: PROPIEDADES FÍSICOQUÍMICAS TEÓRICAS DE LAS LPMOS Y LA POSICIÓN DE LOS MÓDULOS CONSERVADOS EN ESTAS	41
TABLA 17: SECUENCIAS DE REFERENCIA UTILIZADAS PARA LA DETERMINACIÓN DE LA SUBFAMILIA.....	46
TABLA 18: RESUMEN DE LOS RESULTADOS OBTENIDOS DURANTE LA MODELACIÓN DE LAS LPMOS CLONADAS	51
TABLA 19: COMPARACIÓN ENTRE LAS SECUENCIAS DE LAS LPMOS CLONADAS, Y LAS SECUENCIAS UTILIZADAS DE MOLDE PARA SU MODELACIÓN TRIDIMENSIONAL.....	51
TABLA 20: SECUENCIAS DE LPMOS DISPONIBLES EN LA BIBLIOGRAFÍA Y EN LAS BASES DE DATOS.....	59
TABLA 21: PHYLUM DE CADA MICROORGANISMO UTILIZADO PARA LA ELABORACIÓN DEL ÁRBOL FILOGENÉTICO.....	61
TABLA 22: NOMENCLATURA IUPAC PARA PARTIDORES DEGENERADOS.....	71
TABLA 23: SECUENCIAS OBTENIDAS DE LOS rDNA 18S DE LOS HONGOS <i>GLOEOPHYLLUM TRABEUM</i> Y <i>FUSARIUM OXYSPORUM</i>	77
TABLA 24: RESULTADOS COMPLETOS DE PROCESO DE IDENTIFICACIÓN DE LPMO.	78
TABLA 25: SECUENCIAS NUCLEOTÍDICAS COMPLETAS CORRESPONDIENTES A LOS GENES SECUENCIADOS.....	79
TABLA 26: SECUENCIAS PROTEICAS COMPLETAS CORRESPONDIENTES A LOS GENES SECUENCIADOS.	80

Índice de ilustraciones

FIGURA 1: VARIACIÓN EN EL CONSUMO MUNDIAL DE ENERGÍAS RENOVABLES EN LAS ÚLTIMAS CINCO DÉCADAS.....	2
FIGURA 2: DIAGRAMA DE BLOQUES QUE RESUME LAS PRINCIPALES ETAPAS EN EL PROCESO DE PRODUCCIÓN DE BIOETANOL	4
FIGURA 3: ESQUEMA DE LA ESTRUCTURA DEL MATERIAL LIGNOCELULÓSICO.....	5
FIGURA 4: MECANISMO CLÁSICO DE HIDRÓLISIS DE CELULOSA	6
FIGURA 5: SUPERPOSICIÓN DEL SITIO DE UNIÓN DEL IÓN METÁLICO EN DOS ENZIMAS DE TIPO LPMO	8
FIGURA 6: MECANISMOS DE ACCIÓN PROPUESTO PARA LAS LPMOS	9
FIGURA 7: MECANISMO DE DESPOLIMERIZACIÓN DE CELULOSA INVOLUCRANDO TANTO CELULASAS TRADICIONALES COMO LAS LPMO Y LAS CDH	10
FIGURA 8: SECUENCIA Y DETALLE DEL SITIO DE MÚLTIPLE CLONAMIENTO DEL VECTOR pGEM-T EASY DE PROMEGA	15
FIGURA 9: MAPA DEL VECTOR pGEM-T EASY DE PROMEGA.....	16
FIGURA 10: VARIACIÓN EN LA ACTIVIDAD CELULOLÍTICA DE LOS CULTIVOS DE <i>F. OXYSPORUM</i> PARA DIFERENTES MEDIOS DE CULTIVO SUPLEMENTADOS CON PAJA DE TRIGO MOLIDA.....	26
FIGURA 11: VARIACIÓN EN LA ACTIVIDAD CELULOLÍTICA DE LOS CULTIVOS DE <i>F. OXYSPORUM</i> PARA DIFERENTES MEDIOS DE CULTIVO SUPLEMENTADOS CON AVICEL.	27
FIGURA 12: VARIACIÓN EN LA ACTIVIDAD CELULOLÍTICA DE LOS CULTIVOS DE <i>G. TRABEUM</i> PARA DIFERENTES MEDIOS DE CULTIVO SUPLEMENTADOS CON PAJA DE TRIGO MOLIDA.....	29
FIGURA 13: VARIACIÓN EN LA ACTIVIDAD CELULOLÍTICA DE LOS CULTIVOS DE <i>G. TRABEUM</i> PARA DIFERENTES MEDIOS DE CULTIVO SUPLEMENTADOS CON AVICEL.	30
FIGURA 14: VISUALIZACIÓN DE LOS GENES DEL rDNA 18S DE HONGOS EN UNA ELECTROFORESIS EN GEL DE AGAROSA.....	32
FIGURA 15: VISUALIZACIÓN EN ELECTROFORESIS EN GEL DE AGAROSA DEL RESULTADO DE LA EXTRACCIÓN DE RNA DESDE HONGOS.....	34
FIGURA 16: DETALLE DE LOS ALINEAMIENTOS UTILIZADOS PARA EL DISEÑO DE PARTIDORES DEGENERADOS.	36
FIGURA 17: ESQUEMA DE LA POSICIÓN DE LOS PARTIDORES EN RELACIÓN A LOS GENES UTILIZADOS EN LOS ALINEAMIENTOS	36
FIGURA 18: PRODUCTOS DE PCR OBTENIDOS CON EL USO DE PARTIDORES DEGENERADOS Y VISUALIZADOS EN UN GEL DE AGAROSA.....	38
FIGURA 19: PRODUCTOS DE PCR OBTENIDOS CON EL USO DE PARTIDORES ESPECÍFICOS Y VISUALIZADOS EN UN GEL DE AGAROSA	40
FIGURA 20: MAPA DE RESTRICCIÓN DE LAS SECUENCIAS NUCLEOTÍDICAS DE LAS LPMOS CLONADAS.....	43
FIGURA 21: RESIDUOS CONSERVADOS Y MÓDULOS PRESENTES EN LAS SECUENCIAS FoPMO9A, FoPMO9B Y GTPMO9A	44
FIGURA 22: ÁRBOL FILOGENÉTICO OBTENIDO DEL ANÁLISIS DE LAS LPMOS SELECCIONADAS MEDIANTE EL MÉTODO DE XIN LI ET AL.	46
FIGURA 23: ALINEAMIENTO OBTENIDO MEDIANTE LA HERRAMIENTA T-COFFEE DE LAS SECUENCIAS DE LPMOS.....	49
FIGURA 24: ESQUEMA DEL MODELO TRIDIMENSIONAL DE LA PROTEÍNA FoPMO9A.....	52
FIGURA 25: ILUSTRACIÓN ESQUEMÁTICA DE LA ESTRUCTURA DEL SITIO DE UNIÓN A CARBOHIDRATOS 1 (CBM1)	53
FIGURA 26: ILUSTRACIÓN ESQUEMÁTICA DE LA SUPERFICIE DE LAS PROTEÍNAS EN LOS MODELOS OBTENIDOS PARA LAS LPMOS.	54
FIGURA 27: ILUSTRACIONES ESQUEMÁTICAS DEL MODELO TRIDIMENSIONAL OBTENIDO PARA LA PROTEÍNA FoPMO9A	55
FIGURA 28: ILUSTRACIONES ESQUEMÁTICAS DEL MODELO TRIDIMENSIONAL OBTENIDO PARA LA PROTEÍNA FoPMO9B.....	56
FIGURA 29: ILUSTRACIONES ESQUEMÁTICAS DEL MODELO TRIDIMENSIONAL OBTENIDO PARA LA PROTEÍNA GTPMO9A	57
FIGURA 30: ÁRBOL FILOGENÉTICO DE LPMOS CARACTERIZADAS Y DE ORGANISMOS DE REFERENCIA	60
FIGURA 31: ALINEAMIENTO DE SECUENCIAS NUCLEOTÍDICAS CONOCIDAS PARA EL DISEÑO DE LOS PARTIDORES GH61_G1	72
FIGURA 32: ALINEAMIENTO DE SECUENCIAS NUCLEOTÍDICAS CONOCIDAS PARA EL DISEÑO DE LOS PARTIDORES GH61_G2	73
FIGURA 33: ALINEAMIENTO DE SECUENCIAS NUCLEOTÍDICAS CONOCIDAS PARA EL DISEÑO DE LOS PARTIDORES GH61_G4	74
FIGURA 34: MAPAS DE LOS VECTORES DE EXPRESIÓN pPICZ, pPICZA Y pPIC3.5 DE INVITROGEN	75
FIGURA 35: MCS DEL VECTOR DE EXPRESIÓN pPIC3.5	75
FIGURA 36: MCS DEL VECTOR DE EXPRESIÓN pPICZ	76
FIGURA 37: MCS DEL VECTOR DE EXPRESIÓN pPICZA A.	76
FIGURA 38: RESULTADOS ENTREGADOS POR ROBETTA PARA LA MODELACIÓN DE LA PROTEÍNA FoPMO9A	82
FIGURA 39: RESULTADOS ENTREGADOS POR ROBETTA PARA LA MODELACIÓN DE LA PROTEÍNA FoPMO9B	83
FIGURA 40: RESULTADOS ENTREGADOS POR ROBETTA PARA LA MODELACIÓN DE LA PROTEÍNA GTPMO9A	84

Lista de abreviaciones

Abreviación	Significado
AA	Aminoácido
AA9	Familia de Actividades Auxiliares 9
BERIS	Biological & Environmental Research Information System
CAZy	Carbohydrate-Active Enzymes
CBM	Carbohydrate-Binding Module
CDH	Celobiosa Deshidrogenasa
cDNA	DNA complementario
CMC	Carboximetil celulosa
DEPC	Diethylpicrocarbonato
DNS	Acido 3,5-Dinitrosalicílico
dNTP	Desoxirribonucleótidos trifosfato
GH	Glicosil Hidrolasas
GH61	Glicosil Hidrolasa de la familia 61
gDNA	DNA genómico
HEC	Hidroxietilcelulosa
JGI	Joint Genome Institute
LB	Luria Broth
LPMO	Lytic Polysaccharide Monooxygenase
MCS	Sitio de Clonamiento Múltiple
MTep	Millones de Toneladas equivalentes de petróleo.
OD	Densidad Óptica
PCR	Reacción en Cadena de la Polimerasa
PDA	Agar Papa Dextrosa
PDB	Protein Data Bank
pI	Punto Isoeléctrico
PM	Peso Molecular
RCSB	Research Collaboratory for Structural Bioinformatics
RPM	Revoluciones Por Minuto
RNA	Ácido Ribonucleico
rRNA	Ácido Ribonucleico Ribosomal
Ta	Temperatura de alineamiento
WRF	White Rot Fungi
X-Gal	(5-bromo-4-cloro-3-indolil- β -D-Galactopiranosido

I Introducción

I.1 Marco teórico

I.1.1 Antecedentes generales

Desde mediados del siglo XVIII, el mundo vio florecer de forma rápida y acelerada lo que llegaría a conocerse como la revolución industrial. Fue en esta época que la mecanización y la tecnificación de la producción trajeron consigo nuevos productos y la aparición de tecnologías más accesibles para la población. Es a partir de este momento también que el desarrollo de nuevas tecnologías alcanza una velocidad nunca antes vista. Motivada por la industria, la sociedad de consumo y en algunos casos situaciones bélicas, la investigación de nuevas tecnologías se volvería la base para el aumento de la producción en el mundo.

Las ciencias biológicas no se quedan fuera de esta aceleración. Procesos anteriormente artesanales de producción, serán estudiados y usados como piedra angular para la elaboración de nuevos procesos biotecnológicos más complejos. Es así como lentamente nace la biotecnología, un acercamiento nuevo de la biología que pretende dar solución a problemáticas actuales a través del uso del mundo de lo vivo. La comprensión de los procesos biológicos a un nivel molecular, dará la opción de construir sistemas adaptados a las necesidades de la humanidad, creando así procesos que podrán funcionar a la par y en algunos casos superar los más tradicionales métodos químicos.

I.1.1.1 Problema energético mundial

Combustibles fósiles

En las últimas décadas, el mundo ha sido testigo de un aumento constante y acelerado en el consumo energético por parte de las poblaciones humanas. Este consumo, que ha pasado de $8 \cdot 10^9$ a $12.3 \cdot 10^9$ toneladas equivalentes de petróleo en los últimos 20 años, ha sido suplido a través del uso de fuentes energéticas tradicionales (1). Entre estas fuentes se cuentan energías tales como la hidroeléctrica y la nuclear. Sin embargo, los principales exponentes de estas fuentes de energía tradicionales son los combustibles fósiles como el carbón, petróleo y gas natural, que representan más del 85% del consumo total (1).

Estas fuentes energéticas, aunque con procesos productivos establecidos y de relativamente fácil disponibilidad, presentan una gama de impactos negativos en su uso. Debido a la naturaleza de los combustibles fósiles, su aparición en el planeta es el producto de procesos naturales que toman siglos en producirse. El consumo actual sobrepasa por lo tanto con creces la producción natural de estos, volviéndolo una fuente no renovable de energía. Más aún, el consumo indiscriminado y no planificado de estos los han vuelto un bien cada vez más difícil de explotar y cuya producción necesita de esfuerzos cada vez mayores. Como es lógico, al volverse lentamente un bien escaso y de difícil acceso, los combustibles fósiles presentan un valor en constante aumento, generando un impacto negativo en las economías dependientes de estos. Este efecto es todavía más claro en países en donde la producción energética depende en gran parte de la importación de combustibles fósiles, como es el caso de Chile. Esta dependencia energética tiene como consecuencia que el país tenga que asumir las alzas en los precios de los combustibles que los mercados internacionales dictan, además de sufrir los efectos de variaciones e inestabilidades en las economías de los países exportadores. Tal fue el caso del corte en el suministro de gas natural llevado a cabo por Argentina

del año 2004 en adelante, en el que Chile tuvo que cambiar su estructura productiva para pasar al uso de petróleo (2).

Impacto ambiental

Otro punto negativo en el uso de los combustibles fósiles, son los impactos ambientales que éstos producen. Al tratarse de compuestos provenientes del ciclo del carbono, y particularmente de carbono retenido en reservas subterráneas, el consumo de estos implica la liberación al ambiente de elementos de forma poco reversible. Entre estos elementos se encuentra el dióxido de carbono. Este gas ha sido señalado como principal responsable del efecto invernadero, que en resumidas cuentas, consiste en el proceso mediante el cual la radiación solar reflejada es retenida en la Tierra, produciendo un aumento en la temperatura general del planeta (3–5). Esta consecuencia de los gases de efecto invernadero es también conocida como calentamiento global, efecto observado en los últimos 50 años en la tierra. Este fenómeno consta a su vez de varias aristas, se le atribuye el derretimiento de los casquetes polares, además de un cambio climático generalizado. Entre sus efectos se encontrarían asimismo el recrudescimiento de las condiciones climáticas, con temperaturas máximas y mínimas más extremas, además de un aumento en las sequías y un aumento en las precipitaciones (4, 5).

I.1.2 Biocombustibles

Frente a estas múltiples externalidades negativas que genera el uso de combustibles fósiles, el foco de la producción y del desarrollo de nuevas tecnologías ha ido cambiando lentamente hacia nuevas alternativas (Figura 1), particularmente hacia las fuentes de energía renovable (1, 6).

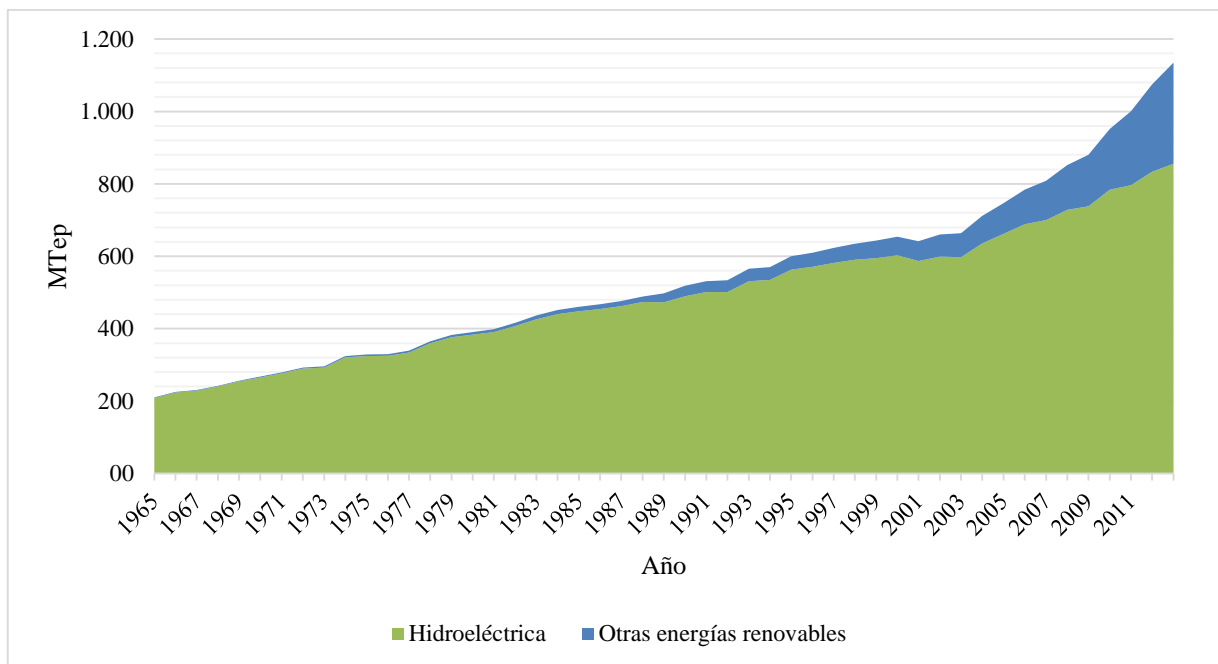


Figura 1: Variación en el consumo mundial de energías renovables en las últimas cinco décadas. El consumo se encuentra expresado en millones de toneladas equivalentes de petróleo (MTep). Se presentan tanto la energía hidroeléctrica como la suma de las energías eólica, geotérmica, solar y proveniente de biomasa y desechos (otras energías renovables). Datos obtenidos de informe anual de BP (1).

Es así como en la actualidad, existe una amplia variedad de fuentes energéticas renovables, entre las que se cuentan la energía solar, eólica y mareomotriz entre otras. Otra opción que se perfila como una de las más viables es la de los biocombustibles. Sin embargo, la aparición de estos en el mundo ha sido un proceso paulatino, y la evolución de estos ha dado lugar a su clasificación en diferentes categorías, dependiendo del tipo de materia prima utilizada para su producción.

I.1.2.1 Biocombustibles de primera y tercera generación

La primera generación de biocombustibles consiste en los procesos de producción de etanol (o bioetanol) a partir de materias primas directamente fermentables. Estos procesos usan como sustrato la fermentación de azúcares simples y almidón, que se obtienen de cosechas de, por ejemplo, granos o caña de azúcar. Al tratarse de azúcares simples, estos procesos no necesitan de pretratamiento intensivos de las materias primas antes de pasar a la fermentación, siendo éste un proceso conocido y bien establecido, similar a los usados para la producción de bebidas alcohólicas. En la actualidad la mayor parte del bioetanol producido en el mundo se realiza a través de este proceso. Los principales exponentes en la producción y uso de estos combustibles son los Estados Unidos y Brasil, en donde la materia prima utilizada es el maíz y la caña de azúcar respectivamente (7).

Sin embargo, la producción de este tipo de combustibles ha sido recibida con críticas por parte de la comunidad y ha presentado inconvenientes. Por una parte se ha criticado que la producción de etanol, haciendo uso de estas materias primas no presenta una reducción considerable en la huella de carbono en relación a los combustibles tradicionales. Se ha postulado incluso que esta podría ser mayor en algunos casos a través de la deforestación indirecta producida en la búsqueda de nuevas tierras cultivables (proceso conocido como uso indirecto de tierras) (8–10). Por otra parte, el uso de granos y cultivos de consumo humano para la producción de combustibles, genera un problema ético de asignación de recursos. Esto se debe a que el desvío de materia prima que podría ser utilizada como alimento para su uso como combustible resulta cuestionable. Un ejemplo práctico de lo que puede ocurrir con este tipo de producciones a nivel masivo fue el aumento en el precio de los granos observados el año 2007 - 2008, atribuido en parte a la producción de bioetanol de primera generación, y que provocó hambrunas en diversos países (11–13).

Es en respuesta a las críticas recibidas por estos biocombustibles, que una segunda y tercera generación de procesos productivos cobra importancia: procesos que usan materias primas que no tienen un uso humano alternativo.

Los biocombustibles de tercera generación corresponden a los procesos cuya materia prima consiste en cultivos especializados para la producción de combustibles, particularmente cultivos de algas. Estos combustibles buscan hacer uso de tanto los azúcares como los aceites presentes en las algas para su transformación a bioetanol y biodiesel (14). Este enfoque ofrece además la ventaja que los cultivos son desarrollados en el mar, y no competirían por terreno con los cultivos actuales.

I.1.2.2 Biocombustibles de segunda generación

Los biocombustibles de segunda generación se diferencian de los anteriormente descritos en el tipo de materia prima que estos utilizan, siendo en este caso desechos vegetales. En un esfuerzo por hacer uso estratégico de los recursos sin afectar a las poblaciones humanas, se desarrollaron procesos para la producción de etanol a partir tanto de desechos de la industria agrícola y forestal, como de desechos vegetales urbanos en general. Estos residuos consisten principalmente en las

partes no comestibles de los cultivos agrícolas, así como restos de la industria maderera y restos de podas. La composición de estos desechos es esencialmente material lignocelulósico, que a su vez se encuentra compuesto principalmente de celulosa, hemicelulosa y lignina. La celulosa consiste en un biopolímero lineal formado por la unión de glucosas a través de enlaces β -1,4-glucosídicos. Estas hebras de celulosa se alinean y unen entre sí mediante interacciones de tipo hidrofóbicas y de Van der Waals, formando regiones tanto cristalinas como amorfas. La hemicelulosa por su parte, es un polímero heterogéneo compuesto de una variedad de pentosas y hexosas, y corresponde al segundo principal componente de la lignocelulosa. Finalmente, la lignina consiste en un heteropolímero formado principalmente por alcoholes aromáticos tales como el alcohol coniferílico y sinapílico, y representa el compuesto más recalcitrante de la madera (13).

Al tratarse la celulosa del polímero más abundante en la tierra, el uso de ésta como fuente energética es una apuesta prometedora. Este también se trataría de un combustible renovable: al liberar al ambiente el carbono que las plantas fijan en su biomasa, este proceso se insertaría en el ciclo del carbono. Se mantendría así, en teoría, un equilibrio entre el carbono liberado y el fijado por las plantas. Por otra parte, se ha calculado que en la actualidad existe suficiente biomasa renovable en los Estados Unidos, para que su uso remplace alrededor del 50% de la demanda suplida por los biocombustibles de primera generación. Más aún, se cree que en este país, hasta mil millones de toneladas de biomasa renovable podrían llegar a cultivarse con el fin de producir biocombustibles. Este volumen de biomasa procesada permitiría suplir el 30% de la demanda de combustibles líquidos para transporte de ese país (15).

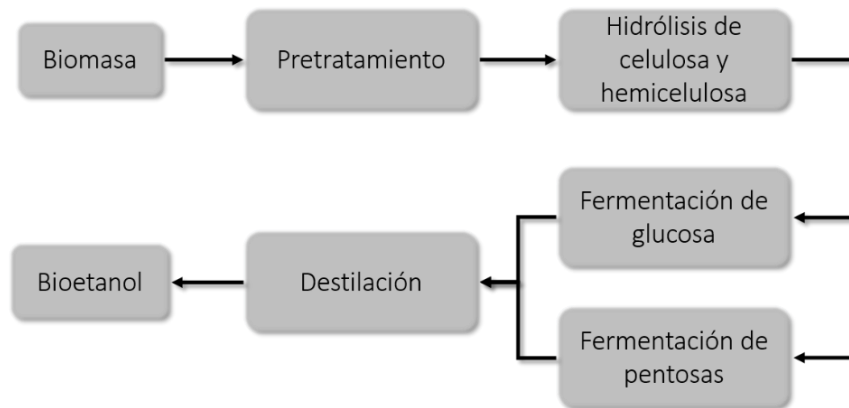


Figura 2: Diagrama de bloques que resume las principales etapas en el proceso de producción de bioetanol a partir de material lignocelulósico.

No obstante, el proceso para la producción de estos biocombustibles (ver Figura 2) no se encuentra libre de inconvenientes. A diferencia de los combustibles de primera generación, los azúcares fermentables en el material lignocelulósico no se encuentran fácilmente accesibles.

En la madera, la celulosa se encuentra incluida en una compleja matriz de hemicelulosa y lignina (Figura 3). La estructura precisa que forma la pared celular en las plantas no ha sido del todo dilucidada. Sin embargo a la fecha se cree que ésta estaría formada por fibras de celulosa que se encontrarían envueltas por una cubierta de hemicelulosa. Ésta última serviría a su vez de elemento de unión con una cubierta de lignina (16). La estructura de los materiales lignocelulósicos representa un obstáculo considerable para el acceso de las enzimas a la celulosa, por una parte por la presencia de la matriz de hemicelulosa y lignina que separa la celulosa del medio, y por otra parte por las regiones cristalinas y no solubles de las microfibras de celulosa. Por consiguiente dos de las principales dificultades en este proceso de producción consisten en la liberación de la celulosa desde esta matriz, y la posterior hidrólisis de la celulosa a azúcares más simples.

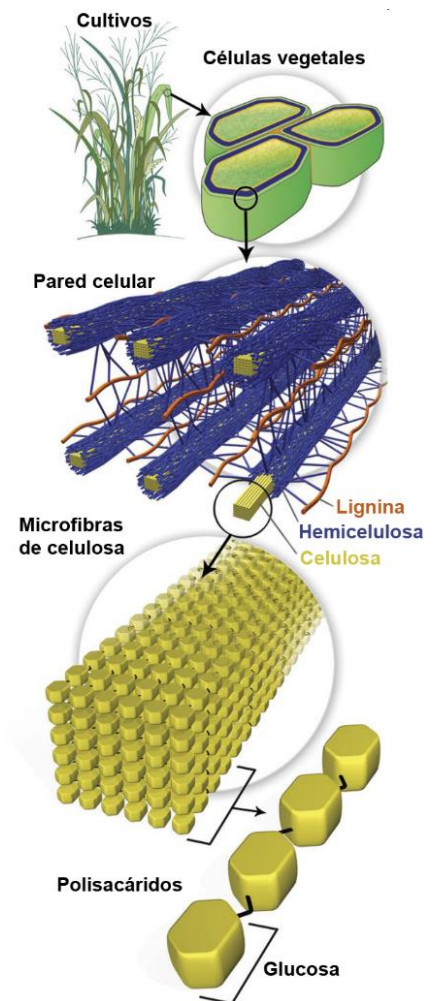
Figura 3: Esquema de la estructura del material lignocelulósico.
Imagen obtenida desde página del BERIS.

Pretratamiento e hidrólisis de la celulosa

El proceso de hidrólisis de la celulosa junto con el pretratamiento de la lignocelulosa, constituyen las etapas más económicamente costosas en el proceso de producción de bioetanol (17). Por sí solas, estas operaciones representan alrededor del 40% del costo total en el proceso de producción de bioetanol (18). Más específicamente, una estimación realizada el 2010, evaluó el costo de las enzimas involucradas en el proceso en alrededor de 0.5 dólares por galón de biocombustible (15). Esto se debe principalmente a las propiedades recalcitrantes del material lignocelulósico, y tiene como consecuencia que a la fecha, la producción de este combustible no sea económicamente atractiva. Muchos esfuerzos han sido por lo tanto concentrados en encontrar formas más eficientes de realizar este proceso. El objetivo de estas investigaciones es obtener un producto atractivo económicamente que pueda competir con los combustibles fósiles en el mercado energético. Se ha estimado además que, de explotar las potenciales mejoras del proceso productivo, el costo de este combustible podría llegar a los 25 dólares el barril, volviéndolo una alternativa interesante (18).

El fin del pretratamiento es lograr la separación de la celulosa de la hemicelulosa y más particularmente de la lignina. A la fecha este proceso puede ser realizado de forma química (tratamiento ácido o alcalino), fisicoquímica (molienda o explosión por vapor) o a través de tratamiento biológico, haciendo uso de organismos degradadores de lignina (13).

Por otra parte, la hidrólisis de la celulosa consiste en la despolimerización de ésta hacia azúcares simples. En las últimas décadas los esfuerzos se han enfocado en encontrar nuevas y mejores celulasas para la degradación de la celulosa y hemicelulosa. Para esta búsqueda, el foco fue rápidamente puesto en los hongos de la madera. Estos organismos, principales responsable de la degradación de madera, presentan una batería de celulasas y otras enzimas que les permiten liberar azúcares simples desde la lignocelulosa para su consumo (19).



Mecanismo clásico de hidrólisis

El estudio tanto de los hongos de pudrición blanca como de los hongos de pudrición parda, ha permitido dilucidar lo que hasta hace algunos años era considerado el mecanismo de acción clásico de las enzimas celulolíticas (Figura 4). Este sistema considera la acción conjunta de una variedad de enzimas en el proceso de liberación de azúcares simples desde material lignocelulósico. En este modelo, las endoglucanasas capaces de hidrolizar la celulosa en sus regiones no cristalinas producirían los extremos libres de celulosa que son el sustrato de las exoglucanasas o celobiohidrolasas. Estas a su vez, una vez disponible su sustrato, serían capaces de liberar dímeros de glucosa (celobiosa) hacia el medio. Estos disacáridos serían finalmente hidrolizados a glucosa a través de la acción de las β -glucosidasas, liberando así azúcares simples e impidiendo la inhibición por producto de las celobiohidrolasas (13, 17, 20, 21).

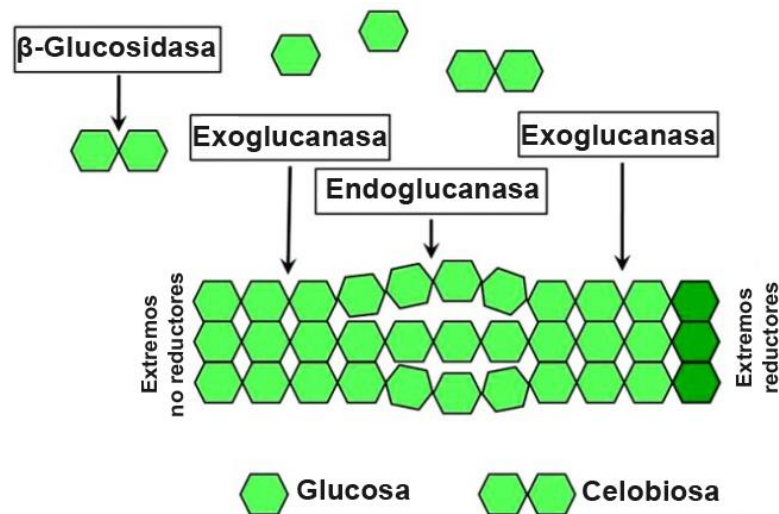


Figura 4: Mecanismo clásico de hidrólisis de celulosa. Imagen obtenida desde trabajo de Waeonukul et al. (22).

Otro mecanismo propuesto, aunque menos aceptado es el del ataque a través de reacciones de Fenton. La naturaleza inaccesible de la celulosa, y la dificultad que este medio presenta para la penetración de las enzimas, ha llevado a algunos investigadores a postular que la hidrólisis se realiza de forma indirecta. Este mecanismo se basaría en la producción por parte de los hongos de radicales libres tales como el hidroxilo ($\cdot OH$), que sería capaz de penetrar la matriz lignocelulósica y producir la oxidación azarosa de la celulosa. Según esta teoría, este radical libre sería producido mediante la oxidación de Fe^{2+} gracias a la acción de H_2O_2 en una reacción de Fenton (17, 19, 21).

A pesar de la actividad conjunta de estas proteínas, la alta capacidad degradativa de algunos hongos no puede ser explicada sólo por la actividad de éstas (17). Es por esto que se han propuesto nuevos mecanismos, tales como la existencia de proteínas disruptivas del sustrato, que facilitarían el acceso de las celulasas al mismo.

Proteínas auxiliares del proceso de hidrólisis

Entre las nuevas ideas que se barajan para explicar la eficiencia del sistema celulolítico de algunos hongos se encuentra la de las proteínas auxiliares. Estas proteínas, en algunos casos enzimas sin

actividad aparente, han sido señaladas como responsables de facilitar el proceso de hidrólisis mediante mecanismos no tradicionales.

Una familia de estas proteínas, conocida como expansin-like (o similares a expansinas) son proteínas que como bien dice su nombre comparten características con las expansinas de plantas. Similares a las expansinas responsables de la flexibilización de la pared celular, cuando las células vegetales aumentan su tamaño, las expansin-like serían secretadas por los hongos para aumentar la susceptibilidad de la pared vegetal al ataque de las celulasas. Entre éstas se han descrito familias tales como las *loosenins* y las *swollenins*, capaces de alterar la estructura del papel filtro y de fibras de algodón sin hidrolizar la celulosa (23–26).

Otra familia que presenta gran interés es la de las llamadas hasta hace un par de años glicosil hidrolasas de la familia 61 (GH61). Estas proteínas, hoy conocidas como monooxigenasas dependientes de cobre (Lytic Polysaccharide Monooxygenase, o LPMO por sus siglas en inglés), han sido uno de los focos de estudio en la última década.

I.1.3 Monooxigenasas Líticas de Polisacáridos (LPMO)

I.1.3.1 Historia

La primera GH61 descrita fue Cell del organismo *Agaricus bisporus* (27). Aunque no se determinó qué actividad poseía, su expresión inducida por la presencia de celulosa en el medio (27, 28), además de que en algunos casos se encontraba asociada a dominios de unión a carbohidratos (CBM), la relacionó a la degradación de ésta. Más tarde, en 1997, la familia GH61 sería creada y el interés por ésta iría en constante aumento (27). El descubrimiento de un gran número de genes de estas enzimas en los genomas de ciertas especies, alcanzando incluso los 36 en *Chaetomium globosum* (17, 29), además de la ausencia de actividad aparente le otorgaron a esta familia un carácter enigmático (30, 31). A pesar de su falta de actividad, se observó más tarde que esta familia de enzimas era capaz de aumentar el efecto de las celulasas. Utilizadas en conjunto con mezclas de celulasas tradicionales, tanto los equipos de Harris (32) como de Dimarogona (33) observaron un incremento de aproximadamente 20% en la hidrólisis de celulosa. Además, el hecho de que esta familia de enzimas sin actividad, mostraba un comportamiento sinérgico al ser utilizada en conjunto con celulasas, atrajo aún más atención sobre sus propiedades (32, 34).

I.1.3.2 Estructura y actividad

Un estudio más profundo de las GH61 revelaría más tarde que estas enzimas no comparten las características clásicas de las otras glicosil hidrolasas (GH) (27). Con la obtención de la primera estructura cristalográfica de una de éstas enzimas se pudo adicionalmente determinar que estas poseían una estructura diferente a la de las otras GH y que carecían del sitio de unión a sustrato que las definía (35). Con las siguientes estructuras cristalográficas se determinó que estas proteínas tenían ciertos caracteres conservados en su estructura. En un primer lugar se observó que se conserva una superposición de hojas β . Este llamado “sándwich” de hojas β es muy similar al presente en las inmunoglobulinas y compone una parte central de las LPMOs. Además, se detectó una superficie plana en una de las caras de las proteínas. La presencia en esta cara, de residuos hidrofóbicos característicos de los sitios de unión a carbohidratos, dio indicios de que se podría tratar del sitio de unión a celulosa de la enzima.

Más adelante, una serie de estudios revelaría que la esquivada actividad de estas enzimas correspondía en realidad a una de tipo monooxigenasa dependiente de un ion metálico y de un agente reductor. Aunque en un principio los análisis de las estructuras de las GH61 llevaron a confusión, se determinó finalmente que el ion necesario para la activación de estas enzimas correspondía a Cobre(II) (21, 32, 35–37). Este ion se encontró asociado en la proteína a tres aminoácidos altamente conservados dentro de la familia: dos histidinas, una de las cuales se encuentra en el extremo N-terminal, y una tirosina. Este sitio de unión, denominado como el “soporte” de histidina, resulta ser una característica compartida transversalmente a lo largo de esta familia (ver Figura 5). Se reveló además, a través de mutagénesis sitio dirigida, que su presencia resulta crítica para la actividad de la enzima (32, 38).

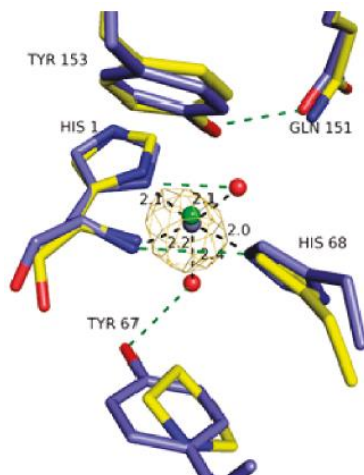


Figura 5: Superposición del sitio de unión del ion metálico en dos enzimas de tipo LPMO. En amarillo se muestra la estructura de TrGH61B de *Trichoderma reesei* (35) y en azul la estructura de la GH61E de *T. terrestris* (32). Imagen obtenida desde trabajo de Harris del 2010 (32).

Estos nuevos avances permitieron finalmente que esta proteína fuera reclasificada fuera de la familia de las GH. En 2013, las GH61 pasaron a formar parte de una nueva categoría de enzimas en la base de datos de enzimas activas sobre carbohidratos (CAZy), la categoría llamada Familia de Actividades Auxiliares 9 (AA9) (39, 40). El nombre de estas proteínas ha ido paulatinamente cambiando además hacia el de monooxigenasas dependientes de cobre (LPMO), nombre que se seguirá usando en este trabajo.

Se han propuesto nuevos mecanismos de acción para esta familia de enzimas, y éstas han sido categorizadas en función del tipo de producto final que éstas producen (17, 36, 41). Al momento de ser escrito este trabajo se proponen dos principales mecanismos de acción de las LPMOs (Figura 6). Ambos mecanismos suponen la oxidación de la cadena de celulosa y su posterior quiebre dejando extremos libres. El primer mecanismo supone que la oxidación ocurre en el carbono 1 de la glucopiranososa, esta oxidación llevaría a una desestabilización del enlace glicosídico, y su posterior quiebre de forma espontánea o catalizada por la LPMO. Este proceso llevaría entonces a la producción de un ácido aldónico y no liberaría extremos reductores desde la celulosa. El segundo mecanismo es similar al primero con la particularidad de que el carbono oxidado es el 4. El posterior quiebre del enlace glicosídico produciría un extremo reductor y la formación de una 4-cetoaldosa en las hebras resultantes.

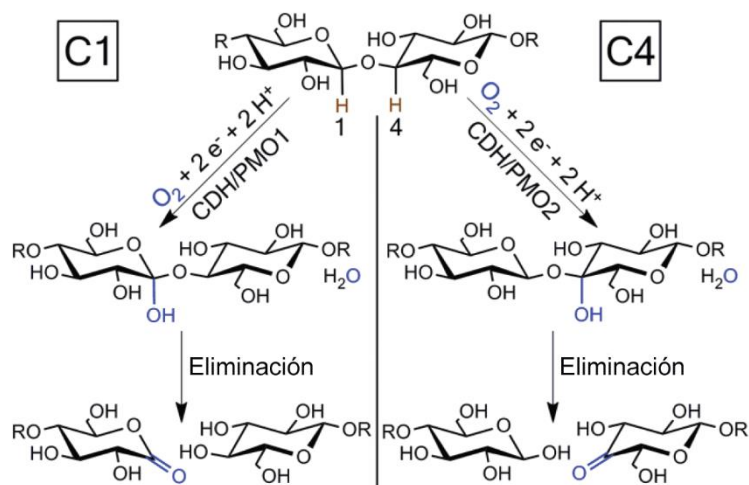


Figura 6: Mecanismos de acción propuesto para las LPMOs. A la izquierda se muestra el mecanismo mediado por la oxidación del carbono 1 de la glucopiranososa, y a la derecha se muestra el mediado por la oxidación del carbono 4. Imagen obtenida del trabajo de William Beeson et al. (41).

Como se señaló anteriormente, se ha determinado a la fecha que la actividad de las LPMOs requiere de la presencia en el medio de ion cobre (II) y de un agente reductor. En experimentos *in vitro*, este agente reductor ha sido añadido en forma de glutatión reducido o de ácido ascórbico (37). Sin embargo se ha postulado que *in vivo*, este donador de electrones podría tratarse de otra enzima actuando sinérgicamente con las LPMOs. La enzima que se ha propuesto para cumplir este rol es la celobiosa deshidrogenasa (CDH). Esta enzima, único flavocitocromo producido por hongos, tendría la capacidad de entregar los electrones necesarios para la actividad de las LPMOs a través de la oxidación de celobiosas. Más aún, se ha demostrado que adicionadas a mezclas de celulasas y LPMOs, estas pueden aumentar en 15 veces la cantidad de monosacáridos liberados desde celulosa pretratada (34).

Finalmente, con los avances realizados en la comprensión de la degradación de la lignocelulosa, sumado a las nuevas proteínas descubiertas que juegan un rol directo o indirecto en esta, el proceso de despolimerización de la celulosa se ha complejizado considerablemente (Figura 7). Esta renovada visión abre nuevas posibilidades para aumentar la eficiencia del proceso, y crea nuevas líneas de investigación que deben ser explotadas.

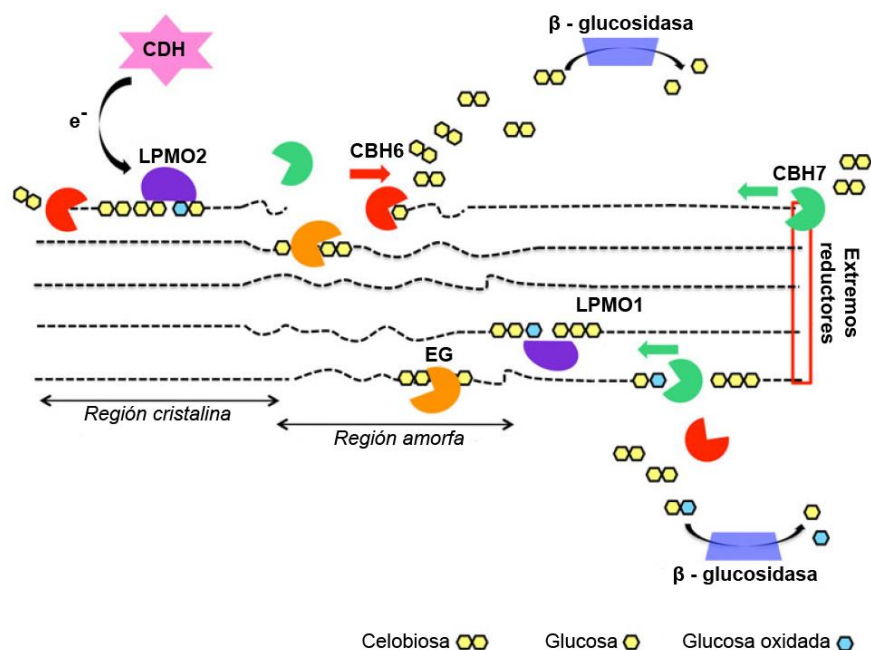


Figura 7: Esquema ilustrando el mecanismo de despolimerización de celulosa involucrando tanto celulasas tradicionales como las LPMO y las CDH. Tanto endoglucanasas (EG) como LPMOs cortan las cadenas de celulosa, liberando extremos libres que son atacados por las celobiohidrolasas (CBH). La celobiosa liberada es hidrolizada por las β -glucosidasas. Imagen obtenida del trabajo de Dimarogona et al. (17).

I.1.4 Hongos productores de celulasas

En este trabajo, la búsqueda de nuevas LPMO se centra en dos hongos: *Fusarium oxysporum* y *Gloeophyllum trabeum*. Estos hongos son de gran importancia para el desarrollo de nuevas tecnologías que permitan mejorar la producción de bioetanol a partir de material lignocelulósico. Ambos han sido estudiados intensamente por su capacidad de degradar la madera, y se ha determinado que los dos poseen la batería de enzimas necesarias para la degradación eficiente de ésta (42–46).

Además de ser un potencial productor de enzimas celulasas, el hongo ascomiceto *F. oxysporum* tiene la particularidad de poder llevar a cabo la fermentación de los azúcares fermentables a etanol, permitiendo así realizar una hidrólisis y una fermentación conjunta (44, 47). Esta opción es atractiva pues eliminaría la necesidad de producir y purificar las enzimas necesarias para la hidrólisis, pudiéndose realizar el proceso inoculando el sustrato directamente con el hongo.

Por su parte el hongo *G. trabeum* es uno de los más estudiados de entre los hongos de pudrición blanca, o WRF por sus siglas en inglés (45). Este tipo de hongos tiene la particularidad de poder degradar tanto la hemicelulosa como la celulosa de la madera, dejando atrás un residuo de lignina modificada. En el caso particular del *G. trabeum*, se ha determinado a la fecha que su capacidad hidrolítica se ve apoyada por reacciones de tipo Fenton, en las que radicales libres producidos por reacciones químicas oxidan la celulosa (17, 19, 28). Además, su uso en procesos de producción de bioetanol ya ha sido analizado en estudios como el de M. L. Rasmussen (48).

Debido a sus propiedades, ambas especies resultan interesantes de ser analizadas en la búsqueda de nuevas LPMOs. La reciente secuenciación de los genomas completos de estos hongos por parte del Joint Genome Institute abre además nuevas posibilidades de investigación. Una búsqueda en

estos genomas de proteínas de tipo LPMO revela la existencia de quince enzimas putativas para *Fusarium oxysporum* y cuatro para *Gloeophyllum trabeum* (49, 50). Por otra parte, al momento de la redacción de este trabajo, la base de datos CAZy cuenta con una secuencia de CDH correspondiente a *F. oxysporum* y tres correspondiente a *G. trabeum* (40). Estos hallazgos y el estudio en mayor profundidad de estas enzimas podrían arrojar luz sobre el mecanismo de hidrólisis de estos hongos, permitiendo así tal vez aumentar la eficiencia del proceso de hidrólisis de celulosa.

1.2 Descripción y justificación del proyecto

En base a lo expuesto anteriormente, se puede determinar que existe una problemática energética a nivel mundial, cuya solución podría radicar en la introducción de nuevos tipos de combustibles. Uno de los enfoques posibles para desarrollar estas nuevas fuentes energéticas es la producción de etanol en base a residuos vegetales. Estos biocombustibles, conocidos como de segunda generación presentan sin embargo en la actualidad, bajas eficiencias y altos costos económicos en su producción. La búsqueda de nuevas formas de producción y de nuevas soluciones biotecnológicas para este proceso, sería entonces un paso fundamental para el surgimiento del bioetanol como una respuesta a esta problemática.

Una de las soluciones que se barajan, es la investigación de los mecanismos que le otorgan a los hongos la capacidad de degradar la madera para su consumo. Esta búsqueda ha revelado en los últimos años nuevos actores que podrían marcar una diferencia considerable en la eficiencia del proceso. Uno de estos actores son las enzimas monooxigenasas líticas de polisacáridos o LPMOs, cuyo mecanismo de acción ha comenzado a ser estudiado recientemente. La capacidad de estas enzimas, de aumentar la eficiencia del proceso de despolimerización de la celulosa al ser agregadas en la solución de reacción, le ha valido ser centro de múltiples estudios en los últimos cinco años.

El propósito del siguiente trabajo fue la identificación de nuevas LPMOs en el transcriptoma de dos hongos: *Fusarium oxysporum* y *Gloeophyllum trabeum*. El trabajo consistió en un primer lugar, en la búsqueda de los medios de cultivo que aumentarían la actividad celulolítica del secretoma de los hongos. Una vez obtenidas, las LPMOs se identificaron mediante técnicas de biología molecular desde el transcriptoma de cada hongo. Una vez identificadas, se clonaron y secuenciaron los genes codificantes de estas proteínas. A continuación estas secuencias se analizaron *in silico* en búsqueda de la información que estas entregan respecto de las proteínas que codifican. Este trabajo se encuentra inserto dentro del proyecto Fondecyt Regular N° 1121088, cuyo fin último es la identificación de proteínas auxiliares para generar mejores mezclas celulolíticas.

I.3 Objetivos

I.3.1 Objetivo general

Identificación, clonación y análisis de nuevas proteínas LPMO desde *G. trabeum* y *F. oxysporum* con miras a mejorar el proceso de producción de bioetanol de segunda generación.

I.3.2 Objetivos específicos

- Identificar medios de cultivo inductores de la producción de celulasas para los hongos de interés.
- Identificar secuencias de consenso de LPMOs conocidas y en base a éstas, diseñar partidores degenerados para PCR.
- Clonar y secuenciar genes que codifican para proteínas LPMO de interés.
- Realizar un análisis comparativo de las secuencias génicas de las LPMOs obtenidas.

II Materiales y métodos

II.1 Materiales

II.1.1 Lista de reactivos y kits utilizados

Tabla 1: Lista de reactivos y kits utilizados en este trabajo.

Reactivo	Nombre	Fabricante
Avicel	Avicel PH-101	Fluka
CMC	CM-cellulose 4N	Megazyme
Gel Extraction Kit	QIAEX II Gel Extraction Kit	Qiagen
Genomic DNA Purification Kit	Wizard® Genomic DNA Purification Kit (#A1120)	Promega
Ladder 100 bp	GeneRuler 100 bp DNA Ladder	Thermo Scientific
Ladder 1 kb	1 Kb Plus DNA Ladder	Invitrogen
Ladder 1 kb Plus	1 Kb Plus DNA Ladder (#10787-018)	Invitrogen
Luria Broth	Luria Broth Base (#241420)	Difco
Miniprep Kit	Axyprep Plasmid Miniprep Kit (#AP-MN-P-250)	Axygen
PCR Clean Up	Gel/PCR DNA Fragments Extraction Kit	Geneaid
Phusion DNA Polymerase	England Biolabs (#M0530S)	Thermo Scientific
Taq DNA Polymerase	GoTaq® Flexi DNA Polymerase (#M829B)	Promega
Transcriptasa reversa	SuperScript™ III Reverse Transcriptase (#18080-044)	Invitrogen
RNase OUT	RNaseOUT™ Recombinant Ribonuclease Inhibitor (#10777-019)	Invitrogen
Vector System	pGEMT-Easy Vector System I (#A1360)	Promega

II.1.2 Equipos

Tabla 2: Lista de equipos utilizados para la realización de este trabajo.

Equipo	Fabricante y modelo
Cámara fotográfica	Kodak easyshare M340
Electroporador	Eppendorf 2510
Incubadoras	Innova 4300 New Brunswick Scientific Max Q 400 barnstead labline
Termocicladores	Eppendorf mastercycler gradient GenMate Series JG

II.1.3 Herramientas computacionales

Tabla 3: Lista de herramientas computacionales utilizadas para la realización de este trabajo.

Herramienta	Fuente	Herramienta	Fuente
Anolea	(51)	Mega6	(52)
BioEdit	(53)	Muscle	(54)
Blast	(55)	NCBI	(56)
CAZy	(40)	NEBcutter V2.0	(57)
ClustalW	(58)	PDB	(59)
CDD	(60)	PyMol	(61)
Compute pI/MW tool	(62)	QMean	(63)
EBI	(64)	Robetta	(65)
FancyGene	(66)	SignalP 4.1	(67)
Gromos	(68)	SwissModel	(69)
I-Tasser	(70)		

II.1.4 Vector de clonación

El vector de clonación utilizado durante este trabajo fue el perteneciente al sistema pGEM[®]-T Easy Vector de Promega. Este vector tiene la particularidad de poseer una timina no apareada en los dos extremos 3' del sitio de clonación. Este extremo está diseñado para facilitar la ligación de fragmentos amplificados usando la Taq Polimerasa, pues ésta tiene como propiedad el añadir una adenina no apareada en el extremo 3' del amplificado. Este vector además posee el gen de la resistencia a la ampicilina, permitiendo así realizar la selección por este antibiótico. Finalmente, este vector tiene en su Sitio de Clonamiento Múltiple (MCS) el gen de la β -galactosidasa. Al ocurrir una clonación exitosa, y en un medio con X-Gal, esto permite seleccionar las colonias de interés por su color blanco. A continuación se presenta el MCS (Figura 8) y el mapa del vector indicado (Figura 9).

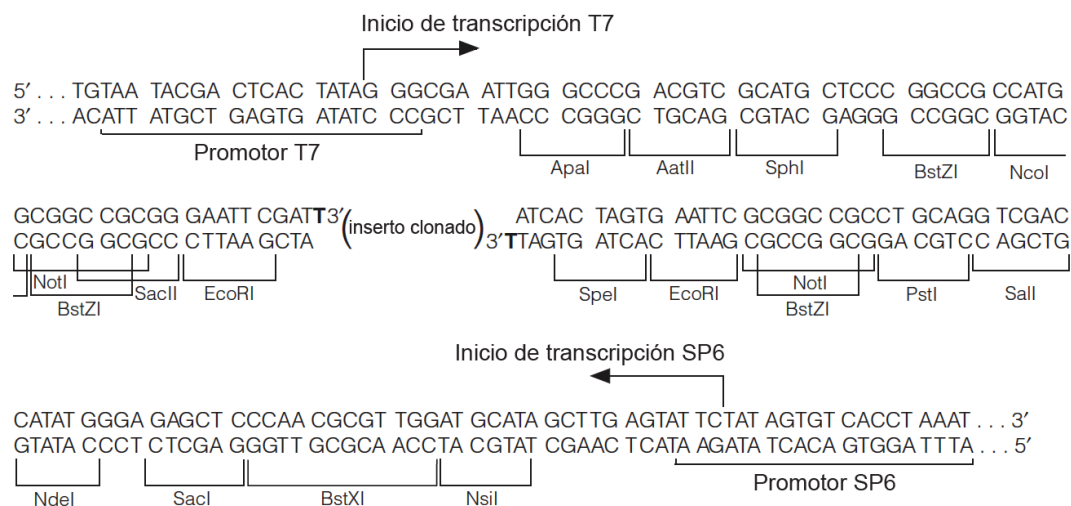


Figura 8: Secuencia y detalle del sitio de múltiple clonamiento del vector pGEM-T Easy de Promega. Se detallan los sitios de restricción y promotores disponibles en el vector.

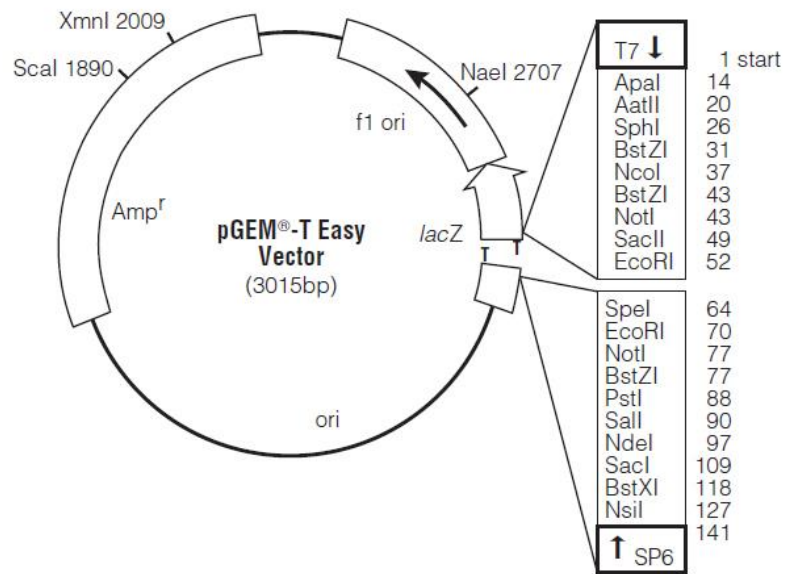


Figura 9: Mapa del vector pGEM-T Easy de Promega. Se detallan los sitios de restricción de todo el vector incluyendo el MCS y el origen de replicación. Se muestran además, esquemáticamente, los genes de resistencia a la ampicilina y de la β -galactosidasa.

II.2 Métodos

II.2.1 Medios de cultivo

II.2.1.1 Medios de cultivo sólido

Los hongos fueron mantenidos para su conservación en placas de PDA que se producían regularmente. La composición de estas placas se encuentra en el Anexo VII.2 y estas eran inoculadas usando un cuadrado de 1 cm² de una placa de PDA incubada con anterioridad.

Para la extracción de RNA se utilizaron placas de medio inductor de celulasas. Con este fin, una vez seleccionados los medios inductores de celulasas más adecuados, estos medios fueron adaptados para cultivos sólidos. Para esto se utilizó la misma composición y protocolo de preparación que para el cultivo líquido añadiéndole a los componentes un 1.5% de agar. En las placas se utilizó solamente paja de trigo como fuente de carbono, la que fue añadida en una concentración de 3% *m/v*.

II.2.1.2 Medios de cultivo líquidos

Con el fin de determinar el medio de cultivo líquido más idóneo para la producción de celulasas, se probaron múltiples medios para cada hongo. Además, se determinó que el protocolo de preparación de estos medios podía influir en la capacidad inductora final de éstos. Un ejemplo de esto es el efecto observado al esterilizar ciertas sales separadas de los otros componentes. También se ensayó esta diferencia, producida tal vez por la precipitación de las sales durante la esterilización.

Otro factor importante en el crecimiento e inducción de la producción de celulasas en los hongos es la fuente de carbono disponible en el medio. Con el fin de evaluar este factor, todos los medios ensayados fueron suplementados ya sea con una fuente de celulosa cristalina (Avicel), o con trigo molido.

En el caso de *Gloeophyllum trabeum* se ensayó tres medios de cultivo mientras que para *Fusarium oxysporum* se ensayó cuatro. Para el primer hongo se probó diferentes composiciones del medio, mientras que para el segundo se probó un mismo medio cambiando otros factores. El primer factor consistió en probar la necesidad de ajustar el pH del medio previo a la esterilización, el segundo factor consistió en esterilizar tanto el extracto de levadura como el CaCl₂ por separado y añadirlo posteriormente al medio. Los medios han sido llamados A, B, C y D, seguido de las iniciales del hongo para el que fueron ensayados: “Gt” para *Gloeophyllum trabeum* y “Fo” para *Fusarium oxysporum*.

Los cultivos se incubaron a 28°C y con agitación baja por a lo menos 13 días. Éstos fueron monitoreados a través de la medición de actividad celulolítica en sus sobrenadantes, usando carboximetil celulosa (CMC) como sustrato. Para la extracción de RNA y producción de DNA complementario (cDNA), los cultivos se detuvieron cuando alcanzaban la fase estacionaria y a fines de su etapa de crecimiento exponencial.

Los medios de cultivo fueron inoculados usando un fragmento de agar de alrededor de 1 cm², extraído desde placas de agar papa dextrosa (PDA) (Anexo VII.2) totalmente cubiertas por micelio del hongo de interés.

II.2.1.3 Ensayo de actividad en medios líquidos

Para evaluar la capacidad inductora de la producción de celulasas de los diferentes medios, se midió la capacidad hidrolítica de las proteínas presentes en el sobrenadante de los cultivos. Esto se realizó tomando una muestra del sobrenadante a diferentes días del cultivo, centrifugándolo a 10,000 RPM por 10 minutos y usando la fase líquida en un ensayo de actividad. La reacción enzimática se llevó a cabo usando 25 μ l de sobrenadante y 50 μ l de sustrato. En este trabajo se utilizó como sustrato una solución de CMC al 1% en buffer acetato de sodio 50 mM pH 5. La reacción se incubó durante 1 hora a 50°C y luego la solución de reacción es ensayada con el método del DNS para medir la concentración de azúcares reductores liberados (71, 72). En esta reacción, 75 μ l del mix de reacción son mezclados con 75 μ l del reactivo DNS e incubados por 10 min a 100°C. Luego estas muestras son traspasadas a una placa de 96 pocillos y se mide su absorbancia a 550 nm (OD_{550}) mediante un lector de placas.

Todos los ensayos fueron realizados en duplicado y usando blancos para determinar la cantidad de azúcares liberados efectivamente por la actividad enzimática. Estos blancos consistían en reacciones en las que el sobrenadante fue reemplazado por un volumen igual de buffer. La actividad CMCasa de los cultivos fue medida en unidades de actividad (U). Cada U corresponde a la cantidad de enzima necesaria para liberar 1 μ mol de equivalente de glucosa por minuto.

II.2.2 Extracción de RNA

Con el fin de realizar la búsqueda de las nuevas proteínas desde genes que se estuvieran efectivamente expresando en los hongos, se procedió a extraer el RNA de estos con el fin de retro transcribirlo y ser usado de molde para las reacciones de PCR.

La extracción de RNA se realizó inicialmente usando biomasa proveniente de hongos crecidos en medios líquidos y usando el reactivo TRIzol® Reagent de Life Technologies. Sin embargo, debido a las dificultades que presenta la extracción desde el micelio de hongos crecidos en medio líquido, se optó por adoptar la extracción de RNA desde placas de medio sólido incorporando las modificaciones descritas por Schumann (73). Siguiendo estas indicaciones, el hongo se creció en placas de un medio de composición idéntica a la de los medios líquidos, con la adición de 1.5% de agar. Posterior a la producción de las placas, el agar se cubrió mediante una membrana de nitrocelulosa previamente esterilizada. La placa se inoculó mediante la colocación de un cubo de agarosa de un cultivo previo en PDA en el centro y sobre la membrana.

Luego de 8 días de cultivo, y cuando el micelio alcanzaba el crecimiento deseado, este era removido mediante el uso de un bisturí. La muestra se pesó y congeló con nitrógeno líquido sobre un mortero previamente tratado con DEPC (Anexo VII.3). Se le añadió entonces un volumen de TRIzol® Reagent equivalente a 10 veces el peso de la muestra y ésta se pulverizó. La muestra fue luego tratada añadiendo 0.2 ml de cloroformo por cada 1 ml de reactivo utilizado, agitando vigorosamente por 15 segundos e incubando a temperatura ambiente por 3 minutos. La muestra fue enseguida centrifugada a 12,000 g por 15 min a 4°C, la fase superior separada a un nuevo tubo y vuelta a tratar con cloroformo de la misma manera. A la fase superior obtenida de la segunda centrifugación se le añadió 0.5 ml de isopropanol por cada 1 ml de reactivo utilizado y se incubó a -20°C por 2 horas. A continuación, ésta se centrifugó a 12,000 g por 10 min a 4°C, y el pellet fue lavado usando 1 ml de etanol frío al 75% por cada 1 ml de reactivo utilizado. La solución se agitó con vórtex y se centrifugó a 7,500 g por 5 min a 4°C, el sobrenadante se eliminó y el pellet se secó por 10-15

minutos a temperatura ambiente. Finalmente el pellet fue resuspendido en 20-50 μ l de agua tratada con DEPC, alicuotado y almacenado a -80°C .

Cabe destacar que todo este proceso debe realizarse usando materiales esterilizados y agua tratada con DEPC (Anexo VII.3). Se debe intentar en todo momento disminuir el efecto de las nucleasas sobre las muestras, con el uso de guantes y manteniendo las muestras en hielo siempre que no se especifique lo contrario.

Con el fin de probar la efectividad del método, se realizaron dos extracciones diferentes para ambos hongos. Ambas extracciones se hicieron siguiendo las modificaciones propuestas por Schumann (73), sin embargo una se hizo a partir de cultivos en medio líquido y otra se realizó a partir de placas cubiertas por la membrana. Se calculó la concentración, integridad y rendimiento del RNA extraído para ambos procedimientos.

II.2.2.1 Verificación de la integridad del RNA extraído

Debido a la labilidad del RNA extraído, se debe verificar la integridad de éste antes de su uso. Con este fin, un volumen de 5 μ l fue corrido en una electroforesis en gel de agarosa al 1.5% y bromuro de etidio al 0.005%. Tanto el gel como el buffer usado durante la electroforesis deben ser hechos con agua tratada con DEPC (Anexo VII.3). Si no ocurrieron problemas de degradación durante la extracción deberían verse dos bandas bien definidas, correspondientes a las subunidades 18S y 28S del RNA ribosomal (rRNA).

El RNA también fue cuantificado mediante la medición de absorbancia a 260 nm y su pureza evaluada en base a la razón entre la absorbancia a 260 nm y la absorbancia a 280 nm.

II.2.3 Síntesis de cDNA

II.2.3.1 Retrotranscripción de RNA a cDNA

La retrotranscripción del RNA a cDNA se realizó utilizando el kit SuperScript™ III Reverse Transcriptase de Invitrogen según las instrucciones del fabricante y añadiendo 5 μ g de RNA total.

II.2.3.2 Purificación del cDNA

Debido a que la retrotranscripción se realiza utilizando un partidador poli-T, el cDNA resultante debe ser purificado. Este paso es crítico, pues los partidores poli-T remanentes en la solución amplificarían de forma inespecífica durante las PCR siguientes.

La purificación se realizó usando el kit Gel/PCR DNA Fragments Extraction Kit de Geneaid, siguiendo las instrucciones del fabricante para PCR Clean Up.

II.2.4 Extracción de DNA genómico

Una vez identificados los genes de interés, y habiéndose determinado por los genomas que estos no presentaban intrones, se procedió a su amplificación y clonación desde DNA genómico (gDNA). El gDNA fue extraído desde hongos crecidos en los medios líquidos seleccionados usando el kit Wizard® Genomic DNA Purification Kit de Promega. Para la extracción se siguieron las instrucciones del fabricante para la extracción desde levaduras, añadiendo un paso de congelación a -80°C previo a la adición de litocasa con el fin de lisar las células.

II.2.5 Diseño de partidores

La amplificación de los genes se realizó en un proceso de dos etapas. En una primera ronda se diseñaron partidores degenerados para poder identificar dentro de la batería de LPMOs de cada hongo, los que se expresan bajo las condiciones de cultivo. En una segunda etapa, se diseñaron partidores específicos para amplificar y clonar los genes seleccionados.

II.2.5.1 Diseño de partidores degenerados

Para el diseño de los partidores degenerados se realizó una búsqueda de las LPMOs presentes en las bases de datos del European Bioinformatics Institute (EBI) (64), del National Center for Biotechnology Information (NCBI) (56) y de la Carbohydrate-Active Enzymes Database (CAZy) (40). Las secuencias obtenidas fueron alineadas en búsqueda de dominios conservados que sirvieran de sitios de unión a partidores. La búsqueda se centró además en la búsqueda de genes parcialmente descritos previamente por un alumno memorista del laboratorio (72). Las secuencias obtenidas fueron alineadas usando ClustalW (58) y visualizadas en el software BioEdit (53). Los partidores fueron diseñados de forma de coincidir con los alineamientos múltiples, lo que significó que éstos contuvieran degeneraciones. Para aumentar las probabilidades de amplificar un gen, se intentó en esta ronda amplificar solamente una fracción de las secuencias de interés en lugar del gen completo.

II.2.5.2 Diseño de partidores específicos

Luego de identificados los genes de LPMOs que se expresan en los hongos a través de la primera ronda de PCRs, se diseñaron partidores para la amplificación del gen completo. Dado que los hongos identificados en la primera etapa no contenían intrones en sus secuencias, se usó la información del genoma de los hongos para el diseño de los partidores y su clonación. Se añadieron además extremos no complementarios con secuencias necesarias para su posterior clonación en vectores de expresión.

II.2.6 Amplificación mediante PCR

II.2.6.1 Amplificación mediante PCR con partidores degenerados

La primera amplificación tuvo como fin buscar genes que efectivamente se expresaran bajo las condiciones de cultivo elegidas. Ésta se realizó con partidores degenerados con el fin de abarcar varios genes a la vez y soslayar mutaciones no previstas. La amplificación se realizó mediante PCR con la siguiente solución de reacción: 1 μ M de cada partidador (directo y reverso), buffer Taq 1X, 0.2 mM de mix de dNTPs, 1.4 mM de MgCl₂, 1 U de Taq polimerasa y 2 μ l de cDNA purificado. El programa utilizado fue el siguiente: 1 ciclo a 95°C por 2 min, 2 ciclos de 95°C por 2 min, 35°C por 30 s y 72°C por 90 s, 30 ciclos de 95°C por 2 min, temperatura de alineamiento (Ta) por 30 s, 72°C por 90 s, y 1 ciclo de 72°C por 5 min.

Los primeros dos ciclos tienen como fin lograr la amplificación menos estricta de secuencias similares a los partidores, y así aumentar el número de copias de hebra molde para los siguientes ciclos. Este paso se hace necesario debido a la probabilidad de que los partidores no coincidan en su totalidad con el molde presente en el cDNA. La temperatura de alineamiento depende del partidador utilizado, y se encuentra detallada para cada uno de estos en la Tabla 13 en la sección III.4.2 de Resultados y Discusión.

II.2.6.2 Amplificación mediante PCR con polimerasa de alta fidelidad

La segunda ronda de amplificaciones fue realizada una vez identificados los genes de interés, y con el fin de amplificar las secuencias completas, y clonar las secuencias en vectores. Para obtener una copia sin errores del gen, ésta fue realizada usando la DNA polimerasa Phusion de alta fidelidad (England Biolabs). Al poseer actividad exonucleasa 3'→5', esta polimerasa detecta errores en la transcripción y es capaz de corregirlos, de esta forma el producto final contiene menos mutaciones que al usar una polimerasa tradicional. La reacción se realizó usando la siguiente solución de reacción: 1 µM de cada partidor (directo y reverso), buffer HF 1X, 0.2 mM de mix de dNTPs, 1 U de Phusion DNA Polymerase, 3% [v/v] de DMSO y 2 µl de gDNA purificado. El programa utilizado fue el siguiente: : 1 ciclo a 98°C por 30 s, 30 ciclos de 98°C por 10 s, Ta por 30 s, 72°C por 30 s, y 1 ciclo de 72°C por 5 min.

En el caso de los partidores FO61G22f y FO61G22r, y debido a las dificultades que la amplificación con partidores con extremos no complementarios produce, el programa de PCR fue adaptado a uno de TouchDown PCR. El programa consiste en 15 ciclos iniciales de Ta decreciente que se realizan con el fin de disminuir la estrictez de la amplificación, sin disminuir excesivamente la especificidad. El programa utilizado fue: 1 ciclo a 95°C por 3 min, 15 ciclos de 95°C por 2 min, Ta (decreciente de 1°C por ciclo desde 50°C a 35°C) por 30 s, 72°C por 90s, 20 ciclos de 95°C por 2 min, 50°C por 30 s, 72°C por 90 s, y 1 ciclo final de 72°C por 5 min. Además de usar este programa, sólo se añadió la DNA Polimerasa a la reacción una vez alcanzados los 95°C iniciales, proceso conocido como Hot Start PCR. La solución de reacción utilizada fue la descrita anteriormente para la Phusion DNA Polymerase.

II.2.6.3 Amplificación de fragmento 18S

Con el fin de determinar molecularmente que los hongos con lo que se trabajó correspondían efectivamente a lo rotulado, se llevó a cabo la amplificación del fragmento 18S de su RNA ribosomal. Este gen corresponde a parte de la subunidad 40S del ribosoma y se caracteriza por poseer una secuencia diferente en cada especie. De esta forma, al secuenciar el 18S de un organismo eucarionte, éste puede ser identificado por su secuencia.

La amplificación se llevó a cabo con los partidores especificados en la Tabla 4. Estos corresponden a partidores específicos para la región correspondiente al 18S en el genoma. La solución de reacción utilizada fue la misma que para los partidores degenerados y se encuentra descrita más arriba (encabezado II.2.6.1). El molde utilizado fue DNA genómico, y se hizo dos pruebas: con el gDNA concentrado y diluido 10 veces. El programa para la amplificación consistió en: 34 ciclos de 94°C por 30 s, 62°C por 1 min, 72°C por 2 min y un ciclo de 72°C por 5 min.

Tabla 4: Secuencia de partidores utilizados para la amplificación del fragmento 18S del RNA ribosomal de hongos.

Nombre partidor	Secuencia (5'→3')
F18S_fwd	ATTGGAGGGCAAGTCTGGTG
F18S_rev	CCGATCCCTAGTCGGCATAG

El proceso de purificación y clonación utilizado fue el mismo que para las etapas de identificación y de amplificación del gen completo. Éste se detalla en los puntos siguientes.

II.2.7 Clonación de genes

Luego de amplificadas las secuencias de los genes de interés, éstas son separadas y purificadas en una electroforesis en gel de agarosa al 1%. Esto se realiza usando el kit QIAEX II Gel Extraction de Qiagen según las instrucciones del fabricante.

II.2.7.1 Ligación

El fragmento es entonces ligado al vector pGEMT-Easy de Promega. Para esto se prepara una solución con 5 μ l de 2X Rapid Ligation Buffer, 0.5 μ l de vector pGEM-T Easy, 1 μ l de ligasa T4 y 3.5 μ l de DNA previamente extraído. Esta solución se deja incubando por 1 hora a temperatura ambiente o toda la noche a 4°C.

II.2.7.2 A-Tailing

En el caso de los genes amplificados usando la DNA polimerasa Phusion, un proceso adicional de adenilación fue necesario previo a la ligación. Este proceso también conocido como “A-tailing” hace uso de la propiedad de la Taq Polimerasa de incorporar una adenina libre al extremo 3’ de la secuencia amplificada, y que es necesario para la ligación del fragmento al vector pGEM-T Easy. La solución de reacción contiene 1 a 5 μ l de DNA, 2 μ l de Buffer 5X de la polimerasa, 0.6 μ l de $MgCl_2$ 25 mM, 1 μ l de dATP 2mM, 1 μ l de Taq polimerasa y agua hasta completar 10 μ l. La solución se incuba finalmente a 70°C por 30 min.

II.2.7.3 Electroporación

El vector es luego transformado en células *E. coli* DH5 α electrocompetentes previamente preparadas (Anexo VII.1) a través de electroporación. El proceso de electroporación consiste en los siguientes pasos y se realiza en ambiente estéril:

- 50 μ l de bacterias se incuban en hielo por 10 min.
- Se añade 5 μ l de solución de ligación a las bacterias y se mezclan suavemente.
- La solución se incuba en hielo por 30 min.
- Las bacterias son traspasadas a una cubeta de electroporación fría.
- La cubeta es puesta en el electroporador y sometida a un pulso de 2200 V.
- Las bacterias se resuspenden en 1 ml de medio LB y se incuban a 37°C por una hora

Luego de transformadas, las células son sembradas en placas de Agar selectivo (Anexo VII.2) e incubadas a 37°C por 16 horas.

II.2.7.4 Verificación de la transformación

La correcta incorporación del gen de interés al vector transformado fue verificada mediante el corte del fragmento ligado al gen desde el plásmido y su posterior visualización. Para esto se realiza una miniprep de un cultivo saturado de 4 ml de las bacterias transformadas. Posteriormente el plásmido es digerido para así obtener el fragmento de interés. En este caso el vector fue digerido usando la enzima de restricción EcoR I, que corta el fragmento en ambos extremos. La solución de reacción para la digestión consta de: 15 μ l de agua, 2 μ l de Buffer EcoR I 10X, 2 μ l de plásmido y 0.5 μ l de EcoR I [10 U/ μ l]. La solución se mezcla suavemente y se incuba por 1 hora a 37°C. El producto de la digestión es entonces visualizado en una electroforesis en gel de agarosa.

II.2.7.5 Secuenciación

La secuenciación de las muestras fue realizada por MacroGen Korea haciendo uso de plásmidos obtenidos mediante una miniprep de los clones transformados. Los resultados fueron analizados haciendo uso del software BioEdit y de las bases de datos de la NCBI.

II.2.8 Análisis bioinformático

Luego de la secuenciación de los genes clonados, éstos fueron analizados haciendo uso de múltiples herramientas bioinformáticas.

II.2.8.1 Mapa de restricción

Una vez obtenidas, y con el fin de determinar qué enzimas de restricción son susceptibles de cortar las secuencias nucleotídicas en lugares no deseados, éstas se analizaron mediante la herramienta en línea NEBcutter V2.0 (57). Ésta se usó para determinar los sitios de corte de todas las enzimas de restricción comercialmente disponibles, y en base a esto se obtuvo un mapa de restricción.

II.2.8.2 Alineamiento de secuencias con otras de la familia AA9

Para probar la existencia de aminoácidos conservados en las secuencias clonadas, se realizó un alineamiento entre las secuencias de LPMO caracterizadas y las obtenidas, usando la herramienta Muscle (54). Se obtuvo los módulos correspondientes a tanto el CBM como el GH61 mediante el uso del CDD (60). Se pudo además obtener una predicción del péptido señal de las proteínas a través del uso del servidor SignalP 4.1 (67).

II.2.8.3 Determinación de la subfamilia de las LPMOs

A pesar de que se trata de una familia recientemente descubierta y de carácter aún enigmático, las LPMOs han intentado ser catalogadas en subfamilias o tipos. Un trabajo interesante en este aspecto es el del grupo de Xin Li (74), que plantea una división según el tipo de producto liberado por la acción de esta enzima. Con el fin de catalogar las LPMOs aquí descritas según el criterio de este grupo, éstas fueron analizadas mediante la misma técnica. En un primer lugar se seleccionaron secuencias correspondientes a LPMOs de *Neurospora crassa* y de otros hongos de referencia. Luego estas secuencias, junto con las obtenidas en este trabajo, fueron cortadas en el extremo N-terminal de forma de eliminar el péptido señal, y dejar en todos los casos una histidina como residuo inicial. A continuación se removió de las secuencias los extremos C-terminales de forma de dejar en todos los casos la secuencia conservada Pro-Gly-Pro como extremo final. Estas secuencias fueron luego alineadas usando la herramienta en línea T-Coffee (75). El resultado de este alineamiento se utilizó enseguida para calcular un árbol filogenético mediante el método de máxima similitud. El árbol fue calculado usando la herramienta Mega6 (52) y graficado con la misma.

II.2.8.4 Modelamiento de la estructura de las LPMOs en base a homología.

Dadas las limitantes impuestas por no poder expresar las proteínas clonadas de forma recombinante, se decidió realizar un análisis bioinformático exhaustivo. Se buscó entonces modelar las proteínas de forma teórica, haciendo uso de las herramientas computacionales creadas por diversos grupos. La modelación fue realizada con el fin de poder predecir la estructura secundaria

y terciaria de las proteínas. Este análisis se realizó como una primera aproximación a las estructuras, sin embargo este tipo de modelo presenta claras limitaciones.

El estudio fue entonces llevado a cabo modelando las proteínas obtenidas en servidores disponibles online. Estos servidores cuentan con diferentes algoritmos para calcular estructuras tridimensionales en base a moldes ya descritos. En forma resumida, estos algoritmos buscan entre las proteínas cuya estructura tridimensional ya haya sido dilucidada por cristalografía, la que sea más similar a la proteína objetivo. Una vez obtenida, la proteína de interés es modelada usando como molde la estructura de ésta, considerando durante el proceso varios factores tales como la energía de la proteína e impedimentos estéricos. Durante este trabajo, las predicciones fueron realizadas ya sea dejando la elección del molde al algoritmo del servidor, o utilizando un molde seleccionado y alineado previamente.

Los servidores utilizados para la modelación fueron SwissModel (69), I-Tasser (70) y Robetta (65). Se analizaron los resultados obtenidos desde estos servidores mediante el uso de diferentes herramientas que entregan un puntaje a la proteína en base a análisis estadísticos de diversos factores. Las herramientas utilizadas fueron Anolea (51), QMean (63), Gromos (68) y el Z-Score (63). El puntaje entregado por estas herramientas, además de un análisis de la integridad de las estructuras obtenidas permitió elegir el modelo más acertado para cada una de las enzimas. Éstas luego fueron visualizadas y sus estructuras analizadas con el uso de PyMol (61).

II.2.8.1 Análisis filogenético

Con el fin de obtener la mayor cantidad de información posible desde los clones obtenidos, se realizó un análisis filogenético de éstos junto con secuencias de LPMOs de otros hongos. Se realizó en primer lugar una búsqueda de todas las LPMOs caracterizadas y disponibles en las bases de datos. Se espera de esta forma poder aproximar las propiedades de las proteínas encontradas a las ya descritas. Se buscó además las secuencias de LPMOs disponibles en los genomas de los hongos basidiomicetos de referencia *Phanerochaete chrysosporium* (76) y *Trametes versicolor* (77) y del hongo ascomiceto *Trichoderma reesei* (78). Finalmente se buscaron todas las LPMO descritas en *Fusarium oxysporum* y *Gloeophyllum trabeum* y disponibles en sus respectivos genomas (49, 50). Las búsquedas fueron realizadas usando las palabras clave IPR005103 y PF03443. Con el fin de evaluar más precisamente la evolución del módulo LPMO, las secuencias fueron analizadas mediante el buscador de dominios conservados (CDD) del NCBI (60) y sólo el módulo GH61 se utilizó para la construcción del árbol filogenético.

Las múltiples secuencias fueron luego alineadas usando la herramienta Muscle (54) y el árbol filogenético construido usando el software Mega6 (52). El árbol fue construido usando el método Neighbor Joining y testeado usando Bootstrapping (500 replicaciones). Las distancias evolutivas fueron calculadas usando el método de corrección de Poisson. Los gaps y deleciones fueron tratados a través de eliminación de a pares. El árbol fue finalmente visualizado utilizando el mismo software Mega6.

III Resultados y Discusión

III.1 Identificación de medios de cultivo inductores de la producción de celulasas

Se ha observado con anterioridad que la producción de LPMOs se encuentra asociada a la producción de celulasas (27, 28). Por otra parte, como los medios inductores de la producción de celulasas se encuentran relativamente bien definidos, se optó por cultivar los hongos en estos. Con el fin de optimizar la expresión de celulasas en los hongos estudiados se procedió a probar diversos medios de cultivo. Se probó un total de 3 medios para *G. trabeum* y 4 medios para *F. oxysporum*. Algunos de estos variaban en composición mientras que en otros se alteró su pH o las condiciones en las que se preparaban. Todos los medios fueron ensayados usando dos fuentes de carbono conocidas por su capacidad de inducir la producción de tanto celulasas como de LPMOs: paja de trigo molido y Avicel (28, 79).

III.1.1 Medios de cultivos para *Fusarium oxysporum*

Para el hongo *Fusarium oxysporum*, se ensayó el mismo medio de cultivo cambiando diferentes factores, tales como el ajuste de pH o la esterilización de sales en forma separada del resto del medio. La composición del medio se encuentra en la Tabla 5. Luego de la esterilización, a este medio se le añaden 7 [ml/l] de la solución de elementos traza (Tabla 6) y 1 [ml/l] de una solución de tiamina a 0.5 [g/l].

Los diferentes medios ensayados y los factores que los diferencian se encuentran resumidos en la Tabla 7.

Tabla 5: Composición de medio de cultivo para *Fusarium oxysporum*.

Compuesto	Concentración [g/l]
(NH ₄)H ₂ PO ₄	2
KH ₂ PO ₄	0.6
MgSO ₄ x 7H ₂ O	0.5
K ₂ HPO ₄ x 3H ₂ O	0.52
CaCl ₂ x 2H ₂ O	0.74
Extracto de levadura	1

Tabla 6: Solución de elementos traza para medios de cultivo.

Compuesto	Concentración [mg/ml]
FeSO ₄ x 7H ₂ O	5
ZnSO ₄ x 7H ₂ O	1.4
MnSO ₄ x H ₂ O	1.6
CoCl ₂ x 6H ₂ O	1.86

Tabla 7: Descripción de los medios de cultivo ensayados para *Fusarium oxysporum*. En todos los casos el medio fue suplementado ya sea con trigo molido o con Avicel. En algunos casos el pH fue ajustado previo a la esterilización, y en el caso del Medio D Fo, el extracto de levadura y el CaCl₂ fueron esterilizados por separado.

Nombre	Medio de cultivo			
	A Fo	B Fo	C Fo	D Fo
Fuente Carbono	Trigo 10% o Avicel 0.5%	Trigo 1%	Trigo 1% o Avicel 0.5%	Trigo 1% o Avicel 0.5%
Ajuste de pH ^A	–	–	6	6
Esterilización de sales por separado ^B	–	–	–	+

^A En caso de ser necesario el pH es ajustado usando ácido fosfórico.

^B El extracto de levadura y el CaCl₂ son esterilizados por separado y añadidos luego al medio.

Durante el desarrollo de los cultivos, estos fueron monitoreados mediante la medición de actividad celulolítica. Al tratarse de proteínas del medio extracelular, las muestras fueron tomadas desde el sobrenadante del cultivo y ensayadas usando CMC como sustrato. Se presentan aquí los resultados para tanto los medios suplementados con paja de trigo (Figura 10) como para los suplementados con Avicel (Figura 11).

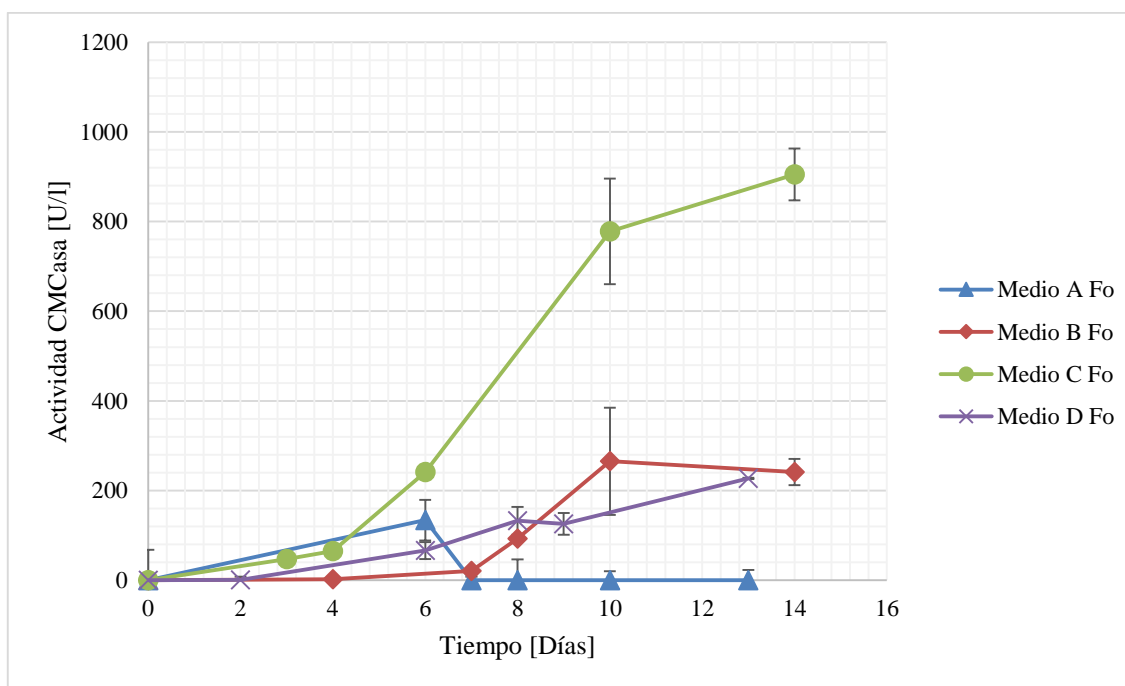


Figura 10: Variación en la actividad celulolítica de los cultivos de *F. oxysporum* para diferentes medios de cultivo suplementados con paja de trigo molido. La actividad es presentada en unidades de actividad por litro de sobrenadante, y fue medida utilizando CMC como sustrato.

De los gráficos de actividad se puede observar que a pesar de que los medios B y C presentaron una actividad creciente en el tiempo, el medio C fue el mejor inductor de la producción de celulasas tanto con Avicel como con paja de trigo.

El medio A, a pesar de resultar discretamente efectivo con Avicel (Figura 11), alcanzando cerca de 120 [U/l] el día 13, no lo fue en el caso de los medios suplementados con paja de trigo (Figura 10), en el que al mismo día no presentó actividad. Este efecto fue probablemente producido por la

concentración de paja utilizada, que en este caso fue de 10% *m/v*. De la abrupta disminución de la actividad de este cultivo, se podría suponer que ésta es producto de la liberación de compuestos fenólicos desde la madera. Se ha demostrado que estos compuestos resultan inhibitorios para la actividad de las celulasas (80). Este efecto puede ser corroborado al comparar la evolución de este medio con la del medio B, cuya composición es idéntica a la del A, salvo que en este caso se utilizó sólo un 1% *m/v* de trigo molido.

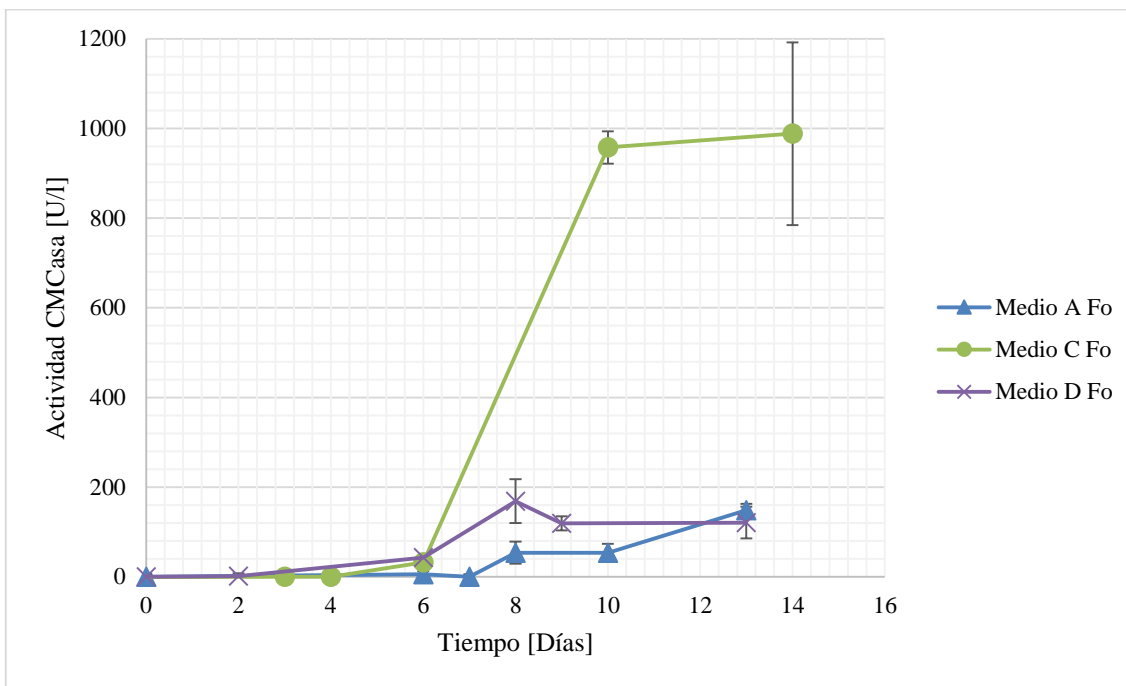


Figura 11: Variación en la actividad celulolítica de los cultivos de *F. oxysporum* para diferentes medios de cultivo suplementados con Avicel. La actividad es presentada en unidades de actividad por litro de sobrenadante, y fue medida utilizando CMC como sustrato.

Por su parte, el medio de cultivo C se diferencia del B en que en el primero, antes de la esterilización del medio, el pH de éste fue ajustado a 6 usando ácido fosfórico. Dado que ambos medios usan la misma composición y la misma concentración de trigo y Avicel, este factor es el único responsable de la diferencia. Se puede observar que el efecto del ajuste de pH es importante, casi quintuplicando la capacidad celulolítica del cultivo. Tanto con trigo como con Avicel, el aumento en la producción comienza a evidenciarse después de una fase lag de 4 días, y parece alcanzar su fase estacionaria al décimo día. En ambos casos se alcanzó una actividad cercana a 1000 [U/l] para este día. Esto podría evidenciar dos posibles efectos. Por una parte, el no ajustar el pH del medio podría estar afectando el crecimiento del hongo, ralentizando así su metabolismo en general. Por otra parte, esta diferencia podría explicarse por la sensibilidad de las celulasas al pH del ambiente en el que se encuentran, viéndose afectada su actividad. Un análisis más detallado del crecimiento de la biomasa, junto con la cuantificación de las celulasas en la solución permitiría analizar en mayor profundidad estos dos factores.

Tanto el medio C como el medio D parecen inducir una mayor producción de celulasas al ser suplementados con trigo. El contenido de celulosa en la paja de trigo es de alrededor del 36%, y su contenido de carbono total ha sido cifrado en alrededor de 40% (81, 82). Esto implicaría que tanto el medio con Avicel como el suplementado con paja de trigo, entregan una concentración similar de carbono al hongo. La diferencia en el comportamiento en estos dos medios podría entonces ser

explicada por la naturaleza más compleja de la madera. Al estar compuesta la paja de trigo por más componentes que el Avicel es posible que una actividad inicial del hongo sobre este sustrato, libere metabolitos capaces de inducir una mayor producción de enzimas celulolíticas. Otra posible explicación de este fenómeno sería que la madera resulta ser un mejor sustrato para el crecimiento del hongo, entregándole a éste micronutrientes no disponibles en el Avicel. Estos nutrientes le permitirían entonces crecer más rápidamente y tener un metabolismo acelerado. Esta tendencia parece ser inversa en el caso del medio C, alcanzándose el máximo más rápidamente en presencia de Avicel que de trigo. Esta diferencia es sin embargo pequeña y podría atribuirse simplemente a diferencias en la manipulación de los hongos.

El experimento con el medio D consistió en una prueba realizada bajo el supuesto de que la esterilización del extracto de levadura y del CaCl_2 en conjunto con el medio podría llevar a la formación de un precipitado. Este efecto, observado en la producción de medios de cultivo para otros hongos fue ensayado esterilizando dichos compuestos de forma separada del resto del medio. Sorprendentemente, este cambio al que se le podrían atribuir resultados positivos, resultó producir un efecto negativo en el cultivo del hongo. Los medios suplementados tanto con Avicel como con paja de trigo mostraron una disminución considerable en la producción de celulasas, pasando de un máximo de 1000 [U/l] a 200 [U/l]. Una posible causa de este fenómeno, podría ser la alteración al pH que producen estos compuestos al ser agregados justo antes de la inoculación, alcanzándose un efecto similar al de los medios cuyo pH no es ajustado.

Luego del análisis de los datos encontrados, se determinó que el medio de cultivo más adecuado para la producción de celulasas desde *F. oxysporum* era el Medio C Fo suplementado con trigo. Se determinó además que la biomasa del cultivo sería cosechada al octavo día, con el fin de obtener la mayor producción de celulasas por parte del hongo, sin llegar a observar los cambios metabólicos característicos de la fase estacionaria.

III.1.2 Medios de cultivo para *Gloeophyllum trabeum*

Para el hongo *Gloeophyllum trabeum* se probó 3 medios (Tabla 8). Un medio simple compuesto únicamente de glucosa (A Gt), un medio complejo (B Gt) en base al trabajo de Chouiter (83), y una modificación de este último (C Gt). En el caso de estos medio no se ajustó el pH previo a la esterilización e inoculación. Al igual que en el caso anterior, todos los medios fueron ensayados usando tanto paja de trigo (Figura 12) como Avicel (Figura 13) como fuente de carbono. En el caso del Medio A Gt, el porcentaje de paja de trigo utilizado fue de 10% *m/v*. La actividad celulolítica fue monitoreada por un mínimo de 13 días utilizando CMC como sustrato.

Tabla 8: Composición de los medios de cultivo ensayados para *Gloeophyllum trabeum*. En los tres casos la composición del medio fue variada y probada con dos fuentes de carbono diferentes: paja de trigo y Avicel.

Medio de cultivo			
Nombre	A Gt	B Gt ^A	C Gt
Composición [g/l]			
Glucosa	1	–	–
KH ₂ PO ₄	–	0.46	0.46
MgSO ₄ x 7H ₂ O	–	0.05	0.005
K ₂ HPO ₄ x 3H ₂ O	–	1	1
Extracto de levadura	–	2	2
Fuente de carbono	Trigo 10% o Avicel 0,5%	Trigo 1% o Avicel 0,5%	Trigo 1% o Avicel 0,5%

^A De acuerdo al medio utilizado por Chouiter (83).

De los gráficos se puede observar que tanto el Medio A como el Medio B no resultaron ser buenos medios inductores de la producción de celulasas. La producción de celulasas en estos medios, no tuvo en ningún caso una tendencia al alza. Cuando sí la tuvo, que es el caso del Medio B con Avicel, este alcanzó un máximo de alrededor de 1400 [U/l]. Aunque esta actividad es mayor que la alcanzada por los cultivos de *F. oxysporum* anteriormente descritos, es bastante inferior al máximo obtenido por el Medio C Gt.

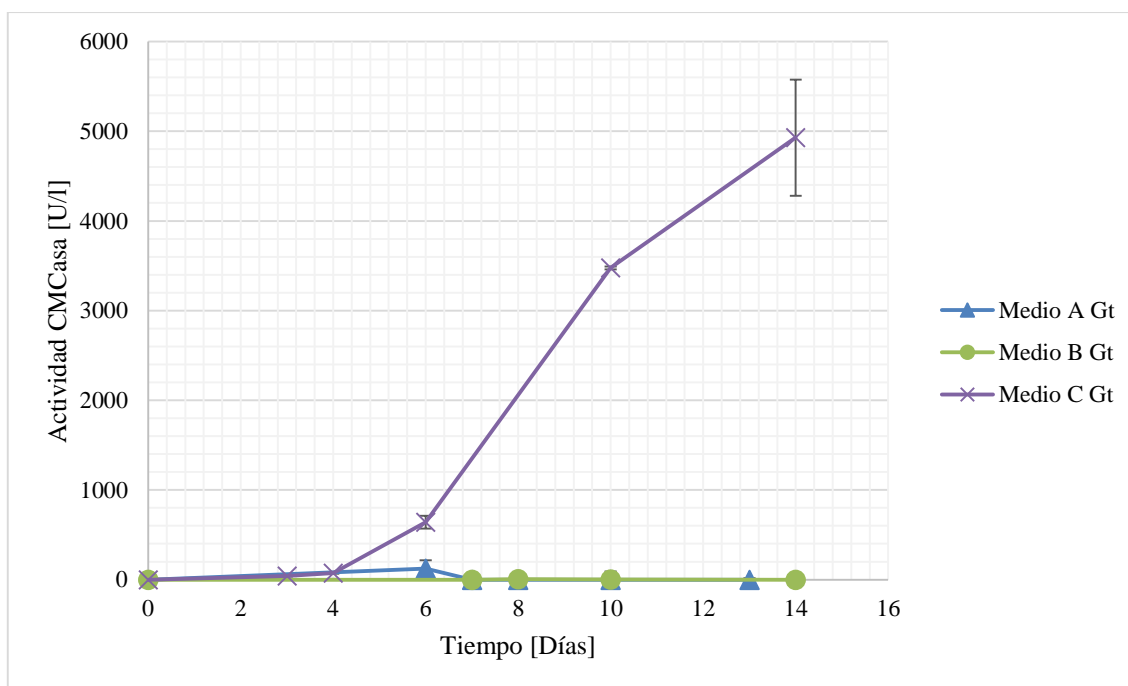


Figura 12: Variación en la actividad celulolítica de los cultivos de *G. trabeum* para diferentes medios de cultivo suplementados con paja de trigo molida. La actividad es presentada en unidades de actividad por litro de sobrenadante, y fue medida utilizando CMC como sustrato.

La poca actividad en el Medio A es probablemente atribuible a que éste era demasiado simple para el cultivo del hongo. Al tratarse de un medio compuesto únicamente por 1 [g/l] de glucosa, el crecimiento del hongo bajo estas condiciones dependía de que éste obtuviera todos los elementos necesarios para su crecimiento desde el trigo o el Avicel. Claramente esto no fue el caso,

debiéndose probablemente a lo recalcitrante que es la lignocelulosa a la degradación, y al hecho de que el Avicel está compuesto únicamente de celulosa. Otra posible explicación de la poca actividad presentada por este medio, es la posibilidad de que la producción o la actividad de las celulasas se vea inhibida por la presencia de glucosa en el medio. Se ha reportado previamente que incluso bajas concentraciones de glucosa (4 [g/l]) pueden presentar un fuerte efecto inhibitorio sobre la actividad de β -glucosidasas y celulasas en general (84).

Por otra parte, se puede apreciar más actividad en el Medio A suplementado con Avicel que en el suplementado con paja de trigo. Este comportamiento apoya la hipótesis ya planteada, que el trigo en altas concentraciones (10% *m/v*) podría liberar compuestos inhibidores del crecimiento o de la actividad celulolítica.

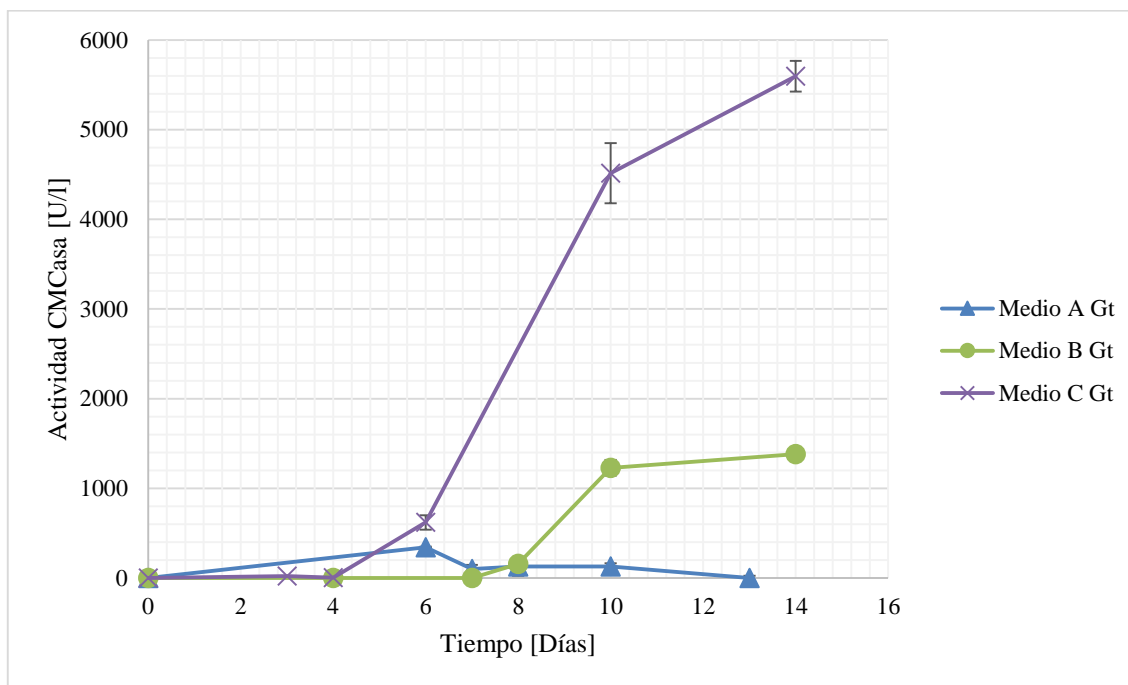


Figura 13: Variación en la actividad celulolítica de los cultivos de *G. trabeum* para diferentes medios de cultivo suplementados con Avicel. La actividad es presentada en unidades de actividad por litro de sobrenadante, y fue medida utilizando CMC como sustrato.

El Medio B Gt por su parte corresponde a un medio anteriormente descrito como el adecuado para la inducción de la producción de celulasas (83) en cultivos del hongo *Ganoderma applanatum*. Este medio sólo presentó inducción de celulasas cuando fue puesto en contacto con Avicel, alcanzando su máximo al décimo día. La falta de actividad en el medio suplementado con trigo puede deberse a múltiples factores. Uno de estos podría ser el efecto inhibitorio que pueden tener sobre las enzimas los compuesto fenólicos liberados por la lignocelulosa al ser degradados (80). Al encontrarse éstos en el sobrenadante al momento de las tomas de muestras, éstos pudieron afectar negativamente los ensayos de actividad. Una posible explicación para la poca actividad que presenta este medio en comparación al Medio C Gt, es que la producción o la actividad de las enzimas necesarias para la degradación del trigo se vieran inhibidas por la presencia de magnesio en el medio. Esto explicaría que sólo algunas de las proteínas que componen la batería enzimática del hongo para degradar lignocelulosa, se vieran expresadas o estuvieran activas en este medio.

La única diferencia que presenta el Medio C Gt con el Medio B Gt es la concentración de magnesio, siendo esta 10 veces menor en el caso del Medio C. Esta diferencia produce una gran brecha en la capacidad inductora de estos medios. Es así como el Medio C alcanza un máximo de alrededor de 5 veces el que alcanza el medio B. El máximo alcanzado es a su vez cinco veces mayor al obtenido de los cultivos de *F. oxysporum*, sugiriendo ya sea, que *G. trabeum* es un productor considerablemente mejor de celulasas que su contraparte, o que las enzimas de éste tienen mayor especificidad por el sustrato utilizado en los ensayos de actividad (CMC). Una comparación más precisa entre la capacidad celulolítica de estos hongos, podría realizarse normalizando la actividad de estos en función de la biomasa producida y las proteínas excretadas al medio extracelular.

Finalmente el Medio C suplementado con trigo fue el elegido para la producción de biomasa con vistas a la extracción de RNA. Nuevamente, estos cultivos fueron cosechados al octavo día con el fin de optimizar la cantidad de mRNA de LPMOs presente en las células.

Aunque estos medios sí resultan útiles para el trabajo realizado, una optimización más profunda de éstos podría asegurar una mayor producción de LPMOs por parte de los hongos. Esto llevaría a su vez a que existiera una mayor probabilidad de que durante el PCR con partidores degenerados, se pudiera amplificar nuevos genes de LPMOs, o genes que se expresan en menor proporción. Un acercamiento interesante para la optimización de estos cultivos, sería la realización de los ensayos de actividad bajo condiciones específicas para pesquisar la actividad de las LPMOs. Como se explicó anteriormente, en los últimos años se ha determinado que existen co-factores tales como cobre y elementos reductores que resultan indispensables para la actividad de la enzima. Realizar ensayos de actividad en presencia de estos elementos, permitiría detectar un aumento de este tipo de proteínas específicamente, y diferenciarlo del aumento de celulasas tradicionales.

III.2 Identificación de los hongos mediante secuenciación del rRNA 18S

Los hongos utilizados en este trabajo se encontraban catalogados por especie de forma previa. No obstante, con el fin de determinar más específicamente la especie de los hongos en cuestión, se decidió secuenciar su gen correspondiente al fragmento ribosomal 18S. Este gen es conocido por presentar alteraciones que son únicas para cada especie, por lo que su secuenciación y comparación con bases de datos de referencia permite identificar de qué organismo se trata.

Los genes del 18S fueron satisfactoriamente amplificados para ambos hongos desde gDNA, como se puede observar en la electroforesis realizada con los productos de PCR (Figura 14). Las bandas obtenidas son de aproximadamente 500 bp. Las secuencias obtenidas para cada uno de estos hongos se resumen en el Anexo VII.6.

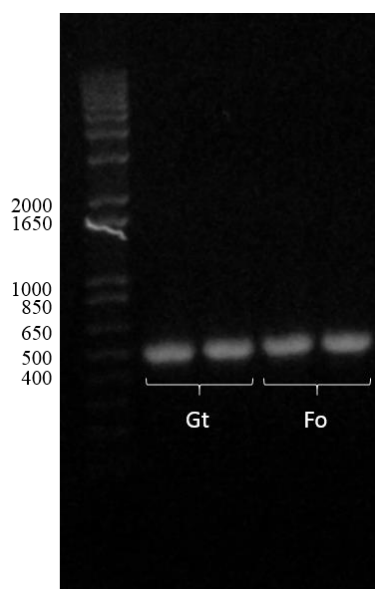


Figura 14: Visualización de la amplificación de los genes del rDNA 18S de hongos en una electroforesis en gel de agarosa. El primer carril corresponde al estándar de peso molecular 1kb Plus DNA Ladder de Invitrogen. Los carriles marcados con un corchete corresponden a las amplificaciones con DNA sin diluir (a la izquierda) y diluido 10 veces (a la derecha), de ambos hongos: *G. trabeum* (Gt), y *F. oxysporum* (Fo).

A continuación, se procedió a analizar estas secuencias y a buscarlas en las bases de datos disponibles a través de Blast (55). Los resultados obtenidos de esta búsqueda se presentan en la Tabla 9 y en la Tabla 10 para los hongos *Gloeophyllum trabeum* y *Fusarium oxysporum*, respectivamente. En estos se puede observar que efectivamente el primer resultado corresponde a lo esperado. Más aún, los resultados siguientes corresponden casi exclusivamente a fragmentos de rDNA 18S de microorganismos del mismo género. Los que no corresponden al mismo género presentan además en casi todos los casos puntajes inferiores a los que sí corresponden. Sin embargo, se puede observar que dentro del mismo género no se observan diferencias entre las diferentes especies, obteniéndose alineamientos de igual puntaje para todos éstos.

Tabla 9: Resultados de la búsqueda de las secuencias del 18S del *Gloeophyllum trabeum* en la base de datos Blast (55). Se presenta el puntaje (Score) e identidad porcentual de los primeros resultados no redundantes.

Microorganismo	Score	Identidad (%)	ID Taxonómico*
<i>Gloeophyllum trabeum</i>	931	100	104355
<i>Gloeophyllum subferrugineum</i>	931	100	941528
<i>Gloeophyllum striatum</i>	931	100	941527
<i>Gloeophyllum abietinum</i>	931	100	180171
<i>Gloeophyllum sepiarium</i>	926	99	40444
<i>Donkioporia expansa</i>	915	99	81044
<i>Coriopsis byrsina</i>	909	99	239205
<i>Neolentinus lepideus</i>	909	99	38799
<i>Veluticeps berkeleyi</i>	909	99	103395

*Identificador de la base de datos taxonómica del NCBI.

Tabla 10: Resultados de la búsqueda de las secuencias del 18S del *Fusarium oxysporum* en la base de datos Blast (55). Se presenta el puntaje (Score) e identidad porcentual de los primeros resultados no redundantes.

Microorganismo	Score	Identidad (%)	ID Taxonómico*
<i>Fusarium oxysporum</i>	929	100	5507
<i>Fusarium solani</i>	929	100	169388
<i>Fusarium verticillioides</i>	929	100	117187
<i>Fusarium proliferatum</i>	929	100	948311
<i>Fusarium fujikuroi</i>	929	100	5127
<i>Ophiocordyceps sinensis</i>	929	100	72228
<i>Fusarium neocosmosporiellum</i>	929	100	1325723
<i>Gliocladiopsis sp.</i>	924	99	1185363
<i>Cylindrocladiella lageniformis</i>	924	99	147879

*Identificador de la base de datos taxonómica del NCBI.

Del análisis de estos resultados, sumados a los obtenidos de las diversas clonaciones descritas más adelante en este trabajo, es posible determinar de forma bastante segura que se trata efectivamente de los hongos en cuestión.

Cabe destacar que este método no probó ser tan específico como se deseaba, y que con su uso por sí sólo no se puede diferenciar entre especies del mismo género. Como se ha reportado con anterioridad, la alta conservación de la secuencia del 18S, que nos permite utilizarlo en la diferenciación de géneros de microorganismos, es a su vez su debilidad, puesto que sus diferencias no son las suficientes como para distinguir entre especies (85).

III.3 Evaluación de método de extracción de RNA desde hongos

El proceso de identificación de las proteínas se realizó a través de la amplificación de sus respectivos genes mediante PCR. En este trabajo, se decidió realizar esta amplificación desde el cDNA obtenido a partir del mRNA de los hongos. Esto asegura que en la identificación se obtengan sólo genes que se estén expresando bajo las condiciones determinadas por los medios de cultivo. Para obtener una adecuada amplificación de los genes de interés, es importante contar con RNA de buena calidad y en concentración suficiente para los procesos siguientes.

La extracción de RNA desde hongos es un proceso complicado que muchas veces no alcanza un resultado adecuado. Con el fin de mejorar este proceso y aumentar así la eficiencia de la búsqueda de nuevas LPMOs, se decidió ensayar la metodología propuesta por Schumann para la extracción de RNA desde hongos (73). En este caso específico se quiso probar el efecto que produce realizar la extracción desde un cultivo sólido con membrana, en lugar de un cultivo líquido.

Luego de realizada la extracción, se analizó la integridad y eficiencia de la extracción. Los resultados se resumen en la Tabla 11. De estos resultados se puede apreciar que la extracción desde placas de medio sólido, cubiertas con membrana, es muy ventajosa. En ambos casos se puede observar que la eficiencia alcanzada es de aproximadamente el doble. De la observación del RNA por electroforesis en gel de agarosa, se pudo además determinar que se mantuvo la integridad de estos para los dos métodos (Figura 15).

Tabla 11: Comparación del rendimiento de la extracción de RNA de hongos desde un cultivo en medio líquido y un cultivo en medio sólido con membrana.

Extracción desde medio líquido				
Hongo	Biomasa ^A [g]	Concentración RNA [µg/ml]	Masa RNA [µg]	Rendimiento [µgRNA/mgBiomasa]
<i>F. oxysporum</i>	0.41	722	28.8	0.07
<i>G. trabeum</i>	0.16	1114	22.3	0.14

Extracción desde medio sólido con membrana				
Hongo	Biomasa ^A [g]	Concentración RNA [µg/ml]	Masa RNA [µg]	Rendimiento [µgRNA/mgBiomasa]
<i>F. oxysporum</i>	0.29	1266	38	0.13
<i>G. trabeum</i>	0.2	3804	76.1	0.38

^A Biomasa inicial utilizada para la extracción

Esta mayor eficiencia en la extracción de RNA usando membranas se debe a varios factores. El principal corresponde a la disminución de la contaminación por agarosa que se produce al separar el cultivo de ésta gracias a la membrana. En cultivos en placa sin membrana y para cultivos líquidos inoculados con cubos de cultivos en PDA, la contaminación por agarosa reduce la eficiencia de la extracción, además de formar un pellet insoluble en la solución de RNA final. Otro factor que mejora la eficiencia, es que en este método no es necesario realizar un proceso de separación, entre por ejemplo, el trigo molido en suspensión en el medio, y el hongo en crecimiento. Este proceso muchas veces produce la rotura del micelio, produciendo la liberación de ribonucleasas que causan la degradación del RNA.

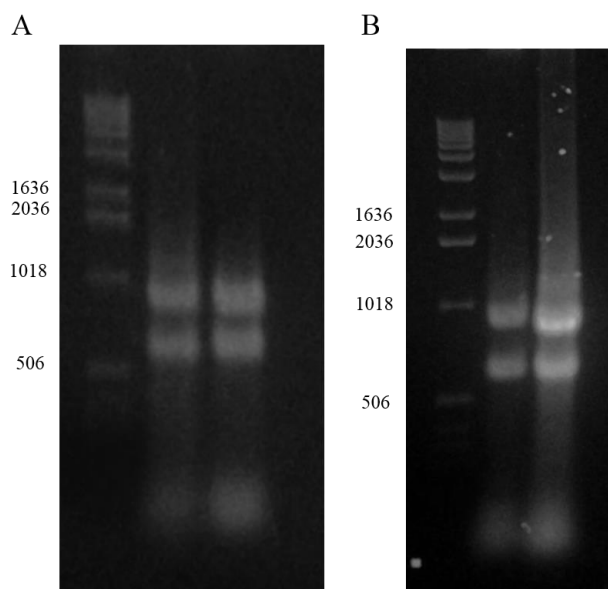


Figura 15: Visualización en electroforesis en gel de agarosa del resultado de la extracción de RNA desde hongos. El gel (A) muestra el resultado de la extracción desde un medio líquido, mientras que (B) desde un medio sólido con membrana. En ambos casos el orden de los carriles de izquierda a derecha es Ladder 1kb DNA de Invitrogen, RNA de *F. oxysporum* y RNA de *G. trabeum*.

III.4 Diseño de partidores

El proceso de clonación de las secuencias fue realizado en dos partes: en un primer lugar se clonaron secuencias parciales de LPMOs usando partidores degenerados, en una segunda etapa, las secuencias de LPMOs seleccionadas fueron clonadas, usando partidores específicos para amplificar su secuencia completa.

III.4.1 Partidores degenerados para la identificación de LPMOs

Para el diseño de los partidores degenerados, se seleccionaron, agruparon y alinearon múltiples secuencias de LPMOs según su semejanza con secuencias parciales ya obtenidas por Alejandro Salinas (72). Estas secuencias parciales (Fof61 y Gaf61) corresponden a fragmentos de LPMOs que se deseaban clonar en su totalidad. Al buscar genes semejantes a éstos se esperaba poder identificar regiones conservadas fuera del fragmento conocido que permitieran diseñar partidores. Al encontrarse en zonas conservadas de estos genes, se esperaba aumentar la probabilidad de amplificar la LPMO de interés desde el cDNA.

Las secuencias seleccionadas se encuentran resumidas en la Tabla 12. Estas secuencias fueron entonces divididas en tres grupos según su semejanza y en base a éstos se pudo realizar alineamientos con el fin de buscar sitios conservados. Los alineamientos obtenidos se presentan de forma resumida en la Figura 16 y los alineamientos completos se encuentran en el Anexo VII.5.1, VII.5.2 y VII.5.3. La posición relativa de los partidores en relación al alineamiento con el que fueron diseñados se muestra en la Figura 17.

Tabla 12: Lista de secuencias nucleotídicas utilizadas en los alineamiento para el diseño de partidores degenerados. Se presentan además los identificadores y las referencias de sus respectivas fuentes.

gi ^A	GenBank	Microorganismo	Referencia
380480975	CCF42122.1	<i>Colletotrichum higginsianum</i>	(56)
346972605	EGY16057.1	<i>Verticillium dahliae</i>	(56)
302405482	XM_003000532.1	<i>Verticillium alfalfae</i>	(56)
302885549	XP_003041666.1	<i>Nectria haematococca</i>	(56)
342883775	EGU84205.1	<i>Fusarium oxysporum</i>	(56)
46119466	XM_384949.1	<i>Fusarium graminearum</i>	(56)
392559558	EIW52742.1	<i>Trametes versicolor</i>	(56)
395329702	EJF62088.1	<i>Dichomitus squalens</i>	(56)
-	Fof61 ^B	<i>Fusarium oxysporum</i>	(72)
-	Gaf61 ^B	<i>Ganoderma applanatum</i>	(72)

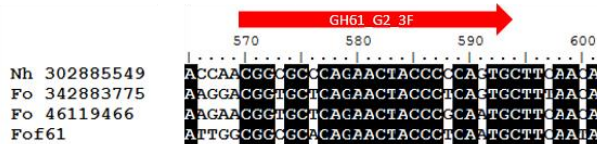
^A Identificador en base de datos NCBI.

^B Identificador de secuencia no registrada en base de datos.

A



B



C



Figura 16: Detalle de los alineamientos utilizados para el diseño de partidores degenerados. Las secuencias utilizadas se encuentran detalladas en la Tabla 12. Los partidores diseñados se marcan con flechas de acuerdo a su emplazamiento en los alineamientos. En el caso del par de partidores GH61_G1_3 se muestran tanto los partidores directo como reverso (A). En el caso del par de partidores GH61_G2 y GH61_G4 sólo se muestra el partidor directo (B y C respectivamente). Las bases diferentes se muestran con fondo blanco mientras que las bases que coinciden en todo el alineamiento se muestran con fondo negro.

Finalmente, este análisis llevó al diseño de tres pares de partidores, dos diseñados para identificar LPMOs de *Fusarium oxysporum* (GH61_G1 y GH61_G2) y uno diseñado para identificar LPMOs en *Gloeophyllum trabeum* (GH61_G4). Éstos se encuentran resumidos en la Tabla 13. Tanto el par de partidores GH61_G2 como el GH61_G4 utilizan como partidor reverso un mismo partidor con un extremo poli-T. Al no encontrar dominios conservados hacia el extremo 3' de estas secuencias se decidió hacer uso del extremo poli-T producido durante la retrotranscripción para amplificar estos genes. Esta particularidad se reflejó en la Figura 17 poniendo los partidores reversos fuera y en el extremo 3' de los alineamientos.

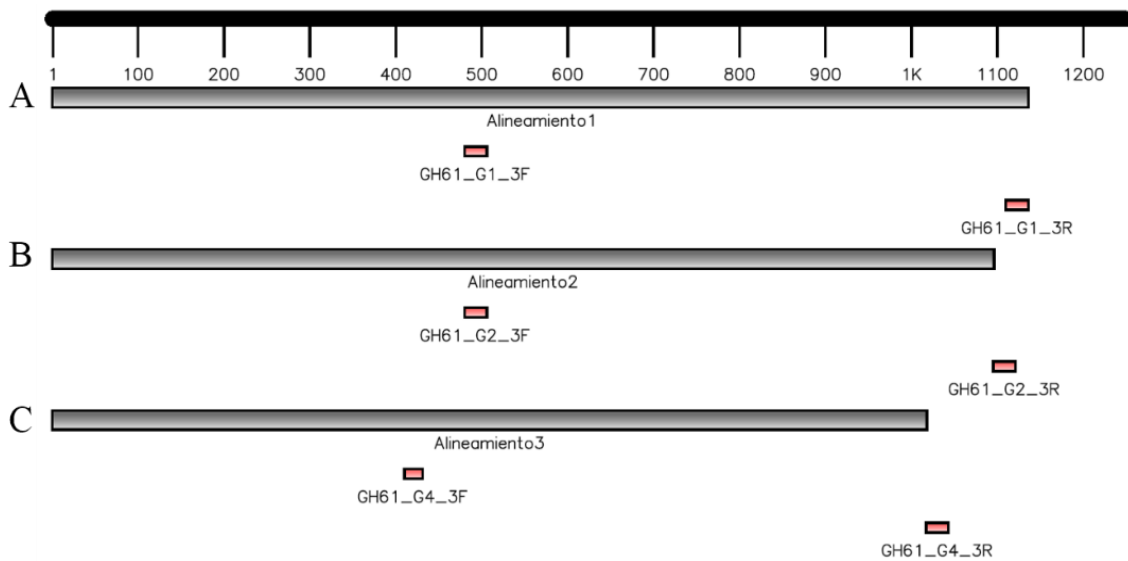


Figura 17: Esquema de la posición aproximada de los partidores en relación a los genes utilizados en los alineamientos. Cada par de partidores se muestra debajo del bloque gris correspondiente al alineamiento desde el que fue diseñado. Se muestran los alineamientos utilizados para diseñar los partidores GH61_G1 (A), GH61_G2 (B) u GH61_G4 (C). Imagen producida usando FancyGene (66).

III.4.2 Partidores específicos para la amplificación de LPMOs

Luego de identificados los genes, se diseñaron partidores complementarios para la amplificación completa de éstos. Se decidió además, con vistas a la posterior clonación de estos genes en un sistema de expresión, incluir en algunos de estos partidores extremos no complementarios conteniendo los sitios de restricción necesarios. Los sitios de restricción añadidos correspondieron al de la enzima EcoR I en el partidor directo, y al de la enzima Not I en el partidor reverso. Éstos fueron añadidos de forma que los genes pudiesen ser clonados ya sea en el plasmidio pPICZ o el pPICZ α del EasySelect™ Pichia Expression Kit de Invitrogen, o también en el vector pPIC 3.5 del Pichia Expression Kit de la misma marca. Asimismo, con el fin de optimizar la expresión de la proteína en estos sistemas, se incluyó en el partidor directo, una modificación del gen en su extremo 5'. Esta secuencia (5'-TATAAATGTCT-3') corresponde a la secuencia de consenso para la traducción que ha sido reportada en levaduras (86). Los mapas de los diferentes vectores de expresión seleccionados pueden ser encontrados en el Anexo VII.6 .

Tabla 13: Lista de partidores degenerados y específicos diseñados, sus secuencias, sus temperaturas de alineamiento (Ta) y el largo esperado del fragmento amplificado.

Nombre partidor	Secuencia* (5'→3')	Ta [°C]	Tamaño amplicón esperado [bp]
GH61_G1_3F	CGGYGCMCAGAACTACCCYCARTGC	67	565
GH61_G1_3R	CTAGTTCATCATGTCTCGCGAGCGTGGC		
GH61_G2_3F	YGCHCAGAACTACCCBCARTGC	60	547
GH61_G2_3R	GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTT		
GH61_G4_3F	TYATCCGSCAYGAGATCATYGC	60	500
GH61_G4_3R	GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTT		
GT61G41f	GCGCGAATTCATAAATGTCTTCCTTCACC	60	782
GT61G41r	GCGGCCGCTTAATTGCCGGTC		
FO61G22f	GCGAATTCTATAAATGTCTTGCAACCTCTC	50	750
FO61G22r	GCGGCCGCTTAGATAGTTCGGAC		
F61_192f	TCAAGATGCCTTCTTTCACCTCCAAG	55	1098
F61_192r	TCTGCTTAGTTCATGAGGTCACGAGC		

*Se utilizó la nomenclatura IUPAC para los nucleótidos degenerados (Anexo VI.4)

III.5 Identificación de genes usando partidores degenerados

III.5.1 Amplificación mediante partidores degenerados

En la primera etapa de identificación, se realizaron PCRs usando los partidores degenerados diseñados. Esto tiene como fin amplificar secuencias similares a las utilizadas para el diseño de los partidores, y poder así, identificar qué enzimas se están expresando.

Se usó para estas amplificaciones cDNA como molde, con el fin de amplificar sólo secuencias correspondientes a genes efectivamente expresados y no solamente presentes en el genoma. Luego de realizados, los productos de PCRs fueron visualizados en una electroforesis en gel de agarosa (Figura 18). Para los pares de partidores GH61_G1, GH61_G2 y GH61_G4, se encontraron bandas de 500 bp, 450 bp y 500 bp respectivamente, lo que coincide aproximadamente con el tamaño esperado (Tabla 13). La amplificación se realizó en el caso de GH61_G1 y GH61_G2 desde cDNA de *Fusarium oxysporum*, y en el caso de GH61_G4 desde cDNA de *Gloeophyllum trabeum*.

Además de las bandas esperadas, se obtuvieron con estos PCRs múltiples bandas o patrones de amplificación inespecíficos, como se puede observar en los casos A y B. Estos resultados son productos del uso de partidores degenerados durante la amplificación. Al tratarse de partidores no específicos, estos se unen a más de un molde dentro del cDNA, produciéndose así una multiplicidad de productos.

Una vez obtenidos, se purificaron y clonaron estos productos de PCR como se describe en Materiales y Métodos en el punto II.2.7.

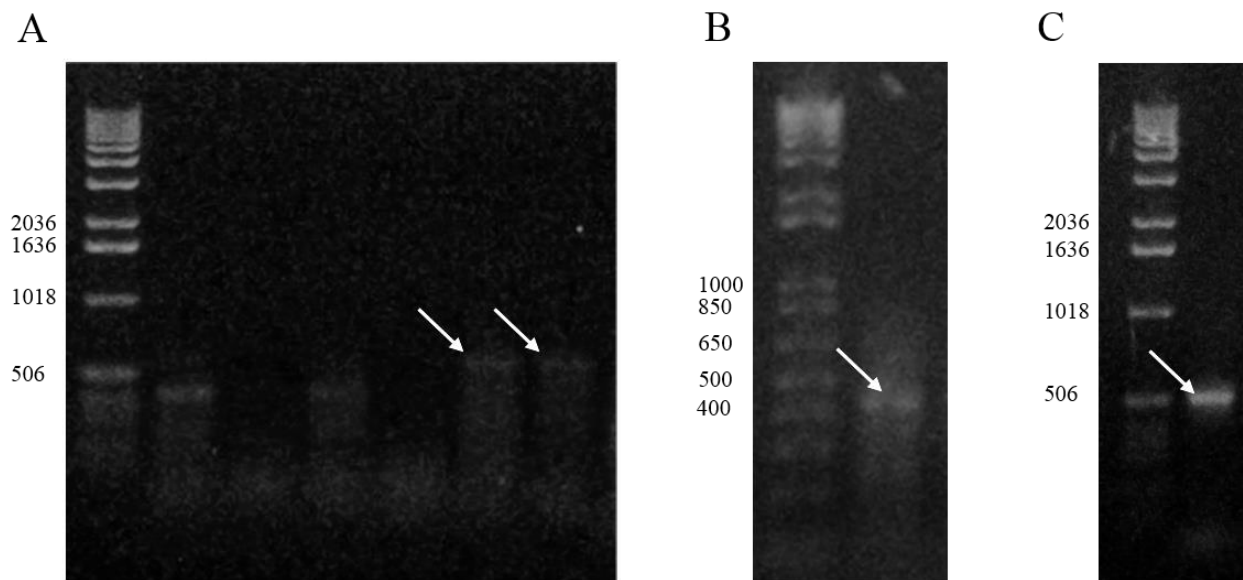


Figura 18: Productos de PCR obtenidos con el uso de partidores degenerados y visualizados en un gel de agarosa. En todas las imágenes el primer carril corresponde al estándar utilizado (1 kb DNA Ladder de Invitrogen en (A) y (C), 1 kb Plus DNA Ladder de Invitrogen en (B)). Los amplicones se encuentran señalados con una flecha y corresponden al uso de los partidores GH61_G1 (A), GH61_G2 (B) y GH61_G4 (C).

III.5.2 Clonamiento en vector pGEM-T Easy y secuenciación

Los fragmentos obtenidos de la amplificación con partidores degenerados fueron clonados en el vector pGEM-T Easy, y posteriormente secuenciados como se describió en Materiales y Métodos.

De la secuenciación se obtuvieron tres clones con fragmentos de interés, los clones FO61_G19, FO61_G22 y GT61_41, correspondientes a los pares de partidores GH61_G1, GH61_G2 y GH61_G4 respectivamente. Las secuencias obtenidas fueron analizadas mediante BioEdit y comparadas con las bases de datos de NCBI (56).

Como era de esperar, las secuencias obtenidas corresponden a fragmentos de LPMOs descritas en los genomas de los microorganismos en cuestión (Tabla 14). Esto se debe a que el diseño de los partidores, como se explicó anteriormente, sólo consideraba la amplificación de fragmentos de la secuencia completa. Se puede además observar que los porcentajes de identidad entre las secuencias obtenidas y las ya presentes en la base de datos son altos, lo que da seguridad de que se trata efectivamente de las LPMOs. Más aún, al analizar las secuencias en búsqueda de dominios conservados, se puede observar que los tres contienen, aunque incompletos, fragmentos del módulo GH61. Las pequeñas diferencias observadas entre las secuencias obtenidas y las descritas, responsables del pequeño porcentaje de diferencia, son productos probablemente de mutaciones

precisas entre las cepas utilizadas o de errores introducidos por la Taq Polimerasa durante la amplificación. Las secuencias completas obtenidas se encuentran en el Anexo VII.6.

Tabla 14: Resultados de la identificación y secuenciación parcial de las LPMO usando partidores degenerados. Las secuencias resultantes fueron buscadas en la base de datos Blastx (55) y se seleccionó el mejor resultado.

Resultado secuenciación		Mejor resultado Blastx			Cobertura (%) ^b	Identidad (%)
Nombre Clon	Microorganismo	Microorganismo	gi ^a	GenBank ^a		
FO61_G19	<i>Fusarium oxysporum</i>	<i>Fusarium oxysporum</i>	342878891	EGU80176.1	47	97
FO61_G22	<i>Fusarium oxysporum</i>	<i>Fusarium oxysporum</i>	342885003	EGU85118.1	23	97
GT61_G41	<i>Gloeophyllum trabeum</i>	<i>Gloeophyllum trabeum</i>	339219008	AEJ35168.1	30	98

^a Identificador en base de datos NCBI

^b Cobertura considerando como 100% la secuencia descrita en Blastx

III.6 Clonación de genes completos usando partidores específicos

III.6.1 Amplificación de genes completos de LPMO

En la segunda etapa de clonación, y luego de tener identificados los genes de interés, se diseñaron partidores específicos para la amplificación completa de estos. Esto permite obtener la secuencia completa de cada gen. Gracias al genoma publicado para estos hongos, se pudo determinar que en los tres casos se trata de genes que no poseen intrones, por lo que éstos pudieron ser amplificados directamente desde DNA genómico.

Después de amplificados, los fragmentos fueron visualizados en una electroforesis en gel de agarosa (Figura 19). En estos geles se pudo observar que efectivamente se obtuvieron amplicones del tamaño esperado. Estos eran de 1098 bp, 750 bp y 780 bp para los genes FoPMO9A, FoPMO9B y GtPMO9A respectivamente (Tabla 13). La amplificación inespecífica, sobretodo en el caso de FoPMO9B se debe probablemente a los partidores utilizados. En este caso, los partidores utilizados contenían pares de bases en sus extremos que no alineaban con el molde. Estos fueron utilizados para incorporar en la secuencia, sitios de restricción y secuencias de consenso necesarias para la clonación en vectores de expresión. Al tener una cola no complementaria, la estrictez de la PCR tuvo que ser disminuida, para así lograr obtener el producto deseado. Esto tuvo como consecuencia la aparición de productos de amplificación no específicos.

Una vez obtenidos, estos productos de PCR fueron purificados y clonados como se describe en Materiales y Métodos en el punto II.2.7.

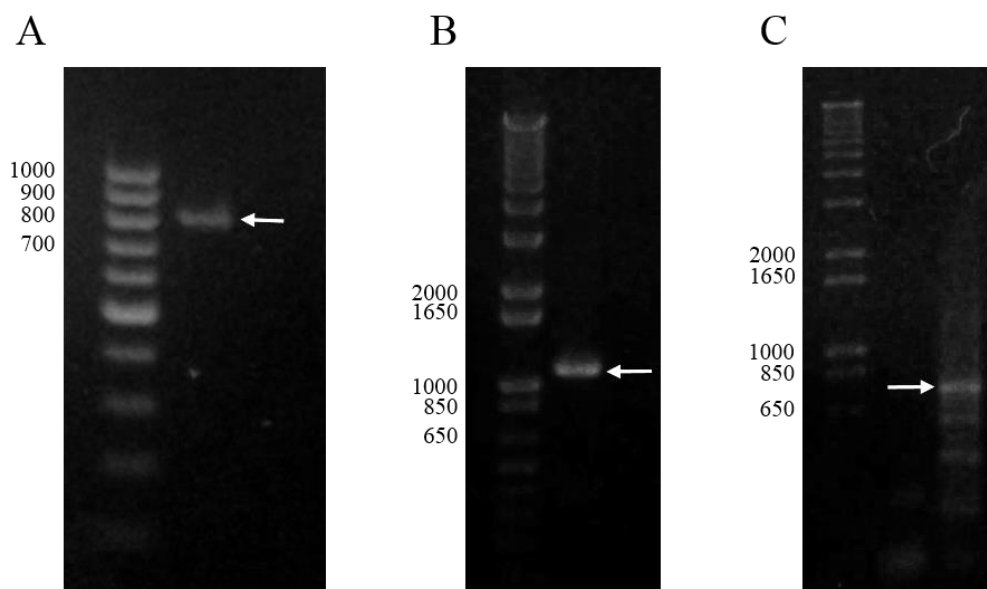


Figura 19: Productos de PCR obtenidos con el uso de partidores específicos y visualizados en un gel de agarosa. En todas las imágenes el primer carril corresponde al estándar utilizado (100 bp DNA Ladder de Thermo Scientific en (A), y 1 kb Plus DNA Ladder de Invitrogen en (B) y (C)). Los amplicones se encuentran señalados con una flecha y corresponden a los genes GtPMO9A, FoPMO9A y FoPMO9B.

III.6.2 Obtención de secuencias de LPMOs completas

De la secuenciación de los clones GtPMO9A, FoPMO9A y FoPMO9B se obtuvieron secuencias correspondientes a genes completos de tres LPMOs, uno correspondiente a *G. trabeum*, y dos correspondientes a *F. oxysporum*. A continuación, estas secuencias fueron comparadas con sus contrapartes anteriormente descritas en los genomas, y que sirvieron de molde para el diseño de los partidores. La información obtenida de este análisis se encuentra resumida en la Tabla 15, las secuencias completas se encuentran en los Anexos VII.9.1 y VII.9.2.

Tabla 15: Comparación de las secuencias obtenidas en este trabajo con las utilizadas como hebra molde para el diseño de partidores. Se entrega el largo de cada proteína además de las diferencias que presentan en relación a las secuencias descritas en el genoma.

Clon	Largo		Comparación con hebra molde			
	bp	AA	Identidad (%)*	Bases diferentes	AA diferentes	Gaps
FoPMO9A	1089	362	98	19	4	0
FoPMO9B	729	242	99	4	1	0
GtPMO9A	759	252	99	3	1	0

*Porcentaje en relación a bases diferentes

Como se puede observar, las diferencias entre ambas secuencias en el caso de FoPMO9B y GtPMO9A son relativamente pequeñas, y sólo una de las mutaciones lleva a un cambio en la secuencia aminoacídica. En el caso del clon FoPMO9A, las diferencias son más marcadas, presentándose 4 aminoácidos diferentes. En ningún caso se observaron mutaciones sin sentido (nonsense mutation). Las diferencias observadas son atribuibles a mutaciones y diferencias entre las cepas de hongo utilizadas en este trabajo con las utilizadas para la secuenciación del genoma. Otra posible explicación es la introducción de errores por parte de la DNA polimerasa, que a pesar de tratarse de una tipo con actividad autocorrectora, no es infalible. El análisis de otros clones

obtenidos a partir de las mismas ampliaciones utilizadas para las LPMOs estudiadas, no reveló diferencias en las secuencias. Esto apoya la idea de que se trataría de mutaciones en la cepa del hongo, aunque no descarta la posibilidad de una mutación introducida en una etapa temprana de la PCR.

Tabla 16: Propiedades fisicoquímicas teóricas de las LPMOs clonadas en este trabajo y la posición de los módulos conservados en estas.

	FoPMO9A	FoPMO9B	GtPMO9A
Propiedades			
pI	5.54	7.12	6.49
PM [kDa]	37.22	25.31	26.65
Posición de módulos conservados			
Péptido señal	Met1 - Ala21	Met1 - Ala19	Met1 - Ala20
GH61	His22 - Ile241	His20 - Ile234	His21 - Ile243
CBM	–	–	–

Se analizó por otra parte, haciendo uso de herramientas bioinformáticas, la presencia de módulos conservados en las secuencias (60). Se calculó además el peso molecular y el punto isoeléctrico (pI) de las proteínas clonadas a través del programa Compute pI/MW tool de ExPasy (62). Se puede observar de los resultados que las tres LPMOs cuentan con un péptido señal en su extremo N-terminal de alrededor de 20 aminoácidos (Tabla 16). Esto es esperable para este tipo de proteínas pues su actividad tiene lugar en el medio extracelular, y sería el rol de este péptido regular la exportación de la proteína hacia éste.

Otra característica esperable es la presencia del módulo GH61 (identificador: pfam03443) en todas las proteínas. Este módulo, que aún conserva el nombre anteriormente dado a estas proteínas, es característico de esta familia y es lo que las define como tal. Finalmente se puede observar que ninguna de las tres LPMOs presenta CBM. Cabe destacar que los cálculos de peso molecular (PM) y de pI corresponden a aproximaciones. En particular, el valor del pI de la proteína en su estado nativo es altamente dependiente de los aminoácidos que se encuentren en su superficie, variando por lo tanto con su estructura.

Es interesante destacar que una búsqueda mediante tBlastn (55) de la secuencia aminoacídica de FoPMO9A, arroja como primer resultado con un Score de 498 y una identidad del 91%, una secuencia proveniente de maíz (*Zea mays*). Esta secuencia (gi: 226529852), obtenida de un análisis de todo el cDNA proveniente de esta planta, sería una proteína en expresión. Al tratarse el maíz de una de las plantas susceptibles de ser atacadas por el hongo *F. oxysporum* (87), es posible elucubrar dos posibles explicaciones para esta semejanza. Es posible por una parte, que la planta utilizada para la secuenciación del cDNA se haya encontrado contaminada con el hongo, produciéndose así la secuenciación de un cDNA que realmente provenía del hongo. O por otra parte, es posible también que la proteína provenga efectivamente del maíz. En este último caso, la similitud que esta presenta con FoPMO9A podría representar una adaptación evolutiva del hongo, particularmente un proceso de evolución convergente. Al asemejar sus LPMOs a las que su organismo hospedero utiliza para los procesos de turgencia, el hongo aumentaría las probabilidades de debilitar la pared celular y favorecer la infección. Otra posibilidad es que esta similitud sea producto de una transferencia horizontal de genes, en las que un gen se transfiere de una especie a otra. De esta

forma, la transferencia del gen de la LPMO del maíz hacia *F. oxysporum*, le habría dado a este último una ventaja evolutiva (85).

III.7 Análisis bioinformático de las secuencias de LPMOs identificadas

III.7.1 Mapa de restricción de las LPMOs

Una vez clonadas, las secuencias de las LPMOs pueden ser utilizadas entre otras cosas para su inclusión en un sistema de expresión. Este trabajo requiere de múltiples pasos que incluyen el corte con enzimas de restricción del gen desde el vector en el que esté clonado. Con el fin de determinar las enzimas de restricción más idóneas para ser utilizadas en este proceso, es importante disponer de una lista con todas las enzimas de restricción susceptibles de cortar la secuencia en lugares no deseados. Es para esto que se analizaron las secuencias de las LPMOs haciendo uso de la herramienta informática NEBcutter V2.0 disponible en línea (57).

Los resultados obtenidos se presentan en forma de un mapa de restricción para cada una de las secuencias estudiadas (Figura 20). Esta búsqueda se realizó considerando todas las enzimas de restricción de tipo II y III disponibles comercialmente. Los resultados sólo incluyen enzimas que cortan la secuencia una única vez. Este análisis permite por ejemplo descartar el uso de algunas de estas enzimas durante el proceso de clonación de la secuencia. Sin embargo, para el uso de una enzima que corte la cadena más de una vez, es necesario realizar un análisis más exhaustivo.

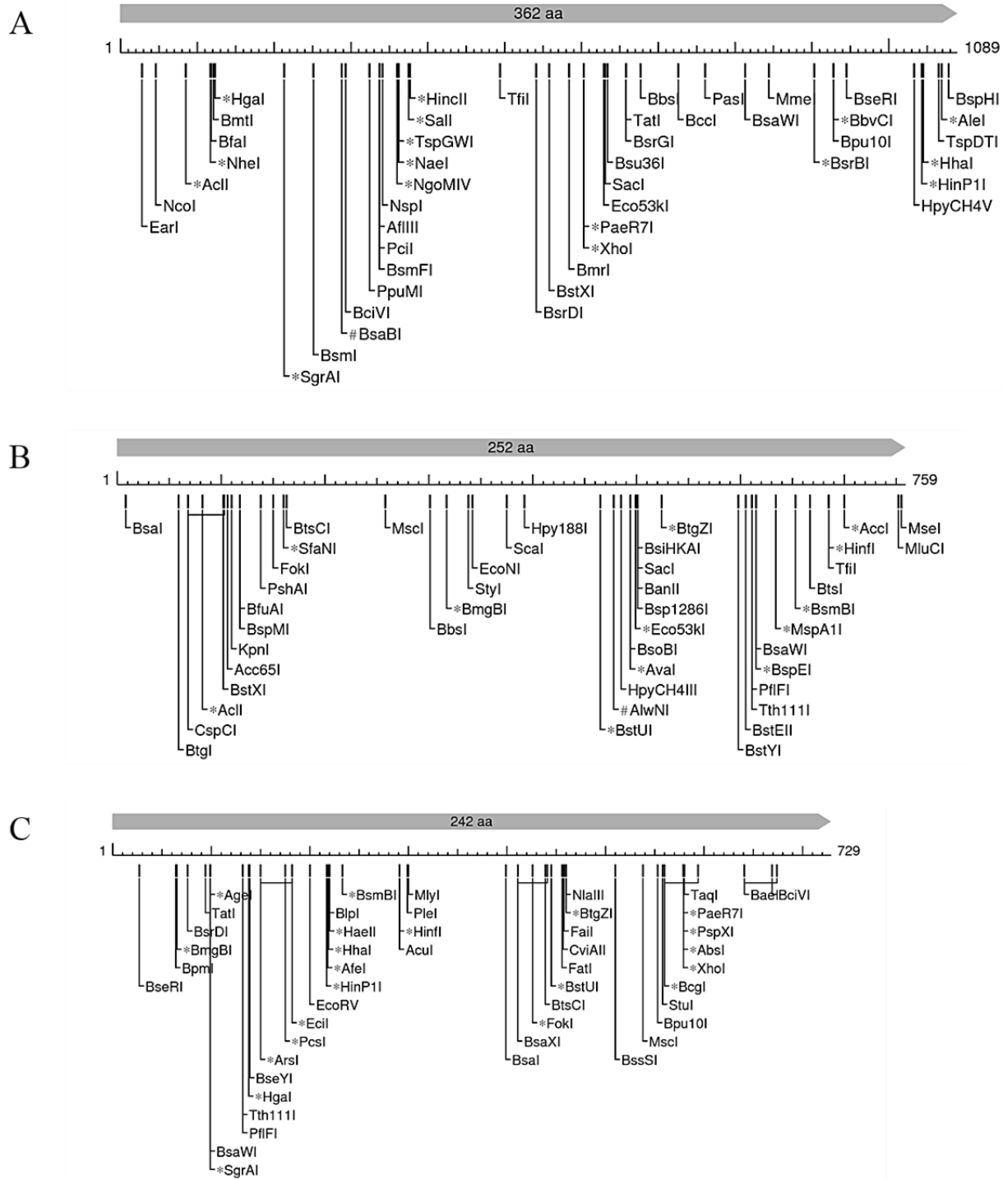


Figura 20: Mapa de restricción de las secuencias nucleotídicas de las LPMOs clonadas. Los mapas corresponden a las secuencias FoPMO9A (A), FoPMO9B (B) y GtPMO9A (C). Se presenta un esquema de la localización aproximada de los sitios de restricción en relación a la proteína (flecha gris) y al gen (eje graduado). Se destacan sitios de corte afectados por metilaciones CpG (*) o por otro tipo de metilaciones (#).

III.7.2 Comparación con otras secuencias de LPMO de la familia AA9

Para determinar la existencia de residuos conservados entre las secuencias encontradas y las ya descritas, se realizó un alineamiento múltiple usando la herramienta Muscle (54). Éste fue construido en base a secuencias de LPMOs ya caracterizadas en la literatura. Se puede de esta forma determinar que las secuencias obtenidas correspondan a LPMOs de la familia AA9, verificándose la existencia de los patrones necesarios.

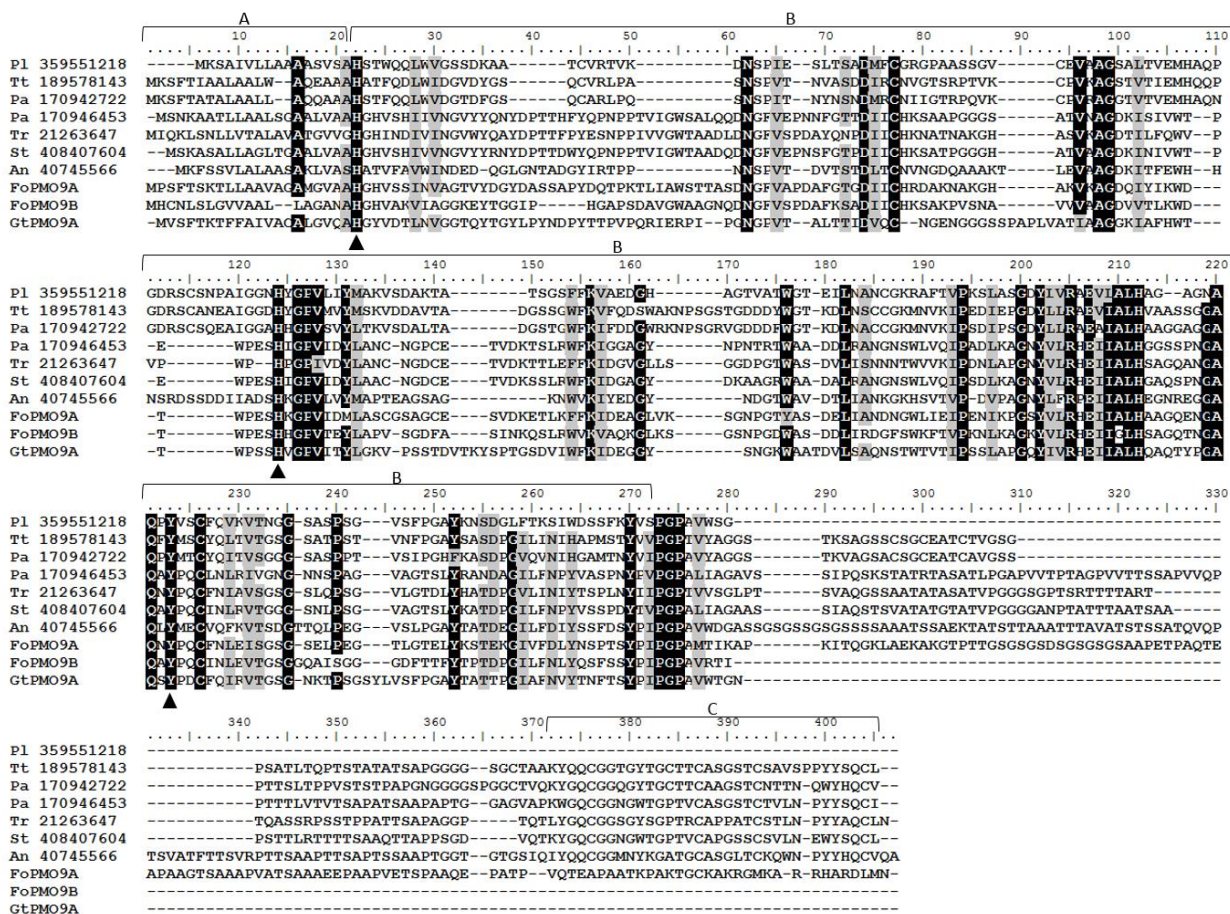


Figura 21: Residuos conservados y módulos presentes en las secuencias FoPMO9A, FoPMO9B y GtPMO9A. Los aminoácidos comunes en las secuencias son presentados con fondo negro y los similares con fondo gris. Los residuos conservados y pertenecientes al sitio de unión del ion cobre se encuentran indicados por un triángulo negro (▲). Los módulos de interés se encuentran marcados con un corchete. Estos corresponden al péptido señal (A), el módulo GH61 (B) y el CBM (C).

Luego de realizada la búsqueda de las LPMOs caracterizadas disponibles en las bases de datos, éstas son alineadas con las secuencias de las LPMOs clonadas en este trabajo. Se puede encontrar un resumen de todas las secuencias obtenidas de la búsqueda en la Tabla 20 más adelante.

Este alineamiento (Figura 21) reveló que las tres nuevas secuencias tienen una estructura similar a las descritas con anterioridad. Se puede observar que varios de los aminoácidos presentes, particularmente en el módulo GH61, son igual o similares entre todas las secuencias. Independientemente de esto, es posible notar una alta variabilidad entre las secuencias, a lo que se podría esperar de un módulo conservado. Esta heterogeneidad es, sin embargo, característica de esta familia de proteínas (36, 41, 88). Del análisis de las secuencias utilizando la

herramienta SignalP 4.1 se pudo observar también que todas las secuencias aquí utilizadas poseen un péptido señal en su extremo N-terminal (67). Éste corresponde probablemente a la señal de exportación hacia el medio extracelular, y no se debiera encontrar en la proteína madura.

Como se observó anteriormente (Tabla 16), de las proteínas secuenciadas, ninguna posee sitio de unión a carbohidrato (CBM), propiedad que sí presentan algunas de las secuencias aquí descritas. Esta falta de CBM no está fuera de lo común, pues como se señaló anteriormente, sólo alrededor de un 20% de las LPMOs se encuentran asociadas a CBM, usualmente de tipo 1 y 18, que en su mayor parte se encuentran presentes en sus extremos carboxilo (27, 39). A pesar de que no se ha demostrado a la fecha que la presencia de CBM en las LPMOs sea un factor crucial para su actividad (32), sí se ha demostrado que su ausencia afecta la habilidad de estas proteínas para unirse al sustrato, así como la especificidad que éstas presentan frente a algunos de estos (31, 89). Estos factores no dejan de ser relevantes, particularmente si se considera el uso industrial de esta enzima. Se puede además observar del alineamiento que la proteína FoPMO9A presenta una mayor extensión, particularmente hacia su extremo C-terminal. A pesar de encontrarse en la posición en donde las otras LPMOs presentan un CBM, el inserto en FoPMO9A no parece serlo. Un análisis de esta secuencia no reveló que poseyera los módulos conservados en estas regiones. Más aún, no se pudo detectar en su secuencia las cisteínas ni las tirosinas clásicas de los CBMs (28).

Finalmente, uno de los puntos más relevantes en este alineamiento, es el probar que los residuos conservados transversalmente en las LPMOs se encuentran también en las secuencias clonadas durante este trabajo. Estos aminoácidos se encuentran marcados en la Figura 21, y corresponden a las dos histidinas y la tirosina responsables de unir el ion cobre a la proteína (32, 36). Este cofactor, en conjunto con un agente reductor, han sido señalados como indispensables para el correcto funcionamiento de la enzima, por lo que la presencia del llamado “soporte” de histidina es un indicio de que podría tratarse de una proteína activa (27, 32, 38).

III.7.3 Determinación de la subfamilia de las LPMOs

Debido a que el interés en estas proteínas es un fenómeno reciente, la familia de las LPMOs sólo recién comienza a descubrirse. Uno de los acercamientos que se ha usado con este fin, es tratar de establecer diferencias entre sus integrantes, para así poder ordenarlas en subfamilias. Un trabajo interesante en este aspecto fue llevado a cabo por Xin Li y su equipo (74). En este trabajo, múltiples LPMOs de *Neurospora crassa* y de otros hongos fueron separadas en subfamilias según el tipo de producto resultante de la acción enzimática. Los investigadores postularon que las LPMOs podían ser separadas en tres subfamilias o tipos. El tipo 1 que corresponde a las enzimas que oxidan la celulosa en el carbono 1, produciendo ácido aldónico. El tipo 2 oxidaría la celulosa en el carbono 4, produciendo una cetoaldosa. Finalmente el tipo 3 correspondería a enzimas menos específicas, capaces de funcionar como una de tipo 1 o 2. Para intentar determinar a qué tipo de proteína corresponden las tres LPMOs descritas en el presente trabajo, las secuencias de éstas fueron analizadas mediante el mismo método postulado por Xin Li. Para esto se utilizaron las mismas secuencias de referencia de este grupo que se encuentran detalladas en la Tabla 17. El resultado de este trabajo corresponde a un alineamiento de las secuencias utilizadas (Figura 23), y a un árbol filogenético en el que se pueden diferenciar los tres grupos de proteínas (Figura 22).

Tabla 17: Secuencias de referencia utilizadas para la determinación de la subfamilia a la que pertenecen las LPMOs clonadas en este trabajo.

gi	GenBank	Nombre	Microorganismo
85078092	NCU08760	NCU08760	<i>Neurospora crassa</i>
28916920	EAA26656.1	NCU03328	<i>Neurospora crassa</i>
85092580	XP_959466.1	NCU02344	<i>Neurospora crassa</i>
85110940	XP_963702.1	NCU00836	<i>Neurospora crassa</i>
85104082	XP_961662.1	NCU01050	<i>Neurospora crassa</i>
85092689	XP_959499.1	NCU02240	<i>Neurospora crassa</i>
85119231	XP_965598.1	NCU02916	<i>Neurospora crassa</i>
85107660	XP_962414.1	NCU07898	<i>Neurospora crassa</i>
85090783	XP_958583.1	NCU05969	<i>Neurospora crassa</i>
85090094	XP_958254.1	NCU07760	<i>Neurospora crassa</i>
262118544	3EJA_A*	TtGH61E	<i>Thielavia terrestris</i>
345531919	3ZUD_A*	TAGH61A	<i>Thermoascus aurantiacus</i>
198443120	2VTC_A*	HjCel61B	<i>Trichoderma reesei</i>

* Identificador en base de datos PDB (59)

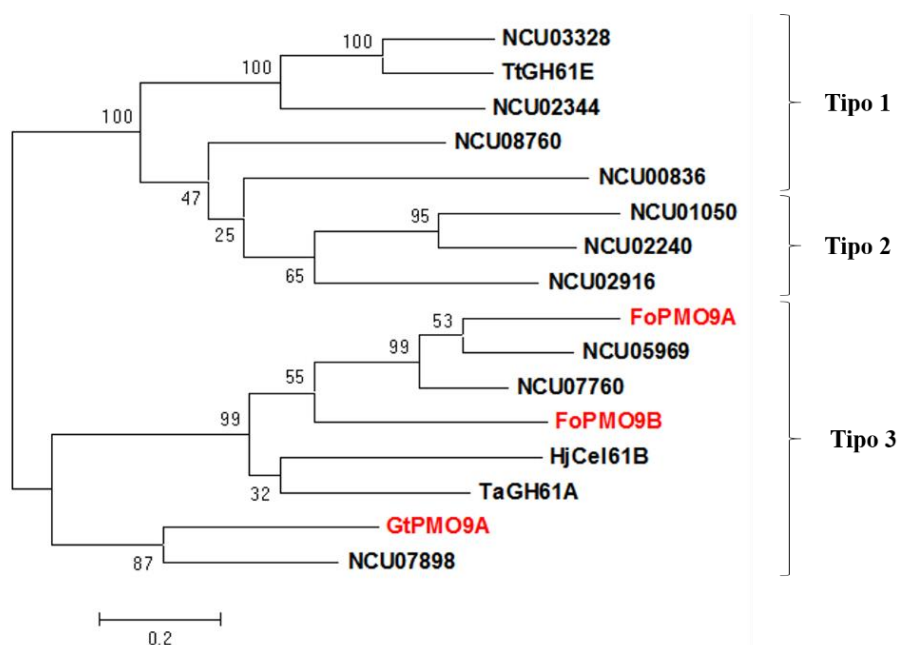


Figura 22: Árbol filogenético obtenido del análisis de las LPMOs seleccionadas mediante el método de Xin Li et al. (74). El árbol se muestra con ramas de largo proporcional a la distancia evolutiva. En corchetes se identifican los 3 grupos a los que estas proteínas pertenecerían. En rojo se muestran las LPMOs clonadas durante este trabajo. El porcentaje de certitud de cada nodo se detalla a un costado de estos.

Del árbol presentado en la Figura 22 se puede observar que se logró obtener una diferencia filogenética entre los distintos tipos de LPMO. Se puede notar también, que aunque existen 3 grupos distintos, el tipo 1 y el tipo 2 parecen estar más estrechamente relacionados entre sí que con el tipo 3. Además, según lo obtenido desde el árbol filogenético, se puede decir que las tres LPMOs clonadas durante este trabajo (FoPMO9A, FoPMO9B y GtPMO9A), corresponderían a LPMOs de tipo 3. De esta observación, se podría entonces esperar que estas proteínas presenten una actividad poco específica, pudiendo oxidar tanto el carbono 1 como el 4 de los residuos de piranosa. De ser así, la actividad de esta proteína podría ser medida a través de la identificación de los productos liberados durante la oxidación. Estos podrían ser ácidos aldónicos en el caso de que se oxide el carbono 1, o se podría intentar medir la aparición de extremos reductores en el caso que se oxide el carbono 4.

Es sin embargo interesante destacar que el resultado presentado en la Figura 22 es ligeramente diferente del obtenido por Xin Li (74). En el material presentado por este equipo, se puede apreciar que existe una clara relación de las secuencias NCU08760 y NCU00836 con las proteínas de tipo 1. Esta relación clara no es visible en el árbol presentado en este trabajo. Se puede notar, que al repetir el experimento añadiendo las tres nuevas LPMOs, tanto NCU08760 como NCU00836 aparecen evolutivamente más cercanas a las de tipo 2. Más aún, el resultado obtenido mediante la metodología de Xin Li parece ser muy sensible a las condiciones en las que el análisis se realiza, presentando resultados diferentes dependiendo de con qué herramienta se realiza el alineamiento, y dependiendo de cómo se calcule el árbol filogenético. Estas diferencias se traducen en, por ejemplo, secuencias que pasan a formar parte de una familia diferente (resultados no mostrados). En el caso específico de las dos secuencias anteriormente señaladas, se decidió mantenerlas en la subfamilia asignada por Xin Li et al., esto debido a que el posterior análisis del alineamiento no arrojó grandes similitudes entre éstas y las del grupo 2. Finalmente, las amplias variaciones que este modelo presenta, hacen de éste una herramienta de análisis preliminar que debe ser considerada con precaución. Una categorización más precisa de las LPMOs podría ser alcanzada a través de la purificación y caracterización de las proteínas *in vivo*.

Del alineamiento de las secuencias utilizadas y su separación según tipo se pueden observar características comunes a cada familia (Figura 23). Es interesante notar que en el caso de la histidina inicial, ésta tiene tendencia a estar flanqueada por un set de aminoácidos conservados y diferentes para cada tipo (Región A, Figura 23). Esto es particularmente visible en el tipo 2 en que el patrón His-Thr-Ile-Phe se mantiene constante en las tres secuencias. Un caso similar ocurre para la segunda histidina que compone el llamado “soporte” del ion cobre (Región B, Figura 23). En este caso se puede además observar la conservación de un patrón Gly-Pro río abajo de la histidina. Por otra parte, se puede apreciar que el patrón Ser-His-Lys-Gly-Pro se encuentra presente en tanto LPMOs de tipo 2 como de tipo 3. Más aún, en algunas secuencias de tipo 3, la lisina se ve reemplazada por una histidina. Al ser ambos aminoácidos de carga positiva, se podría esperar que este reemplazo no cause un cambio estructural mayor en la proteína, manteniéndose cierta similitud entre los dos tipos de LPMOs. Un análisis similar puede ser realizado con la tirosina perteneciente al sitio activo (Región C, Figura 23). Por una parte se puede observar que además de la tirosina, existe una glutamina conservada en todas las secuencias. Por otra parte, se puede notar que en todas las proteínas de tipo 3 existe una prolina conservada después de la tirosina. Esta prolina es reemplazada en las de tipo 1 y 2 por una metionina, leucina o una isoleucina, todos aminoácidos hidrofóbicos. Finalmente, se destaca en la figura mediante un rectángulo amarillo la inserción de una decena de aminoácidos en las LPMOs de tipo 3, particularidad que ya había sido observada en trabajos previos (37), y que forma un loop en la superficie de la proteína.

Estas similitudes y diferencias dentro de las regiones que flanquean los aminoácidos del sitio activo marcan el funcionamiento de la proteína. Como se destaca en el trabajo de Xin Li, aunque la posición del soporte de histidina es bastante conservada, el funcionamiento de éste se ve afectado por los aminoácidos aledaños. Si bien el mecanismo de acción del cobre sería el mismo, la orientación del sitio activo puede determinar que la oxidación se realice en el carbono 1 o 4 en la cadena de celulosa. Más aún, los residuos hidrofóbicos presentes en la superficie activa de la enzima, pueden afectar la unión de esta con el sustrato, uniéndose por ejemplo de forma lineal con las hebras de celulosa o cruzadas entre varias de estas (36). Ésta alta variación de la superficie activa de la enzima se podría explicar por la alta variabilidad estructural de la celulosa encontrada en la naturaleza. De esta forma, la batería de diferentes LPMOs que algunos hongos presentan, serían una adaptación para la utilización de múltiples sustratos.

A pesar de que el mecanismo de acción de las LPMOs, su actividad y su estructura recién comienzan a dilucidarse, análisis de este tipo junto a subcategorizaciones tentativas pueden ayudar a arrojar un poco de luz sobre estas problemáticas. El estudio de los aminoácidos cercanos a los residuos que componen el sitio activo nos puede ayudar a entender la especificidad de una proteína sobre un sustrato en particular, o a diferenciar variaciones en los mecanismos de acción de las proteínas de esta familia.

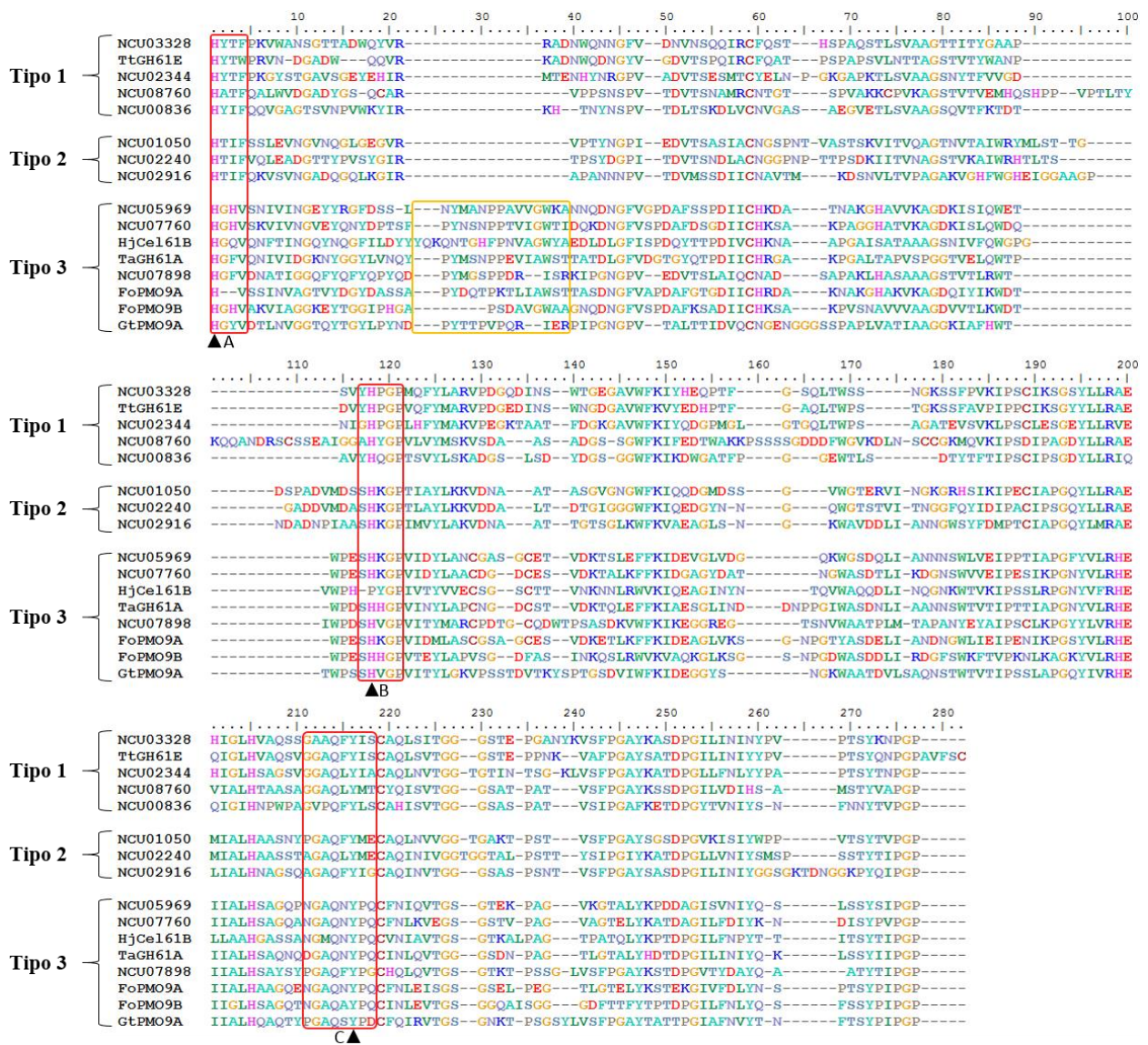


Figura 23: Alineamiento obtenido mediante la herramienta T-Coffee de las secuencias de LPMOs detalladas en la Tabla 17. Las secuencias se encuentran separadas por el tipo a que corresponden según el análisis realizado. Los residuos conservados pertenecientes al sitio activo de la enzima se encuentran marcados por un triángulo negro (▲). Las regiones que rodean estos aminoácidos se encuentran marcadas con un rectángulo rojo: A y B en el caso de las histidinas, C para la prolina. Se marca además mediante un rectángulo amarillo una región presente únicamente en las LPMOs de tipo 3.

III.7.4 Estructura tridimensional de las LPMOs identificadas

La obtención de la estructura terciaria de una proteína de forma precisa es un trabajo complicado que requiere de la realización de estudios de cristalografía, entre otros. No obstante, existen métodos informáticos que permiten obtener predicciones de la estructura de una proteína mediante la comparación de éstas con algunas de las ya dilucidadas. Este análisis, aunque impreciso, puede ser mejorado ya sea mediante un estudio crítico e informado de las estructuras obtenidas, o a través de la optimización computacional del modelo.

En el caso del trabajo aquí presentado, la determinación de la estructura 3D de las proteínas obtenidas consiste en una simple aproximación inicial al problema. El trabajo fue realizado

mediante el uso de servidores educacionales dispuestos para este fin (tales como Robetta, I-Tasser y SwissModel) y debe ser por lo tanto analizada críticamente.

De los diferentes servidores disponibles para la modelación de la estructura de la proteína se obtuvieron múltiples modelos. Luego de seleccionar los que presentaban la secuencia completa de las enzimas, éstos fueron analizados mediante el uso de herramientas computacionales. Estas están destinadas a medir estadísticamente qué tan probable es que la estructura obtenida sea correcta. Las herramientas utilizadas fueron Anolea (51), Gromos (68), Q-Mean (63) y Z-Score (63). Diferentes herramientas basan sus cálculos en diferentes parámetros, por lo que es importante contar con los resultados de varias de ellas, para tener una idea más precisa respecto del modelo. La explicación del funcionamiento a grandes rasgos de estas herramientas puede ser encontrada en el Anexo VII.10.1.

En la Tabla 18 se puede encontrar un resumen de los resultados obtenidos para cada uno de los modelos. Se aprecia de estos, que los modelos con mejor puntuación corresponden en todos los casos al Modelo 1 entregado por el servidor Robetta (65). Estos resultados fueron además confirmados mediante la evaluación con las herramientas Anolea y Gromos. En el Anexo VII.10 se muestran los resultados de éstas dos para los modelos elegidos. Más aún, los modelos de Robetta obtuvieron consistentemente mejores puntajes en todos los modelos producidos, en comparación con los arrojados por I-Tasser (70). Cabe destacar que en los resultados no se muestran los modelos obtenidos a través del uso de SwissModel (69). Esto se debe a que en la totalidad de los casos los modelos arrojados por este servidor no contenían la proteína completa, eliminando en muchos casos la histidina inicial de la proteína, que como ya se dijo, tiene gran importancia en la funcionalidad de la proteína. Este problema no se pudo solucionar a través de cambios en la configuración o en la información de entrada del programa. Es interesante acotar que la calidad de los modelos en estos casos está relacionada con el tiempo que éste toma en ser calculado. Las modelaciones más rápidas fueron las de SwissModel, tomando alrededor de dos horas, seguidas por las de I-Tasser, que demoraban alrededor de 2 días. Por último, las de Robetta, que obtuvieron mejores resultados, se demoraban de dos a tres semanas en completarse.

Para las proteínas de *F. oxysporum*, tanto en el caso de Robetta como en el de I-Tasser, el molde elegido para la modelación fue la cadena A de la LPMO de *Thermoascus aurantiacus* presentada el 2011 por Quinlan et al. (38), registrada con el identificador 3ZUD en la Protein Data Bank (PDB) de RCSB (59). En el caso de *G. trabeum*, los modelos se realizaron usando como molde la cadena A de la LPMO3 del hongo ascomiceto *Neurospora Crassa* presentada el 2012 por Li y Beeson (36), con identificador 4EIS en la PDB. La evaluación de los moldes utilizados mediante las mismas herramientas utilizadas para evaluar las modelaciones entrega una idea de la similitud entre ambos (Tabla 18).

Tabla 18: Resumen de los resultados obtenidos durante la modelación de las LPMOs clonadas. Se detallan los puntajes tanto de QMean como de Z-Score para cada uno de los modelos obtenidos. Se presentan además los resultados obtenidos al evaluar los moldes con las mismas herramientas.

Proteína	Servidor	Modelo	QMean	Z-Score	Proteína	Servidor	Modelo	QMean	Z-Score
<i>FoPMO9A</i>					<i>GtPMO9A</i>				
	Robetta	Modelo 1	0.638	-1.58		Robetta	Modelo 1	0.673	-1.02
		Modelo 2	0.621	-1.77			Modelo 2	0.652	-1.25
		Modelo 3	0.612	-1.88			Modelo 3	0.632	-1.47
		Modelo 4	0.643	-1.51			Modelo 4	0.616	-1.64
		Modelo 5	0.614	-1.86			Modelo 5	0.646	-1.32
	I-Tasser	Modelo 1	0.466	-3.59		I-Tasser	Modelo 1	0.503	-2.85
		Modelo 2	0.463	-3.62			Modelo 2	0.491	-2.90
		Modelo 3	0.441	-3.89			Modelo 3	0.570	-2.13
		Modelo 4	0.457	-3.70			Modelo 4	0.464	-3.27
		Modelo 5	0.437	-3.93			Modelo 5	0.518	-2.70
<i>FoPMO9B</i>					<i>Moldes</i>				
	Robetta	Modelo 1	0.745	-0.28			3ZUD	0.68	-0.888
		Modelo 2	0.719	-0.55			4EIS	0.697	-0.697
		Modelo 3	0.692	-0.84					
		Modelo 4	0.680	-0.97					
		Modelo 5	0.716	-0.60					
	I-Tasser	Modelo 1	0.560	-2.26					
		Modelo 2	0.551	-2.36					
		Modelo 3	0.565	-2.20					
		Modelo 4	0.534	-2.53					
		Modelo 5	0.578	-2.063					

Del alineamiento entre las secuencias utilizadas de molde y las secuencias de las LPMOs clonadas, se puede observar que tanto en el caso de FoPMO9B y como en el de GtPMO9A, los moldes cubren la casi totalidad de la proteína (Tabla 19). La región no cubierta, de alrededor del 9%, corresponde en estos casos al péptido señal que no se encuentra en los moldes.

Tabla 19: Comparación entre las secuencias aminoacídicas de las LPMOs clonadas, y las secuencias utilizadas de molde para su modelación tridimensional.

LPMO	Molde		Score	Cobertura [%]	Identidad [%]
	GenBank	gi*			
FoPMO9A	3ZUD_A	345531919	238	65	52
FoPMO9B	3ZUD_A	345531919	206	91	47
GtPMO9A	4EIS_A	388326915	227	91	53

*Identificador en las bases de datos de la NCBI

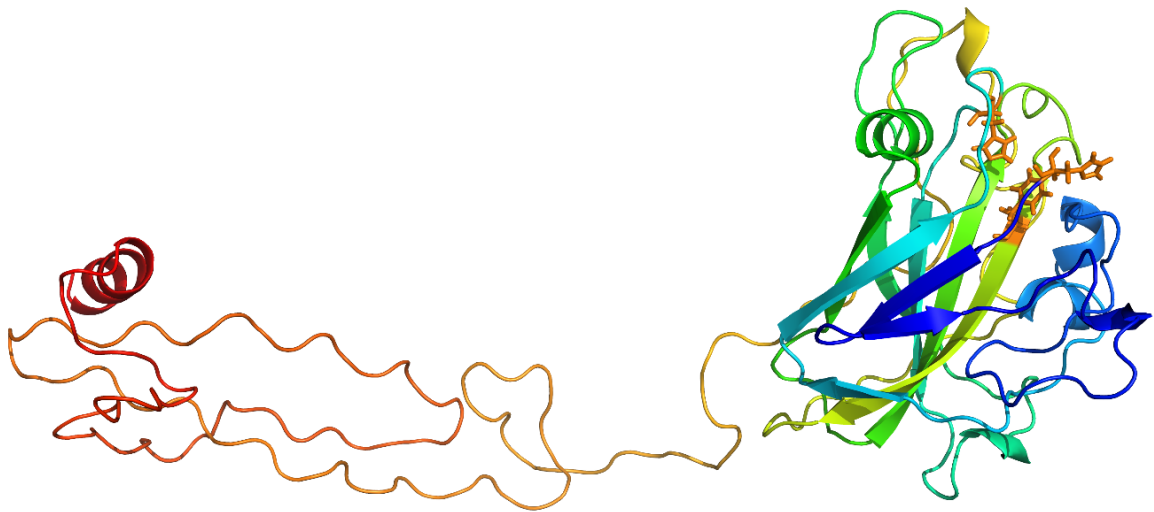


Figura 24: Esquema del modelo tridimensional de la proteína FoPMO9A. Se puede observar a la derecha el módulo GH61 correspondiente a la LPMO. Hacia la izquierda se puede observar la extensión del extremo C-terminal de la proteína. Los residuos pertenecientes al sitio de unión del ion cobre (II) se destacan en anaranjado.

En el caso de FoPMO9A, se puede notar que la cobertura es menor que en los otros dos casos (Del alineamiento entre las secuencias utilizadas de molde y las secuencias de las LPMOs clonadas, se puede observar que tanto en el caso de FoPMO9B y como en el de GtPMO9A, los moldes cubren la casi totalidad de la proteína (**Tabla 19**). **La región no cubierta**, de alrededor del 9%, corresponde en estos casos al péptido señal que no se encuentra en los moldes.

Tabla 19). Esta diferencia se debe a la inserción que esta proteína presenta hacia el extremo N-terminal. Como se describió anteriormente, aunque esta sección corresponde a la localización de los CBMs en otras proteínas, la secuencia de éste no fue reconocida como tal ni comparte las mismas propiedades. Este inserto fue de todas formas modelado junto con el resto de la proteína (Figura 24). Al no poseer un molde en el que basar la modelación, no es de esperar que la estructura de esta región de la proteína corresponda a la estructura real. El análisis de esta sección no reveló el patrón de cisteínas característico de los CBM ni una presencia particular de residuos hidrofóbicos (28). La búsqueda de esta secuencia a través de Blast permitió determinar que esta secuencia efectivamente se encuentra en múltiples otras LPMOs, por lo que se puede suponer que se trata de una región funcional de la proteína, y no de una anomalía introducida en el proceso de clonación.

La descripción y análisis de los modelos obtenidos se hará a continuación de forma simultánea para las tres LPMOs modeladas. Tres imágenes que ilustran los aspectos importantes para cada modelo han sido obtenidas a través del uso de la herramienta PyMol (61). Éstos se presentan en la Figura 27, Figura 28 y Figura 29 para las proteínas FoPMO9A, FoPMO9B y GtPMO9A, respectivamente.

En las tres modelaciones obtenidas se puede apreciar que se conservan los elementos hasta ahora descritos como importantes en este tipo de enzimas, vale decir, la estructura general de la proteína consistente en un “sándwich” de hojas β similar a la presente en las inmunoglobulinas, además de la presencia de entre 2 a 5 α -hélices (17, 27, 32, 35, 36, 90). Todos estos elementos fueron mantenidos durante la modelación y son claramente visibles en la sección A de las imágenes presentadas.

En todos los modelos es visible además la estructura formada por dos histidinas y una tirosina correspondiente al “soporte” de histidina descrito en las LPMOs. Éste está compuesto por los residuos His22-His107-Tyr195, His20-His100-Tyr188 y His21-His106-Tyr195 para las proteínas FoPMO9A, FoPMO9B y GtPMO9A, respectivamente. Esta es la estructura responsable de unir el ion cobre (II) indispensable para la actividad de la proteína y que se encuentra asociado a su actividad oxidativa. La observación en mayor detalle de este sitio de unión (sección B de las imágenes), revela que los residuos que la componen no se encuentran tan estrechamente unidos, como se puede observar en las estructuras determinadas por cristalografía (Figura 5). Esta anomalía puede ser explicada por el hecho de que estas estructuras son aproximadas y no han pasado por un proceso de refinación. Otro factor importante a considerar, es que durante la modelación por los softwares utilizados no se considera la presencia del ion cobre. La ausencia de éste y del efecto que pueda tener sobre los residuos también explicaría la orientación inesperada de los aminoácidos del soporte de histidina. La posición de los aminoácidos puede por lo tanto no ser la que se presenta en la proteína silvestre. De todas formas, la presencia del sitio de unión del ion cobre reafirma la posibilidad de que se trate de una proteína activa una vez clonada y expresada.

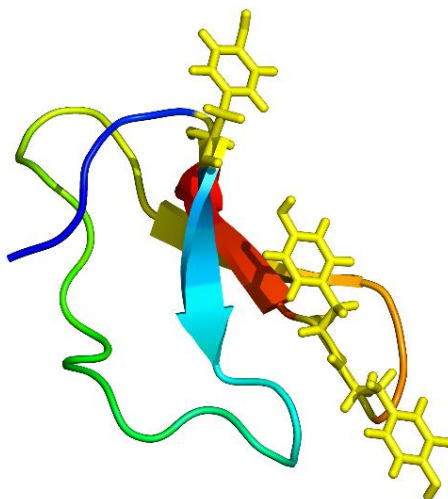


Figura 25: Ilustración esquemática de la estructura del sitio de unión a carbohidratos 1 (CBM1). Se resaltan en amarillos los residuos correspondientes a tirosinas que se encuentran en su superficie de unión. El extremo N-terminal se destaca en color azul, mientras que el C-terminal se destaca en color rojo.

Otro aspecto interesante de notar en la estructura es una superficie plana que se forma en la cara de la proteína que contiene el sitio de unión a cobre. Esta formación se ve además acompañada por la presencia de residuos hidrofóbicos como lo son tirosina, triptófano y fenilalanina (sección D de las imágenes). La presencia de estos anillos aromáticos, en ocasiones paralelos a la superficie de la proteína, ya ha sido descrita en otras LPMOs caracterizadas y debido además a la gran similitud que estos presentan con el sitio de unión del CBM1 (Figura 25), se ha postulado que éste podría ser el punto de asociación de estas proteínas con la celulosa (32, 36, 89, 91).

Finalmente, se realizó un alineamiento entre los modelos obtenidos y las estructuras utilizadas como molde para su modelación (sección C en las imágenes). En éstas se puede observar que las proteínas mantienen una estructura similar a las obtenidas mediante cristalografía. No se observan grandes variaciones en la estructura en relación a su molde, con excepción de la proteína FoPMO9B. En ésta (Sección C, Figura 28) si se puede observar la desaparición de una α -hélice en

el extremo inferior izquierdo de la imagen. Ésta diferencia se debe a una región presente en el molde que no se encontraba en la proteína clonada.

Por otra parte se decidió ilustrar el inserto, que según el análisis de las subfamilias (Encabezado III.7.3), se encuentra presente en las LPMOs de tipo 3. Esta región marcada con un recuadro amarillo en la Figura 23, fue destacada en la superficie del modelo obtenido (Figura 26). Se puede observar que como era de esperar, este inserto se presenta en aproximadamente la misma localización en los tres modelos. El hecho de que este inserto se encuentre hacia uno de los extremos de la superficie plana de la proteína, podría indicar una posible relación con su actividad, mecanismo de acción o especificidad de sustrato. Sin embargo, no hay a la fecha estudios que hayan caracterizado la función de este inserto en particular.

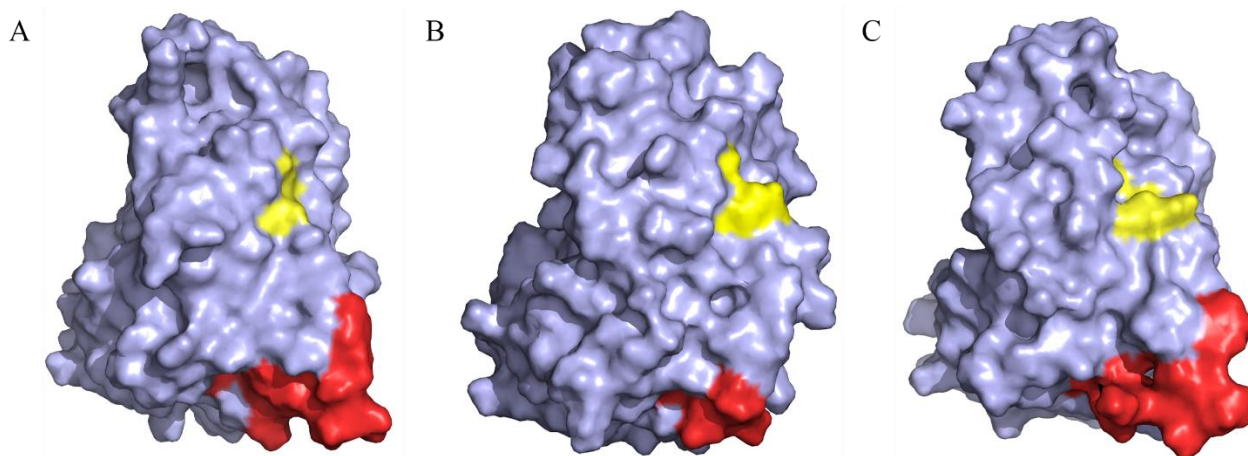


Figura 26: Ilustración esquemática de la superficie de las proteínas en los modelos obtenidos para las LPMOs FoPMO9A (A), FoPMO9B (B) y GtPMO9A (C). Se destacan en amarillo los residuos pertenecientes al sitio de unión de cobre (II), y en rojo el inserto presente en las LPMOs de tipo 3 y señalado con un recuadro amarillo en la Figura 23.

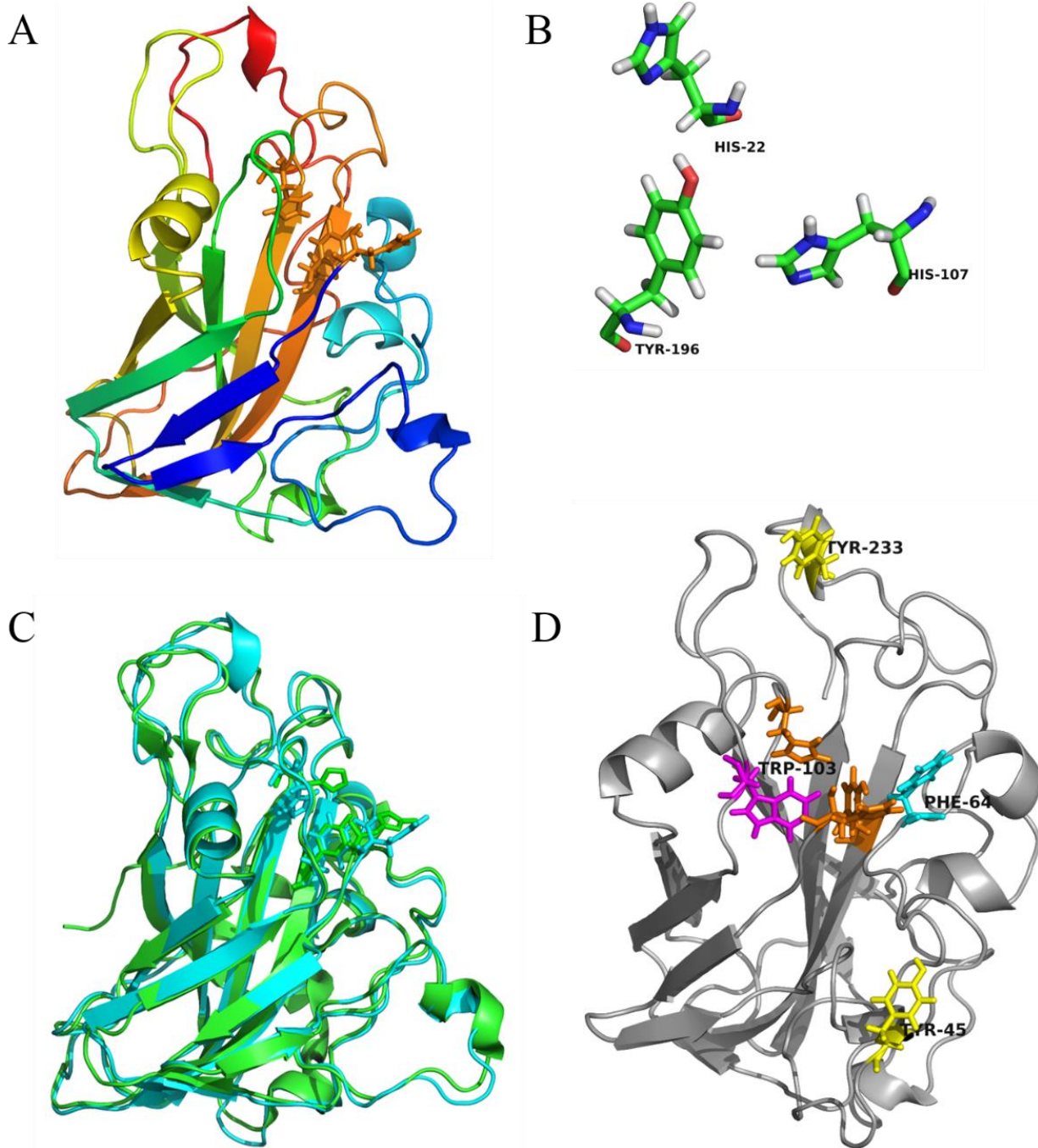


Figura 27: Ilustraciones esquemáticas del modelo tridimensional obtenido para la proteína FoPMO9A. En (A) se muestra la sección correspondiente al módulo GH61, resaltando en anaranjado los residuos del sitio de unión del ion cobre. En este caso no se muestra la inserción en el extremo C-terminal. En (B) se muestra un detalle de la posición de los residuos conservados del sitio de unión. En (C) se presenta una superposición entre el molde (3ZUD_A) utilizado para modelar esta proteína (verde) y el resultado de la modelación (cian). Finalmente en (D) se muestran los residuos hidrofóbicos presentes en la superficie plana de la proteína, posibles responsables de unirse a los carbohidratos: en amarillo la tirosina, en cian la fenilalanina y en magenta el triptófano. Se presentan además las histidinas del sitio de unión del ion cobre en anaranjado.

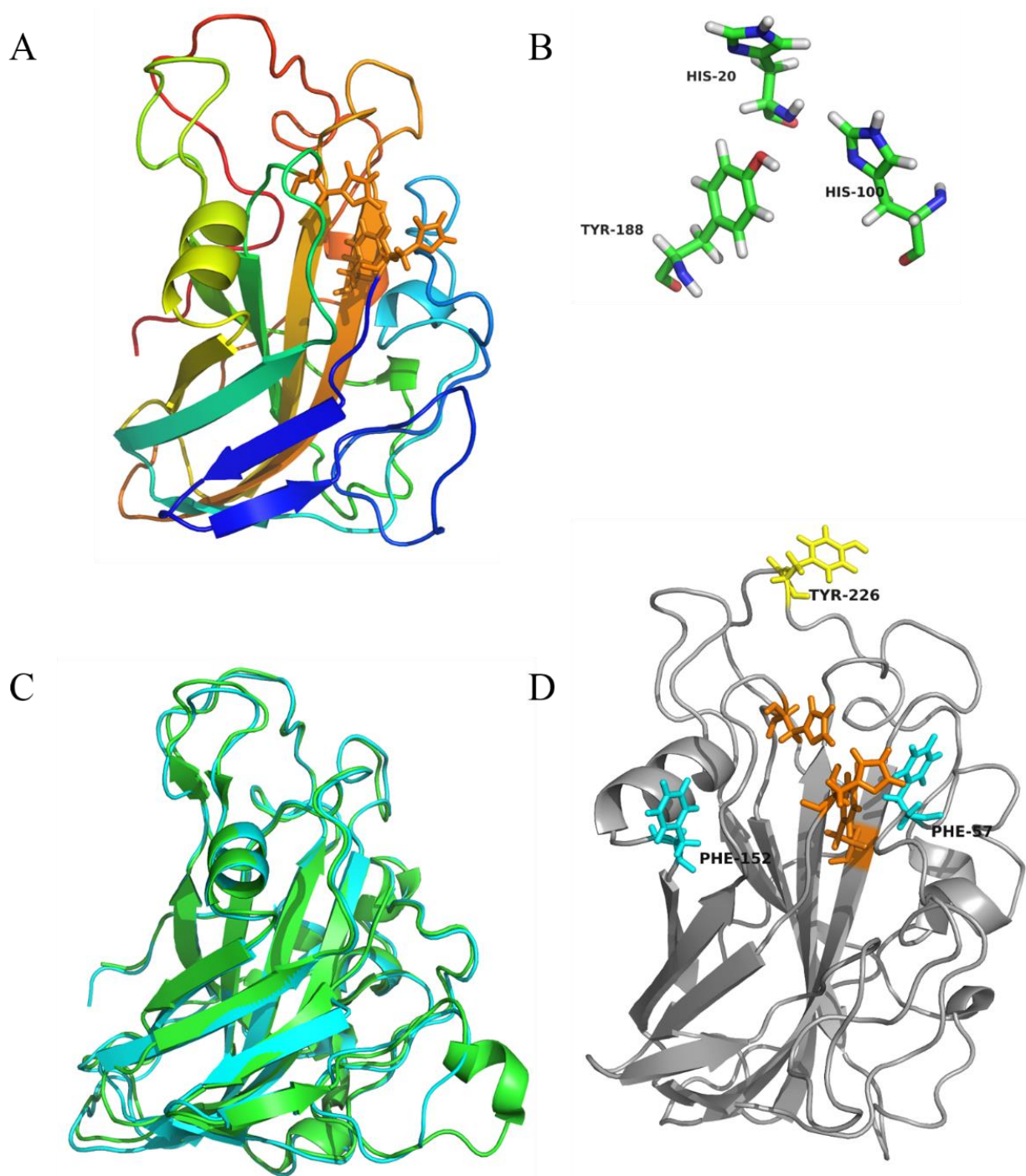
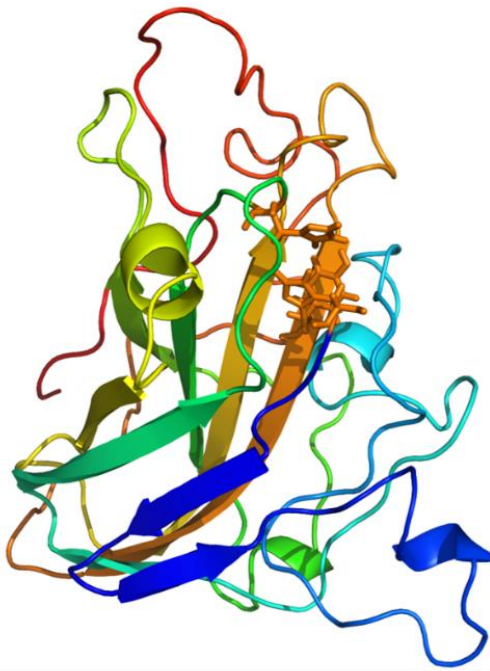
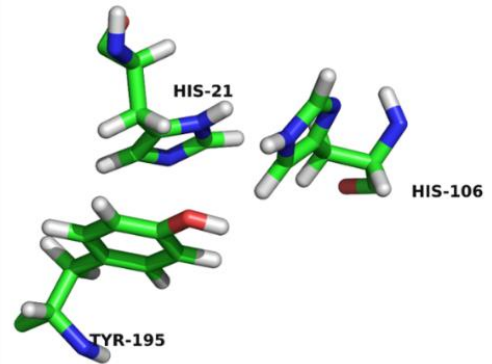


Figura 28: Ilustraciones esquemáticas del modelo tridimensional obtenido para la proteína FoPMO9B. En (A) se muestra la proteína completa resaltando en anaranjado los residuos del sitio de unión del ion cobre. En (B) se muestra un detalle de la posición de los residuos conservados del sitio de unión. En (C) se presenta una superposición entre el molde (3ZUD_A) utilizado para modelar esta proteína (verde) y el resultado de la modelación (cian). Finalmente en (D) se muestran los residuos hidrofóbicos presentes en la superficie plana de la proteína, posibles responsables de unirse a los carbohidratos: en amarillo se presenta la tirosina y en cian la fenilalanina. Se presentan además las histidinas del sitio de unión del ion cobre en anaranjado.

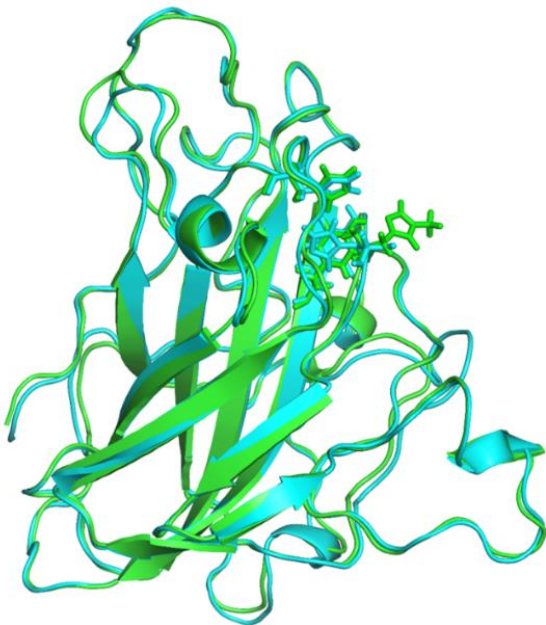
A



B



C



D

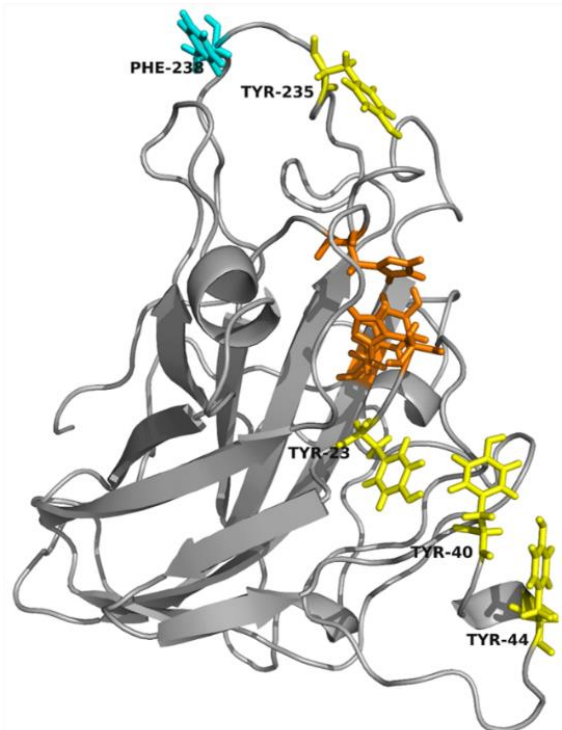


Figura 29: Ilustraciones esquemáticas del modelo tridimensional obtenido para la proteína GtPMO9A. En (A) se muestra la proteína completa resaltando en anaranjado los residuos del sitio de unión del ion cobre. En (B) se muestra un detalle de la posición de los residuos conservados del sitio de unión. En (C) se presenta una superposición entre el molde (3EIS_A) utilizado para modelar esta proteína (verde) y el resultado de la modelación (cian). Finalmente en (D) se muestran los residuos hidrofóbicos presentes en la superficie plana de la proteína, posibles responsables de unirse a los carbohidratos: en amarillo se presenta la tirosina y en cian la fenilalanina. Se presentan además las histidinas del sitio de unión del ion cobre en anaranjado.

III.7.5 Análisis filogenético de las LPMOs identificadas

Para la caracterización funcional de la LPMO es necesario disponer de la proteína purificada. Esto se puede lograr mediante la purificación de la proteína a partir del medio de cultivo o a partir de la proteína producida en forma recombinante. Sin embargo, estos experimentos van más allá del alcance del trabajo aquí expuesto. Es por esto que, con el fin de obtener información sobre la posible actividad y propiedades de estas proteínas, se procedió a compararlas con otras cuyas propiedades ya están caracterizadas. En este caso se procedió a construir un árbol filogenético que permitiese relacionar evolutivamente las LPMOs clonadas con las ya descritas.

Se encontró un total de 11 secuencias de LPMOs descritas y caracterizadas en la bibliografía consultada y en las bases de datos. Adicionalmente, se buscaron las proteínas descritas para hongos de referencia; se encontró una secuencia de LPMO en el genoma del hongo ascomiceto *Trichoderma reesei*, además de diez y dieciocho en los hongos basidiomicetos *Phanerochaete chrysosporium* y *Trametes versicolor*, respectivamente. Finalmente, se buscó en el genoma de *F. oxysporum* y *G. trabeum* otras LPMOs diferentes de las secuenciadas, encontrándose catorce y tres respectivamente. Todas estas secuencias, sus identificadores y sus referencias se encuentran detallados en la Tabla 20.

Luego de identificadas, las secuencias fueron utilizadas para la construcción de un árbol filogenético como se describe en Materiales y Métodos (II.2.8.1). El árbol fue editado y visualizado utilizando el software Mega6 (Figura 30).

Se puede observar en el árbol que las LPMOs tienden a agruparse en sus ramas según se trate de proteínas provenientes de ascomicetos o de basidiomicetos. Esto es visible por ejemplo, en varias ramas que sólo poseen secuencias pertenecientes a *P. chrysosporium*, *T. versicolor* y *G. trabeum*, todos basidiomicetos. Esta segmentación se repite además en las proteínas de hongos ascomicetos, con ramas en las que sólo se tienen proteínas pertenecientes a *F. oxysporum*, *T. reesei* y *T. aurantiacus* entre otros. En la Tabla 21 se puede encontrar una lista detallada que describe cuáles de estos hongos corresponden a basidiomicetos y cuales a ascomicetos. De esta segregación, se puede intuir que estas líneas evolutivas tuvieron una divergencia temprana, generándose así una marcada diferencia entre las secuencias en la actualidad.

Por otra parte, a pesar de esta separación entre phyla, se puede destacar que dentro de cada especie existe una amplia variabilidad entre las secuencias de sus diferentes LPMOs. Esto es observable en el hecho de que para cada especie, las LPMOs se encuentran divididas en múltiples ramas. Esta divergencia evolutiva puede ser asociada a una especialización de cada enzima para un sustrato específico. Esta característica, sumada a la gran cantidad de genes diferentes de LPMOs presentes en algunas especies, apoyaría la teoría de que estas proteínas han evolucionado para entregarle al hongo una mayor variedad de sustratos utilizables .

Tabla 20: Secuencias de LPMOs disponibles en la bibliografía y en las bases de datos. En la primera sección se resumen las secuencias de LPMOs con actividad caracterizada, se entrega además el nombre e identificador en la base de datos del NCBI. En la segunda sección se entregan los identificadores de las secuencias obtenidas desde los genomas de los hongos de interés, tales como los presenta el JGI.

gi^a	GenBank Accession	Nombre	Organismo	Ref.
370987906	BAL43430.1	PcGH61D	<i>Phanerochaete chrysosporium</i>	(37)
158489801	ABW56451.1	TaGH61A	<i>Thermoascus aurantiacus</i>	(32, 34, 38)
189578143	ACE10231	TtGH61B	<i>Thielavia terrestris</i>	(32, 34)
189578146	ACE10234	TtGH61E	<i>Thielavia terrestris</i>	(32, 34)
408407604	G2Q9T3.1	StCel61a	<i>Sporotrichum thermophile</i>	(33)
21263647	O14405.1	Cel61A	<i>Trichoderma reesei</i>	(30)
40745566	EAA64722.1	AN1602.2	<i>Aspergillus nidulans</i>	(92)
74623591	Q96WQ9.1	AkCel61	<i>Aspergillus kawachii</i>	(31, 93)
359551218	AEV53599.1	PIEGL1	<i>Pyrenochaeta lycopersici</i>	(94)
170946453	CAP73254.1	PaGH61A	<i>Podospora anserina</i>	(95)
170942722	CAP68375.1	PaGH61B	<i>Podospora anserina</i>	(95)

Id^b	Organismo	Ref.	Id^b	Organismo	Ref.
1436	<i>Fusarium oxysporum</i>	(49)	45893	<i>Gloeophyllum trabeum</i>	(50)
2877	<i>Fusarium oxysporum</i>	(49)	103762	<i>Gloeophyllum trabeum</i>	(50)
3692	<i>Fusarium oxysporum</i>	(49)	69702	<i>Gloeophyllum trabeum</i>	(50)
4925	<i>Fusarium oxysporum</i>	(49)	120961	<i>Trichoderma reesei</i>	(78)
8789	<i>Fusarium oxysporum</i>	(49)	24196	<i>Trametes versicolor</i>	(77)
8974	<i>Fusarium oxysporum</i>	(49)	34184	<i>Trametes versicolor</i>	(77)
9416	<i>Fusarium oxysporum</i>	(49)	39560	<i>Trametes versicolor</i>	(77)
11576	<i>Fusarium oxysporum</i>	(49)	40255	<i>Trametes versicolor</i>	(77)
12766	<i>Fusarium oxysporum</i>	(49)	36998	<i>Trametes versicolor</i>	(77)
15085	<i>Fusarium oxysporum</i>	(49)	69029	<i>Trametes versicolor</i>	(77)
16904	<i>Fusarium oxysporum</i>	(49)	81252	<i>Trametes versicolor</i>	(77)
1727	<i>Fusarium oxysporum</i>	(49)	114738	<i>Trametes versicolor</i>	(77)
12517	<i>Fusarium oxysporum</i>	(49)	162601	<i>Trametes versicolor</i>	(77)
12425	<i>Fusarium oxysporum</i>	(49)	162729	<i>Trametes versicolor</i>	(77)
41123	<i>Phanerochaete chrysosporium</i>	(76)	43058	<i>Trametes versicolor</i>	(77)
42504	<i>Phanerochaete chrysosporium</i>	(76)	45192	<i>Trametes versicolor</i>	(77)
42616	<i>Phanerochaete chrysosporium</i>	(76)	51004	<i>Trametes versicolor</i>	(77)
4967	<i>Phanerochaete chrysosporium</i>	(76)	53180	<i>Trametes versicolor</i>	(77)
31049	<i>Phanerochaete chrysosporium</i>	(76)	62497	<i>Trametes versicolor</i>	(77)
41563	<i>Phanerochaete chrysosporium</i>	(76)	71674	<i>Trametes versicolor</i>	(77)
41650	<i>Phanerochaete chrysosporium</i>	(76)	143379	<i>Trametes versicolor</i>	(77)
121193	<i>Phanerochaete chrysosporium</i>	(76)	147827	<i>Trametes versicolor</i>	(77)
129325	<i>Phanerochaete chrysosporium</i>	(76)			
10320	<i>Phanerochaete chrysosporium</i>	(76)			

^a Bases de datos del NCBI

^b Bases de datos de genómicas del JGI

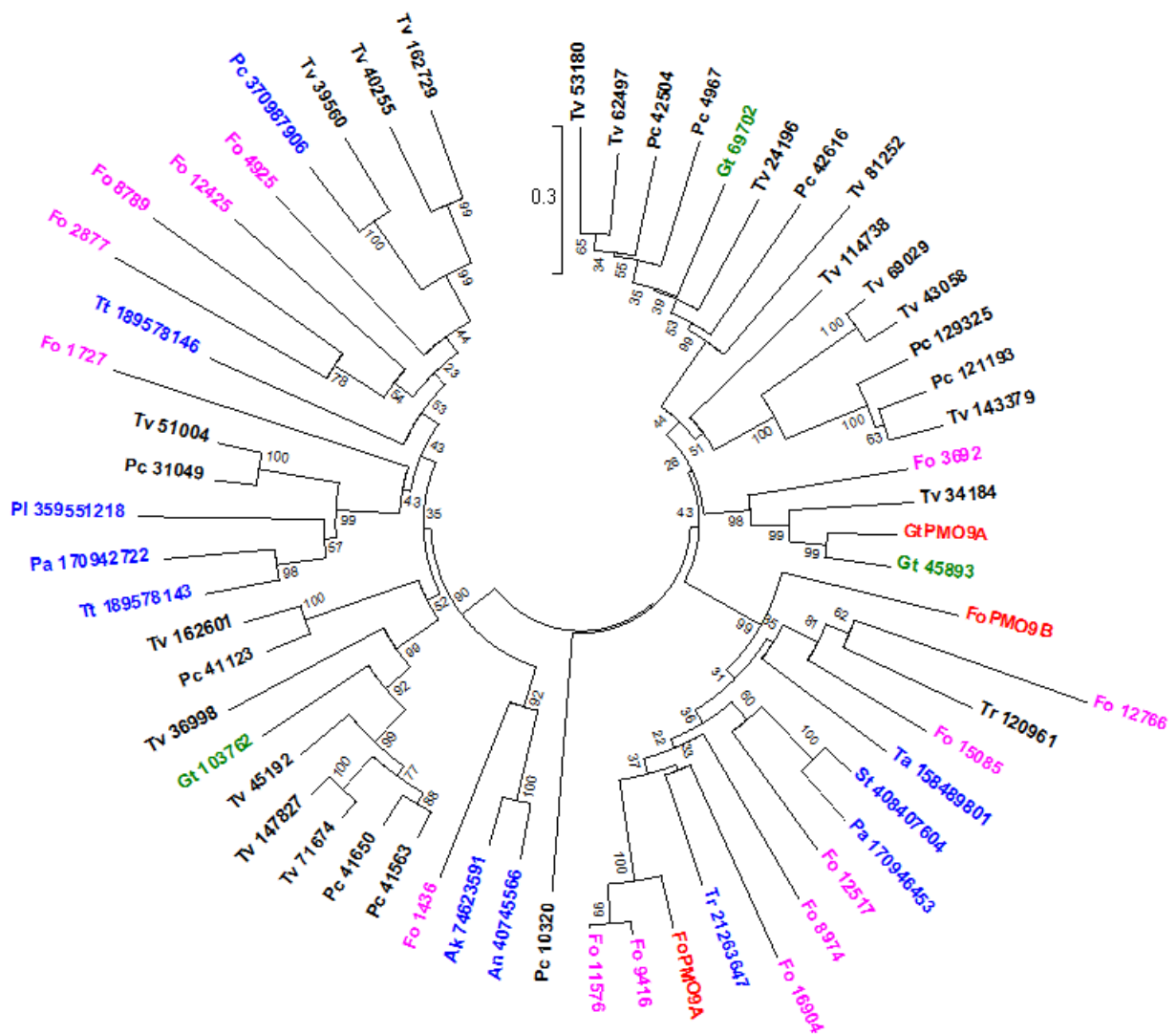


Figura 30: Árbol filogenético de LPMOs caracterizadas y de organismos de referencia. Las secuencias utilizadas corresponden al módulo GH61 de las LPMOs resumidas en la Tabla 20 y se presentan con su identificador e iniciales del organismo. El árbol se muestra con ramas de largo proporcional a la distancia evolutiva. En rojo se presentan las proteínas estudiadas en este trabajo, en azul las proteínas caracterizadas a la fecha, y en fucsia y verde las LPMOs presentes en el genoma de *F. oxysporum* y *G. trabeum* respectivamente. El porcentaje de certeza de cada nodo se detalla a un costado de estos.

A partir del árbol filogenético, en el caso de GtPMO9A, se puede observar una tendencia similar. Por una parte, una de las 3 LPMOs (Gt 45893) de este organismo es sumamente cercana evolutivamente a esta secuencia, compartiendo un 76% de la secuencia con ésta. Esto sugiere que existe cierto grado de especialización por parte del *G. trabeum* hacia cierto tipo de sustratos. Sin embargo, por otra parte se puede observar que las otras dos LPMOs de este hongo Gt 69702, y más claramente Gt 103762, son considerablemente distantes de las primeras dos.

La adición de proteínas de actividad caracterizada al árbol filogenético tuvo como objetivo el poder deducir, según su cercanía evolutiva, qué tipo de propiedades presentarían las LPMOs secuenciadas. Sin embargo, se observa que sólo en el caso de FoPMO9A existe una proteína

aceptablemente cercana en el árbol (Tr 21263647) como para poder suponer una posible similitud en sus características.

Bajo este punto de vista, y en base a la caracterización llevada a cabo por Karlsson el año 2001 (30), esta proteína debería presentar características similares a las de la Cel61A. En este trabajo, Karlsson encuentra que esta enzima (Tr 21263647), sí presenta actividad sobre algunos sustratos tales como CMC e hidroxietilcelulosa (HEC). Es interesante resaltar que en los ensayos realizados por este autor, la enzima sí logra reducir la viscosidad en estos sustratos y disminuir su tamaño, sin embargo, esta actividad no se ve corroborada por la aparición de extremos reductores en solución. Esta diferencia podría ser explicada por los mecanismos de acción de esta proteína propuestos una década más tarde (95). En éstos se postula que si bien el mecanismo oxidativo de la enzima permite cortar la cadena glicosídica, este corte no libera necesariamente extremos reductores. Esto se traduciría entonces en el efecto observado. En base a los resultados de este equipo y la cercanía filogenética que estas secuencias presentan, se podría esperar que FoPMO9A presente un mecanismo de tipo 1, vale decir que la oxidación se lleve a cabo en el carbono 1, produciéndose un ácido aldónico en lugar de un extremo reductor.

En general en los ensayos realizados por Karlsson, la actividad de la Cel61A fue varios órdenes de magnitud inferior a los medidos por el control (la celobiohidrolasa Cel7A). Esta baja actividad puede no obstante, explicarse debido a la ausencia de cofactores indispensables para la actividad de la Cel61A. Sólo recientemente se ha descubierto que la presencia de ion cobre (II) en solución, además de la presencia de un agente reductor, permiten activar la enzima. Tal vez la repetición de experimentos similares a los de Karlsson con la adición de estos cofactores, entregaría resultados más esclarecedores respecto de la actividad de estas enzimas.

El hecho de que no se encuentren proteínas caracterizadas cercanas a las aquí clonadas en el árbol filogenético, no nos permite estimar qué características tendrán éstas enzimas una vez expresadas o purificadas. No obstante, este resultado no resulta desalentador, pues entrega la posibilidad de caracterizar un nuevo tipo de proteína con, tal vez, propiedades diferentes a las ya estudiadas.

Tabla 21: Phylum de cada microorganismo utilizado para la elaboración del árbol filogenético presentado en la Figura 30.

Phylum	
Ascomiceto	Basidiomiceto
<i>Fusarium oxysporum</i>	<i>Phanerochaete chrysosporium</i>
<i>Thermoascus aurantiacus</i>	<i>Gloeophyllum trabeum</i>
<i>Thielavia terrestris</i>	<i>Trametes versicolor</i>
<i>Sporotrichum thermophile</i>	
<i>Trichoderma reesei</i>	
<i>Aspergillus nidulans</i>	
<i>Aspergillus kawachii</i>	
<i>Pyrenochaeta lycopersici</i>	
<i>Podospora anserina</i>	

IV Conclusiones

Se concluye que el trabajo realizado cumplió con los objetivos planteados inicialmente.

Se logró efectivamente determinar medios de cultivo inductores de la producción de celulasas, y las condiciones de preparación idóneas para este fin. En el caso de *Gloeophyllum trabeum*, el mejor medio inductor correspondió a una variación del medio presentado por Chouiter (83), en el que la concentración de magnesio se reduce a un 10% del original. En el caso de *Fusarium oxysporum*, el medio más idóneo fue alcanzado a través del ajuste de pH del medio a 6 previo a la esterilización. Ambos medios presentaron los mejores resultados al ser suplementados con un 1% de paja de trigo. Se determinó además que el porcentaje de paja de trigo y la esterilización de sales por separado son factores que afectan la capacidad inductora de los medios utilizados, debiéndose tomar en cuenta a la hora de buscar nuevos medios. Los hongos cultivados en este medio presentaron mayor producción de celulasas en el medio extracelular, asegurando de esta forma una mayor probabilidad de identificar LPMOs desde su RNA. Luego de identificados, estos medios también pudieron ser utilizados en la preparación de medios de cultivo sólidos.

Por otra parte, a través de la amplificación, clonación y secuenciación de los genes del rRNA 18S se pudo identificar con relativa exactitud el género de los hongos con los que se trabajó. A pesar de no tratarse de una herramienta particularmente precisa para diferenciar especies dentro de un mismo género, el posterior análisis de las proteínas clonadas permitió determinar con bastante seguridad que los hongos correspondían a *Fusarium oxysporum* y *Gloeophyllum trabeum*.

Durante este trabajo también se pudo determinar una metodología más eficiente para la extracción de RNA desde hongos. El uso de membranas entre el medio sólido y el micelio del hongo permitieron realizar una mejor extracción de RNA. Al incorporar esta modificación en la metodología de extracción, se pudo disminuir el efecto negativo de dos factores: el primero es la contaminación por agarosa, que disminuye la eficiencia de la extracción y produce un pellet insoluble en la solución final. El segundo es la degradación del RNA que se produce durante la separación del micelio del trigo presente en el medio de cultivo líquido. Ésta se produce a causa de la liberación de ribonucleasas y disminuye la integridad y eficiencia del método. El control de ambos factores permitió finalmente duplicar el rendimiento en la extracción.

Haciendo uso de las bases de datos disponibles y de herramientas de alineamiento, se pudo identificar secuencias conservadas presentes en las secuencias de múltiples LPMOs. Tales bloques conservados se encuentran en los módulos GH61 de las enzimas, y tienen mayor similitud dentro de un mismo género de hongo. A partir de estos se diseñaron partidores degenerados con el fin de detectar transcritos de LPMO, en la forma de cDNA, producto de la expresión de estos genes en los hongos. La amplificación con el uso de estos partidores fue lograda a través del estudio de las variables que afectan el alineamiento de los partidores al molde, y a través del re-diseño de los programas de PCR. La aplicación de estas modificaciones, en conjunto con la técnica de extracción de RNA, permitieron la amplificación de fragmentos de genes correspondientes a tres LPMOs: dos de *F. oxysporum* y una de *G. trabeum*.

Habiendo sido identificados, y con la información obtenida desde los genomas de los hongos, se diseñaron partidores para la amplificación completa de los genes. Estos partidores fueron diseñados además con extremos no complementarios, consistentes en las secuencias necesarias para su posterior clonación en un vector de expresión. A pesar de las complicaciones que la adición de estos extremos acarrea, la amplificación pudo ser llevada a cabo a través del uso de técnicas de

Touchdown y Hot Start PCR. De esta amplificación se obtuvo entonces las secuencias completas correspondientes a las tres LPMOs anteriormente identificadas.

El último objetivo se logró completar a través del uso de una variedad de herramientas bioinformáticas. Éstas se usaron para analizar en profundidad las secuencias obtenidas, obteniéndose así estimaciones acerca de la funcionalidad de estas proteínas. Se obtuvo el mapa de restricción de los genes clonados, permitiendo así evaluar el tipo de enzimas de restricción susceptibles de ser utilizadas en el proceso de clonación y expresión. Se evaluó el tipo de familia al que las proteínas pertenecen, determinándose que son todas de la subfamilia 3. A partir de esto se puede esperar que los productos de la actividad oxidativa de estas enzimas sean ácido aldónico o cetoaldosas. Del alineamiento de las secuencias aminoacídicas con las de LPMOs conocidas, se pudo observar que las proteínas clonadas cuentan con los residuos hasta ahora descritos como importantes para la actividad de estas enzimas, y que forman el “soporte” de histidina. Más aún, el modelamiento de la estructura tridimensional aproximada de las LPMOs, permitió determinar que el emplazamiento de estos residuos seguiría el patrón observado hasta ahora, formándose las estructuras conservadas en esta familia de enzimas, como por ejemplo el sitio de unión de ion cobre. Finalmente, del análisis filogenético de las proteínas en relación a las ya descritas, se pudo observar que FoPMO9B y GtPMO9A no presentan gran similitud con las anteriormente caracterizadas. Esto permite suponer que se trata de proteínas con propiedades diferentes a las LPMOs descritas con anterioridad, lo que ratifica la importancia de su estudio. FoPMO9A por su parte si es filogenéticamente cercana a una proteína ya descrita, por lo que se podría esperar que su mecanismo de acción fuese de tipo 1, vale decir, que esta oxide el carbono 1 de la celulosa.

Finalmente, del trabajo realizado se puede concluir que las secuencias obtenidas corresponden a enzimas LPMOs probablemente activas, y susceptibles de ser utilizadas en la mejora de un proceso de despolimerización de celulosa. El siguiente paso lógico en esta línea de trabajo consistiría en la transformación de estos genes en un sistema de expresión, para su posterior caracterización funcional, y el estudio del potencial rol auxiliar de estas enzimas sobre la degradación de lignocelulosa.

V Recomendaciones

Las siguientes son recomendaciones entregadas para el desarrollo de futuros trabajos o para la continuación de este:

- Desarrollar un ensayo de actividad específico para la detección de la actividad de las LPMOs aquí identificadas. Esto se podría lograr a través del análisis de las concentraciones ideales de ion cobre (II) o de agente reductor en la solución de reacción. Estos ajustes, sumados a ensayos similares a los realizados por Karlsson et al. o Westereng et al. permitirían evidenciar de forma más clara la actividad de estas enzimas (30, 37, 96).
- Determinar medios de cultivos que induzcan específicamente la producción de LPMOs. Haciendo uso de un ensayo de actividad que mida la actividad de éstas, se pueden buscar condiciones de cultivo que induzcan la producción de LPMOs y no de celulasas en general.
- Buscar una forma más precisa de identificar molecularmente los hongos utilizados, tal vez a través del uso de otro marcador molecular diferente al 18S.
- Establecer una estrategia de purificación de la proteína desde el medio extracelular.
- Optimizar los modelos de la estructura tridimensional de las proteínas obtenidas. Este estudio permitiría obtener información más precisa de la estructura y su relación con la actividad.
- Analizar en profundidad la actividad de estas enzimas sobre material lignocelulósico, los productos liberados por su acción, y la capacidad sinérgica que presente con celulasas tradicionales.

VI Bibliografía

1. **British Petroleum.** 2014. BP Statistical Review of World Energy June 2014.
2. **Kozulj R.** 2005. Crisis de la industria del gas natural en Argentina. Naciones Unidas - CEPAL, Santiago de Chile.
3. **Galbe M, Zacchi G.** 2002. A review of the production of ethanol from softwood. *Appl. Microbiol. Biotechnol.* **59**:618–28.
4. **Heritage W.** 2007. Climate Change and World Heritage.
5. **Kerr R.** 2007. Global Warming Is Changing the World. *Science* (80-.). **316**:188–190.
6. **Hoekman SK.** 2009. Biofuels in the U.S. – Challenges and Opportunities. *Renew. Energy* **34**:14–22.
7. **RFA.** 2013. World Fuel Ethanol Production.
8. **Fargione J, Hill J, Tilman D, et al.** 2008. Land clearing and the biofuel carbon debt. *Science* **319**:1235–8.
9. **Searchinger T, Heimlich R, Houghton R a, et al.** 2008. Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science* **319**:1238–40.
10. **Lapola DM, Schaldach R, Alcamo J, et al.** 2010. Indirect land-use changes can overcome carbon savings from biofuels in Brazil. *Proc. Natl. Acad. Sci. U. S. A.* **107**:3388–93.
11. **Rosegrant M.** 2008. Biofuels and Grain Prices : Impacts and Policy Responses.
12. **Mueller S a., Anderson JE, Wallington TJ.** 2011. Impact of biofuel production and other supply and demand factors on food price increases in 2008. *Biomass and Bioenergy* **35**:1623–1632.
13. **Dashtban M, Schraft H, Qin W.** 2009. Fungal Bioconversion of Lignocellulosic Residues ; Opportunities & Per- spectives. *Int. J. Biol. Sci.* **5**:578–595.
14. **Lee R a., Lavoie J-M.** 2013. From first- to third-generation biofuels: Challenges of producing a commodity from a biomass of increasing complexity. *Anim. Front.* **3**:6–11.
15. **Kim TH, Kim TH.** 2014. Overview of technical barriers and implementation of cellulosic ethanol in the U.S. *Energy* **66**:13–19.
16. **Yang B, Dai Z, Ding S-Y, et al.** 2011. Enzymatic hydrolysis of cellulosic biomass. *Biofuels* **2**:421–450.
17. **Dimarogona M, Topakas E, Christakopoulos P.** 2013. Recalcitrant polysaccharide degradation by novel oxidative biocatalysts. *Appl. Microbiol. Biotechnol.* **97**:8455–65.
18. **Yang B, Wyman CE.** 2008. Pretreatment: the key to unlocking low-cost cellulosic ethanol. *Biofuels, Bioprod. Biorefining* **2**:26–40.

19. **Kenneth AJ, Houtman CJ, Ryan ZC, et al.** 2001. Pathways for Extracellular Fenton Chemistry in the Brown Rot Basidiomycete *Gloeophyllum trabeum*. *Appl. Environ. Microbiol.* **67**:2705–2711.
20. **Quinlan RJ, Sweeney MD, Lo Leggio L, et al.** 2011. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. U. S. A.* **108**:15079–84.
21. **Phillips CM, Beeson WT, Cate JH, et al.** 2011. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. *ACS Chem. Biol.* **6**:1399–406.
22. **Ratanakhanokchai K, Waeonukul R, Pason P, et al.** 2013. *Paenibacillus curdlanolyticus* Strain B-6 Multienzyme Complex: A Novel System for Biomass Utilization, p. 369–394. *In* Matovic, MD (ed.), *Biomass Now – Cultivation and Utilization*. InTech.
23. **Saloheimo M, Paloheimo M, Hakola S, et al.** 2002. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *Eur. J. Biochem.* **269**:4202–4211.
24. **Quiroz-Castañeda RE, Martínez-Anaya C, Cuervo-Soto LI, et al.** 2011. Loosenin, a novel protein with cellulose-disrupting activity from *Bjerkandera adusta*. *Microb. Cell Fact.* **10**:8.
25. **Quiroz-castañeda RE, Folch-mallol JL.** 2011. Plant cell wall degrading and remodeling proteins: current perspectives. *Biotechnol. Apl.* **28**.
26. **Baccelli I, Luti S, Bernardi R, et al.** 2013. Cerato-platanin shows expansin-like activity on cellulosic materials. *Appl. Microbiol. Biotechnol.* **98**:175–84.
27. **Leggio L Lo, Welner D, Maria L De.** 2012. A structural overview of GH61 proteins degrading polysaccharide monooxygenases – fungal cellulose. *Comput. Struct. Biotechnol. J.* **2**.
28. **Yakovlev I, Vaaje-Kolstad G, Hietala AM, et al.** 2012. Substrate-specific transcription of the enigmatic GH61 family of the pathogenic white-rot fungus *Heterobasidion irregulare* during growth on lignocellulose. *Appl. Microbiol. Biotechnol.* **95**:979–90.
29. **Joint Genome Institute USD of E.** 2004. *Chaetomium globosum* Genome Draft.
30. **Karlsson J, Saloheimo M, Siika-Aho M, et al.** 2001. Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *Eur. J. Biochem.* **268**:6498–507.
31. **Koseki T, Mese Y, Fushinobu S, et al.** 2008. Biochemical characterization of a glycoside hydrolase family 61 endoglucanase from *Aspergillus kawachii*. *Appl. Microbiol. Biotechnol.* **77**:1279–85.
32. **Harris P V, Welner D, McFarland KC, et al.** 2010. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry* **49**:3305–16.
33. **Dimarogona M, Topakas E, Olsson L, et al.** 2012. Lignin boosts the cellulase performance of a GH-61 enzyme from *Sporotrichum thermophile*. *Bioresour. Technol.* **110**:480–7.

34. **Langston J a, Shaghasi T, Abbate E, et al.** 2011. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl. Environ. Microbiol.* **77**:7007–15.
35. **Karkehabadi S, Hansson H, Kim S, et al.** 2008. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *J. Mol. Biol.* **383**:144–54.
36. **Li X, Beeson WT, Phillips CM, et al.** 2012. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure* **20**:1051–61.
37. **Westereng B, Ishida T, Vaaje-Kolstad G, et al.** 2011. The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS One* **6**:e27807.
38. **Quinlan RJ, Sweeney MD, Lo Leggio L, et al.** 2011. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc. Natl. Acad. Sci. U. S. A.* **108**:15079–84.
39. **Levasseur A, Drula E, Lombard V, et al.** 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**:41.
40. **Cantarel.** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**:233–238.
41. **Beeson WT, Phillips CM, Cate JHD, et al.** 2012. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J. Am. Chem. Soc.* **134**:890–2.
42. **Christakopoulos P, Goodenough PW, Kekos D, et al.** 1994. Purification and characterisation of an extracellular beta-glucosidase with transglycosylation and exo-glucosidase activities from *Fusarium oxysporum*. *Eur. J. Biochem.* **224**:379–85.
43. **Christakopoulos P, Nerinckx W, Kekos D, et al.** 1996. Purification and characterization of two low molecular mass alkaline xylanases from *Fusarium oxysporum* F3. *J. Biotechnol.* **51**:181–9.
44. **Panagiotou G, Kekos D, Macris BJ, et al.** 2003. Production of cellulolytic and xylanolytic enzymes by *Fusarium oxysporum* grown on corn stover in solid state fermentation. *Ind. Crops Prod.* **18**:37–45.
45. **Cohen R, Suzuki MR, Hammel KE.** 2005. Processive Endoglucanase Active in Crystalline Cellulose Hydrolysis by the Brown Rot Basidiomycete *Gloeophyllum trabeum*. *Appl. Environ. Microbiol.* **71**:2412–2417.
46. **Mansfield SD, Saddler JN, Gu GM.** 1998. Characterization of endoglucanases from the brown rot fungi *Gloeophyllum sepiarium* and *Gloeophyllum trabeum*. *Enzyme Microb. Technol.* **23**:133–140.
47. **Ali SS, Khan M, Fagan B, et al.** 2012. Exploiting the inter-strain divergence of *Fusarium oxysporum* for microbial bioprocessing of lignocellulose to bioethanol. *AMB Express* **2**:16.
48. **Rasmussen ML, Shrestha P, Khanal SK, et al.** 2010. Sequential saccharification of corn fiber and ethanol production by the brown rot fungus *Gloeophyllum trabeum*. *Bioresour. Technol.* **101**:3526–33.

49. **Joint Genome Institute USD of E.** *Fusarium oxysporum* v1.0 Genome Draft. 2011.
50. **Joint Genome Institute USD of E.** 2013. *Gloeophyllum trabeum* v1.0 Genome Draft.
51. **Melo F, Feytmans E.** 1998. Assessing protein structures with a non-local atomic interaction energy. *J. Mol. Biol.* **277**:1141–52.
52. **Tamura K, Peterson D, Peterson N, et al.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**:2731–9.
53. **Hall T.** 1999. BioEdit a userfriendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**:95–98.
54. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**:1792–7.
55. **NCBI.** 2012. Blast.
56. **NCBI.** The National Center for Biotechnology Information.
57. **Vincze T.** 2003. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Res.* **31**:3688–3691.
58. **Larkin M a, Blackshields G, Brown NP, et al.** 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**:2947–8.
59. **Berman HM, Westbrook J, Feng Z, et al.** 2000. The Protein Data Bank. *Nucleic Acids Res.* **28**:235–42.
60. **Marchler-Bauer A, Lu S, Anderson JB, et al.** 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* **39**:D225–9.
61. **Schrödinger LLC.** The PyMOL Molecular Graphics System. 1.3.
62. **Wilkins MR, Gasteiger E, Bairoch A, et al.** 1999. Protein identification and analysis tools in the ExPASy server., p. 531–52. *In* The Proteomics Protocols Handbook.
63. **Benkert P, Biasini M, Schwede T.** 2011. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**:343–50.
64. **EBI.** European Bioinformatics Institute.
65. **Bakerlab, University of Washington.** 2011. Robetta.
66. **Rambaldi D, Ciccarelli FD.** 2009. FancyGene: dynamic visualization of gene structures and protein domain architectures on genomic loci. *Bioinformatics* **25**:2281–2.
67. **Nordahl Petersen T, Brunak S, Von Heijne B, et al.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**:785–786.

68. **Gunsteren W.** 1999. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem.* **103**:3596–3607.
69. **Arnold K, Bordoli L, Kopp J, et al.** 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**:195–201.
70. **Zhang Y.** 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**:40.
71. **Lorenz Miller G.** 1959. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.*
72. **Salinas A, Vega M, Lienqueo ME, et al.** 2011. Cloning of novel cellulases from cellulolytic fungi: heterologous expression of a family 5 glycoside hydrolase from *Trametes versicolor* in *Pichia pastoris*. *Enzyme Microb. Technol.* **49**:485–91.
73. **Schumann U, Smith N a, Wang M-B.** 2013. A fast and efficient method for preparation of high-quality RNA from fungal mycelia. *BMC Res. Notes* **6**:71.
74. **Li X, Beeson WT, Phillips CM, et al.** 2012. Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. *Structure* **20**:1051–61.
75. **Notredam C, Higgins D, Heringa J.** 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**:205–217.
76. **Joint Genome Institute USD of E.** 2005. *Phanerochaete chrysosporium* v2.0 Genome Draft.
77. **Joint Genome Institute USD of E.** 2010. *Trametes versicolor* Genome Draft.
78. **Joint Genome Institute USD of E.** 2008. *Trichoderma reesei* v2.0 Genome Draft.
79. **Rytioja J, Hildén K, Hatakka A, et al.** 2014. Transcriptional analysis of selected cellulose-acting enzymes encoding genes of the white-rot fungus *Dichomitus squalens* on spruce wood and microcrystalline cellulose. *Fungal Genet. Biol.*
80. **Ximenes E, Kim Y, Mosier N, et al.** 2010. Inhibition of cellulases by phenols. *Enzyme Microb. Technol.* **46**:170–176.
81. **Teagasc.** 2010. Straw for Energy.
82. **Lindedam J, Bruun S, Jørgensen H, et al.** 2010. Cellulosic ethanol: interactions between cultivar and enzyme loading in wheat straw processing. *Biotechnol. Biofuels* **3**:25.
83. **Chouiter R, Roy I, Bucke C.** 2008. Optimisation of β -glucuronidase production from a newly isolated *Ganoderma applanatum*. *J. Mol. Catal. B Enzym.* **50**:114–120.
84. **Xiao Z, Zhang X, Gregg DJ, et al.** 2004. Effects of sugar inhibition on cellulases and beta-glucosidase during enzymatic hydrolysis of softwood substrates. *Appl. Biochem. Biotechnol.* **113-116**:1115–26.
85. **Madigan M, Martinko J, Dunlap P, et al.** 2009. *Brock Biology of Microorganisms*, 12th ed. Pearson International Edition.

86. **Romanos M a, Scorer C a, Clare JJ.** 1992. Foreign gene expression in yeast: a review. *Yeast* **8**:423–88.
87. **Liu B, Giesler LJ, Jackson-Ziems TA, et al.** 2012. Major Fusarium Diseases on Corn, Wheat, and Soybeans in Nebraska.
88. **Horn SJ, Vaaje-Kolstad G, Westereng B, et al.** 2012. Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* **5**:45.
89. **Boraston AB, Bolam DN, Gilbert HJ, et al.** 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382**:769–81.
90. **Hemsworth GR, Davies GJ, Walton PH.** 2013. Recent insights into copper-containing lytic polysaccharide mono-oxygenases. *Curr. Opin. Struct. Biol.* **23**:660–8.
91. **Ekwe E, Morgenstern I, Tsang A, et al.** 2013. Non-Hydrolytic Cellulose Active Proteins: Research Progress and Potential Application in Biorefineries. *Ind. Biotechnol.* **9**:123–131.
92. **Bauer S, Vasu P, Persson S, et al.** 2006. Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proc. Natl. Acad. Sci. U. S. A.* **103**:11417–22.
93. **Hara Y, Hinoki Y, Shimoi H, et al.** 2003. Cloning and sequence analysis of endoglucanase genes from an industrial fungus, *Aspergillus kawachii*. *Biosci. Biotechnol. Biochem.* **67**:2010–3.
94. **Valente MT, Infantino A, Aragona M.** 2011. Molecular and functional characterization of an endoglucanase in the phytopathogenic fungus *Pyrenochaeta lycopersici*. *Curr. Genet.* **57**:241–51.
95. **Bey M, Zhou S, Poidevin L, et al.** 2013. Cello-Oligosaccharide Oxidation Reveals Differences between Two Lytic Polysaccharide Monooxygenases (Family GH61) from *Podospora*. *Appl. Environ. Microbiol.* **79**.
96. **Westereng B, Agger JW, Horn SJ, et al.** 2013. Efficient separation of oxidized cello-oligosaccharides generated by cellulose degrading lytic polysaccharide monooxygenases. *J. Chromatogr. A* **1271**:144–52.

VII Anexos

VII.1 Preparación de bacterias *E. coli* DH5 α electrocompetentes

Inocular 1 litro de medio LB usando 10 ml de un cultivo saturado de bacterias *E. coli* DH5 α preparado con antelación (cultivo overnight). Incubar el medio de cultivo a 37°C con agitación vigorosa (250 RPM aprox.) hasta que alcance una OD₆₀₀ de 0.5. Detener la incubación colocando el cultivo en hielo por 30 minutos. Centrifugar el cultivo 10 min a 4000 RPM y 4°C, remover el sobrenadante y resuspender el total del pellet en 1 litro de glicerol al 10%. Además de trabajar en ambiente estéril, se debe poner especial atención durante todo el proceso de realizar todas las manipulaciones en hielo y de mantener el glicerol frío (4°C). El cultivo se vuelve a centrifugar por 10 min a 4000 RPM y 4°C, retirando el sobrenadante y resuspendiendo el pellet en 500 ml de glicerol 10%. El proceso de centrifugado se repite dos veces más, siempre disminuyendo el volumen de resuspensión a un décimo del volumen inicial, hasta que se alcanza la concentración de células deseadas (e. g. 2 ml final). El cultivo es entonces alicuotado en volúmenes de 50 μ l o según lo que sea necesario y almacenado a -80°C.

VII.2 Medios de cultivo sólidos

Durante el transcurso del trabajo se usaron múltiples medios de cultivo sólidos. A continuación se detalla la composición de cada uno de éstos.

- Medio agar papa dextrosa
Agar 3.9%
- Medio LB Agar
Agar 1.5%, medio LB 2.5%.
- Medio LB Agar selectivo
Agar 1.5%, medio LB 2.5%, Ampicilina 100 [μ g/ml], X-Gal 80 [μ g/ml], IPTG 0.5 mM.

VII.3 Tratamiento de implementos mediante DEPC

Uno de los principales problemas a la hora de extraer RNA consiste en la acción de nucleasas en la solución. Con el fin de reducir el efecto de éstas, tanto los materiales como el agua utilizada deben ser tratados mediante DEPC previo a su uso. En el caso de los materiales, éstos deben ser remojados en agua con DEPC al 0.1% toda la noche o 1 hora a 37°C. Luego los materiales deben ser cubiertos con papel o aluminio y esterilizados a 120°C por 20 min. Al agua a utilizar durante los protocolos se le debe agregar 0.1% de DEPC y debe esterilizarse previo a su uso.

VII.4 Nomenclatura IUPAC para partidores degenerados

Tabla 22: Nomenclatura IUPAC para partidores degenerados. Se entrega el significado para cada código utilizado.

Código	Bases	Código	Bases
R	A, G	W	A, T
Y	C, T	H	A, C, T
M	A, C	B	C, G, T
K	G, T	V	A, C, G
S	C, G	D	A, G, T
		N	A, C, G, T

VII.5 Alineamientos para el diseño de partidores degenerados

VII.5.1 Alineamiento para el diseño de los partidores GH61_G1

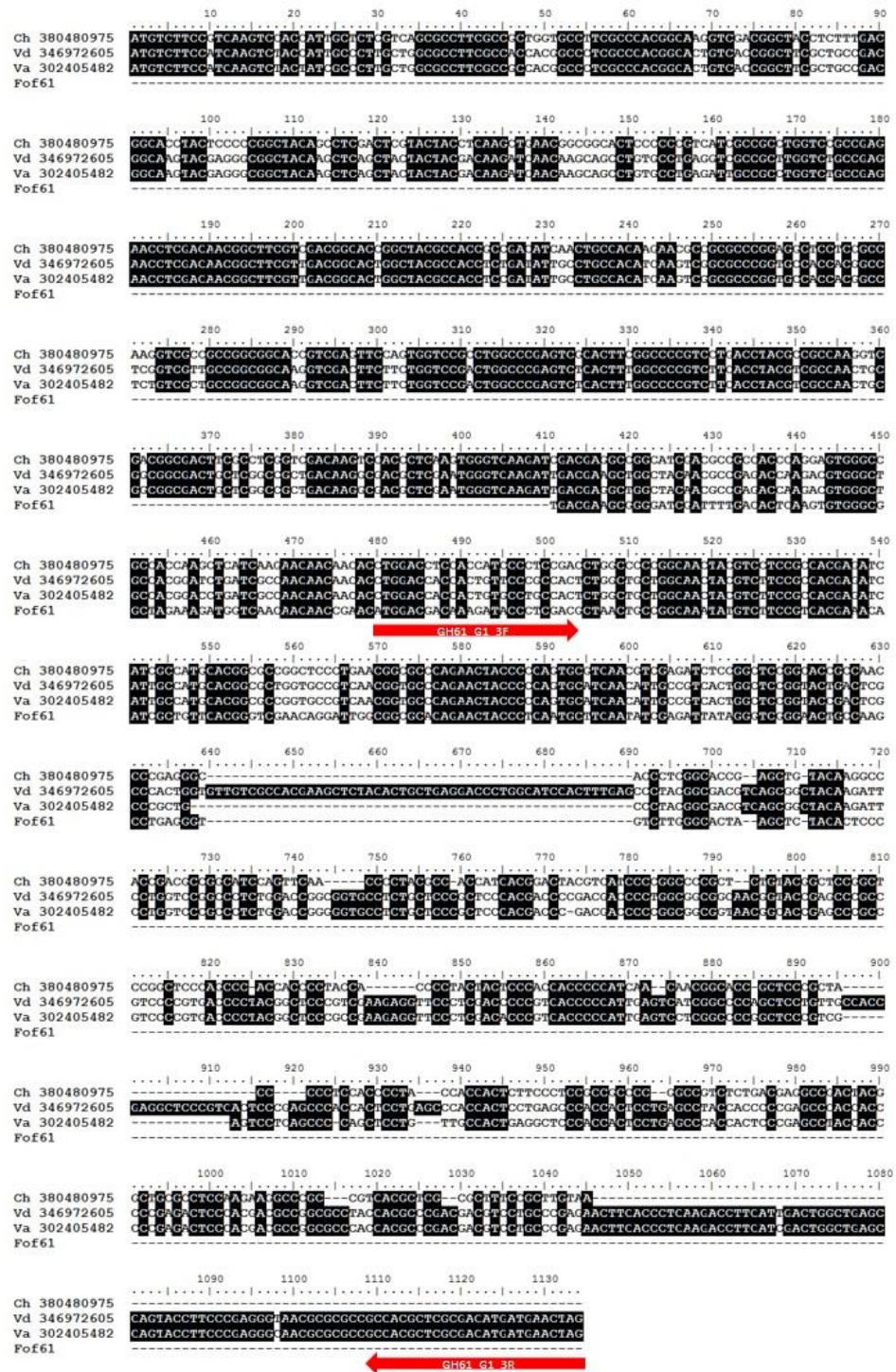


Figura 31: Alineamiento completo de secuencias nucleotídicas conocidas para el diseño de los partidores degenerados GH61_G1. Las flechas indican el emplazamiento y dirección de amplificación diseñado para cada partidor.

VII.5.2 Alineamiento para el diseño de los partidores GH61_G2



Figura 32: Alineamiento completo de secuencias nucleotídicas conocidas para el diseño de los partidores degenerados GH61_G2. La flecha indica el emplazamiento y dirección de amplificación para el partidior directo. En este caso el partidior reverse corresponde al partidior poli-T y no se encuentra dibujado.

VII.5.3 Alineamiento para el diseño de los partidores GH61_G4

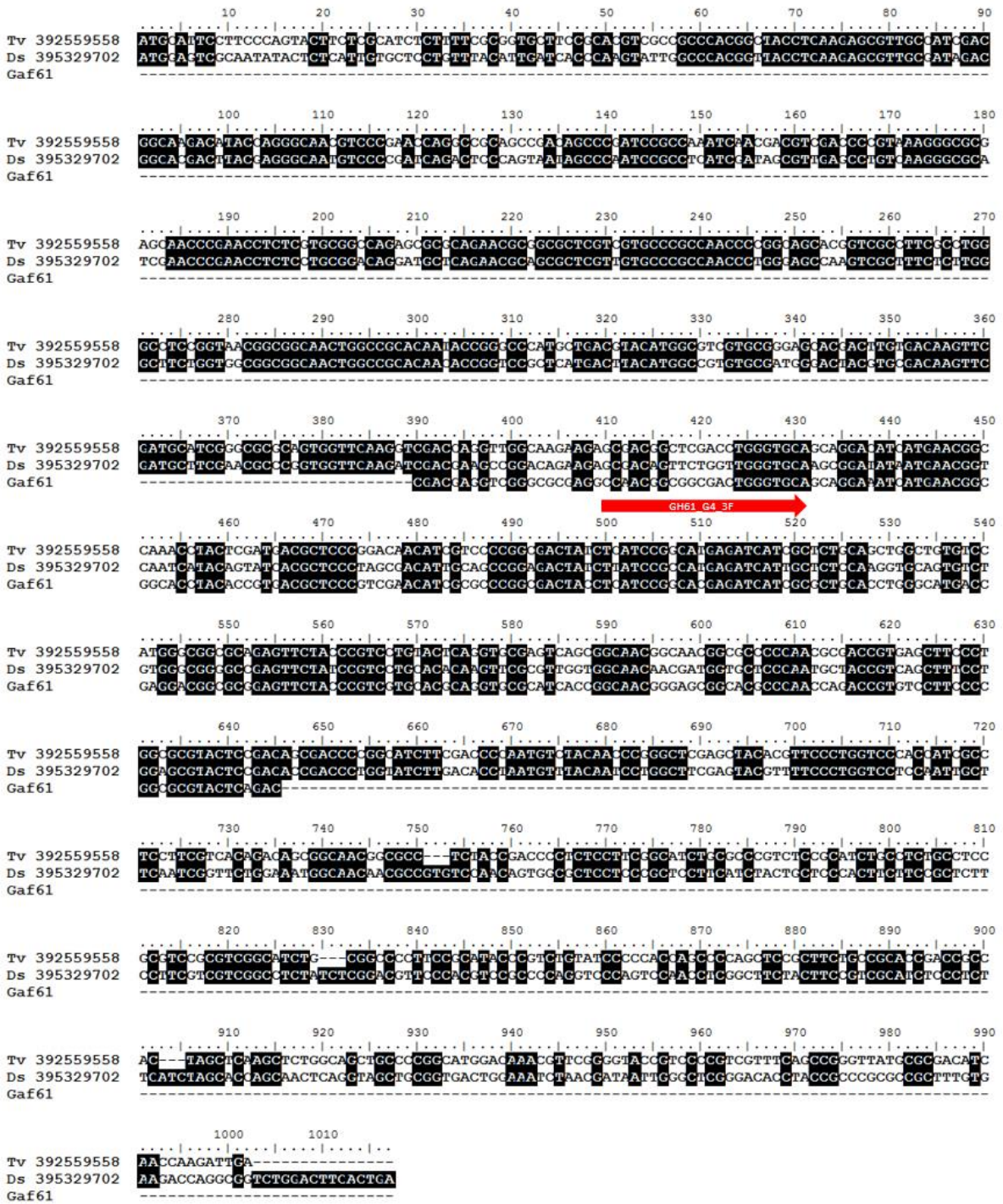


Figura 33: Alineamiento completo de secuencias nucleotídicas conocidas para el diseño de los partidores degenerados GH61_G4. La flecha indica el emplazamiento y dirección de amplificación diseñado para el partidor directo. En este caso el partidor reverse corresponde al partidor poli-T y no se encuentra dibujado.

VII.6 Mapas de los vectores de expresión seleccionados

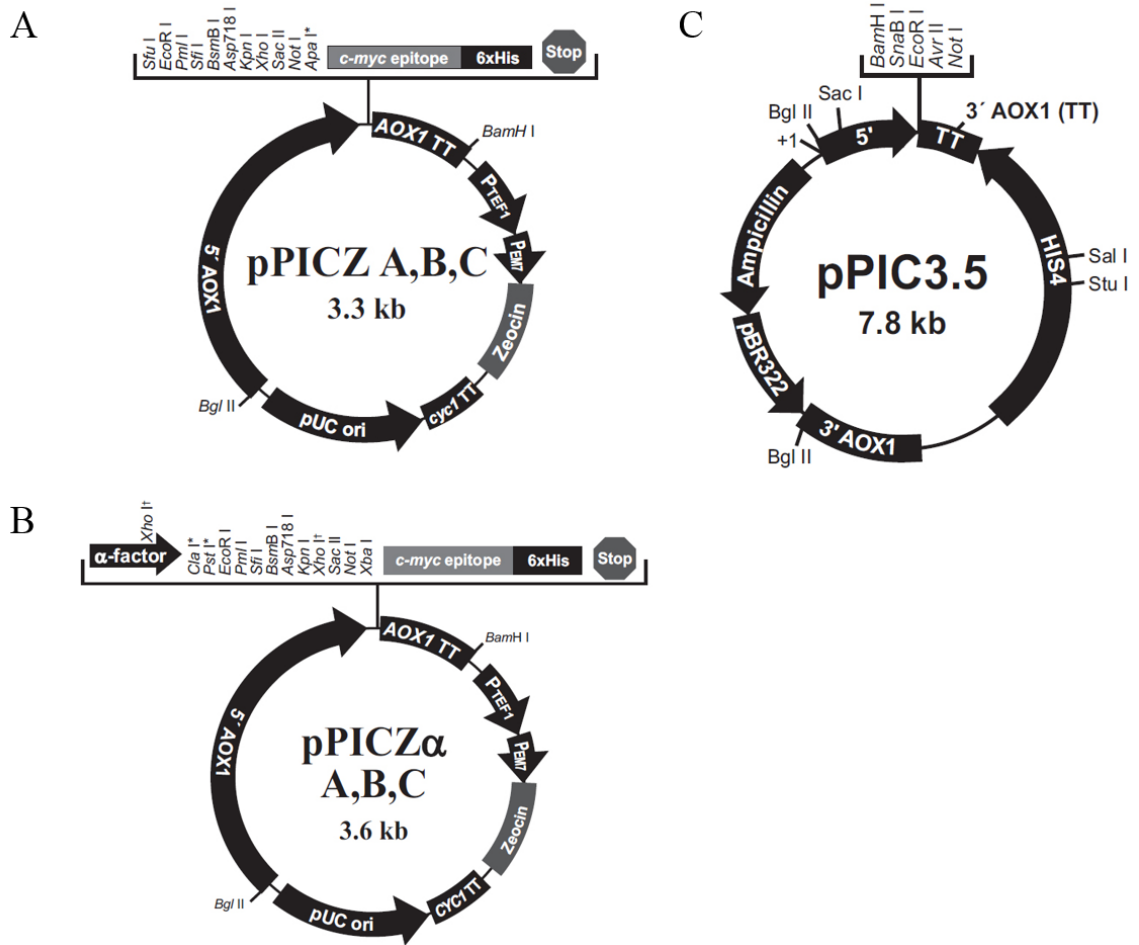


Figura 34: Mapas de los vectores de expresión pPICZ (A), pPICZα (B) y pPIC3.5 (C) de Invitrogen. En el caso de pPICZ el sitio de restricción entre Not I y el epítipo *myc* varía a Apa I, Xba I o SnaB dependiendo de si se trata del vector A, B o C respectivamente. En el caso de pPICZα, Pst I y Cla I se encuentran sólo en las versiones B y C respectivamente.

```

AOX1 mRNA 5'end (824)
TTATCATCAT TATTAGCTTA CTTTCATAAT TCGGACTGGT TCCAATTGAC AAGCTTTTGA TTTTAACGAC
                    5' AOX 1 Primer Site (855-875)
                    BamHI   SnaB I EcoR I   Avr II
TTTTAACGAC AACTTGAGAA GATCAAAAAA CAACTAATTA TTCGAAGGAT CCTACGTAGA ATTCCTTAGG
Not I
GCGGCCGCGA ATTAATTCGC CTTAGACATG ACTGTTCCTC AGTTCAAGTT GGGCACTTAC GAGAAGACCG
                    3' AOX 1 Primer Site (1055-1075)
GTCTTGCTAG ATTCTAATCA AGAGGATGTC AGAATGCCAT TTGCCTGAGA GATGCAGGCT TCATTTTTGA
                    AOX1 mRNA 3' end (1146) ↓
TACTTTTTTA TTTGTAACCT ATATAGTATA GGATTTTTTT TGTCATTTTG TTTCTTC
    
```

Figura 35: MCS del vector de expresión pPIC3.5

5' end of AOX1 mRNA

811 AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG TTCCAATTGA

5' AOX1 priming site

871 CAAGCTTTTG ATTTTAACGA CTTTTAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT

Sfu I EcoR I Pml I Sfi I BsmB I Asp718 I Kpn I Xho I

931 ATTCGAAACG AGGAATTCAC GTGGCCACG CCGCCGCTC GGATCGGTAC CTCGAGCCGC

Sac II Not I Apa I myc epitope

991 GGCGGCCGCC AGCTT GGGCCC GAA CAA AAA CTC ATC TCA GAA GAG GAT CTG
 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

Polyhistidine tag

1042 AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA GTTTAGCCT TAGACATGAC
 Asn Ser Ala Val Asp His His His His His His ***

1098 TGTTCCTCAG TTCAAGTTGG GCACTTACGA GAAGACCGGT CTGCTAGAT TCTAATCAAG

3' AOX1 priming site

1158 AGGATGTCAG AATGCCATTT GCCTGAGAGA TGCAGGCTTC ATTTTGTATA CTTTTTTATT

3' polyadenylation site

1218 TGTAACCTAT ATAGTATAGG ATTTTTTTTG TCATTTTGT

Figura 36: MCS del vector de expresión pPICZ. El sitio de restricción entre Not I y el epítipo *myc* varía a Apa I, Xba I o SnaB dependiendo de si se trata del vector A, B o C respectivamente.

5' end of AOX1 mRNA

811 AACCTTTTTT TTTATCATCA TTATTAGCTT ACTTTCATAA TTGCGACTGG TTCCAATTGA

5' AOX1 priming site

871 CAAGCTTTTG ATTTTAACGA CTTTTAACGA CAACTTGAGA AGATCAAAAA ACAACTAATT

931 ATTCGAAACG **ATG** AGA TTT CCT TCA ATT TTT ACT GCT GTT TTA TTC GCA GCA
 Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala

983 TCC TCC GCA TTA GCT GCT CCA GTC AAC ACT ACA ACA GAA GAT GAA ACG GCA
 Ser Ser Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala

α -factor signal sequence

1034 CAA ATT CCG GCT GAA GCT GTC ATC GGT TAC TCA GAT TTA GAA GGG GAT TTC
 Gln Ile Pro Ala Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe

1085 GAT GTT GCT GTT TTG CCA TTT TCC AAC AGC ACA AAT AAC GGG TTA TTG TTT
 Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe

α -factor priming site Xho I'

1136 ATA AAT ACT ACT ATT GCC AGC ATT GCT GCT AAA GAA GAA GGG GTA TCT CTC
 Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu

Kex2 signal cleavage EcoR I Pml I Sfi I BsmB I Asp718 I

1187 GAG AAA AGA GAG GCT GAA GC GAATTCAC GTGGCCACG CCGGCCGTC TCGGATCGGT
 Glu Lys Arg Glu Ala Glu Ala

Ste13 signal cleavage

Kpn I Xho I Sac II Not I Xba I c-myc epitope

1244 ACCTCGAGCC GCGGCGGCC GCCAGCTTTC TA GAA CAA AAA CTC ATC TCA GAA GAG
 Glu Gln Lys Leu Ile Ser Glu Glu

polyhistidine tag

1299 GAT CTG AAT AGC GCC GTC GAC CAT CAT CAT CAT CAT CAT TGA GTTTGTAGCC
 Asp Leu Asn Ser Ala Val Asp His His His His His His ***

1351 TTAGACATGA CTGTTCTCA GTTCAAGTTG GCACTTACG AGAAGACCGG TCTTGCTAGA

3' AOX1 priming site

1411 TTCTAATCAA GAGGATGCA GAATGCCATT TGCTGAGAG ATGCAGGCTT CATTTTTGAT

3' polyadenylation site

1471 ACTTTTTTAT TTGTAACCTA TATAGTATAG GATTTTTTTT GTCATTTTGT TTCTTCTCGT

Figura 37: MCS del vector de expresión pPICZ α A.

VII.7 *Secuencias del rRNA 18S de los hongos*

Tabla 23: Secuencias obtenidas de la secuenciación de los rDNA 18S de los hongos *Gloeophyllum trabeum* y *Fusarium oxysporum* utilizado en el trabajo.

Fusarium oxysporum

ATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATT
AAAGTTGTTGTGGTTAAAAAGCTCGTAGTTGAACTTGGGCCTGGCTGGCCGGTCCGCCT
CACCGCGTGTACTGGTCCGGCCGGGCCTTCCCTCTGTGGAACCCCATGCCCTTCACTGG
GTGTGGCGGGGAAACAGGACTTTTACTGTGAAAAAATTAGAGTGCTCCAGGCAGGCCTA
TGCTCGAATACATTAGCATGGAATAATAGAATAGGACGTGTGGTTCTATTTTGTGGTTT
CTAGGACCGCCGTAATGATTAATAGGGACAGTCGGGGGCATCAGTATTCAATTGTCAGA
GGTCAAATCCTTGGATTTATTGAAGACTAACTACTGCGAAAAGCATTGCCAAGGATGTTT
TCATTAATCAGGAACGAAAGTTAGGGGATCGAAGACGATCAGATACCGTCGTAGTCTTA
ACCATAAACTATGCCGACTAGGGATCG

Gloeophyllum trabeum

ATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATT
AAAGTTGTTGCAGTTAAAAAGCCCGTAGTTGAACTTCAGACCTGGCTGGGGCGTCTGCCT
AACGGTATGTACTGTCTGGCCGGGTCTTACCTCTTGGTGAGCCGGCATGCCCTTCACTGG
GTGTGTCCGGGAACCAGGACTTTTACCTTGAGAAAATTAGAGTGTTCAAAGCAGGCCTA
TGCCCGAATACATTAGCATGGAATAATAAAATAGGACGTGCGGTTCTATTTTGTGGTTT
CTAGAGTCGCCGTAATGATTAATAGGGATAGTTGGGGGCATTAGTATTCAGTTGCTAGAG
GTGAAATCCTTGGATTTACTGAAGACTAACTACTGCGAAAAGCATTGCCAAGGATGTTT
CATTAATCAAGGACGAAGGTTAGGGGATCGAAAACGATCAGATACCGTTGTAGTCTTAA
CAGTAACTATGCCGACTAGGGATCGG

VII.8 *Secuencias parciales de LPMOs*

Tabla 24: Resultados completos de proceso de identificación de LPMO.

Nombre Clon	Secuencia
FO61_G19	TCGGTGCCAGAACTACCCCCAGTGCTTCAACCTCGAGATCTCTGGTTCCGGCTCTGAGCTCCC TGAGGGTACCCTCGGAACTGAGCTGTACAAGTCTACCGAGAAGGGTATTGTCTTCGATCTCTA CAACAGCCCTACCAGCTACCCCATCCCCGGCCCTGCTATGACTATCAAGGCCCAAGATTAC CCAGGGCAAGCTCGCTGAGAAGGCCAAGGGTACCCCACTGGCTCTGGCTCCGGTCCGA CTCTGGTTCTGGCTCTGGCTCTGCCGCCCTGAGACCCCGCCAGACTGAGGGTCCCGCCGCT GGTACCTCCGCTGCCGCTCCTGTTGCCACTTCCGCTGCCGCTGAGGAGCCCGCTGCCCCGCTGC AGACATCCCCTGCTGCTCAAGAGCCTGCCACTCCTGTCCAGACTGAGGGTCCCTGCTGTACCA AGCCCGCCAAGACTGGCTGCAAGGCCAAGCGCGGCATGAAGGCCCGCCACGCTCGCGAC ATGATGAACTAGA
FO61_G22	TGCACAGAACTACCCGAGTGCATCAACCTCGAGGTTACTGGCAGTGGCGGCCAGGCTATCAG CGGTGGTGGCGACCTCACTACCTTCTACACTCCACCGACCCTGGTATCCTCTTCAACCTCTAC CAGTCTTTCTCTCCTACCCATTCCCGGCCCTGCTGTGCGAACTATCTAAGACGATTCCGAGG TGGTGCAGACAAGTGGTTTCTAGTCCCAGGTCGAGCTTGGGACCAAAGAGCTGGGCAAAGGG CATGCCTTAGGCTCTGATATTCATTGGTTGCCGCTAGTCGGTTCACGTTTTACAAGCAGCAGA GCTGTGTTGTTTATGTATCTACATTTATGACCCGTACAAAGTTACCTTATGCAATGAACATATT GAATTCAAAAAAAAAAAAAAAAAAAAAAAAAAGTACTAGTCGACGCGTGGCCA
GT61_G41	TTATCCGGCATGAGATCATCGCTTGCACCAGGCGCAGACGTATCCGGGCGCCAGAGCTACC CCGACTGCTTCCAGATCCGGGTGACCGGGTCCGGAACAAGACCCCAAGCGGAAGCTACCTC GTCTCTTTCCCGGTGCTTACACTGCCACGACCCCGGAATCGCTTCAACGTCTACACCACT TCACCTCGTACCCGATCCCTGGCCCCGCCGCTTGGACCGCAATTAAGTGATCAGATAGCGAG TCCTCAAGGCGGACGGTCGAGTGTGCTCACCCGTGTACTIONAGGACTCGTGACTTGAGCGGGC ATCGTGCTCAGCATGTTCAAGCATTGGATTGGATTAATTGACTAGTTCGAATCTCGATATGAAT TTCTGTCTTCTGATGCTTCTTCAGATGTGTGGGCTTCCGGATAGTTAAACTGCTGTACATAT GTAAATTATGGAGAATGTGATTTGCTGTGCCTTCAAGCTAAAAAAAAAAAAAAAAAAGTACT AGTCGACGCGTGGCCA

VII.9 *Secuencias completas de las LPMOs*

VII.9.1 *Secuencias nucleotídicas de las LPMOs identificadas*

Tabla 25: Secuencias nucleotídicas completas correspondientes a los genes secuenciados.

Nombre Clon	Secuencias nucleotídicas
FoPMO9A	<p>ATGCCTTCTTTACCTCCAAGACTCTTCTCGCCGCCGTTGCCGGTGCCATGGGCGTTGCTGCC CACGGTCACGTTTCTCTATCAACGTTGCTGGCACCGTCTACGACGGCTACGACGCTAGCTCT GCTCCCTACGACCAGACCCCAAAGACCCTCATCGCCTGGTCCACCACCGCTTCCGATAACGG TTTTGTGCGCCCTGATGCCTTCGGCACCGGCGACATCATCTGCCACCGCGATGCTAAGAATG CCAAGGGCCACGCCAAGGTCAAGGCCGGTGACCAGATCTACATCAAGTGGGATACCTGGCC CGAGAGCCACAAGGGACCTGTCATTGACATGTTGGCCTCTTGGCGTTCTGCCGGCTGTGAGT CCGTCGACAAGGAGACCCTCAAGTTCTTCAAGATCGACGAGGCTGGTCTTGTCAAGTCTGGT AACCCCGTACTTATGCCTCTGACGAGCTGATTGCCAACGACAACGGCTGGCTCATCGAGAT TCCCGAGAACATCAAGCCCGGTTCTTACGTTCTTCGTCACGAGATCATTGCTCTCCACGCCGC TGGTCAGGAGAACGGTGCTCAGAACTACCCCAAGTGTCTTCAACCTCGAGATCTCTGGTTCCG GCTCTGAGCTCCCTGAGGGTACCCTCGGAAGTGTGAGCTGTACAAGTCTACCGAGAAGGGTATT GTCTTCGATCTCTACAACAGCCCTACCAGCTACCCCATCCCGGCCCTGCTATGACTATCAAG GCCCCAAGATTACCCAGGGCAAGCTCGCTGAGAAGGCCAAGGGTACCCCAACTGGCT CTGGCTCCGGTTCGACTCTGGTCTGGCTCTGGCTCTGCCGCCCTGAGACCCCGCCAG CTGAGGCTCCCGCCGCTGGTACCTCCGCTGCCGCTCCTGTTGCCACTTCCGCTGCCGCTGAGG AGCCCGCTGCCCCGCTCGAGACATCCCTGCTGTCTAAGAGCCTGCCACTCCTGTCCAGACT GAGGCTCCTGCTGCTACCAAGCCCGCCAAGACTGGCTGCAAGGCCAAGCGCGGCATGAAGG CCCGCCGTCACGCTCGTGACCTCATGAACTAA</p>
FoPMO9B	<p>ATGCACTGCAACCTCTCTCTCGGTGTCGTGGCTGCTCTCCTCGCTGGAGCCAACGCCACGGT CACGTCGCAAAGGTCATTGCCGGTGGCAAGGAGTACACCGTGGTATTCCCCACGGTGTCTCC CTCTGACGCTGTCGGCTGGGCGCTGGCAACCAGGACAACGGCTTCGTTTACCCGATGCTT TCAAGTCCGCCGATATCATCTGCCACAAGAGCGCTAAGCCCGTCTCCAACGCCGTCGTTGTT GCCGCTGGTGATGTTGTCCTCTGAAGTGGGACACCTGGCCCGAGTCTCACCACGGCCCCGT CACTGAGTACCTCGTCCCGTCAGCGGTGACTTTGCCAGCATCAACAAGCAGAGCCTCCGAT GGTCAAGGTGCTCAGAAGGGTCTCAAGTCTGGCAGCAACCCCGTACTGGGCCTCCGA CGATCTCATCCGCGATGGCTTCTCATGGAAGTTCACCGTCCCAAGAACCTCAAGGCCGGTA AATACGTTCTCCGACACGAGATCATCGGCCTTACTCTGCTGGCCAGACCAACGGTGCTCAG GCCTACCCCAAGTGCATCAACCTCGAGGTTACTGGCAGTGGCGGCCAGGCTATCAGCGGTGG TGGCGACTTCACTACCTTCTACTACTCCACCGACCTGGTATCCTCTTCAACCTCTACCAGTC TTTCTCTTCTACCCCATTCCCGGCCCTGCTGTCCGAATATCTAA</p>
GtPMO9A	<p>ATGGTCTCCTTACCAAGACATTTCTCGCCATCGTTGCCTGTGCGCTCGGCGTCCAAGCCAC GGGTACGTTGATACCTCAACGTTGGTGGCACCCAGTACACTGGGTACCTGCCTTACAACGA CCCGTACACGACCCCGTCCCGAGCGCATCGAGCGTCCATCCAGGGAAACGGCCCGTCA CAGCCCTAACTACAATCGACGTTCAATGCAACGGGGAGAATGGCGGAGGCAGCTCTCCGGC GCCCTCGTGGCCACTATTGCGGCGGGTGGCAAGATCGCCTTCCATTGGACCACCTGGCCGT CTTCGCACGTCGGTCCCGTAATCACTACCTTGGCAAGGTGCCGAGCAGCACCGACGTTACC AAGTACTCGCCACCGGCTCGGACGTCATCTGGTCAAGATCGACGAAGGTGGATACTCTAA CGGCAAGTGGGCGCCACCGACGTCCTTCCGCGCAAAAACAGCACCTGGACGGTCACTATCC CGAGCTCGTTGCTCCGGGGCAGTACATCGTCCGCCACGAGATCATCGCCTTGCACCAGGCG CAGACGTATCCGGGCGCCAGAGTACCCCGACTGCTTCCAGATCCGGGTGACCGGGTCCGG AAACAAGACCCCGAGCGGAAGCTACCTCGTCTCCTTCCCGGTGCTTACTACTGCCACGACCC CCGGAATCGTTTCAACGCTACACCAACTTACCTCGTACCCGATCCCTGGCCCCGCCGCTT GGACCGCAATTA</p>

VII.9.2 Secuencias proteicas de las LPMOs identificadas

Tabla 26: Secuencias proteicas completas correspondientes a los genes secuenciados.

Nombre Clon	Secuencia proteica
FoPMO9A	MPSFTSKTLLAAVAGAMGVAAHGHVSSINVAGTVYDGYDASSAPYDQTPKTLIAWSTTASDNG FVAPDAFGTGDIIICHRDAKNAKGHAKVKAGDQIYIKWDTWPESHKGPVIDMLASCGSAGCESV DKETLKFFKIDEAGLVKSGNPGTYASDELIANDNGWLEIPENIKPGSYVLRHEIILHAAGQENG AQNYPQCFNLEISGSGSELPEGTLGTELYKSTEKGIVFDLYNSPTSYPGPAMTIKAPKITQGKLA EKAKGTPTTGS GSGSDSGSGSGSAA PETPAQTEAPAAGTSA AAPVATSAAAEEPAAPVETSPA AQ EPATPVQTEAPAATKPAKTGCKAKRGMKARRHARDLMN
FoPMO9B	MHCNLSLGVVAALLAGANAHGHVAKVIAGGKEYTGGIPHGAPSDAVGWAAGNQDNGFVSPDA FKSADIICHSKSAKPVSNVVAAGDVVTLKWDTPESHHPVTEYLAPVSGDFASINKQSLRW VKVAQKGLKSGSNPGDWASDDLIRDGFSWKFTVPKNLKAGKYVLRHEIIGLHSAGQTNGAQAY PQCINLEVTGSGGQAISGGDFTFYTPDTPGILFNLYQSFSSYPIGPVAVRTI
GtPMO9A	MVSFTKTFFAIVACALGVQAHG YVDLNVGGTQYTG YLPYNDPYTTPVPQRIERPIPGNGPV TAL TTIDVQCNGENGGGSSPAPLVATIAAGGKIAFHWTTPSSHVGPVITYLGKVPSSDVTKYSP TG SDVIWFKIDEGGYSNGKWAATDVLSAQNSTWTVTIPSSLAPGQYIVRHEIILHQAQTYPGAQSY PDCFQIRVTGSGNKTPSGSYLVSPFGAYTATTPGIAFNVYTNFTSYPIGPVAVWTGN

VII.10 Resultados de la modelación tridimensional

VII.10.1 Herramientas de evaluación de modelos tridimensionales de proteínas

Descripción, a grandes rasgos, del funcionamiento de las herramientas computacionales utilizadas para evaluar las estructuras tridimensionales.

Anolea: Esta herramienta permite evaluar la calidad del empaquetamiento de la proteína. Ésta evalúa las interacciones energéticas de cada átomo pesado en la cadena proteica. Estos cálculos se realizan en el ambiente no local del átomo, vale decir todos los átomos que se encuentren a una distancia inferior a 7 Å, y que no se encuentren a menos de 11 residuos de distancia en la secuencia. El resultado se grafica en base a la “energía” de cada aminoácido: en verde cuando la energía es favorable (negativa) o en roja cuando no lo es (positiva).

Gromos: Es una herramienta destinada para su uso en la simulación de dinámicas moleculares. Su uso en las estructuras tridimensionales permite obtener un cálculo de la energía de cada aminoácido. El resultado de éste se gráfica de igual forma que el entregado por Anolea.

QMean: Esta herramienta es la combinación de varias evaluaciones estructurales basadas en potenciales. Ésta se calcula para cada residuo y para todo el modelo, y su valor va de 0 a 1 con mayores valores para mejores modelos. Entre los factores considerados se incluyen:

- La evaluación de la torsión del aminoácido en base al efecto de los que lo flanquean.
- El efecto energético de aminoácidos situados lejos de éste.
- La relación que existe entre hidrofobicidad de los aminoácidos y su posición al interior de la proteína.
- La concordancia que existe entre la secuencia y la estructura secundaria.

Z-Score: Este puntaje consiste en una comparación del QMean obtenido para el modelo, en relación a los obtenidos para estructuras tridimensionales conocidas de alta resolución. Representa una medición de la calidad general del modelo en base a qué tan parecido es éste a estructuras nativas. Modelos de baja calidad debiesen presentar puntajes negativos inferiores, mientras que los de buena calidad debiesen ser más cercanos a cero.

VII.10.2 Resultados para el mejor modelo obtenido para FoPMO9A.

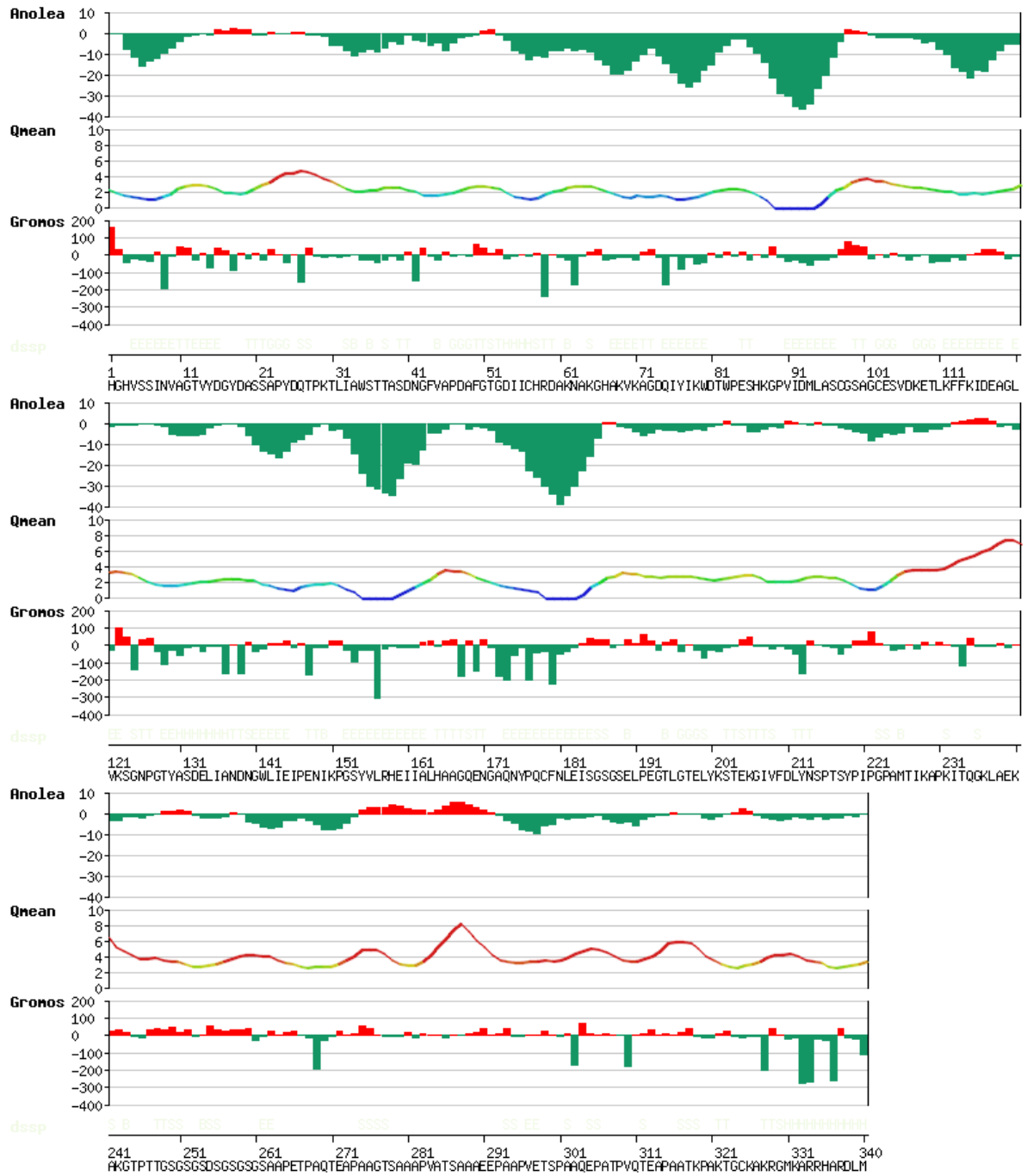


Figura 38: Resultados obtenidos para el Modelo 1 entregado por Robetta para la modelación de la proteína FoPMO9A. En el gráfico se presentan los resultados de Anolea, QMean y Gromos a lo largo de toda la proteína.

VII.10.3 Resultados para el mejor modelo obtenido para FoPMO9B.

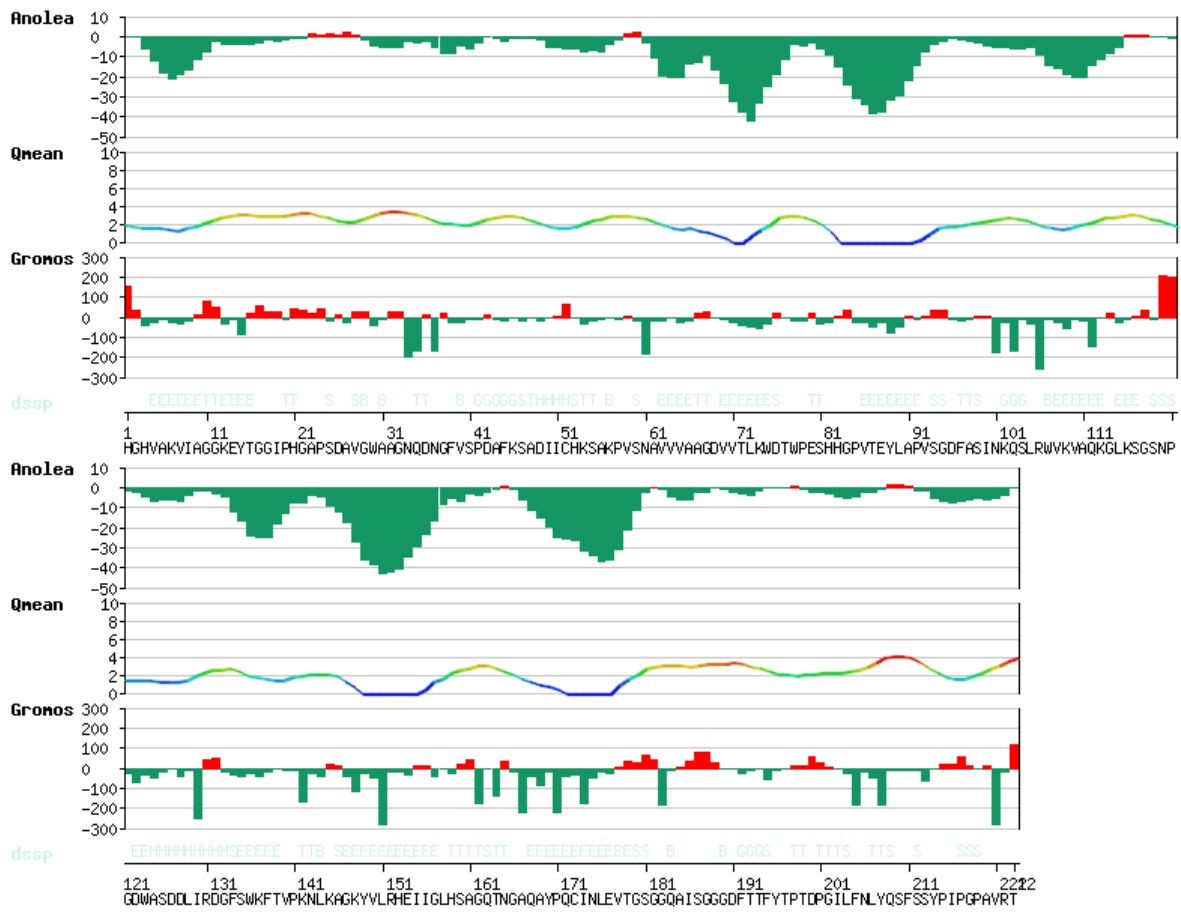


Figura 39: Resultados obtenidos para el Modelo 1 entregado por Robetta para la modelación de la proteína FoPMO9B. En el gráfico se presentan los resultados de Anolea, QMean y Gronos a lo largo de toda la proteína.

VII.10.4 Resultados para el mejor modelo obtenido para GtPMO9A.

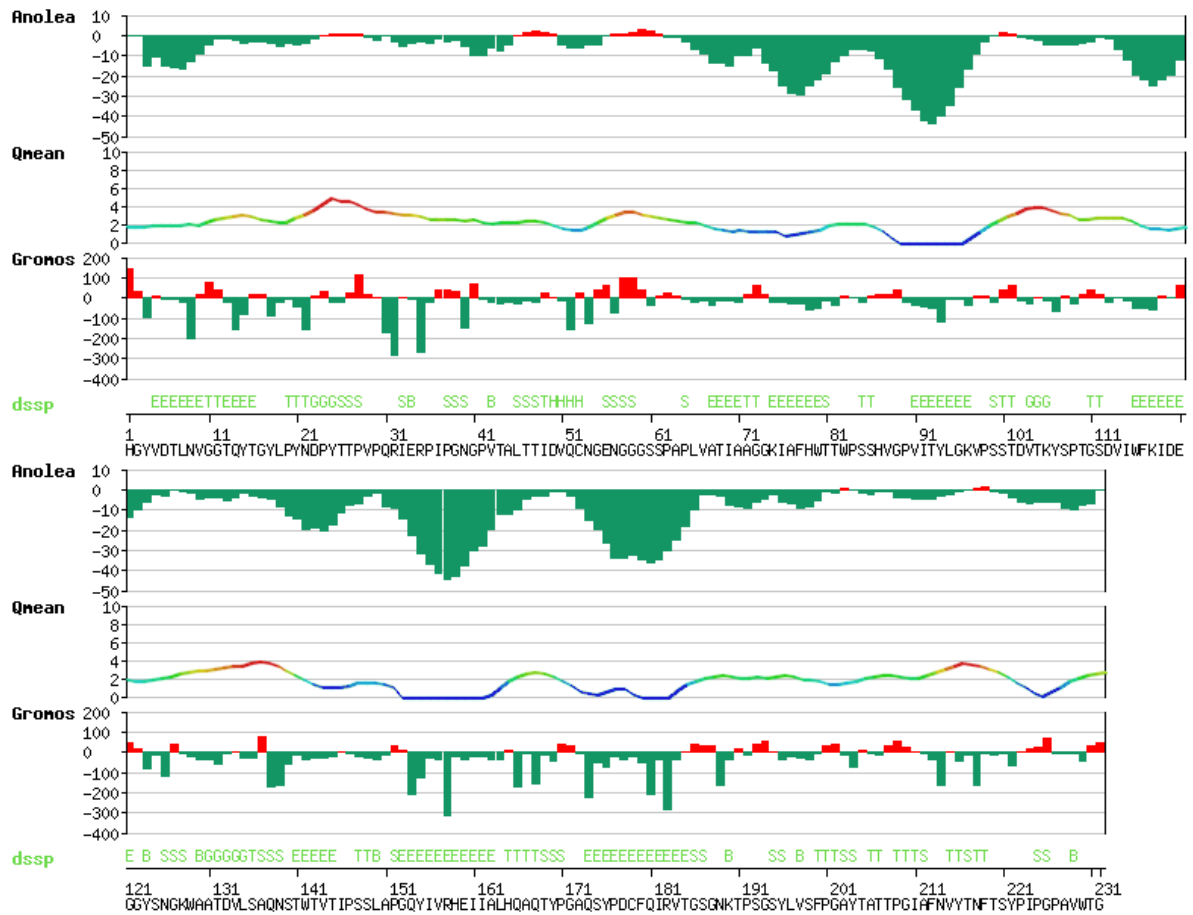


Figura 40: Resultados obtenidos para el Modelo 1 entregado por Robetta para la modelación de la proteína GtPMO9A. En el gráfico se presentan los resultados de Anolea, QMean y Gromos a lo largo de toda la proteína.