



ESCUELA DE PREGRADO – FACULTAD DE CIENCIAS – UNIVERSIDAD DE CHILE

“Análisis teórico-bioinformático de una enzima con  
potencial actividad  $\alpha$ -glucuronidasa del hongo  
*Gloeophyllum trabeum* y clonamiento del gen”

Seminario de Título entregado a la Universidad de Chile en cumplimiento parcial de los  
requisitos para optar al Título de Ingeniero en Biotecnología Molecular

**Joaquín Antonio Francis Schell**

Directora de seminario de título: María Oriana Salazar Aguirre

Co-directora de seminario de título: Luciana Lehuedé

Profesora patrocinante: Claudia Renate Stange Klein

Agosto 2022

Santiago – Chile



## INFORME DE APROBACIÓN SEMINARIO DE TÍTULO

Se informa a la Escuela de Pregrado de la Facultad de Ciencias, de la Universidad de Chile que el Seminario de Título, presentado por el **Sr(ta)**.

**Joaquín Antonio Francis Schell**

**“Clonamiento y análisis teórico-bioinformático de una enzima con potencial actividad  $\alpha$ -glucuronidasa proveniente del hongo *Gloeophyllum trabeum*”**

Ha sido aprobado por la Comisión de Evaluación, en cumplimiento parcial de los requisitos para optar al Título de Ingeniero en Biotecnología Molecular.

Dra. Oriana Salazar Aguirre  
**Directora Seminario de Título** \_\_\_\_\_

Dra. Luciana Lehuedé  
**Codirectora Seminario de Título** \_\_\_\_\_

Dra. Claudia Stange Klein  
**Profesora patrocinante** \_\_\_\_\_

**Comisión Revisora y Evaluadora**

Dr. Ricardo Cabrera  
**Presidente Comisión:** \_\_\_\_\_

Dr. Marcelo Baeza  
**Evaluador:** \_\_\_\_\_

Santiago de Chile, Agosto 2022

## Biografía



Nací el 28 de mayo del año 97, un día lluvioso. He tenido una vida bastante tranquila donde gracias al apoyo de mi familia y seres queridos he tenido la libertad de explorar todo lo que capte mi interés. Me gusta saber de todo, y puedo pasarme horas enteras leyendo o investigando cualquier cosa que me parezca curiosa, y por ello desde una temprana edad desarrollé un interés por las ciencias básicas y aplicadas. La biología en particular siempre me pareció (y me sigue pareciendo) fascinante, y siento que estamos en una época ideal para poder seguir aprendiendo y descubriendo cosas

nuevas y sorprendentes

Tengo un abanico de aficiones, empezando por el manga, anime y juegos, hasta las bicicletas y generos musicales del más amplio espectro. En parte siento que gracias a esa variedad he podido hallar matices en la vida que me han permitido crecer como persona, y a encontrar valor en cada conversación y vínculo que tengo con otras personas.

## **Dedicatoria**

A mi gata mamita, que me acompañó y aconsejó durante toda la carrera

## **Agradecimientos**

Agradezco a la profesora Oriana, por su paciencia y presencia durante toda mi estadía en el laboratorio, sin su guía y apoyo no hubiera podido enfrentar todos los percances que fueron surgiendo durante este trabajo.

Agradezco a Luciana por su excelente calidad como científica y su apoyo que me inspiró a seguir adelante.

Agradezco a Stephanie por todos los consejos, saberes y buena onda que hicieron todo mucho más fácil.

Agradezco a Coni, Nico, Benja, Vale y Cata por hacer del laboratorio un espacio ameno en donde trabajar.

## Tabla de contenido

Biografía -----	ii
Dedicatoria-----	iii
Agradecimientos -----	iv
Índice de tablas -----	vii
Índice de figuras -----	viii
Lista de abreviaturas-----	ix
Resumen-----	x
Abstract-----	xi
Introducción -----	1
1. Características generales del xilano y sus aplicaciones -----	1
2. Degradación de xilano en la naturaleza -----	2
2.1 Hongos como fuente de enzimas con actividad sobre xilano-----	3
2.1.1 Presencia de enzimas con actividad celulasa y hemicelulasa -----	3
2.1.2 Hongos de pudrición blanca-----	4
2.1.3 Hongos de pudrición parda-----	5
2.1.4 Hongos de pudrición blanda -----	7
2.2 Enzimas celulasas y hemicelulasas-----	7
2.2.1 Glicosil hidrolasas implicadas en la degradación de celulosa y hemicelulosa -----	7
2.2.2 Glicosil hidrolasas como enzimas auxiliares de celulasas y hemicelulasas-----	8
3. Características generales de las $\alpha$ -glucuronidasas -----	11
4. Utilización de enzimas glicosil hidrolasas en la producción industrial de xilooligosacaridos (XOS)-----	14
5. Motivación, propuesta y objetivos -----	16
5.1 Objetivo general -----	18
5.2 Objetivos específicos -----	18
Materiales y métodos-----	20
1. Técnicas de alineamiento-----	20
2. Obtención, validación y refinamiento del modelo por homología -----	21
3. Análisis de modelos de proteínas-----	22
4. Cultivo de hongos y extracción de RNA-----	22
4.1 Mantenición de la cepa fúngica-----	22

4.2 Cultivo de <i>G. trabeum</i> para la extracción de RNA-----	23
4.3 Extracción total de RNA desde micelio -----	23
5. Síntesis de cDNA -----	24
6. Amplificación y aislamiento de la secuencia de DNA codificante del gen de Gt115 -----	25
6.1 Diseño de partidores -----	25
6.2 Amplificación por PCR -----	25
7. Clonamiento de la secuencia de DNA codificante aislado -----	26
7.1 Ligación al vector pJET 1.2 Blunt-----	26
7.2 Transformación del producto de ligación en <i>Escherichia coli DH5<math>\alpha</math></i> -----	28
8. Evaluación de colonias recombinantes -----	28
8.1 Extracción de plásmido desde cultivo de bacterias -----	28
8.2 Digestión con enzimas de restricción -----	29
8.3 Secuenciación de plásmidos para cada colonia-----	29
Resultados-----	30
1. Análisis del sitio activo y estructura general de las enzimas $\alpha$ -glucuronidasas GH67-----	30
2. Análisis del sitio activo y estructura general de las enzimas $\alpha$ -glucuronidasas GH115 ----	34
3. Búsqueda de de enzimas glucuronidasas mediante alineamiento masivo con BLASTp ---	41
4. Obtención y refinamiento de un modelo por homología -----	42
5. Análisis del sitio activo y estructura general del modelo por homología de Gt115 -----	47
6. Identificación y aislamiento de la secuencia para la glucuronidasa Gt115 -----	51
7. Clonamiento y análisis de la secuencia aislada -----	56
Discusión y proyecciones-----	60
Conclusiones-----	74
Referencias-----	76
Anexo-----	79

## Índice de tablas

Tabla 1. Enzimas fúngicas relevantes en la descomposición de madera, junto con su clasificación, el número de ellas que está caracterizada, y sus aplicaciones actuales. ....	4
Tabla 2. Componentes de medio mínimo para hongos suplementado con alfalfa .....	23
Tabla 3. Partidores diseñados para aislar la secuencia de DNA codificante de Gt115 desde cDNA total.....	25
Tabla 4. Primeros cinco resultados del alineamiento en BLASTP de Gt115 sobre las secuencias no redundantes de todas las bases de datos .....	41
Tabla 5. Únicos cuatro resultados del alineamiento en BLASTP de Gt115 sobre las secuencias de la base de datos Protein Data Bank (PDB) .....	42
Tabla 6. Resultados del análisis del modelo de homología de Gt115 sin refinar obtenido con MolProbity.....	44
Tabla 7. Resultados de MolProbity obtenidos del análisis del modelo de Gt115 proveniente del servidor I-TASSER después de cinco ciclos de optimización. ....	45
Tabla 8. Rotaciones $\omega$ medidas para los residuos de unión a xilano en los sitio +1 y -1 para el modelo de Gt115 y la enzimas AxyAgu115A (6NPS), SdeAgu115A (4ZMH) BoAgu115A (4C90) y BtGH115A (5BY3) .....	50
Tabla 9. Distancia entre los residuos de la diada de ácidos catalíticos AC1 (Acido aspártico que va primero en la secuencia) y AC2 (Acido glutámico que va segundo en la secuencia) con respecto a la base catalítica BC (Ácido aspártico que se encuentra en el loop móvil característico de la enzima). ....	51
Tabla 10. Grupos de residuos conservados catalíticamente relevantes en posición y/o función para las enzimas de la familia GH67 de las que actualmente se dispone un modelo 3D obtenido experimentalmente. ....	81
Tabla 11. Grupos de residuos conservados catalíticamente relevantes en posición y/o función para las enzimas de la familia GH115 de las que actualmente se dispone un modelo 3D obtenido experimentalmente. ....	81
Tabla 12. Análisis de MolProbity de residuos catalíticamente relevantes de la proteína Gt115 (XP_007865968).....	82



## Índice de figuras

Figura 1. Estructura esquemática del xilano.....	2
Figura 2. Mecanismo de inversión para las enzimas $\alpha$ -glucuronidasas. ....	12
Figura 3. Enzimas en vista superficie representantes de las familias de $\alpha$ -glucuronidasas GH67 y GH115.....	13
Figura 4. Secuencia en formato FASTA de la enzima Gt115 de <i>Gloeophyllum trabeum</i> .....	20
Figura 5. Programa de PCR para el aislamiento de la secuencia de DNA codificante para la enzima Gt115.....	26
Figura 6. Mapa de restricción y características principales del vector pJET 1.2/Blunt en su forma abierta .....	27
Figura 7. Sitio catalítico de la enzima GlcA67A de <i>Pseudomonas cellulosa</i> en conformación abierta (A) (1GQI) y en conformación cerrada (B) (1GQL).....	31
Figura 8. Estructura tridimensional de GlcA67A de <i>Pseudomonas cellulosa</i> (1GQI). ....	32
Figura 9. Modelo del sitio activo de enzimas $\alpha$ GH67 con su estructura tridimensional experimentalmente resuelta.....	33
Figura 10. Cavidad catalítica de la enzima $\alpha$ GH115 BoAguA115 (4C91) de <i>B. ovatus</i> .....	35
Figura 11. Modelo en vista caricatura de enzimas $\alpha$ GH115 de 4 dominios.....	36
Figura 12. Alineamiento en vista caricatura entre BoAgu115A y BtGH115.....	36
Figura 13. Modelo de enzima SdeAgu115A (SdeAgu115A).. ....	38
Figura 14. Surco catalítico de BoAgu115A (4C91). ....	39
Figura 15. Vista del bolsillo catalítico de BoAgu115A.....	40
Figura 16. Gráfico de Ramachandran para los modelos por homología de Gt115 sin refinar y refinado.....	46
Figura 17. Modelo por homología refinado obtenido para Gt115 (XP_007865968).. ....	47
Figura 18. Distancias de los residuos de Trp de los sitios +1 y -1 de las enzimas AxyAgu115A y Gt115. ....	48
Figura 19. Comparación en los ángulos $\omega$ de los residuos de triptófano de los sitios de unión al sustrato +1 y -1 de Gt115 ( <i>G. trabeum</i> ) y AxyAgu115A ( <i>A. xylanus</i> ).....	49
Figura 20. Bolsillo catalítico de Gt115.....	50
Figura 21. Producto de la extracción de RNA total de <i>G. trabeum</i> , analizado en un gel de agarosa/formaldehido 1%.....	53
Figura 22. Electroforesis del resultado del RTPCR realizado para aislar secuencias codificantes de Gt115. En agarosa 1%. ....	54
Figura 23. Gel de agarosa 1% donde se cargó la reacción de PCR por triplicado a 60°C del cDNA total de <i>G. trabeum</i> . ....	55
Figura 24. Electroforesis en gel de agarosa 1% de los plasmidios recombinantes 4P15XC1, 4P15XC2 y 4P15XC4 intactos y digeridos con XmnI.....	57
Figura 25. Alineamiento de los resultados de secuenciación obtenidos para los plásmidos de las colonias 4P15XC1, 4P15XC2, 4P15XC4, el vector pJET1.2/blunt vacío y un constructo hipotético del vector pJET1.2/blunt y la secuencia codificante para Gt115. ....	58
Figura 26. Alineamiento por Clustal $\Omega$ de las secuencias de GlcA67 y AguA.....	79
Figura 27. Alineamiento por Clustal $\Omega$ de las secuencias de BtGH115, AxyAgu115A, BoAgu115A, Gt115 y SdeAgu115A. ....	80

## Lista de abreviaturas

PDA	Medio agar papa dextrosa
BML	Biomasa lignocelulósica
$\alpha$ -Araf	Arabinofuranosa
GA	Ácido Glucurónico
MeGA	Ácido Metil glucurónico
AA	Ácido acético
GX	Glucuronoxilanos
AX	Arabinoxilanos
GAX	Glucuronoarabinoxilanos
AGX	Arabinoglucuronoxilanos
XOS	Xilooligosacáridos
HPB	Hongos de pudrición blanca
HPP	Hongos de pudrición parda
HPBI	Hongos de pudrición blanda
PCR	Reacción en cadena de la polimerasa
cDNA	Ácido desoxirribonucleico complementario
TAE	Tampón Tris-acetato-EDTA

## Resumen

Este seminario tuvo el objetivo de estudiar la estructura de una potencial  $\alpha$ -glucuronidasa del hongo *G. trabeum* usando métodos bioinformáticos, así como también clonar la secuencia de DNA codificante de su gen correspondiente en *E. coli*. Se realizó una revisión bibliográfica y un análisis del sitio activo de la familias de glicosil hidrolasas GH67 y GH115. Para la primera familia, tanto el mecanismo catalítico como especificidad de sustrato se encuentran bien caracterizados, sin embargo, para la familia GH115 tales características no se encuentran totalmente definidas. Los factores determinantes del mecanismo catalítico y la especificidad de sustrato de la familia GH115 podrían asociarse a cómo se mueve un loop móvil del sitio catalítico y a la composición y forma del surco catalítico. Se obtuvo un modelo por homología monomérico de una enzima con posible actividad  $\alpha$ -glucuronidasa de *G. trabeum* (XP\_007865968.1), denominada en este trabajo como Gt115, el cual fue refinado hasta que la geometría de la mayoría de los residuos se encontrara sobre del estándar mínimo geométrico para su confiabilidad. El modelo de Gt115 mostró una estructura de 5 dominios y un sitio catalítico con aminoácidos conservados en posiciones idénticas propios de la familia GH115. Gt115 muestra un surco catalítico largo cuyos residuos clave presentan rotaciones  $\omega$  no observadas en sus contrapartes bacterianas, lo que podría implicar afinidad por sustratos más largos y una interacción con el sustrato distinta a lo visto en otras  $\alpha$ -glucuronidasas de la misma familia.

Desde un cultivo de *G. trabeum* crecido en alfalfa se aisló un fragmento de DNA del tamaño correspondiente a la secuencia de DNA codificante para Gt115. El fragmento se intentó ligar al vector pJET 1.2/Blunt y posteriormente clonar el producto de ligación en

*Escherichia coli* DH5 $\alpha$  sin éxito, hecho atribuido principalmente a una ligación no exitosa, y a que el marcador de selección secundaria del vector de clonación se encontraba no funcional. Para confirmar la expresión del gen de Gt115 en *G. trabeum* se deberá secuenciar el fragmento de ADN de 3000 pb aislado por RtPCR.

### **Abstract**

The aim of this seminar was to study the structure of a potential  $\alpha$ -glucuronidase from the fungus *G. trabeum* using bioinformatics methods, as well as to clone the transcript of its corresponding gene in *E. coli*. A literature review and active site analysis of the GH67 and GH115 families of glycosyl hydrolases was performed. For the first family, both the catalytic mechanism and substrate specificity are well characterized, however, for the GH115 family such characteristics are not fully defined. The determinants of the catalytic mechanism and specificity of the GH115 family could be associated with how a mobile loop moves through the catalytic site and the composition and shape of the catalytic groove. A monomeric homology model of an enzyme with possible  $\alpha$ -glucuronidase activity from *G. trabeum* (XP\_007865968.1), referred to in this work as Gt115, was obtained and refined until the geometry of most residues was above the minimum geometric standard for reliability. The model of Gt115 showed a 5-domain structure and a catalytic site with conserved amino acids in identical positions characteristic of the GH115 family. Gt115 presents a long catalytic groove whose key residues present  $\omega$ -rotations not observed in its bacterial counterparts, which could imply affinity for longer substrates and an interaction with the substrate different from that seen in other  $\alpha$ -glucuronidases of the same family.

A DNA fragment of the size corresponding to the sequence of the transcript coding for Gt115 was isolated from a culture of *G. trabeum* grown on alfalfa. The fragment was attempted to ligate into the pJET 1.2/Blunt vector and subsequently clone the ligation product into *Escherichia coli* DH5 $\alpha$  without success, a fact mainly attributed to unsuccessful ligation, and to the secondary selection marker of the cloning vector being non-functional. To confirm the expression of the Gt115 gene in *G. trabeum*, the 3000 bp DNA fragment isolated by RtPCR should be sequenced.



# Introducción

## 1. Características generales del xilano y sus aplicaciones

La biomasa lignocelulósica (BML) es el componente principal de la materia seca de las plantas y es la materia prima más abundante en la tierra, estando compuesta principalmente de celulosa (21% a 55%), hemicelulosa (5% a 50%) y lignina (5% a 35%) mezcladas en distintas proporciones dependiendo del tipo de planta (Yousuf y col., 2020).

La hemicelulosa específicamente corresponde en promedio a un tercio de toda la BML, la cual está compuesta principalmente por polisacáridos unidos por enlaces  $\beta$ -(1 $\rightarrow$ 4) y  $\beta$ -(1 $\rightarrow$ 3), incluyendo xiloglucanos, xilanos, mananos y glucomananos (Scheller y Ulvskov, 2010). El xilano particularmente constituye el 20% total de las paredes celulares primarias y es el principal componente de la hemicelulosa, compuesto por unidades de  $\beta$ -D-xilopiranosas, formando polímeros entre 70 y 200 unidades, estando estos sustituidos por residuos de arabinofuranosa ( $\alpha$ -Araf), ácido glucurónico (GA)/metil-glucurónico (MeGA), ácido acético (AA) entre varios otros (Collins y col., 2005) como se puede observar en la Figura 1.

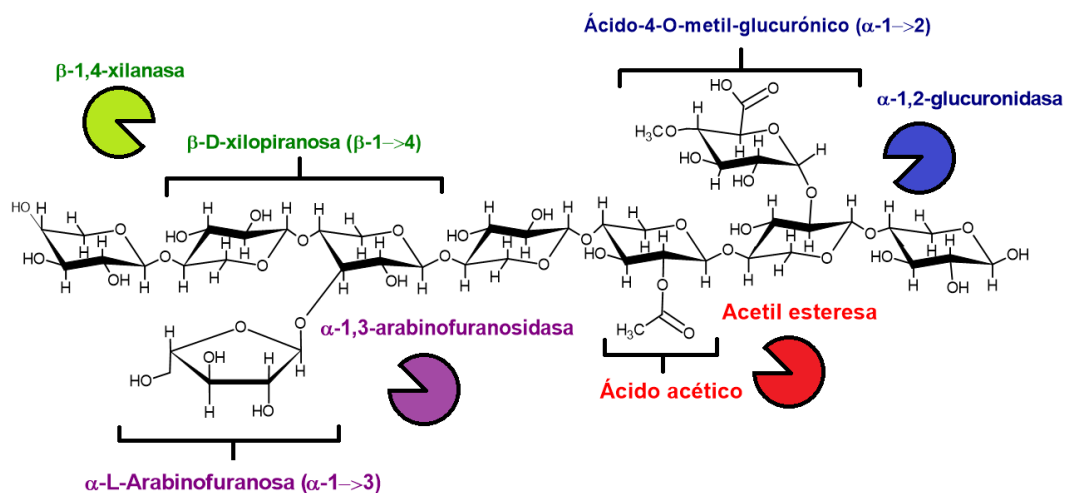


Figura 1. Estructura esquemática del xilano. Cada componente está indicado por colores junto a la enzima que hidroliza el enlace que lo conecta a la cadena principal.

El xilano se puede dividir en varias subclases según el tipo y grado de sustitución, pudiendo encontrarse en forma de homoxilanos, glucuronoxilanos (GX), arabinoxilanos (AX) o combinaciones complejas de varios sustituyentes en forma de arabinoglucuronoxilanos (AGX), glucuronoarabinoxilanos (GAX) y heteroxilanos (Santibáñez y col., 2021).

Hoy en día se han hallado varios usos para el xilano tales como la elaboración de películas poliméricas, hidrogeles, nanopartículas, biocombustibles, suplementos alimenticios, endulzantes y prebióticos (Fu y col., 2019; Santibáñez y col., 2021).

## 2. Degradación de xilano en la naturaleza

La degradación de xilano en los seres vivos tiene como función principal la obtención de azúcares más simples que sirvan como fuente de carbono y energía para su supervivencia. Se ha reportado que un gran espectro de bacterias, son capaces de degradar principalmente cadenas cortas de xilano (Uffen, 1997), incluyendo bacterias



termofílicas (Sizova y col., 2011), intestinales (Zhang y col., 2014) y marinas (Sung y col., 2021) entre varias otras. Por otro lado, para hongos, la degradación de xilano es fundamental para la supervivencia de hongos filamentosos saprofitos y algunas levaduras, que contienen una gran variedad de enzimas que actúan que facilitan la descomposición de todos los componentes de la madera (Uffer, 1997; Li y col., 2022). Dada la adaptación dirigida a la degradación de madera, los hongos saprofitos se consideran como una fuente rica de enzimas con actividad celulosa y hemicelulosa, que poseen características deseables para enzimas industriales útiles para procesos como el blanqueamiento de papel, obtención de biocombustibles y generación de compuestos farmacológicos (Li y col., 2022), de ello los hongos saprofitos en particular serán el foco de este seminario.

## **2.1 Hongos como fuente de enzimas con actividad sobre xilano**

### **2.1.1 Presencia de enzimas con actividad celulosa y hemicelulosa en hongos**

Las glicosil hidrolasas cumplen un rol fundamental en los hongos descomponedores de madera. Los hongos secretan hacia el espacio extracelular estas enzimas para degradar la pared celular de las plantas para obtener carbohidratos más simples que puedan ser utilizados como fuente de carbono y energía. Según reportan Rafiei y col. (2021), la mayoría de las enzimas secretadas encargadas de la degradación de la pared celular (Cell wall degrading enzymes o CWDEs por sus siglas en inglés) pertenecen al grupo de las glicosil hidrolasas, grupo que comprende enzimas como quitinasas, endoglucanasas, exoglucanasas, xilanasas y amilasas. Estas enzimas actúan en conjunto para la

penetración de la pared celular, expansión dentro del tejido vegetal, virulencia, patogenia silente y la inducción de la respuesta inmune.

En la Tabla 1 se compilan varios tipos de enzimas con roles en la descomposición de madera y sus aplicaciones:

Tabla 1. Enzimas fúngicas relevantes en la descomposición de madera, junto con su clasificación, el número de ellas que está caracterizada, y sus aplicaciones actuales.

Enzimas implicadas en la fitopatogénesis	Clasificación EC	Total de enzimas caracterizadas <sup>1</sup>	Aplicaciones
Endoglucanasas	3.2.1.4	107	Sacarificación e hidrólisis de BML <sup>2</sup> , Detergente textil <sup>3</sup>
Exoglucanasas	3.2.1.91	40	Sacarificación e hidrólisis de BML <sup>2</sup> , Detergente textil <sup>3</sup>
$\beta$ -Glicosidasas	3.2.1.21	37	Sacarificación e hidrólisis de BML <sup>2</sup> , Detergente textil <sup>3</sup>
Quitinasas	3.2.1.14	36	Fungicida, Insecticida, Producción de proteína unicelular <sup>4</sup>
$\alpha$ -Amilasas	3.2.1.1	19	Conversión y Sacarificación de almidón, Detergente textil <sup>5</sup>
Endo- $\beta$ -1,4-Xilanasas	3.2.1.8	173	Producción de XOS <sup>6</sup> y extracción de arabinoxilano de productos alimenticios <sup>7</sup>
$\alpha$ -L-Arabinofuranosidasas	3.2.1.55	36	Enzima auxiliar en la hidrólisis de xilano <sup>6</sup>
$\alpha$ -Glucuronidasas	3.2.1.131	7	Enzima auxiliar en la hidrólisis de xilano <sup>6</sup>

1. Base de datos mycoCLAP (Strasser y col., 2015) 2. (Soccol y col., 2011) 3. (Ejaz y col., 2021). 4. (Singh y col., 2021) 5. (De Souza y De Oliveira, 2010) 6. (Santibañez y col., 2021) 7. (Bhardwaj, Kumar y Verma, 2019)

### 2.1.2 Hongos de pudrición blanca

Los hongos de pudrición blanca (HPB) son una subdivisión de los hongos descomponedores de madera, su principal característica se halla en que al momento de atacar el tejido vegetal estos tienen como blanco todos los constituyentes de la madera

incluyendo celulosa, hemicelulosa y lignina, dejando una materia blanca, siendo usualmente celulosa o hemicelulosa enriquecida o parcialmente degradada (Morrell y Zabel, 2020).

Existen dos categorías generales para estos hongos, distinguibles por su mecanismo de acción, por un lado, están los HPB de tipo 1, los cuales realizan un ataque simultáneo a todos los componentes de la madera haciendo uso de lignina peroxidasas, celulasas y hemicelulasas, generalmente degradando maderas duras. Por otro lado, los HPB de tipo 2 atacan de manera secuencial a la madera concentrándose en una primera fase de deslignificación para después atacar los carbohidratos de la pared celular generalmente prefiriendo maderas blandas (Langer y col., 2021).

La degradación completa de lignina ha sido observada tanto en hongos de pudrición blanca de tipo 1 como de tipo 2, por ello contienen en su genoma una gran cantidad de genes para enzimas capaces de degradar lignina, principalmente peroxidasas y lacasas. Por otro lado, las enzimas hidrolíticas presentes en HPB generalmente corresponden a endoglucanasas, exoglucanasas y  $\beta$ -glucosidasas para celulosa (Kim y col., 2016) y endoxilanasas,  $\beta$ -xilosidasas y  $\alpha$ -arabinofuranosidasas para hemicelulosa (Daniel, 2014; Strasser y col., 2015). Actualmente debido a la dificultad de cultivo y la ausencia general de características como termoestabilidad y resistencia al pH las enzimas de este grupo no han sido comercialmente exploradas (Peralta y col., 2017).

### **2.1.3 Hongos de pudrición parda**

Los hongos de pudrición parda (HPP) son una subdivisión de los hongos descomponedores de madera que se distinguen por su preferencia por maderas blandas y

por la degradación selectiva de celulosa y hemicelulosa de la pared celular, donde la lignina en vez de ser completamente degradada como es el caso de los hongos de pudrición blanca es ampliamente desmetilada, dejando el tejido vegetal con un tono café y textura marchita debido a la pérdida de celulosa y hemicelulosa (Langer y col., 2021).

El mecanismo de acción de los HPP consiste en una etapa de degradación no enzimática donde la celulosa y la hemicelulosa se despolimerizan por medio de reacciones oxidativas de Fenton y cambios de pH a través de la secreción de metabolitos (Daniel, 2016), que da paso a la segunda etapa de tipo enzimático, donde se secretan enzimas hidrolasas que degradan completamente los carbohidratos que quedaron despolimerizados y accesibles en la primera fase (Floudas, 2021)

Los hongos de pudrición parda contienen en su mayoría genes codificantes para endoglucanasas, pectinasas y hemicelulasas, presentando una preferencia por la degradación de hemicelulosa, donde se ha observado que hasta el momento solamente *Coniphora puteana* posee un sistema completo para degradar celulosa incluyendo  $\beta$ -glicosidasas y celobiohidrolasas (Daniel, 2016).

Krah y col. (2018) han sugerido que las enzimas de los HPP tienen especificidades de sustrato variadas, y que se pueden dividir en generalistas y especializados en maderas blandas, donde incluso se ha observado que hongos como *Gloeophyllum trabeum* son capaces de degradar algunos pastos (Kraffenberger y Schilling, 2013), lo que apunta a los HPP como una fuente de posibles enzimas generalistas capaces de degradar diferentes tipos de tejido vegetal. Al igual que los HMB, los HPP aún no han sido explorados comercialmente debido a la falta de estudios, difícil cultivo y ausencia de

características deseables en las enzimas actualmente caracterizadas (Peralta y col., 2017).

#### **2.1.4 Hongos de pudrición blanda**

Los hongos de pudrición blanda (HPBI) es una clasificación intermedia principalmente de hongos ascomicetos que exhiben estrategias de descomposición de madera que se encuentran entre los HPB y los HPP, creciendo dentro de las paredes celulares de la madera, dejando pequeñas cavidades que dependiendo del hongo presentan tanto el blanqueamiento visto en los HPB como la textura café y quebradiza de los HPP (Langer y col., 2021). Los HPBI crecen rápidamente en ambientes altos en humedad, usualmente en temperaturas muy frías o calientes para HPB o HPP (Stokland y col., 2012).

Sista y col. (2017) reportaron que entre los hongos más populares para el estudio de enzimas involucradas en la degradación de BML los HPBI poseen una mayor cantidad de genes relacionados con estas enzimas que los HPB y HPP. Actualmente existen celulasas comerciales provenientes de HPBIs, particularmente de *Aspergillus niger* (Sigma Aldrich: C1184) y *Trichoderma sp.* (Sigma Aldrich: C1794).

## **2.2 Enzimas celulasas y hemicelulasas**

### **2.2.1 Glicosil hidrolasas implicadas en la degradación de celulosa y hemicelulosa**

Como se ha mencionado antes la celulosa es uno de los componentes principales de las paredes celulares vegetales pudiendo formar parte de estas hasta en un 55%. La celulosa está compuesta por unidades de D-glucosa unidas por enlaces glicosídicos  $\beta$ -1,4, que en la madera se encuentran en altos grados de polimerización ( $\approx 20.000$  DP) donde forman estructuras cristalinas y amorfas altamente empacadas (Habibi y col., 2010). La

degradación enzimática de la celulosa se lleva a cabo a grandes rasgos por 3 tipos de enzimas, las Celobiohidrolasas (EC 3.2.1.91) que cortan los extremos de cadenas largas, Endoglucanasas (EC 3.2.1.4) que cortan en posiciones internas de los polímeros y  $\beta$ -Glucosidasas (EC 3.2.1.21) que cortan cadenas cortas de glucooligosacáridos a glucosa (Bornscheuer y col., 2014). La degradación desde celulosa ocurre por medio de una sinergia entre las celulasas que actúan en los extremos y en las posiciones internas (endo-exo) donde las endoglucanasas que cortan de manera aleatoria generan extremos nuevos para que las celobiohidrolasas los corten, permitiendo la solubilización progresiva de la celulosa desde los polímeros altamente empaçados (Jalak y col., 2012).

El caso de la hemicelulosa resulta ser más complicado, la diversidad en su composición (xiloglucanos, xilanos, mananos y glucomanos) requiere una mayor cantidad de enzimas que permitan el corte y la solubilización de sus componentes, sin embargo, debido a que el xilano es el componente principal de la hemicelulosa se considera el mecanismo degradación enzimática de este polímero como el principal. Entre las enzimas xilanasas más relevantes se encuentran las Endoxilanasas (EC 3.2.1.8) las cuales cortan al azar residuos de xilano en posiciones internas, las familias GH10 y GH11 particularmente tienen actividad generalista en xilano, la primera prefiriendo XOS, cadenas cortas y/o ramificadas y la segunda cadenas largas sin sustituciones y xilano insoluble (Santibañez y col., 2021). La degradación de xilano soluble y XOS se lleva a cabo por  $\beta$ -Xilosidasas (EC 3.2.1.37) que cortan estos polímeros a unidades de xilosa (Houfani y col., 2020).

### **2.2.2 Glicosil hidrolasas como enzimas auxiliares de celulasas y hemicelulasas**

Para la degradación de celulosa la participación de otras glicosil hidrolasas se observa limitada a la sinergia Exo-Endo asociada a las celobiohidrolasas y endoglucanasas descrita anteriormente por Jalak y col. (2012), sin embargo, para el caso de la hemicelulosa, la mayoría de los polímeros de esta están altamente sustituidos por moléculas de ácido glucurónico y metil glucurónico (GA/MeGA), de arabinofuranosa ( $\alpha$ -Araf) y grupos acetilo AA, los cuales debido a impedimento estérico dificultan el acceso de endoxilanasas y  $\beta$ -xilosidasas a la cadena de xilano y por ello el corte de la cadena (Santibañez y col., 2021).

Tanto GA/MeGA como  $\alpha$ -Araf sustituyen al xilano en distintos tipos de combinaciones y proporciones formando los distintos tipos de polímeros sustituidos mencionados en el punto 1 de esta sección. Para las coníferas, xilanos como el glucuronoxilano (GX) y glucuronoarabinoxilano (GAX) representan un porcentaje menor en las paredes secundaria y primaria, mientras que para las dicotiledóneas el GX representa hasta un 30% de la pared secundaria mientras que para las paredes de gramíneas y otras monocotiledóneas los GAX representan un porcentaje de hasta 40% para paredes primarias y hasta 50% para paredes secundarias (Scheller y Ulvskov, 2010). Los arabinoxilanos (AX) por otra parte se presentan como los polisacáridos más comunes encontrados en la hemicelulosa de plantas herbáceas, donde se ha reportado que pueden llegar a un porcentaje de 16,4% peso/peso en salvado de trigo (Zhou y col., 2016; Marcotuli y col., 2016).

Para AX y GAX la hidrólisis de los residuos de  $\alpha$ -Araf se lleva a cabo por enzimas  $\alpha$ -arabinofuranosidasas (EC 3.2.1.55) que cortan residuos de arabinosa o arabinofuranosa

de la cadena de xilano. Beldman y col. (1997) originalmente clasificaron las enzimas arabinofuranosidasas en 3 tipos, el tipo (A) que no puede degradar polímeros, el tipo (B) que degrada polímeros y el tipo AHX que son enzimas con actividad específica para arabinoxilano. Estas últimas pertenecen a las familias GH43 y GH62, mientras que las del tipo A y B pertenecen a la familia GH51.

Para GX y GAX la hidrólisis de GA/MeGA es ejecutada por enzimas  $\alpha$ -glucuronidasas (EC 3.2.1.131) que cortan los residuos de GA/MeGA que sustituyen la cadena de xilano. Existen dos familias dentro de esta clasificación, la familia GH67 y GH115 cuya diferencia principal radica en que la primera familia corta residuos de GA/MeGA exclusivamente del extremo reductor de la cadena de xilano, mientras que la segunda puede realizar cortes en posiciones internas (Rogowsky y col., 2014).

Para  $\alpha$ -arabinofuranosidasas la sinergia con xilanasas se ha observado para enzimas provenientes de hongos como *Apergillus niger* (Wan Azalee y col., 2016), *Aspergillus nidulans* (Pereira y col., 2018) y *Aspergillus hortai* (Terrone y col., 2020) aumentando el rendimiento de hidrólisis entre 1,36 veces a 2,6 veces para diferentes sustratos. Por otro lado, para  $\alpha$ -glucuronidasas (EC 3.2.1.131) la sinergia con xilanasas se ha observado en una enzima de la familia GH67 proveniente de *Aspergillus fumigatus*, que presenta sinergia con endoxilanasas de la familia GH10 en la hidrólisis de oligosacáridos de madera de Haya (Rosa y col., 2013). En la misma línea una enzima de la familia GH115 proveniente del hongo *Schizophyllum commune* también demostró sinergia con una endoxilanasas de la familia GH10 en la digestión de GAX, aumentando 5 veces el



rendimiento de la hidrólisis de xilano a xilobiosa en comparación con la endoxilanasas sola (McKee y col., 2016)

A pesar de su importancia como enzimas auxiliares las  $\alpha$ -glucuronidasas están relativamente poco caracterizadas a nivel global, la base de datos CaZy (Drula y col., 2022) reporta solo 34 enzimas caracterizadas con esta actividad, 26 para la familia GH67 y 8 para la familia GH115, en comparación, existen anotadas 1232 enzimas con actividad  $\alpha$ -L-arabinofuranosidasa. Debido a ello se ha considerado relevante avanzar en la búsqueda y caracterización de estas enzimas, particularmente la familia GH115 cuya capacidad de cortar residuos en posiciones internas de la cadena de xilano es atractiva para la hidrólisis enzimática industrial de hemicelulosa (Santibañez y col., 2021)

### **3. Características generales de las $\alpha$ -glucuronidasas**

Las enzimas  $\alpha$ -glucuronidasas (EC 3.2.1.131) actúan en la cadena de xilano hidrolizando el enlace  $\alpha$ -(1 $\rightarrow$ 2) que conecta la cadena principal con el GA/MeGA a través de un mecanismo de inversión (Figura 2), consistente en que un ácido catalítico protona el carbono anomérico del azúcar, y una base catalítica activa una molécula de agua cercana que interactúe con el carbono protonado, formando un estado de transición que se resuelve en la escisión del GA/MeGA de la cadena de xilano (Dodd y Cann, 2009),

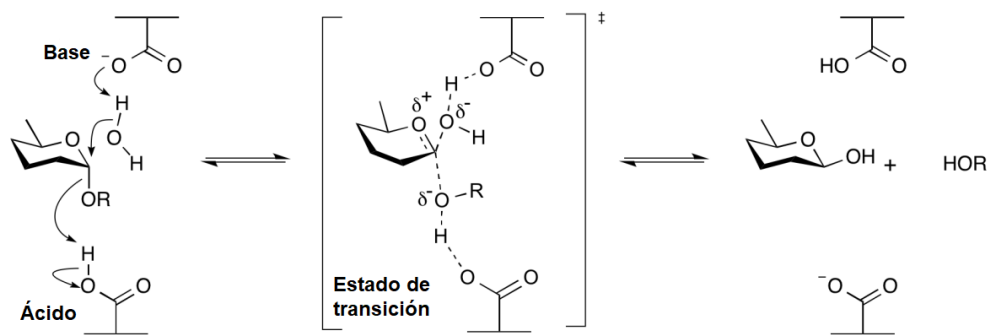


Figura 2. Mecanismo de inversión para las enzimas  $\alpha$ -glucuronidasas. (Williams (2013). Inverting glucosidase mechanism [Esquema] [https://www.cazypedia.org/index.php/File:Inverting\\_glucosidase\\_mechanism.png](https://www.cazypedia.org/index.php/File:Inverting_glucosidase_mechanism.png).

Las  $\alpha$ -glucuronidasas que actúan sobre xilano se dividen en dos familias principales, la familia GH115 y la familia GH67 (Figura 3), siendo la principal diferencia funcional entre las dos familias que la primera es capaz de cortar residuos de GA/MeGA que se encuentran tanto en medio de la cadena de xilano como en el extremo no reductor de esta, mientras que las segundas solo son capaces de hacer el corte en el extremo no reductor (Rogowski y col., 2014).

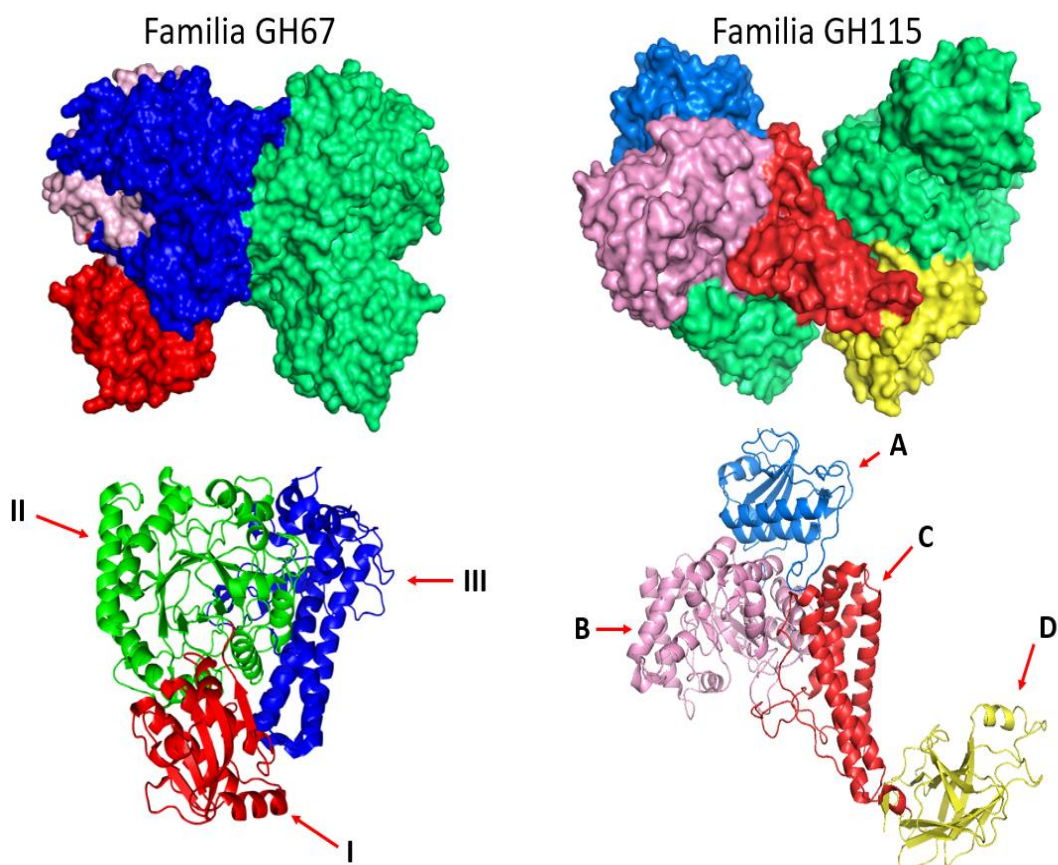


Figura 3. Enzimas en vista superficie y cartoon mostrando la estructura de dominios de dos representantes de las familias de  $\alpha$ -glucuronidasas GH67 y GH115. A la izquierda se observa GlcA67A (1GQI) de *Pseudomonas cellulosa* y a la derecha BoAgu115A (4C90) de *Bacteroides ovatus*. Con flechas se señalan los dominios de cada enzima.

La familia GH67 es un grupo de enzimas homodiméricas de 3 dominios que tienen actividad sobre xilooligosacáridos y polímeros cortos de xilano, tanto cuando son secretadas al espacio extracelular como cuando están unidas a la membrana (Nagy y col., 2002). Estas se presentan en conjunto con otras glicosidohidrolasas en situaciones donde se requiera la degradación de un sustrato vegetal, generalmente mostrando sinergia con  $\beta$ -xilanasas y  $\beta$ -xilosidasas que cortan cadenas largas de xilano generando XOS sustituidos por MeGA que posteriormente son cortados por las enzimas de la

familia GH67 (Jia y col., 2014). Actualmente la única  $\alpha$ -glucuronidasa comercial que existe pertenece a la familia GH67 (Megazyme: K-AGLUA).

A diferencia de la familia GH67, la familia GH115 tiene preferencia por sustratos más largos de xilano, incluyendo GX y GAX de varios tipos de árboles (Wang y col. 2016; Yan y col., 2021) y en algunos casos por oligómeros de arabinogalactano (Aalbers y col., 2015). Al igual que su preferencia de sustrato, la familia GH115 es diversa en su arquitectura de dominios y estructura cuaternaria, donde se han reportado enzimas homodiméricas con 4 dominios (Rogowski y col., 2014) y 5 dominios (Wang y col., 2016; Yan y col., 2021) así como una enzima monomérica de 4 dominios (Aalbers y col., 2015). Hasta ahora todas las estructuras resueltas de enzimas de la familia GH115 provienen de bacterias (Drula y col., 2022) , por lo que se posee muy poco conocimiento sobre sus contrapartes fúngicas que podrían poseer características estructurales y funcionales distintas dadas las adaptaciones de los hongos a la vida saprofita, por esa razón la obtención y análisis de estructuras 3D de  $\alpha$ -glucuronidasas GH115 se tiene como un objetivo prioritario en orden de profundizar el conocimiento de la diversidad, preferencia de sustrato y potencial biotecnológico de estas enzimas.

#### **4. Utilización de enzimas glicosil hidrolasas en la producción industrial de xilooligosacaridos (XOS)**

Se ha observado que los xilooligosacaridos (XOS), polímeros de xilano de bajo peso molecular, han mostrado un amplio abanico de usos y aplicaciones, tales como su potencial prebiótico para *Bifidobacteria sp.* y *Lactobacillus sp.* (Na y Kim, 2007; Lin y col., 2016) y anticancerígeno (Aachary y col., 2015), servir como endulzantes y agentes

espumantes bajos en calorías (Imaizumi y col., 1991; Abdo y col., 2021), además de ser moléculas altamente estables ante el calor y la presión (Silva y col., 2020; Silva y col., 2022) lo que los hace una biomolécula muy atractiva cuya producción aun representa una variedad de desafíos (Santibañez y col., 2021).

Para la producción industrial de XOS se ha optado principalmente por una estrategia de dos pasos, donde en el primer paso la BML se pretrata con agentes básicos, ácidos o con altas presiones para separar la celulosa y la lignina de la hemicelulosa, la cual en el segundo paso procede a tratarse con enzimas xilanolíticas que den como producto XOS (Santibañez y col., 2021). El pretratamiento de hidrólisis alcalina sumado a la hidrólisis enzimática ha sido preferido por sobre otras debido a que el pretratamiento alcalino produce una menor cantidad de productos secundarios tóxicos como el furfural y el hidroximetil furfural (HMF) (Cardoso y col., 2020), además de que la estrategia completa contamina menos que otras técnicas usadas para el mismo propósito (Aachary y Prapulla, 2010).

Para este método la hidrólisis de la cadena de xilano es llevada a cabo por enzimas endo1-4- $\beta$ -xilanasas (E.C. 3.2.1.8) representadas principalmente por las familias GH10 y GH11 que se utilizan como enzimas libres (Poletto y col., 2020), inmovilizadas (Rajagopalan y col., 2016) o secretadas directamente por un organismo por medio de fermentación *in situ* (Amorim y col., 2019). Como se ha demostrado antes, la utilización de enzimas auxiliares como  $\alpha$ -glucuronidasas y  $\alpha$ -L-arabinofuranosidasas aumenta el rendimiento de la hidrólisis de xilano por endoxilanasas (McKee y col., 2016; Terrone y col., 2020), por ello el uso de estas enzimas auxiliares presenta potencial para el

mejoramiento de la producción de XOS, permitiendo no solo un mayor rendimiento a menor costo sino que también una digestión ajustada a cada sustrato variando la proporción de cada una de estas enzimas en la mezcla (Santibañez y col., 2021)

## **5. Motivación, propuesta y objetivos**

Hoy en día los xilooligosacaridos (XOS) se consideran una biomolécula muy atractiva debido a su amplio espectro de usos y aplicaciones. La mayoría de los métodos utilizados para la obtención de XOS requieren de dos pasos, primero, un pretratamiento con altas presiones o agentes alcalinos/ácidos para aislar y romper la hemicelulosa, y segundo, un tratamiento con enzimas xilanasas que hidrolizan el xilano obteniendo XOS como resultado (Santibañez y col., 2021). Para estos métodos el uso de xilanasas solas ha resultado efectivo en la producción de XOS (Bragatto y col., 2013), sin embargo, la mayoría de los xilanos obtenido del pretratamiento alcalino, si bien roto en polímeros más cortos hidrolizables por las xilanasas, se encuentran altamente sustituidos por decoraciones de GA/MeGA y  $\alpha$ -Araf que impiden la acción de las xilanasas sobre la cadena de xilano (Scheller y Ulskov, 2010).

El uso de enzimas auxiliares que corten las decoraciones de xilano ( $\alpha$ -arabinofuranosidas y  $\alpha$ -glucuronidasas) se ha probado como una estrategia viable tanto para el aumento del rendimiento de la hidrólisis de xilano a XOS y xilosa (McKee y col., 2016; Terrone y col., 2020). A pesar de su potencial uso como enzimas auxiliares, el número de  $\alpha$ -glucuronidasas caracterizadas es relativamente bajo en comparación con otras enzimas glicosil hidrolasas (Tabla 1), estando la mayor cantidad de conocimiento y aplicaciones en la familia GH67 (Zhou y col., 2018). En contraste, el conocimiento

sobre las enzimas familia GH115 es limitado y concierne casi totalmente a enzimas de origen bacteriano (Yan y col., 2021). A pesar de lo anterior, el hecho de que su flexibilidad les permita cortar residuos de GA/MeGA ubicados en posiciones internas de la cadena de xilano las coloca en un posición más ventajosa sobre las enzimas GH67, lo que les otorga una posible ventaja en la producción de XOS.

Tomando en cuenta el poco conocimiento que se tiene sobre las enzimas de la familia GH115, particularmente las de origen fúngico, sumado a su posible potencial biotecnológico en la producción de XOS en este seminario se propuso caracterizar bioinformáticamente y clonar la secuencia de DNA codificante de un potencial enzima  $\alpha$ -glucuruonidasa (XP\_007865968), denominada provisoriamente como Gt115, proveniente del hongo modelo de pudrición parda *Gloeophyllum trabeum* disponible en el Centro de Biotecnología y Bioingeniería (CEBIB). Esta enzima se ha anotado temporalmente como perteneciente a la familia GH115 (Floudas y col., 2012) y se ha detectado tanto en el transcriptoma como secretoma del hongo, pero sin haberse demostrado actividad  $\alpha$ -glucuruonidasa (Presley y Schilling, 2017; Presley y col., 2020).

En base a lo anteriormente dicho se propone realizar una investigación en tres etapas, siendo la primera etapa una revisión completa de las características funcionales y estructurales de las familias de enzimas GH67 y GH115 donde se busca compilar información sobre las características distintivas de cada familia como son su geometría de su sitio activo, arquitectura de dominios, topología, mecanismo catalítico, residuos catalíticos y especificidad de sustrato que sirvan como criterios de distinción que permitan diferenciar de forma más completa a las dos familias, valiéndose tanto de

literatura como el análisis de los modelos 3D de las enzimas de cada familia que se encuentran actualmente disponibles.

Una vez obtenida la información de las dos familias de  $\alpha$ -glucuronidasas se propone en una segunda etapa realizar un análisis bioinformático de Gt115 utilizando herramientas de alineamiento, modelamiento por homología y análisis proteico, comparando Gt115 con las enzimas modeladas y/o caracterizadas de las dos familias de  $\alpha$ -glucuronidasas mencionadas anteriormente y utilizando los criterios de distinción obtenidos en la primera etapa para conocer la pertenencia de Gt115 a una de las dos familias, así como también predecir sus características fisicoquímicas, su actividad y su especificidad de sustrato.

En una tercera y última etapa, se buscará aislar el transcrito de Gt115 desde *Gloeophyllum trabeum* e insertarlo en un vector plasmidial y obtener clones recombinantes de *Escherichia coli* que contengan este transcrito para su futura expresión en la levadura *Pichia pastoris*.

### **5.1 Objetivo general**

Obtener una predicción sobre la especificidad de sustrato y posible actividad  $\alpha$ -glucuronidasa de la enzima putativa Gt115 del hongo *Gloeophyllum trabeum* y clonar su secuencia de ADN codificante en *Escherichia coli*.

### **5.2 Objetivos específicos**

1. Obtener información sobre los rasgos estructurales y funcionales característicos de las enzimas pertenecientes a las familias GH67 y GH115 usando fuentes



bibliográficas y herramientas bioinformáticas para contrastarlos con la características halladas en Gt115

2. Predecir la posible pertenencia de Gt115 a las familias GH67 y GH115 analizando la secuencia del gen codificante para Gt115 por medio de herramientas bioinformáticas
3. Generar un modelo por homología de la estructura tridimensional de la proteína hipotética  $\alpha$ -glucuronidasa Gt115 del hongo *G. trabeum* para el análisis de su sitio activo, topología, arquitectura de dominios y la predicción de su actividad catalítica y especificidad de sustrato
4. Aislar y clonar en *E. coli* la secuencia de DNA codificante para la proteína hipotética  $\alpha$ -glucuronidasa del hongo *G. trabeum*

# Materiales y métodos

## 1. Técnicas de alineamiento

La secuencia proteica de la enzima Gt115 (Figura 4) fue identificada en el proteoma de la cepa de *Gloeophyllum trabeum* ATCC 11539, el cual se encuentra disponible en la base de datos BioProject bajo el código de acceso PRJNA64553. La secuencia fue utilizada sin modificaciones.

```
>XP_007865968.1 hypothetical protein GLOTRDRAFT_121308
MWAAYASWCVLLTLLTPAVRAIGQATCVSFSASSGVTSTFAVVSSGRAAPLLLSSEWPG
VQRAAGDFAADVQRVTGILPTLANFTSASGSGSGSSALAPIIVGTLGQSSLIAQVNNNTG
LDVSSIEGQWEAFLTREVQNPLPGISSAYVMIGADKRGTIYALYDHSEQMGVSPWYWWADV
TTRNSEIYFASSGCQHGAPTVMKYRGI FLNDEAPALMNWAMEKYTNGTGSAAATNSPLQSPYYT
GLFELLRLKANYLWPAMWSSAFYVDDPLNPYYADLYGIVMGTSHEEPMTRSTPVEWDIYGT
GPWDYSVNQQAIYNFWVQGAERSRNYETLYTVGMRGAGDLPLAEGQNIQLLEQVVSDQRGIL
ENVFNGTGKTI TEI PQMWCLYKEVQGYEDGMTVPDDITLMWTDDNWGNIRRFPLSSEMNR
GGAGVYYHFDYVGDPRNFKWITTTQIAKTYEEMSLAVDREATRIWIVNVGDMKPYEMDIEFF
ITLGWNSSVWNLNLDNFVTSWAEREFDVPTSTALQITGIIGNYTRMNARRPELMNSTVYT
LTDYREADTVLSEWQSLVQQSTQIYNLPAAYQPAFYELVHHTV LASANLNAMWIYAGVNNL
MGSQGRYAVNNYADLVQELFEQDYEYEQYHTLLDGKWDHMDQTHVVYVYWQEPPTNTMPM
VTRMQTKKQNIAGVMRVVPEGLGVWPGDNQYNCAQGYSCPNTMTLDPYIPVGNRYIDIGA
GGPAPFSWASTNVSWLNVSTTSGYIDSNNLETRVYITADWSQVSGLEYGMVTFNATTPNPP
TPGQTVLSQPVYFMANNTSPTSGFSGFVEGDGGISIEAAHTSRNTSVAGITWTEIPNYGKTV
SGVTPWPERTGDNGGNFTAGTGPSIEYDFYFNTIGQAGNVTVTYVSPSLNANGDDRPLGFA
IGIDDATPQSEYFVPYATPGQLPPQWDGVGWVADAIISVPTNFTIAPGAHTLKIWMIEPAV
VVQKIVIDTGGVRPSYLGPPESITV
```

Figura 4. Secuencia en formato FASTA de la enzima Gt115 de *Gloeophyllum trabeum*

La secuencia de Gt115 se alineó por medio de la herramienta BLASTP (Boratyn y col., 2012) en dos búsquedas distintas. En una primera instancia el alineamiento se hizo sobre todas las bases de datos de proteínas no redundantes, para esto se utilizaron los parámetros por defecto de BLAST correspondientes a un valor límite de E 0,05 (número de alineamientos obtenidos por coincidencia si la base de datos fueran secuencias al

azar) y word size de 6 (número mínimo de caracteres para que haya un emparejamiento entre dos secuencias), modificando la matriz de alineamiento a una BLOSUM90 con el objetivo de obtener resultados con mayor porcentaje de identidad a la secuencia ingresada. En una segunda instancia el alineamiento se hizo sobre la base de datos Protein Data Bank (PDB) (Berman y col., 2000), usando los parámetros por defecto de BLAST, pero excluyendo aquellas secuencias provenientes de muestras ambientales. De cada búsqueda se seleccionaron las 5 primeras secuencias con los menores valores E y el mayor puntaje de BLAST.

La herramienta Clustal  $\Omega$  (Madeira y col., 2022) se usó para generar alineamientos de los resultados obtenidos de BLAST y JALVIEW (Waterhouse y col., 2009) para generar las representaciones graficas.

## **2. Obtención, validación y refinamiento del modelo por homología**

La secuencia completa de Gt115 fue ingresada el servidor I-TASSER (Roy y col., 2008; Zhang, 2010) sin asignar ningún tipo de restricción y sin excluir modelos de otras enzimas. El modelo fue seleccionado según su puntaje C, que indica la confianza del modelo.

Para la validación se usó el servidor MolProbity (Williams y col., 2018) para evaluar los parámetros geométricos del modelo obtenido como ángulos  $\phi/\psi$ , calidad de los rotámeros, solapamiento estérico y desviaciones del carbono  $\beta$  todo para comprobar la calidad del modelo obtenido.

Para el refinamiento del modelo se utilizó el servidor YASARA (Krieger y col., 2009) para realizar una minimización de energía, para después ingresar el resultado al servidor Galaxy Refine (Heo y col., 2013; Lee y col., 2016) para hacer un refinamiento más fino y dirigido sobre las torsiones  $\phi/\psi$ . Este proceso se repitió 5 veces hasta que el modelo presentó parámetros de validación geométricos aceptables incluyendo torsiones de Ramachandran permitidas mayores al 90%, rotámeros favorecidos mayores al 98% (Laskowski y col., 1993) y Puntuación-Z de la distribución de Ramachandran entre -2 y 2 (Sobolev y col., 2020).

### **3. Análisis de modelos de proteínas**

El programa PyMol versión 2.4.1 (Schrödinger y DeLano, 2020) fue usado para visualizar todos los modelos de proteínas estudiados en este trabajo, así como también medir las distancias y ángulos de los aminoácidos por medio de la herramienta measure.

## **4. Cultivo de hongos y extracción de RNA**

### **4.1 Mantención de la cepa fúngica**

La cepa de *Gloeophyllum trabeum* fue donada por el profesor René Carmona de la Facultad de Ciencias Agronómicas de la Universidad de Chile. Desde la placa original se cortó un cuadrado de agar con micelio de 1cm x 1cm y se colocó sobre la superficie de una placa de medio PD agar fresco, donde fue cultivado por una semana a 27°C en oscuridad. Este proceso se repitió cada 2 semanas en orden de tener micelio fresco para los experimentos.

#### 4.2 Cultivo de *G. trabeum* para la extracción de RNA

Se eligieron alfalfa comercial (*Medicago sativa*) y aserrín de maqui (*Aristotelia chilensis*) proveniente de madera obtenida de un bosque renoval del sur de Chile como fuente de carbono para cultivar el hongo *G. trabeum*, e inducir la transcripción del gen codificante para la enzima Gt115. Se molió cada una de las fuentes de carbono con una batidora de mano y se tamizó en un tamiz de 1mm, recuperando partículas <1 mm. Estas partículas fueron autoclavadas a 120°C por 20 min y agregadas en un 2% m/v al siguiente medio mínimo de hongos (Tabla 2):

Tabla 2. Componentes de medio mínimo para hongos suplementado con alfalfa

Componente	Cantidad
KH <sub>2</sub> PO <sub>4</sub>	0,46 g/L
MgSO <sub>4</sub>	0,005 g/L
K <sub>2</sub> HPO <sub>4</sub>	1 g/L
Extracto de levadura	2 g/L
Alfalfa molida/Aserrín de maqui <1 mm	2%

Para facilitar la posterior extracción de RNA se colocó una membrana de nitrocelulosa con tamaño de poro de 0,45 µm sobre el agar, y el micelio se cultivó a 27°C en oscuridad sobre la superficie de la membrana por 2 semanas o hasta que el micelio hubiera alcanzado un crecimiento de 2 cm de radio desde la fuente original.

#### 4.3 Extracción total de RNA desde micelio

Para la extracción de RNA se tomó como base el protocolo propuesto Schumann y col. (2013) con la única variación de que el reactivo TRIzol se agregó en una proporción de 1 mL por cada 100 mg de micelio molido. La pureza del RNA extraído se comprobó por medio de la obtención de la razón A<sub>260</sub>/A<sub>280</sub> medida por un espectrofotómetro, donde un

valor de 1,8 hacia arriba indica un nivel alto de pureza sin contaminación de proteínas. El término  $A_{260}$  indica la absorbancia a 260 nm de DNA en solución y  $A_{280}$  la absorbancia a 280 nm de proteínas en la solución.

La integridad del RNA se observó por electroforesis cargando 1 mL del extracto de RNA puro en un gel de agarosa-formaldehído 1% corriendo el gel en tampón Tris-acetato-EDTA (TAE) libre de nucleasas a 150 V por 10 minutos, comprobando que puedan observarse las bandas correspondientes a los rRNA 28S y 18S del RNA ribosomal.

## **5. Síntesis de cDNA**

Para la síntesis de cDNA se usó el protocolo estándar de Thermo Fisher para la enzima Retrotranscriptasa SuperScript IV (2015a). En un tubo estéril de 1,5 mL se agregaron los siguientes componentes:

1. 1  $\mu$ L de Oligo d(T)<sub>20</sub> 50  $\mu$ M
2. 1  $\mu$ L de DNTPs 10 mM
3. RNA templado (5  $\mu$ g máximo)
4. Agua libre de nucleasas hasta completar 13  $\mu$ L

La mezcla se incubó por 5 min a 65°C y se dejó 1 min en hielo. Seguidamente se preparó el buffer de reacción mezclando en un tubo estéril de 1,5 mL los siguientes componentes

1. 5  $\mu$ L de Buffer SSIV
2. 1  $\mu$ L de DTT 100 mM
3. 1  $\mu$ L Transcriptasa reversa SuperScript IV 200 U/ $\mu$ L

El contenido de los dos tubos anteriores se mezcló y agitó, llevando a cabo la reacción a una temperatura de 55°C por 30 min y subiendo a 80°C por 15 min para inactivar la enzima.

## **6. Amplificación y aislamiento del transcrito del gen de Gt115**

### **6.1 Diseño de partidores**

Para aislar la secuencia de DNA codificante para Gt115 se diseñaron un par de partidores usando la herramienta primer BLAST (Ye y col., 2012) tomando como referencia la secuencia de RNA mensajero disponible en la base de datos NCBI (XM\_007867777.1), los partidores generados se muestran en la Tabla 3.

Tabla 3. Partidores diseñados para aislar la secuencia de DNA codificante de Gt115 desde cDNA total

Nombre del partidor	Sentido	Secuencia 5' a 3'
GH115_F	Forward	TACGCGTCTTGGTGCGTC
GH115_R	Reverse	ACTCTCAGGAGGACCGAGATAG

### **6.2 Amplificación por PCR**

Para la reacción de PCR se seleccionó la DNA polimerasa Phusion y se usó el protocolo sugerido por NEB (2021) generándose el siguiente programa mostrado en la Figura 5:

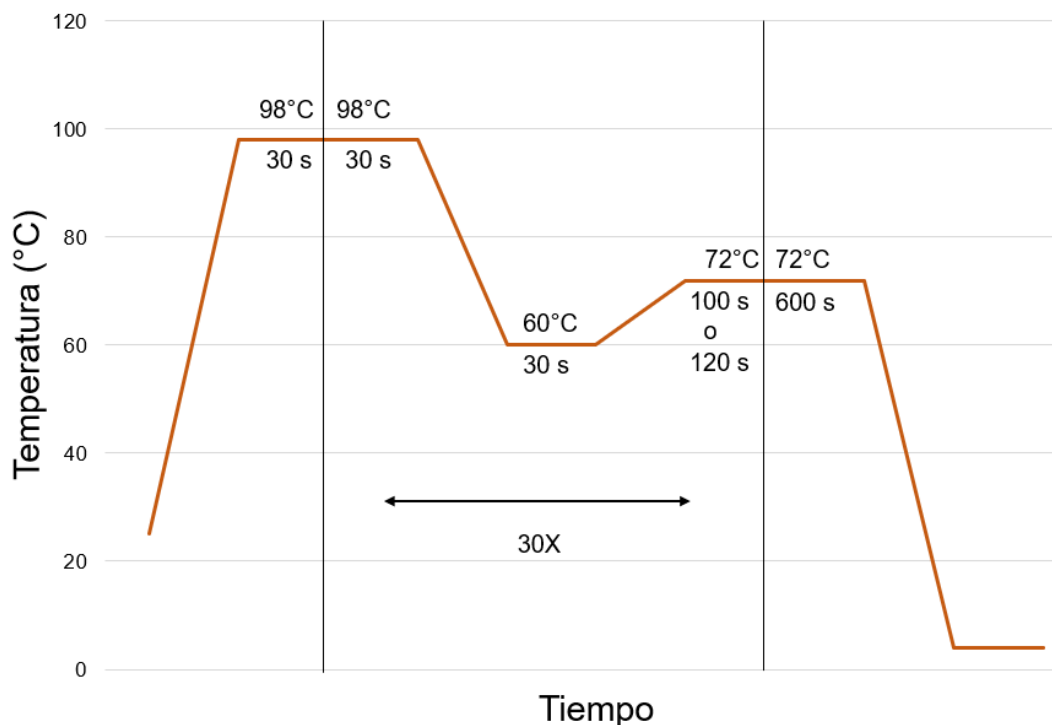


Figura 5. Programa de PCR para el aislamiento de la secuencia de DNA codificante para la enzima Gt115. La presencia del producto de PCR se observó por electroforesis en un gel de agarosa 1% (p/v) corrido por 50 minutos a 80 V en tampón Tris-acetato-EDTA (TAE), cargando de 1 a 5  $\mu\text{L}$  de la reacción de PCR de 20  $\mu\text{L}$  en el gel.

De manera separada una reacción de PCR posterior fue cargada en su totalidad en un gel de agarosa 0,8% (p/v) y corrido en las mismas condiciones anteriores, aislando el fragmento de DNA objetivo cercano a 3000 pb, purificándolo desde el gel usando el kit de extracción de geles GeneJet de Thermo Fisher (2015b).

## 7. Clonamiento de la secuencia de DNA codificante aislada

### 7.1 Ligación al vector pJET 1.2 Blunt

Para clonar el fragmento de DNA obtenido por PCR se utilizó el vector pJET 1.2 Blunt (Figura 6) (Thermo Fisher, 2018) que contiene un gen de resistencia a la ampicilina. Por



medio del gen letal *Eco47I/T7* que impide el crecimiento de colonias no recombinantes, por lo tanto, solo las células que tengan un vector que posea un inserto que interrumpa este gen podrán crecer en el medio selectivo.

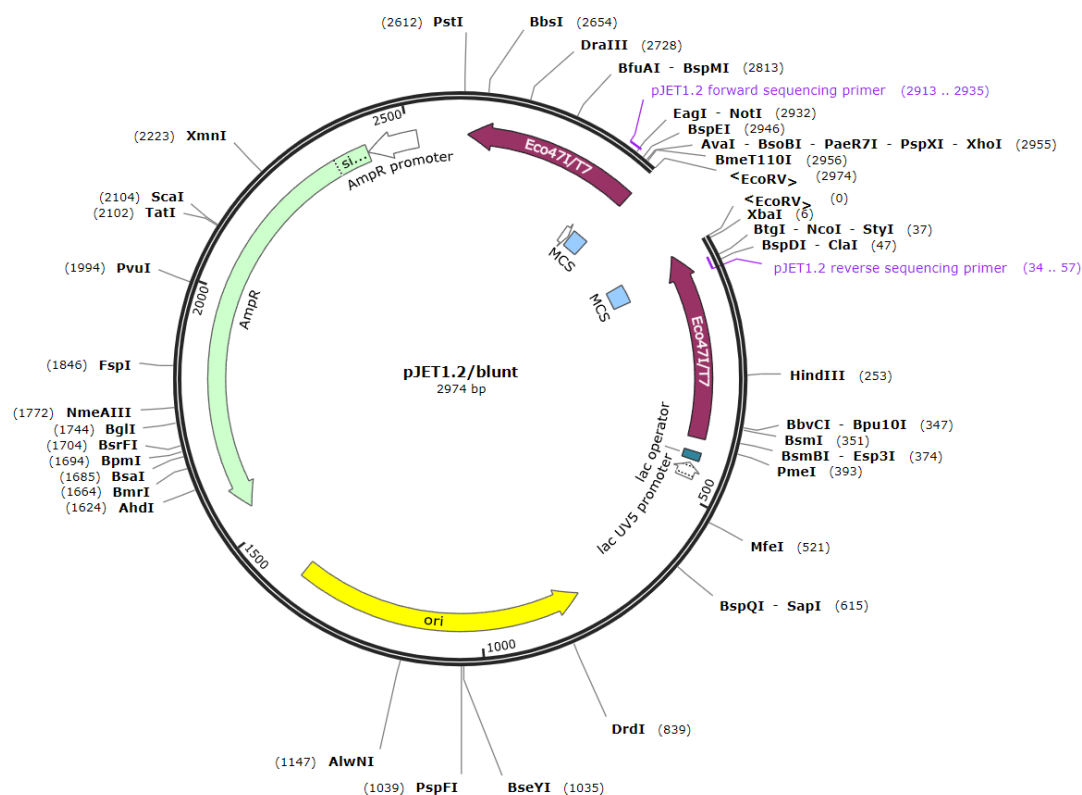


Figura 6. Mapa de restricción y características principales del vector pJET 1.2/Blunt en su forma abierta. Se utilizó el protocolo propuesto por Thermo Fisher (2012) mezclando en un tubo a un volumen total de 20  $\mu$ L los reactivos listados a continuación:

1. 4  $\mu$ L de buffer de ligación 5X
2. 1  $\mu$ L de DNA ligasa T4
3. 2  $\mu$ L de vector pJET 1.2/Blunt
4. 6  $\mu$ L de producto de PCR purificado
5. 7  $\mu$ L de agua estéril libre de nucleasas

La razón inserto:vector correspondió a la recomendada de 3:1 y la reacción fue incubada por 24 horas a 14°C.

## **7.2 Transformación del producto de ligación en *Escherichia coli* DH5α**

El procedimiento de producción y transformación de células electrocompetentes de *E. coli* DH5α fue adaptado del protocolo propuesto por Sambrook y col. (2001) con pequeñas modificaciones. Para la producción de células electrocompetentes se usó glicerol al 15% y todas las centrifugaciones se llevaron a cabo a 7.000 RPM.

Las células electrocompetentes fueron extraídas de su almacenamiento a -80°C y se descongelaron en hielo. Seguidamente una vez parcialmente descongeladas al tubo se le agregaron 2 µL de la mezcla de ligación diluida 5 veces, se agitó gentilmente y se transfirió a una cubeta de electroporación estéril, donde se aplicó un pulso de 2,50 kV usando el electroporador Gene Pulser Xcell Microbial System (1652662). Las células electroporadas fueron rápidamente transferidas a un tubo eppendorf de 1,5 mL con 500 µL de medio LB estéril para ser recuperadas por 2 horas a 37°C. Posterior a la recuperación se centrifugaron las células a 10.000 RPM y se descartaron 400 µL de medio, resuspendiendo las células en los 100 µL restantes. El volumen total de resuspensión se sembró en placas de medio LB agar con ampicilina 100µg/mL, siendo las colonias transformantes aquellas que puedan crecer en el medio selectivo.

## **8. Evaluación de colonias recombinantes**

### **8.1 Extracción de plásmido desde cultivo de bacterias**

De las placas de selección de ampicilina se seleccionaron al menos 2 colonias de bacterias, picándose con una punta plástica de pipeta estéril e inoculándose en tubos de

ensayo que contenían 5 mL medio LB líquido con una concentración de ampicilina de 1  $\mu\text{g/mL}$ . Los cultivos resultantes fueron usados en su totalidad para extraer el plásmido usando el Kit de Miniprep GeneJet de Thermo Fisher (2015b). La pureza del plásmido fue evaluada obteniendo la razón de absorbancias  $A_{260}/A_{280}$ , siendo un valor superior a 1,8 indicador de un plásmido puro sin contaminación proteica.

### **8.2 Digestión con enzimas de restricción**

Para comprobar la presencia de inserto se digirieron los plásmidos extraídos con la enzima *XmnI* en un reacción de 10  $\mu\text{L}$  donde se mezcló 1  $\mu\text{L}$  de enzima de restricción, 2  $\mu\text{L}$  de DNA, 1  $\mu\text{L}$  de tampón CutSmart de NEB y 6  $\mu\text{L}$  de agua estéril, siendo la reacción llevada a cabo por 1 hora a 37°C seguida por una fase de inactivación a 80°C por 15 minutos. Los resultados de la digestión fueron separados por electroforesis y evaluados en un gel de agarosa 1% corrido en tampón TAE a 80V por 50 minutos.

### **8.3 Secuenciación de plásmidos para cada colonia**

Para evaluar la presencia o ausencia de inserto dentro del plásmido se enviaron los productos de extracción al servicio de secuenciación de la Universidad Católica de Chile y los resultados evaluados usando el programa SnapGene (Insightful Science).

## Resultados

Este seminario tuvo como objetivo el análisis bioinformático y aislamiento de la secuencia de DNA codificante de una enzima con posible actividad  $\alpha$ -glucuronidasa llamada Gt115. Conforme a eso el trabajo se enfocó en tres etapas: La primera correspondió a un análisis bioinformático de las características geométricas, topología, arquitectura de dominios y sitios activos de las dos familias conocidas de  $\alpha$ -glucuronidasas GH115 y GH67, con el objetivo de recabar criterios de distinción que permitan clasificar a Gt115 dentro de una de las dos familias; La segunda, la generación de un modelo por homología que permita analizar la características geométricas y catalíticas de Gt115, así poder obtener detalles sobre su potencial actividad, especificidad de sustrato y sus posibles interacciones con este; La tercera y última, a el aislamiento de la secuencia de DNA codificante de Gt115 desde el hongo *Gloeophyllum trabeum*. En adelante se presentarán los resultados obtenidos de cada una de las tres etapas, así como también detalles de las estrategias que se usaron para obtenerlos.

### 1. Análisis del sitio activo y estructura general de las enzimas $\alpha$ -glucuronidasas

#### **GH67**

Las enzimas  $\alpha$ -glucuronidasas de la familia GH67 ( $\alpha$ GH67) están presentes en todos los dominios de la vida, identificándose más de 1150 enzimas repartidas entre bacterias, arqueas y hongos según la base de datos CaZy (Drula y col., 2022). A pesar de ello, solo 26 están caracterizadas y de ese total solo 3 tienen una estructura tridimensional resuelta de manera experimental, una correspondiente a *Pseudomonas cellulosa* y las otras dos (prácticamente idénticas) de dos cepas distintas de *Geobacillus stearothermophilus*.

Las enzimas  $\alpha$ GH67 son poco flexibles en su preferencia por sustrato, en general cortando residuos de MeGA en el extremo no reductor de XOS o polímeros de xilano cortos (Nagy y col., 2002). La poca flexibilidad en la unión del sustrato se debe principalmente a que el sitio activo de la enzima se encuentra en un bolsillo profundo e hidrofóbico (Figura 7) que puede albergar en su interior solo el residuo de MeGA junto al extremo no reductor de la cadena de xilano. El extremo reductor se ancla longitudinalmente en la parte más externa del bolsillo, donde se encuentran residuos aromáticos representados en forma de superficie en las zonas naranja y cian de la Figura 7 (Nurizzo y col., 2002).

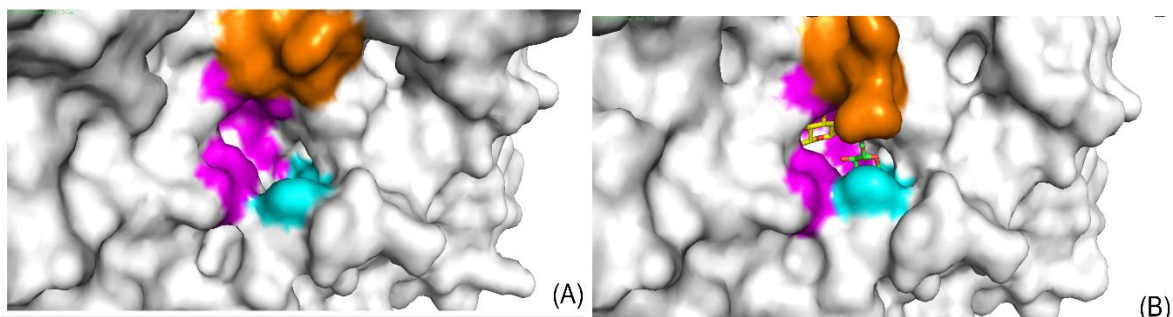


Figura 7. Sitio catalítico de la enzima GlcA67A de *Pseudomonas cellulosa* en conformación abierta (A) (1GQI) y en conformación cerrada (B) (1GQL). Los sitios de unión a xilano están coloreados en cian y naranja, el bolsillo catalítico coloreado en magenta. En B se muestran los productos de la hidrólisis ácido glucurónico (GlcA) en amarillo, y xilopiranososa (Xyl) en verde.

El análisis de la estructura tridimensional de las  $\alpha$ GH67 que han sido obtenidas experimentalmente sugiere que estas enzimas funcionan como homodímeros con una conformación de mariposa (Figura 8). Cada cadena del homodímero posee tres dominios, un dominio N-terminal formado por un  $\beta$ -sándwich y dos a cuatro  $\alpha$ -hélices, seguido de un dominio central caracterizado por poseer un dominio comúnmente conocido como barril TIM ( $\beta/\alpha$ )<sub>8</sub>, donde se encuentra su sitio activo y la mayoría de los residuos de unión a xilano. El C-terminal, por último, está compuesto principalmente de

un paquete  $\alpha$ -helices que se superponen sobre el dominio central y el bolsillo catalítico (Golan y col., 2004; Nurizzo y col., 2002).

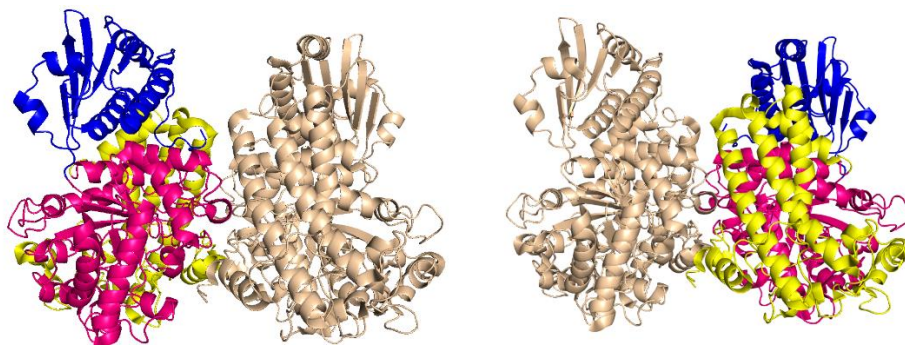


Figura 8. Estructura tridimensional del homodímero de GlcA67A de *Pseudomonas cellulosa* (1GQI) en formato caricatura con dos vistas giradas en 180° horizontalmente, en la cadena A el dominio N-terminal está coloreado en azul, el central catalítico de rosado y el C-terminal de amarillo, la cadena B está coloreada de color trigo. Modelo obtenido de Nurizzo y col., 2002.

La reacción catalítica se lleva a cabo mediante un mecanismo de inversión de la configuración del carbono anomérico. El sitio activo se ubica en el centro del barril TIM del dominio central donde la función y posición de los aminoácidos implicados está casi completamente conservada tanto en las enzimas de *P. cellulosa* como en *G. stearothermophilus* evidenciado en el alineamiento entre Glc67A y AguA en la figura 26 de la sección 1 del anexo, cuyos residuos conservados fueron anotados en la tabla 10 de la sección 2 del anexo. El bolsillo catalítico tiene dos subsitios para el anclaje de la cadena de xilano y un sitio de corte que se encuentra entre medio de estos dos, como se puede observar en la Figura 9:

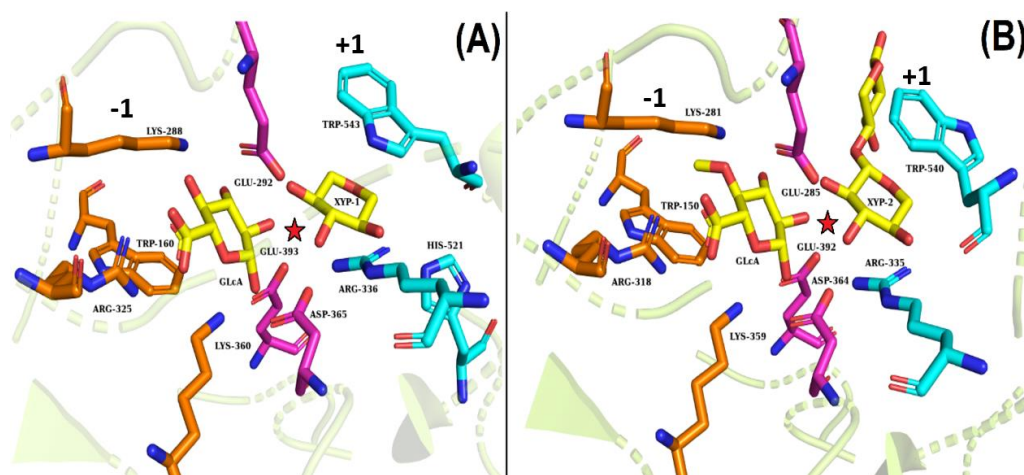


Figura 9. Modelo del sitio activo de enzimas  $\alpha$ GH67 con su estructura tridimensional experimentalmente resuelta. (A) Sitio activo de GlcA67A de *P. cellulosa* (1GQL). Los residuos del sitio -1 estan marcados en cian, los catalticos marcados en magenta, los pertenecientes al sitio +1 en naranja y los residuos producto del sustrato cortado (GlcA+xilopiranos) en amarillo. El sitio de corte se indica con un estrella roja (B) Sitio activo de AguA de *G. stearothermophilus*(1K9D).

El sitio +1 se encuentra en la parte ms externa de la enzima y contiene la mayora de los residuos de unin a xilano y es donde se une el extremo reductor de la cadena, este sitio contiene uno o ms residuos aromticos de Trp y uno o dos residuos con carga positiva correspondientes a His y/o Arg, que permiten el anclaje del extremo no reductor al sitio activo. Por otra parte, en la parte ms profunda del bolsillo se encuentra el sitio -1 donde se une el GA/MeGA por interacciones por residuos cargados positivamente, generalmente siendo Arg y Lys, y un residuo aromtico, generalmente Trp, que orienta el GA/MeGA (Golan y col., 2004; Nurizzo y col., 2002). El sitio de corte se encuentra entre los sitios -1 y +1 descritos anteriormente y contiene los residuos catalticos correspondientes a los cidos carboxlicos Asp y Glu.

## 2. Análisis del sitio activo y estructura general de las enzimas $\alpha$ -glucuronidasas

### **GH115**

Las enzimas  $\alpha$ -glucuronidasas de la familia GH115 ( $\alpha$ GH115), al igual que las  $\alpha$ GH67, están repartidas en todos los dominios de la vida, identificándose 1186 hasta la fecha según la base datos CaZy. Del total solo 8 están caracterizadas y de ellas solo 4 poseen su estructura 3D dilucidada experimentalmente, todas pertenecientes a bacterias, siendo estas *Amphibacillus xylanus*, *Bacteroides ovatus*, *Bacteroides thetaiotaomicron* y *Saccharophagus degradans* el alineamiento múltiple de las enzimas GH115 provenientes de estos organismos se puede encontrar detallado en la figura 27 de la sección 1 del anexo y sus aminoácidos conservados anotados en la tabla 11 de la sección 2 del anexo.

En comparación con las enzimas  $\alpha$ GH67, las  $\alpha$ GH115 poseen la capacidad de cortar residuos de GA/MeGA no solo del extremo no reductor de la cadena de xilano, sino que también de posiciones internas, esto debido a que en vez de tener un bolsillo catalítico profundo poseen un cavidad ancha (Figura 10) que es capaz de unir al xilano de manera superficial, pudiendo cortar a los sustituyentes que se encuentren en medio de la cadena (Rogowski y col., 2014).



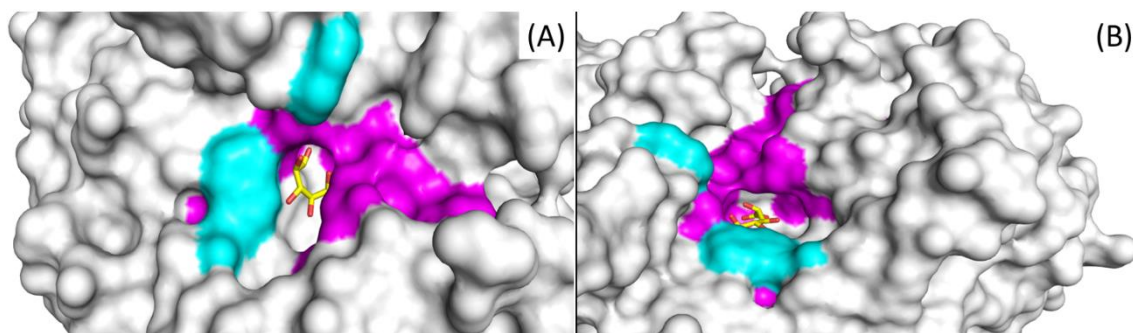


Figura 10. Cavity catalítica de la enzima  $\alpha$ GH115 BoAguA115 (4C91) de *B. ovatus*, los residuos del bolsillo catalítico están coloreados en magenta, los relacionados con la unión a la cadena de xilano en cian y la molécula asociada de GlcA en estado abierto se encuentra en el bolsillo catalítico coloreada en amarillo. En el panel A se muestra la cavity vista desde arriba y en el panel B, la cavity se muestra vista desde un lado.

Estructuralmente las enzimas  $\alpha$ GH115 son diversas. En los casos de BoAgu115A de *B. ovatus* y BtGH115 de *B. thetaiotaomicron* se ha observado una estructura de 4 dominios (Figura 11), que se organiza en un dominio N-terminal  $\alpha/\beta$  compuesto por 2  $\alpha$ -hélices y 6 hebras- $\beta$  (A), un dominio catalítico en forma de un barril TIM (B), un dominio intermedio (C) compuesto por un paquete de 5  $\alpha$ -hélices, y un dominio C-terminal (D) compuesto por un  $\beta$ -sándwich. Sin embargo, a pesar de la similitud estructural se ha observado que BoAguA115A adopta una estructura de homodímero en forma de mariposa (Figura 11A), mientras que BtGH115 se presenta como un monómero (Figura 11B). Esto se debe principalmente a dos razones: La primera se relaciona con el hecho que los dominios C y D de BoAgu115A presentan rigidez y estabilidad, lo que permite una mayor cantidad de puntos de contacto entre los dos dominios; En segundo lugar, el dominio D de BtGH115 se repliega hacia el interior, en dirección al centro catalítico (Figura 12), lo que disminuye el área de contacto que tendría con el otro monómero (particularmente con el centro catalítico de la otra cadena como es el caso de

BoAguA115A), por lo tanto haciendo más difícil la dimerización (Rogowski y col., 2014; Aalbers y col., 2015).

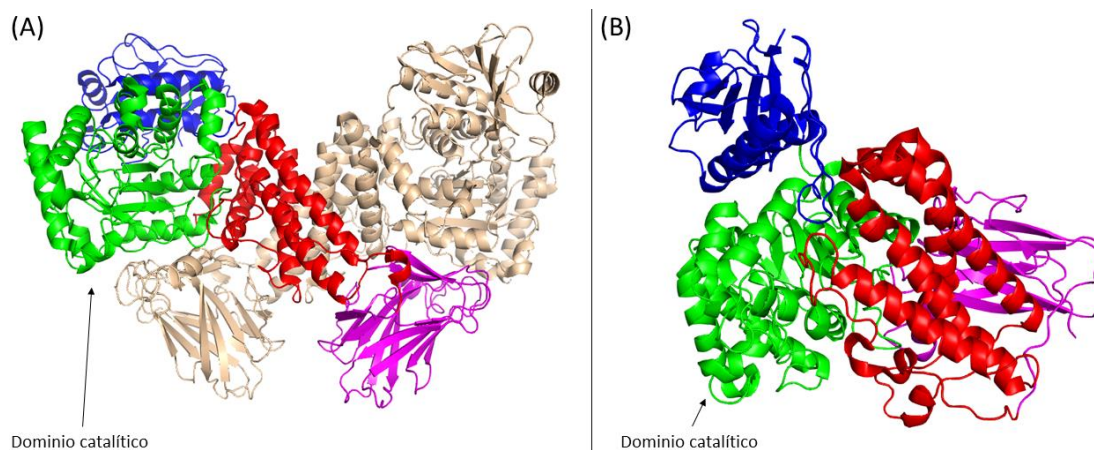


Figura 11. Modelo en vista caricatura de enzimas  $\alpha$ GH115 de 4 dominios (A) Enzima BoAgu115A (4C91), el dominio A está coloreado azul, el B (indicado como el dominio catalítico) coloreado en verde, el C coloreado en rojo, el D en magenta y la otra cadena coloreada en trigo. (B) Enzima BtGH115 (5BY3), el dominio A está coloreado azul, el B (indicado como el dominio catalítico) coloreado en verde, el C coloreado en rojo y el D en magenta.

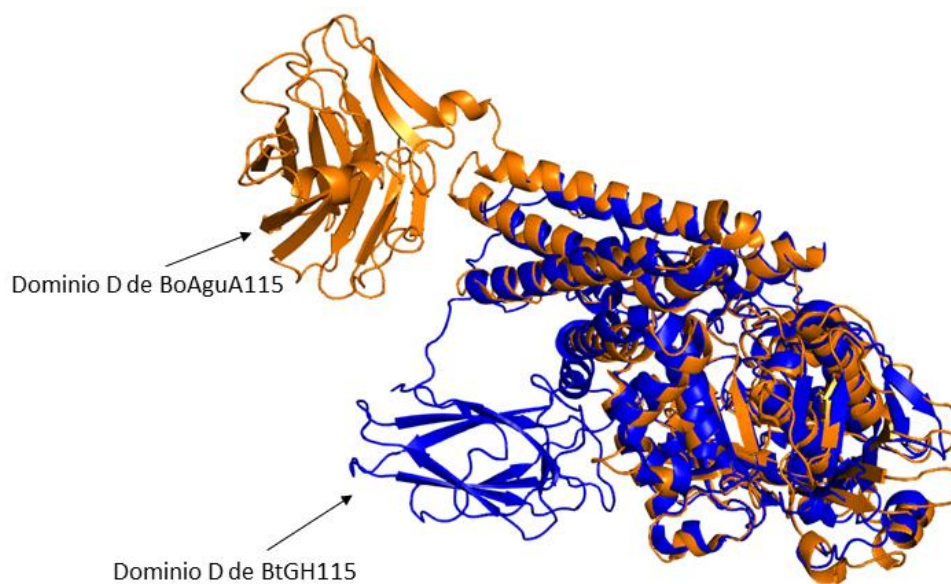


Figura 12. Alineamiento en vista caricatura entre BoAgu115A (coloreado en naranja) y BtGH115 (coloreado en azul) mostrando como difiere el dominio D de posición entre las dos enzimas. Modelos obtenidos de Aalbers y col., 2015 y Rogowski y col., 2014.

En el caso de AxyAgu115A y SdeAgu115A de *Amphibacillus xylanus* y *Saccharophagus degradans*, respectivamente, se observa una estructura de 5 dominios (Figura 13) similar a las observadas en BoAgu115A y BtGH115, con la diferencia de que estas enzimas poseen un fragmento adicional de aproximadamente 119 aminoácidos, ubicado entre los dominios C y D. Este inserto, denominado dominio C<sup>+</sup>, corresponde a un barril- $\beta$  con 8 hebras antiparalelas (Wang y col., 2016). En BoAgu115A las interacciones entre los protómeros se dan entre los paquetes de  $\alpha$ -hélices de los dominios C, y entre el dominio D y el dominio B'<sup>1</sup> de la otra cadena. En contraste, para AxyAgu115A y SdeAgu115A la presencia del dominio C<sup>+</sup> tiene como consecuencia que la interacción superficial entre los dominios C y C' se reduzca a un contacto de 5 aa entre dos hebras- $\beta$  antiparalelas y a contactos entre el dominio C<sup>+</sup> de una cadena con los dominios B' y C' de la otra (Rogowski y col., 2014; Wang y col., 2016).

---

<sup>1</sup> Para distinguir entre la notación de dominios de una proteína homodimérica se asigna un apóstrofe (') para marcarla como el dominio correspondiente a la otra cadena del dímero.

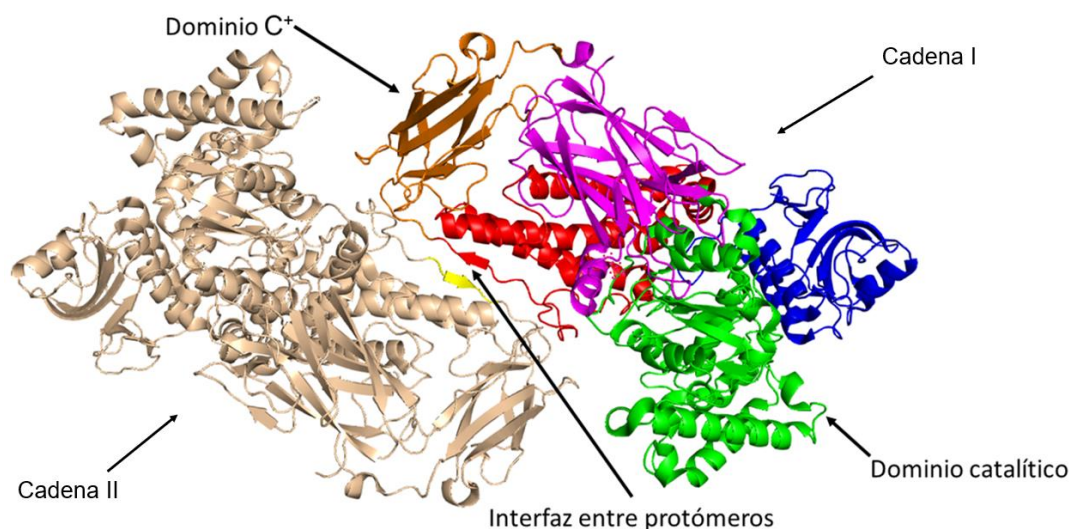


Figura 13. Modelo de enzima SdeAgu115A (SdeAgu115A). En la cadena (I) el dominio A está coloreado en azul, el B en verde, el C en rojo, el C<sup>+</sup> en naranja y el D en magenta. La cadena (II) está coloreada en trigo a excepción de la hebra- $\beta$  que interactúa con su análoga antiparalela de la cadena (I) coloreada en amarillo. La interfaz entre protómeros, el dominio catalítico y el dominio C<sup>+</sup> están indicados con flechas negras.

Funcionalmente, el mecanismo catalítico de las enzimas  $\alpha$ GH115 está parcialmente dilucidado (Wang y col., 2016). Para el caso de BoAgu115A, SdeAgu115A y AxyAgu115A, estas presentan una preferencia por sustratos GAX y XOS sustituidos con MeGA y comparten la gran mayoría de los residuos catalíticamente relevantes (Tabla 11 en el Anexo). La cavidad catalítica de estas enzimas posee un sitio +1 compuesto por al menos un residuo aromático de Trp, que sirve de plataforma hidrofóbica que une la cadena de xilano, y un sitio -1 que presenta las mismas características, los dos sitios de BoAgu115A se presentan en la Figura 14. El residuo de Trp que se encuentra en el sitio +1 es especialmente relevante ya que se ha observado que al ser reemplazado por un residuo de alanina reduce la eficiencia catalítica (kcat/KM) cerca de 2800 veces y por ello cumple un rol fundamental en unión y especificidad de sustrato (Aalbers y col., 2015; Rogowski y col., 2014).

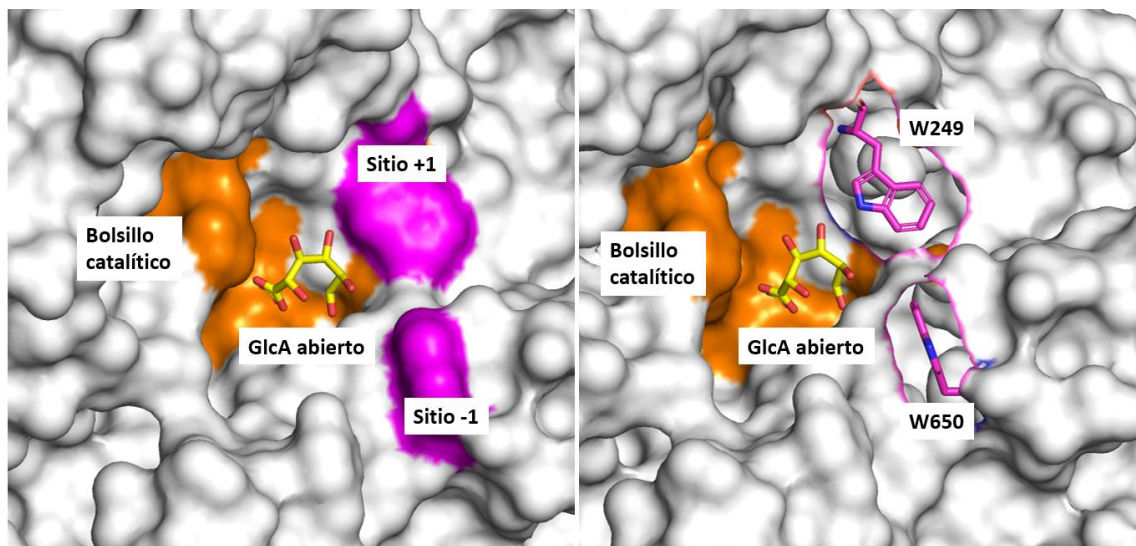


Figura 14. Surco catalítico de BoAgu115A (4C91). Los sitios +1 y -1 y sus respectivos residuos aminoacídicos se señalan en morado, el bolsillo catalítico se señala en naranja, y una molécula de GlcA abierto dentro del bolsillo catalítico se señala en amarillo.

En el bolsillo catalítico (Figura 15) se encuentra la diada de ácidos catalíticos Glu y Asp (D206 y E207 en BoAgu115A), que en la secuencia lineal ocupan posiciones consecutivas, y la base catalítica correspondiente a un residuo de Glu ubicado en un loop altamente móvil (D332 en BoAgu115A), que, según sugieren Yan y col., (2021), al momento de unir al sustrato se estabiliza en una posición más cercana a este para lograr la hidrólisis del residuo de MeGA de la cadena de xilano. En el sitio catalítico también se encuentran varios residuos con carga positiva, entre ellos His, Lys y Arg, cuya principal función es la unión a GA/MeGA y la modulación de pKa de la base catalítica. Otros residuos con carga negativa, principalmente Asp, sirven de moduladores de pKa para la diada de ácidos catalíticos (Rogowski y col., 2014).

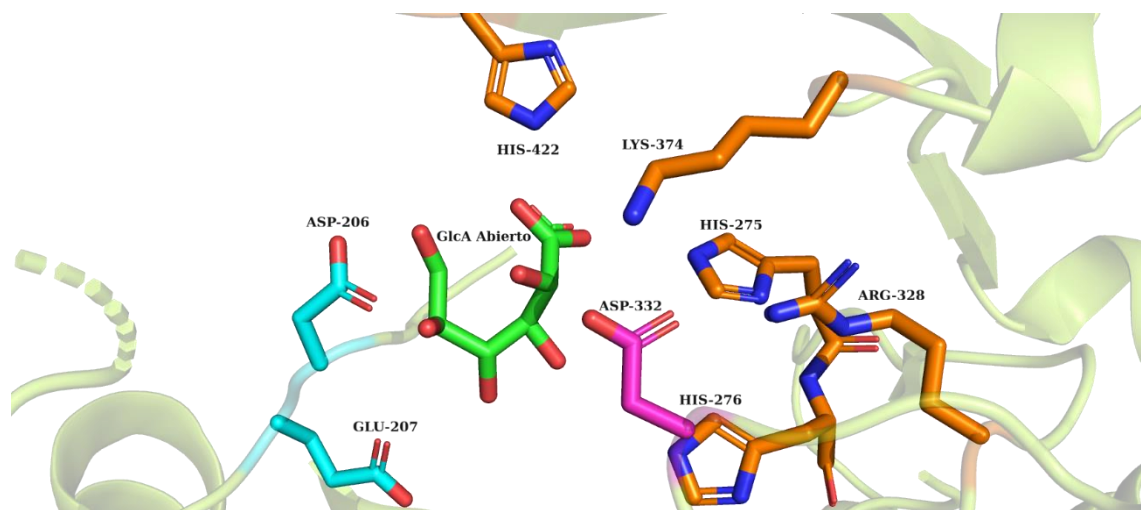


Figura 15. Vista del bolsillo catalítico de BoAgu115A. Mediante un alineamiento con 4C91 se le ha agregado una molécula de GlcA en su conformación abierta simulando las interacciones que tendría con todos los residuos modelados. El GlcA está coloreado en verde, la diada de ácidos catalíticos en cian, la base catalítica en magenta y los residuos asociados a la unión de GlcA en naranja.

Un caso distinto es el de BtGH115A, ya que esta enzima se diferencia de otras pertenecientes a la familia por el hecho de que se observó que no tiene actividad sobre AGX y/o XOS, sino que sobre polímeros de arabinogalactano (AG) presente en la goma arábiga. La diferencia funcional que posee BtGH115A se encuentra en los residuos de unión a sustrato, particularmente en una sustitución de Trp (descrito anteriormente en las otras enzimas) por His en el sitio +1, y en la presencia de un loop en el sitio -1 que desplaza otro residuo de Trp encargado de la unión a xilano. Estos dos cambios juntos provocan una cavidad catalítica más angosta que le otorga especificidad contra oligómeros cortos de arabinogalactanos y su presencia sirve como criterio para poder indicar la especificidad de sustrato de las enzimas de la familia GH115 (Aalbers y col., 2015).

### 3. Búsqueda de enzimas glucuronidasas mediante alineamiento masivo con BLASTp

Los resultados del primer alineamiento BLASTp de Gt115 sobre proteínas no redundantes de todas las bases de datos (Tabla 4) muestran que esta proteína está relacionada con otras de hongos basidiomicetos saprofitos. En todos los casos, a excepción de *H. sulcata*, las enzimas obtenidas están anotadas como pertenecientes a la familia GH115. Adicionalmente, La herramienta de identificación de dominios conservados de BLAST (Marchler-Bauer y col., 2017) identificó en Gt115 dos dominios conservados de la familia GH115 con un valor E muy cercano a 0, uno entre los residuos 203 y 551, y otro entre los residuos 832 y 1013, correspondientes al dominio catalítico y C-terminal observado antes respectivamente.

Tabla 4. Primeros cinco resultados del alineamiento en BLASTp de Gt115 sobre las secuencias no redundantes de todas las bases de datos

Especie	Acceso	% de Identidad	% de Cobertura	Valor E	Puntaje
<i>Heliocybe sulcata</i>	TFK48219.1	86,234	100	0.0	1901
<i>Neolentinus lepideus</i>	KZT21286.1	84,971	100	0.0	1900
<i>Hydnomerulius pinastri</i>	KIJ64385.1	66,434	98	0.0	1449
<i>Phlebiopsis gigantea</i>	KIP03027.1	64,370	99	0.0	1440
<i>Gelatoporia subvermispora</i>	EMD34896.1	64,110	100	0.0	1429

Para alineamiento sobre las proteínas de la base de datos PDB (Tabla 5) se obtuvieron solo cuatro resultados, correspondientes a las enzimas de la familia GH115 cuyas estructuras 3D habían sido obtenidas experimentalmente, correspondientes a las bacterias *Amphibacillus xylanus* (AxyAgu115A), *Bacteroides ovatus* (BoAgu115A), *Bacteroides thetaiotaomicron* (BtGH115A) y *Saccharophagus degradans*

(SdeAgu115A). En todos los casos se observa un porcentaje de identidad cercano o mayor al 30%, además de poseer un porcentaje de cobertura superior al 60% en todos los casos. Al igual que en el alineamiento anterior, la herramienta de dominios conservados de BLAST identificó en todas el dominio C-terminal y catalítico característicos de la familia GH115.

Tabla 5. Únicos cuatro resultados del alineamiento en BLASTP de Gt115 sobre las secuencias de la base de datos Protein Data Bank (PDB)

Especie	Acceso PDB	% de Identidad	% de Cobertura	Valor E	Puntaje
<i>Saccharophagus degradans</i>	4ZMH	40,82	97	0.0	723
<i>Bacteroides ovatus</i>	4C90	32,36	90	4e-151	471
<i>Amphibacillus xylanus</i>	6NPS	28,94	95	5e-122	397
<i>Bacteroides thetaiotamicron</i>	5BY3	32,56	60	3e-93	315

Posterior al alineamiento por BLAST la secuencia de Gt115 fue alineada con la enzimas encontradas en la búsqueda anterior con el programa Clustal  $\Omega$ , el alineamiento completo se encuentra en la figura 27 en la sección 1 del anexo y los residuos conservados anotados en la tabla 11 de la sección 2 del anexo.

#### 4. Obtención y refinamiento de un modelo por homología

Si bien el análisis realizado a la secuencia de Gt115 logró identificar características propias de la familia GH115, está consecuentemente se divide en distintas categorías con su propias características y propiedades como fue mencionado en el punto 2 de esta sección. El análisis de la secuencia si bien puede identificar a grandes rasgos dominios y residuos conservados entre enzimas no permite observar la geometría de la cadena polipeptídica, posicionamiento de aminoácidos en el sitio activo, ni su posible estructura



cuaternaria, por ello, se consideró pertinente generar un modelo por homología que permita identificar su posible estructura terciaria y cuaternaria y las propiedades del surco catalítico que son factores clave para identificar qué tipo de enzima GH115 es Gt115.

Se procedió a generar el modelo por homología se utilizó la herramienta I-TASSER (Roy y col., 2008; Zhang, 2010) que generó cinco modelos donde se eligió el de mayor confianza basado en su C-score (0,20). Los primeros cuatro resultados de estructuras conocidas obtenidas por alineamiento fueron coincidentemente dos de las proteínas obtenidas del alineamiento BLASTP detallado en la Tabla 5, correspondientes a las cadenas A de SdeAgu115A y AxyAgu115A Asimismo, se pudo observar que los primeros cuatro resultados de un análisis de homología estructural correspondían también a los resultados de la Tabla 5, siendo la cadena A de AxyAgu115A la proteína que estructuralmente más se asemeja al modelo obtenido de Gt115.

Para validar el modelo que se obtuvo de Gt115 se le sometió a un análisis usando el servidor de validación de modelos de proteínas MolProbity (Williams y col., 2018) el cual analiza la geometría del modelo 3D, generando un gráfico de Ramachandran observable en la figura 16 (Ramachandran y col., 1963) que visualiza las torsiones de los ángulos dihedros  $\psi/\phi$  de cada residuo de la proteína, clasificándolas en regiones energéticamente favorecidas, permitidas y no permitidas, además de proveer otros parámetros de validación como la calidad de los rotámeros (basado en la rotación  $\chi$  de las cadenas laterales), solapamiento estérico y las desviaciones de la posición del carbono  $\beta$  de cada residuo comparadas con su posición ideal. Usualmente el gráfico de

Ramachandran y la Calidad de los rotámeros son los criterios que tienen más peso en el momento de evaluar la calidad del modelo ya que generalmente no se restringen durante la generación del modelo y el refinamiento, y que representan las posiciones de la cadena principal y de las cadenas laterales respectivamente (Wlodawer, 2017).

El análisis por MolProbity (Tabla 6) del primer modelo indica que no cumple con ninguno de los criterios considerados para la validación, siendo notable un alta cantidad de posiciones no permitidas en el gráfico de Ramachandran y de rotámeros pobres. Conforme a esto, se decidió realizar una refinamiento del modelo de Gt115 en orden de obtener una estructura más confiable para el análisis posterior.

Tabla 6. Resultados del análisis del modelo de homología de Gt115 sin refinar obtenido con MolProbity. Los resultados por sobre el estándar de validación están marcados en verde, los cercanos al estándar de validación en amarillo y los que no cumplen el estándar de validación en rojo.

Contactos de todos los átomos	Sobrelapamiento estérico (> 0.4 Å por cada 1000 átomos)	17.42		Percentil 40 <sup>1</sup>
Geometría de la proteína	Rotameros pobres	132	15.55%	Meta: <0.3%
	Rotameros favorecidos	590	69.49%	Meta: >98%
	Posiciones no permitidas	111	10.94%	Meta: <0.05%
	Posiciones favorecidas	752	74.09%	Meta: >98%
	Puntuación-Z de la distribución de Ramachandran	-5.21 ± 0.21		Meta: abs(Z score) < 2
	Puntuación de MolProbity	3.45		Percentil 10 <sup>2</sup>
	Desviaciones del Cβ >0.25Å	106	11.43%	Meta: 0

1. Corresponde al percentil en el que está el modelo según su calidad comparándolo con otros modelos de resolución<sup>2</sup> similar 2. Puntuación que combina los criterios de sobrelapamiento estérico, calidad de rotámeros y el gráfico de Ramachandran

El refinamiento del modelo de Gt115 se realizó de manera secuencial, empezando primero por una minimización de energía usando el servidor YASARA (Krieger y col., 2009), para después ingresar el resultado en el servidor Galaxy Refine (Heo y col., 2013;

<sup>2</sup> Para un modelo por homología la resolución de este depende de los modelos templado que se usen (Roy y col., 2008; Zhang, 2010), tanto SdeAgu115A y AxyAgu115A fueron usadas como templado para generar Gt115, por ello tendría una resolución equivalente cercana a 2 Å.

Lee y col., 2016), donde se ejecuta un refinamiento usando un método de perturbación y relajación secuencial. Se espera que estos pasos generen una estructura más cercana a una conformación nativa.

La estrategia de usar dos herramientas distintas se consideró debido a que el servidor YASARA optimiza solo los rotámeros de cada residuo aminoacídico y no los ángulos  $\phi/\psi$  relacionadas con las posiciones del gráfico de Ramachandran, por lo que se usó el servidor dedicado el refinamiento de modelos de proteínas Galaxy Refine como un segundo paso para optimizar las torsiones  $\phi/\psi$  del modelo, además de refinar los rotámeros derivados del servidor Yasara eligiendo los de mayor probabilidad. Como resultado final del refinamiento (Tabla 7) se observa una mejoría notable en el solapamiento estérico, posiciones permitidas de aminoácidos, la calidad de los rotámeros y la puntuaciones Rama-z y Molprobity (que reúnen todas las características anteriores en una sola puntuación) si se compara con el modelo sin refinar.

Tabla 7. Resultados de MolProbity obtenidos del análisis del modelo de Gt115 proveniente del servidor I-TASSER después de cinco ciclos de optimización. Los resultados por sobre el estándar de validación están marcados en verde, los cercanos al estándar de validación en amarillo y los que no cumplen el estándar de validación en rojo.

Contactos de todos los átomos	Solapamiento estérico (> 0.4 Å por cada 1000 átomos)	7.35		Percentil 85 <sup>1</sup>
Geometría de la proteína	Rotámeros pobres	3	0.35%	Meta: <0.3%
	Rotámeros favorecidos	839	98.82%	Meta: >98%
	Posiciones no permitidas	20	1.97%	Meta: <0.05%
	Posiciones favorecidas	949	93.50%	Meta: >98%
	Puntuación-Z de la distribución de	-1.33 ± 0.23		Meta: abs(Z)

	Ramachandran			score) < 2
	Puntuación de MolProbity		1,83	Percentil 84 <sup>2</sup>
	Desviaciones del C $\beta$ >0.25Å		35	0.65%
				Meta: 0

1. Corresponde al percentil en el que está el modelo según su calidad comparándolo con otros modelos de resolución similar 2. Puntuación que combina los criterios de solapamiento estérico, calidad de rotámeros y el gráfico de Ramachandran

Los gráficos de Ramachandran para el modelo sin refinar y el modelo refinado se presentan en la Figura 16:

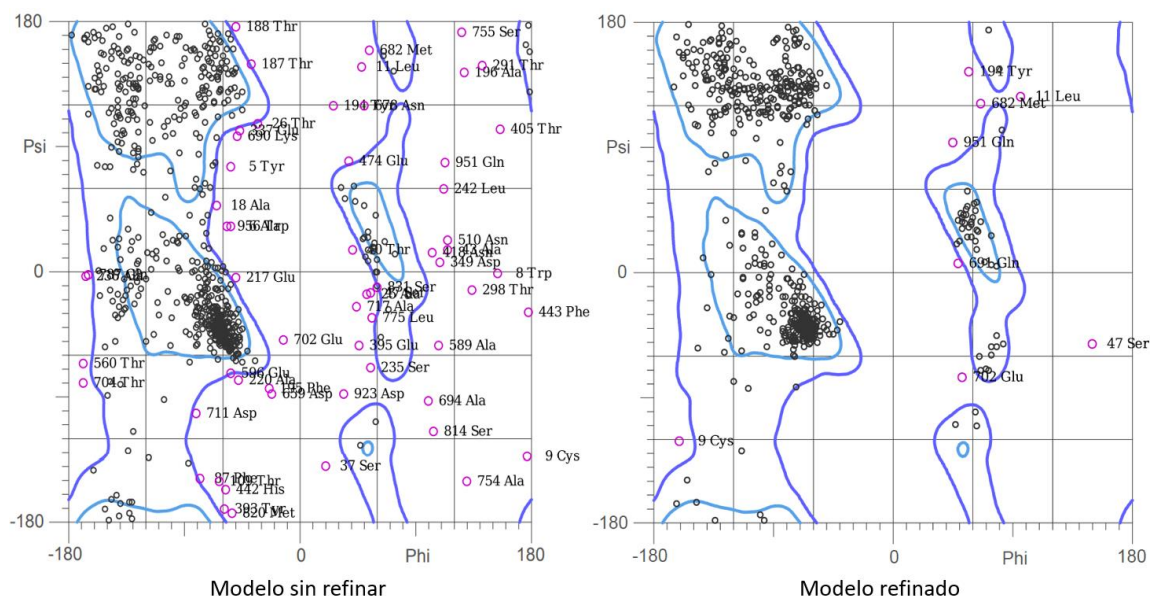
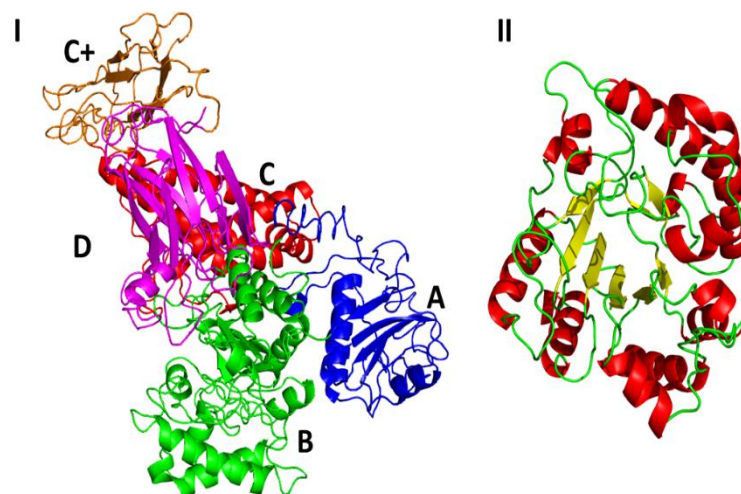


Figura 16. Gráfico de Ramachandran para los modelos por homología de Gt115 sin refinar (derecha) y refinado (izquierda). Los aminoácidos en posiciones no permitidas se señalan con círculos morados.

Hecho esto, usando el gráfico de Ramachandran obtenido de MolProbity se comprobó que los residuos conservados no estuvieran en posiciones no permitidas o que presentaran otros problemas geométricos (Tabla 12 en el anexo), observando que la mayoría de ellos presentan ángulos  $\phi/\psi$  favorables o permitidos en el gráfico de Ramachandran menos el residuo W708 cuyas torsiones no están permitidas. Adicionalmente, los residuos W267, E294, E295, R345, W708, Y714 presentan solapamientos estéricos ligeramente por sobre el corte de 0,4 Å.

## 5. Análisis del sitio activo y estructura general del modelo por homología de Gt115

Se puede observar en el modelo de Gt115 obtenido, que la proteína posee 5 dominios (Figura 17I), donde se distingue un dominio A (aminoácidos 1-202), B (aminoácidos 203-505), C (aminoácidos 506-686), C+ (aminoácidos 687-829) y D (aminoácidos 830-



1017), por ello pertenecería a la subclasificación de glucuronidasas de 5 dominios.

Figura 17. Modelo por homología refinado obtenido para Gt115 (XP\_007865968). En (I) se observa la estructura completa donde cada dominio está indicado con su correspondiente nombre según el consenso de enzimas GH115 de 5 dominios. En (II) se ve solo dominio B donde se encuentra el surco y bolsillo catalítico.

En una observación más cercana, el dominio que debería contener el surco y bolsillo catalítico presenta una estructura de Barril TIM deformado (Figura 17II) ya que existe un número de mayor de  $\alpha$ -hélices más cortas que rodean el barril central formado por hebras- $\beta$ . Se puede notar que la tercera y cuarta hebra- $\beta$  del barril son más cortas de lo normal, lo que es una característica propia de las enzimas GH115, similar a lo que se observa en SdeAgu115A (Wang y col., 2016).

Respecto al sitio de unión del sustrato, en el modelo de Gt115 se muestra un surco catalítico que correspondería a la cavidad ancha donde se aloja la cadena de xilano. El

sitio activo de  $\alpha$ -glucuronidasas de la familia GH115 usualmente presenta subsitios denominados -1 y +1, que son los que interactúan con el sustrato por medio de residuos de Trp que funcionan como plataforma hidrofóbica que permiten anclar la cadena de xilano por medio de interacciones de apilamiento interactuando con los residuos de xilosa que la componen.

Si se compara Gt115 con AxyAgu115A (la estructuralmente más cercana), se puede notar que el que surco catalítico es más largo en Gt115 que en AxyAgu115A (Figura 18) siendo la distancia aproximada (de anillo aromático a anillo aromático entre los residuos de Trp) que separa los subsitios +1 y -1 es casi el doble en Gt115 (18,4 Å) comparada a AxyAgu115A (9,5 Å). La misma diferencia puede observarse con respecto a SdeAgu115A (10,5 Å) y BoAgu115A (10,1 Å), sin embargo, para el caso de BtGH115A la distancia es similar (15,8 Å) entre los residuos equivalentes de His (sitio +1) y Phe (sitio -1).

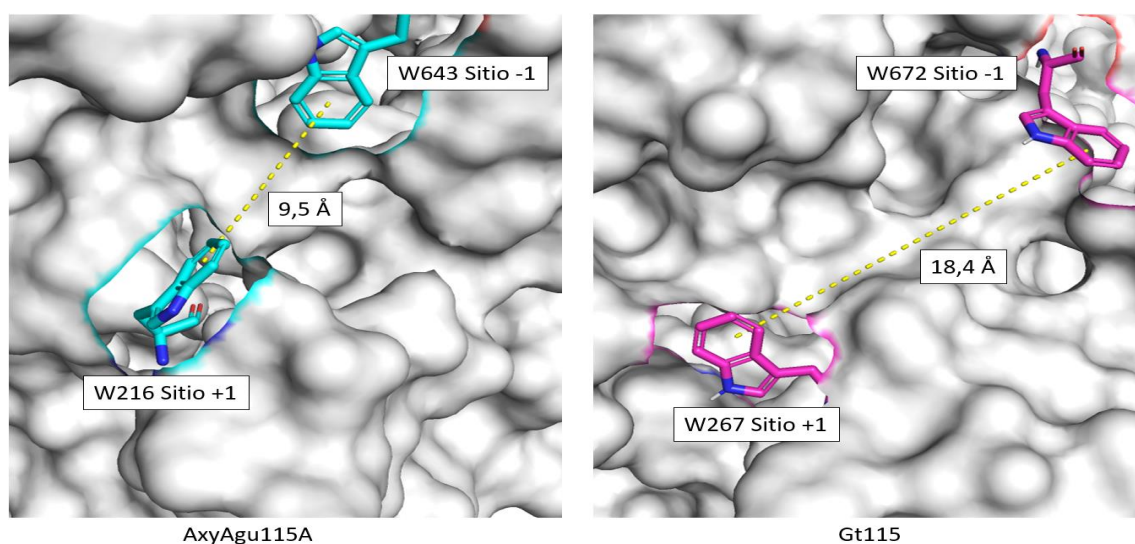


Figura 18. Distancias de los residuos de Trp de los sitios +1 y -1 de las enzimas AxyAgu115A y Gt115. La distancia fue medida en Armstrong de anillo aromático a anillo aromático. Los sitios +1 y -1 de AxyAgu115A están rotulados y coloreados cian y los de Gt115 en morado.

Se observa también que la rotación  $\omega$  de los residuos de Trp<sup>3</sup>, es diferente para cada sitio (Figura 19). W267 En el sitio +1 de Gt115 posee una rotación de 81,3°, mientras que para su equivalente W216 en AxyAgu115A posee un ángulo de 8,1°. Lo mismo ocurre para los residuos del sitio -1, donde W672 de Gt115 tiene una rotación  $\omega$  de -9,7° mientras que W643 de AxyAgu115A tiene una de -91,6°.

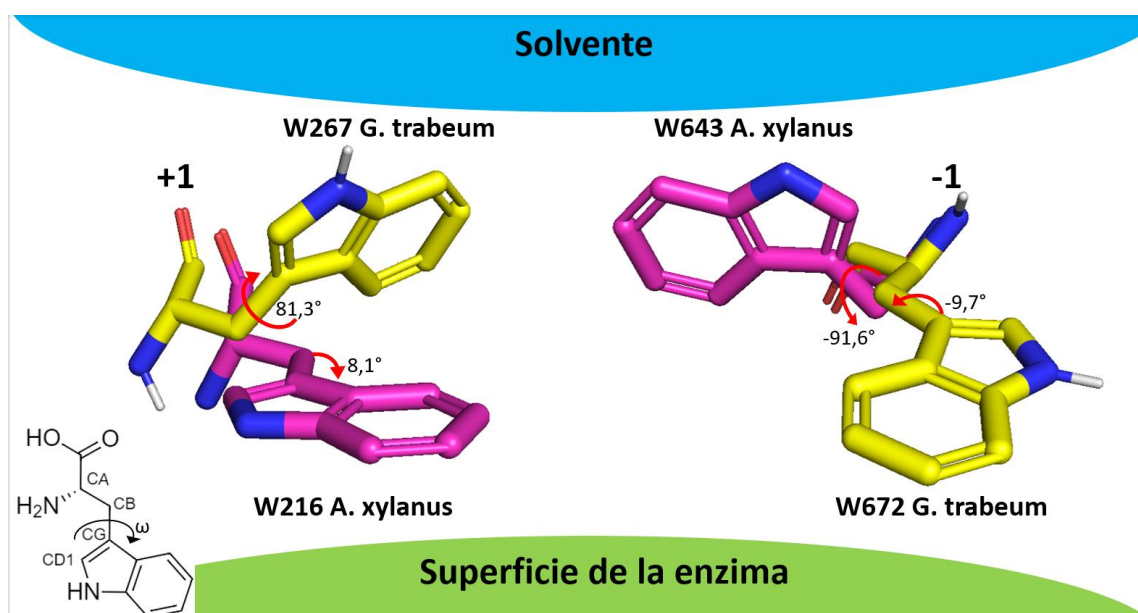


Figura 19. Comparación en los ángulos  $\omega$  de los residuos de triptófano de los sitios de unión al sustrato +1 y -1 de Gt115 (*G. trabeum*) y AxyAgu115A (*A. xylanus*). Para definir la orientación de las caras de los anillos aromáticos en la parte inferior se indica la superficie de la enzima y en la parte superior la parte donde se encontraría el solvente en condiciones biológicas. Los ángulos  $\omega$  para cada aminoácido se muestran con flechas rojas.

En la Tabla 8 se anotan las rotaciones  $\omega$  para los residuos de triptófano análogos para W267 y W672 de Gt115:

<sup>3</sup> El ángulo  $\omega$  corresponde al ángulo diedro entre el carbono  $\alpha$  y el carbono CD1 del anillo del Trp. No debe confundirse con el ángulo  $\omega$  del enlace peptídico, en este caso la anotación  $\omega$  se utiliza para ser consistente con los estudios que analizan esta torsión específica que sería equivalente a la torsión  $\chi_2$  usada para los rotámeros de triptófano.

Tabla 8. Rotaciones  $\omega$  medidas para los residuos de unión a xilano en los sitio +1 y -1 para el modelo de Gt115 y la enzimas AxyAgu115A (6NPS), SdeAgu115A (4ZMH) BoAgu115A (4C90) y BtGH115A (5BY3)

Enzima	Aminoácido sitio +1	Aminoácido sitio -1	Rotación $\omega$ sitio +1	Rotación $\omega$ sitio -1
AxyAgu115A	W216	W643	8,1°	-91,6°
SdeAgu115A	W256	W656	14,5°	-89,5°
BoAgu115A	W249	W650	25°	-87,2°
BtGH115A	H216	-	25,5°	-
Gt115	W267	W672	81,3°	-9,7°

Al analizar el modelo del bolsillo catalítico de Gt115 (Figura 20) se pudo observar una disposición similar a la de BoAgu115A vista anteriormente (Figura 15).

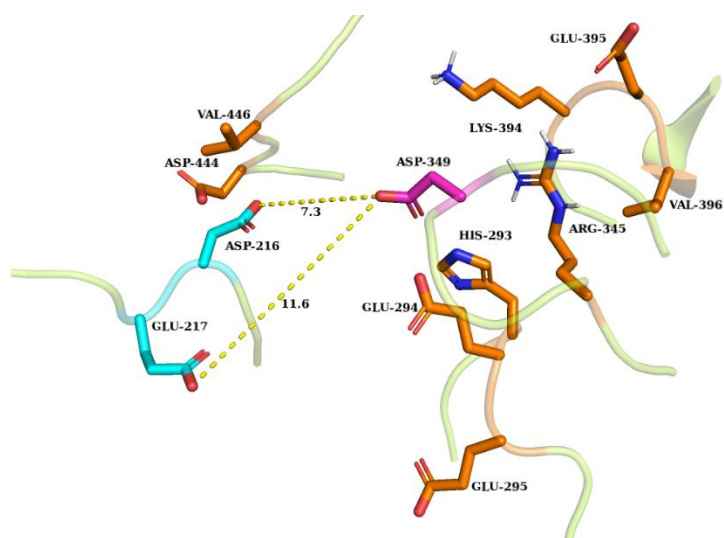


Figura 20. Bolsillo catalítico de Gt115. La diada de ácidos catalíticos se muestra en cian, la base catalítica en magenta y los residuos de unión a GA/MeGA en naranja. La distancia entre la diada y la base catalítica está señalada por líneas amarillas discontinuas.

En la tabla 9 se muestran las distancias entre la diada de ácidos catalíticos y la base catalítica para cada enzima bacteriana de la familia GH115 y para Gt115. Según indican Davies y Henrissat (1995) la distancia entre residuos catalíticos de enzimas con mecanismo de inversión es de 10 Å, lo que se cumple para D216 de Gt115, como



también para las otras enzimas con la excepción de SdeAgu115A cuyo loop móvil donde se encuentra la base catalítica (321-333) se encuentra más separada de la diada de ácidos catalíticos que lo visto en las otras enzimas.

Tabla 9. Distancia entre los residuos de la diada de ácidos catalíticos AC1 (Acido aspártico que va primero en la secuencia) y AC2 (Acido glutámico que va segundo en la secuencia) con respecto a la base catalítica BC (Ácido aspártico que se encuentra en el loop móvil característico de la enzima).

Enzima	Distancia AC1/BC (Å)	Distancia AC2/BC (Å)
AxyAgu115A	9,5	12,8
SdeAgu115A	22,3	22,7
BoAgu115A	9,0	12,6
BtGH115A	10,3	14,3
Gt115	7,4	14,1

A modo de síntesis el análisis del surco catalítico del sitio activo revela que Gt115 posee características propias de la familia GH115, sin embargo, presenta rotaciones  $\omega$  en aminoácidos esenciales para la unión a xilano distintas a las observadas en otras enzimas en aminoácidos esenciales para la unión, así como también un mayor distancia entre sus sitios de unión al sustrato +1 y -1 lo que implica un surco catalítico más largo. La estabilidad de las rotaciones  $\omega$  en el tiempo quedan por ser comprobadas con resultados experimentales o bien simulaciones de dinámica molecular ya que la optimización del modelo resultante podría corresponder a un mínimo local de energía y no a un mínimo absoluto.

## 6. Identificación y aislamiento de la secuencia para la glucuronidasa Gt115

Para inducir la expresión de celulasas y hemicelulasas en *G. trabeum* se han usado sustratos vegetales como madera o bagazo (Presley y Schilling, 2017; Valadares y col. 2019), por ello se eligió cultivar a *G. trabeum* en placas de medio mínimo para hongos

suplementado con alfalfa molida comercial o aserrín de maqui estéril, sobre membranas de nitrocelulosa que permitiera una fácil extracción del micelio.

En las pruebas iniciales después de 2 semanas solo se observó crecimiento significativo del hongo en las placas con alfalfa + membrana y agar expuesto sin membrana (radio de crecimiento  $> 1$  cm), mientras que en las que contenían aserrín no se observó crecimiento, por lo que la alfalfa fue elegida la fuente de carbono para el crecimiento de *G. trabeum* e inducción de la producción de glicosil hidrolasas.

Después de 2 semanas de cultivo se procedió a hacer la extracción de RNA desde el micelio crecido en alfalfa. De tal extracción se obtuvieron 20  $\mu$ L de RNA en solución en un concentración de 288 ng/ $\mu$ L y con una relación  $A_{260}/A_{280}$  de 2,02 indicando que el RNA extraído era de buena calidad sin contaminación de proteínas. Seguidamente, del RNA se tomaron 5  $\mu$ L y se corrió en un gel de agarosa/formaldehido 1% en orden de ver la integridad del RNA, como se muestra en la Figura 21:

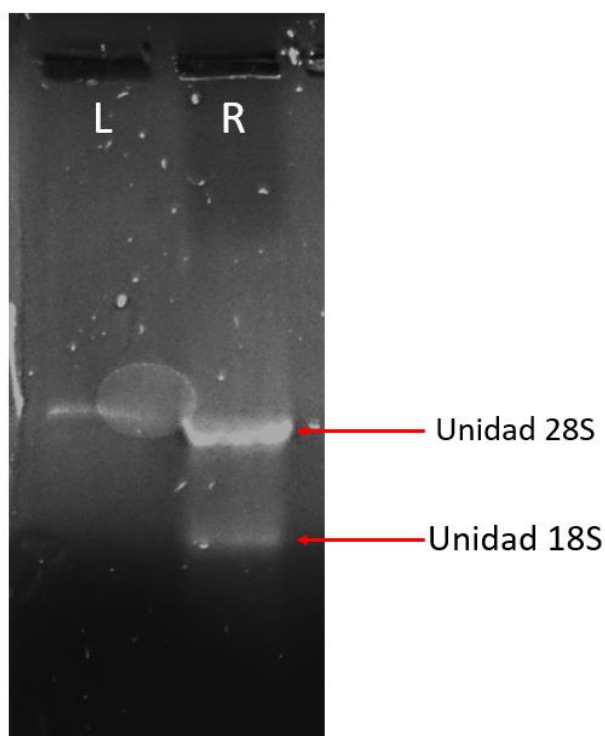


Figura 21. Producto de la extracción de RNA total de *G. trabeum*, analizado en un gel de agarosa/formaldehído 1%. En el carril L se encuentra un estándar de 1 kb y en el R la muestra de 5  $\mu$ L de RNA total extraído. Se señalan las bandas correspondientes a los rRNA 28S y 18S

En el gel se observan las dos bandas correspondientes a los rRNA 28S y 18S, evidenciando la calidad del proceso de extracción. Del estándar de peso molecular se distingue una sola banda a la altura de la unidad 28S debido a las condiciones desnaturalantes del gel producidas por el formaldehído.

El RNA total obtenido se usó para sintetizar cDNA, obteniéndose en una concentración de 332 ng/ $\mu$ L con una razón de absorbancia de  $A_{260}/A_{280}$  de 1,83, estando por sobre el estándar de pureza de 1,80. Desde el cDNA obtenido se realizó PCR en gradiente obteniendo un producto cercano a las 3 kb del mismo largo predicho para el gen buscado, de acuerdo con la secuencia de ADN. Además, se observaron varios productos

inespecíficos menores a 500 pb que aparecen de manera variable según la temperatura de annealing, como se puede observar en la Figura 22:

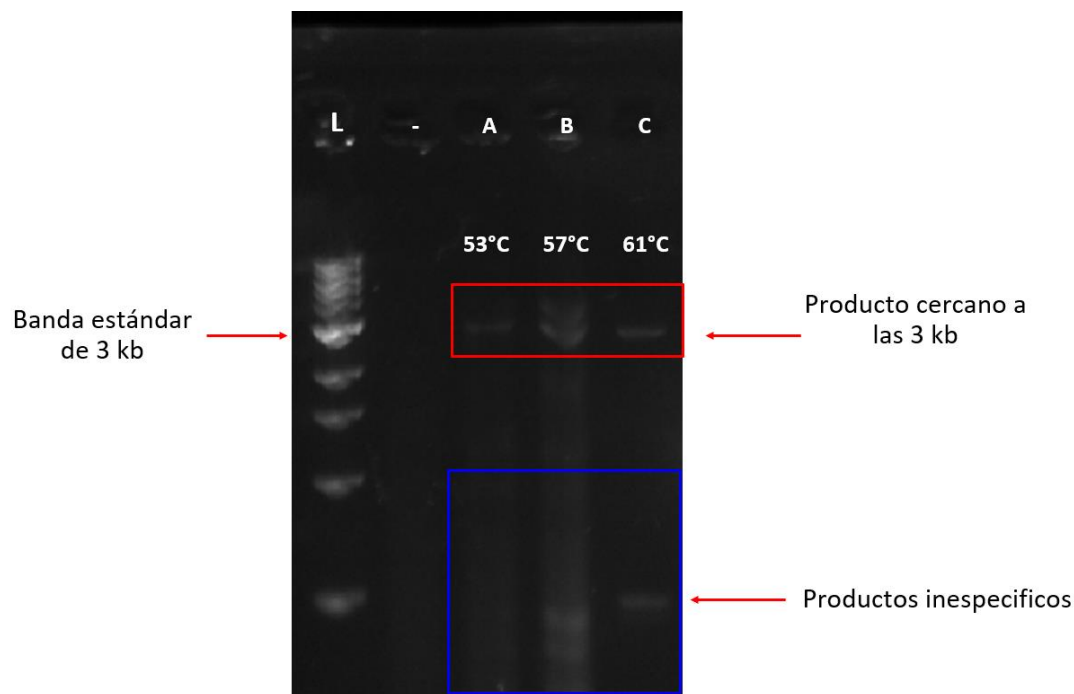


Figura 22. Electroforesis del resultado del RTPCR realizado para aislar secuencias codificantes de Gt115. En agarosa 1%. Se cargaron las muestras de PCR de gradiente a partir de cDNA total de *G. trabeum*. L: Estándar de peso molecular de 1kb; -: Control negativo de PCR; A: Reacción de PCR desde cDNA con temperatura de annealing de 53°C; B: Reacción de PCR desde cDNA con temperatura de annealing de 57°C; C: Reacción de PCR desde cDNA con temperatura de annealing de 61°C. Se indican la banda de peso molecular de 3kb, los productos cercanos a 3kb y los productos inespecíficos.

La secuencia de DNA codificante para Gt115 mide 3054 pb y el producto de PCR buscado resultante del annealing con los partidores diseñados debería medir 3030 pb. Para esta reacción de PCR se distingue un producto cercano a 3 kb para las temperaturas de alineamiento de 53°C y 61°C, mientras que para la temperatura de annealing de 57°C se distinguen dos productos, uno cercano a los 3 kb y otro a los 4 kb. La menor cantidad de productos inespecíficos se nota en la reacción hecha a 61°C y de ello se eligió esta como la temperatura de annealing para las siguientes reacciones. La evidencia anterior

permite relacionar al producto de PCR de 3 kb con el producto de PCR derivado de la secuencia de DNA codificante buscada.

La reacción de PCR desde cDNA se repitió por triplicado con una temperatura de annealing de 60°C, donde 5 µL de la reacción fueron cargados a un gel de agarosa 1% que se pudo observar en la Figura 23:

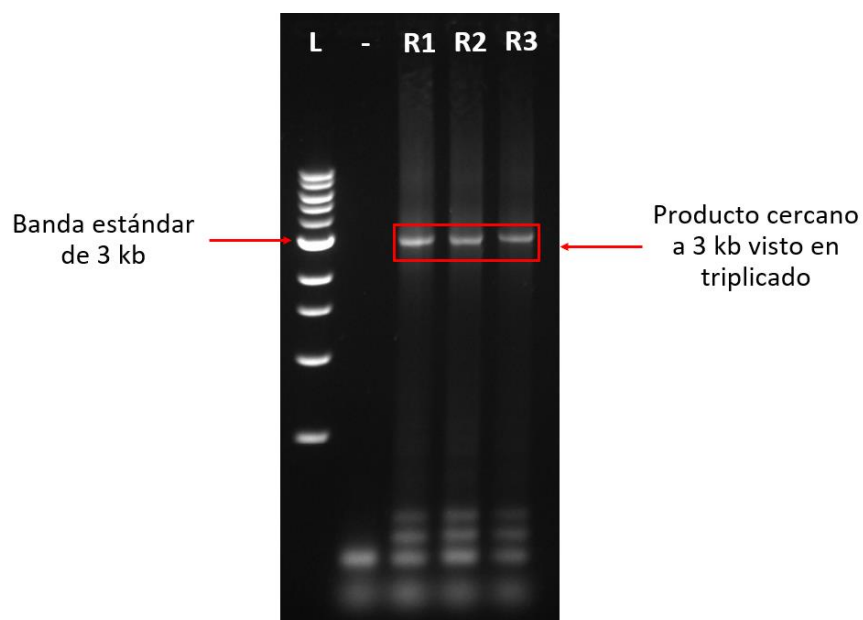


Figura 23. Gel de agarosa 1% donde se cargó la reacción de PCR por triplicado a 60°C del cDNA total de *G. trabeum*. L: Estándar de tamaño de 1 kb; CN, Control negativo de PCR; R1 a R3: Productos de la reacción de PCR en triplicado. La banda de 3 kb y los productos relevantes de la reacción se indican con flechas y cuadros.

Se pudo notar con mayor nitidez y consistencia la presencia clara de un producto cercano a los 3 kb. El volumen restante de las reacciones fue cargado en un gel de agarosa 0,8% y purificado obteniendo 50 µL de DNA a una concentración de 16,8 ng/µL y razón de absorbancia  $A_{260}/A_{280}$  de 1,72, que a pesar de estar debajo del corte de pureza se evaluó como suficiente para los siguientes pasos. Desde esta muestra se realizaron

otras reacciones de PCR para obtener mayores cantidades del fragmento aislado y posteriormente para los protocolos de ligación y transformación.

### **7. Clonamiento y análisis de la secuencia aislada**

Se usó el vector pJET1.2/blunt del Kit CloneJet de Thermo Fisher para la clonación de la secuencia de DNA aislada. Se realizó una reacción de ligación entre el vector en una razón molar de 3:1 Inserto:Vector. El producto de tal reacción se diluyó 5 veces y se agregó a las células competentes para la electroporación. Luego de la transformación se obtuvieron 6 colonias que crecieron en el medio de selección. De estas 3 se propagaron en medio líquido y se cultivaron toda la noche. Los clones fueron llamados 4P15XC1, 4P15XC2 y 4P15XC4. El cultivo de cada clon se usó para la extracción del plásmido. Los plásmidos puros extraídos tenían una concentración de 54,9, 50,2 y 79,4 ng/ $\mu$ L, mientras que la razón  $A_{260}/A_{280}$  fue de 1,9, 2,2 y 1,7 respectivamente.

Los plásmidos resultantes fueron digeridos con la enzima de restricción *XmnI*, que tiene sitios de corte tanto dentro del plásmido como dentro de la secuencia codificante de Gt115. Por lo tanto, en el caso de que la secuencia de DNA codificante esté efectivamente ligada a pJET1.2/blunt se deberían observar en un gel de agarosa dos fragmentos, uno cercano a los 4100 pb y otro cercano a los 1900 pb, siendo el tamaño total del plasmidio recombinante de 6004 pb. El plásmido extraído y la digestión para las colonias 4P15XC1, 4P15XC2 y 4P15XC4 se muestran a continuación en la Figura 24:

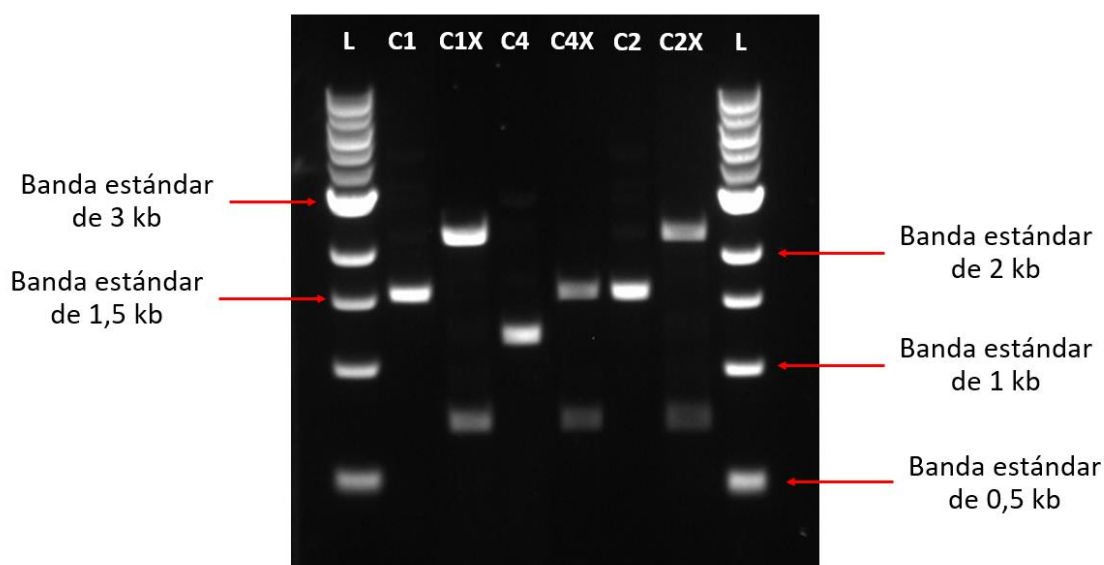


Figura 24. Electroforesis en gel de agarosa 1% de los plasmidios recombinantes 4P15XC1, 4P15XC2 y 4P15XC4 intactos y digeridos con *XmnI*. L: Escala de tamaño molecular de 1kb; C1 y C1X: Plasmidio 4P15XC1 sin cortar, y cortado con *XmnI* C2 y C2X: Plasmidio 4P15XC2 sin cortar: y cortado con *XmnI*; C4 y C4X plasmidio 4P15XC4 sin cortar y cortado con *XmnI*.

De los resultados de la digestión se puede notar que de las digestiones de los plasmidios derivados de las colonias 4P15XC1 y 4P15XC2 se liberaron dos fragmentos. En los pocillos C1X y C2X, se obtuvo un fragmento cercano a las 2200 pb y otro cercano a las 700 pb. Los dos fragmentos suman un valor aproximado al tamaño del vector pJET1.2/blunt no recombinante (2974 pb). Para el caso del plasmidio de la colonia 4P15XC4, se puede notar que tiene un tamaño menor comparado con los plasmidios de las otras dos colonias, lo que se puede deducir del tamaño de los fragmentos producto de la digestión con *XmnI*. Se observan dos fragmentos, uno cercano a las 1600 pb y otro cercano a las 700 pb

Los resultados de la digestión de los plasmidios obtenidos de las colonias sugieren que los plasmidios son no recombinantes. Para comprobar esta sospecha, los plasmidios se

enviaron a secuenciar. Las secuencias obtenidas fueron procesadas para recortar los bordes que presentarían baja calidad de lectura. Las secuencias resultantes se muestran en la sección 3 del Anexo. Seguidamente, usando el programa Clustal  $\Omega$  se alinearon las secuencias de 4P15XC1, 4P15XC2 y 4P15XC4 con la secuencia de un constructo hipotético formado por el vector pJET1.2/blunt y un inserto con la secuencia del codificante de Gt115 obtenido del genoma de *G. trabeum*. El resultado de este alineamiento se puede observar gráficamente en la Figura 25:

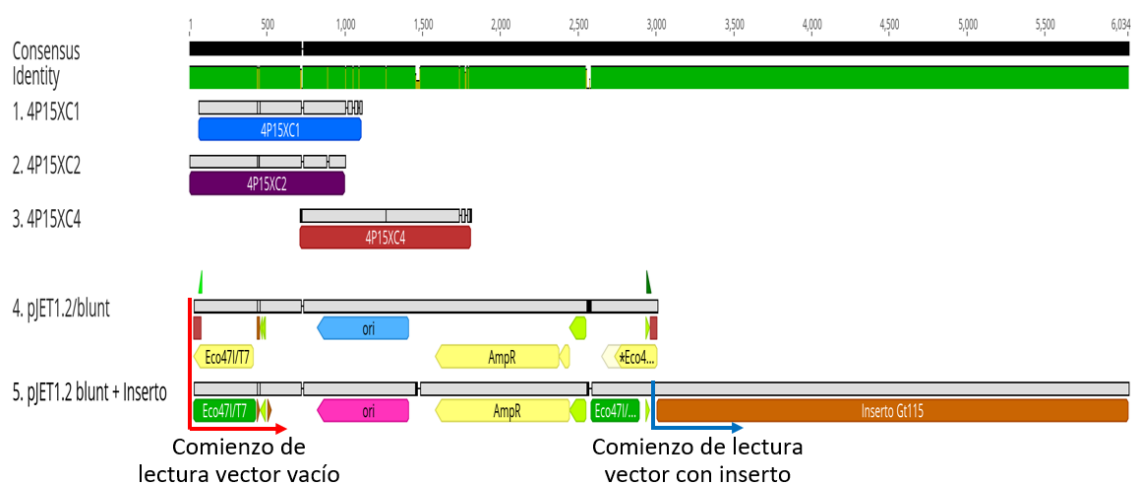


Figura 25. Alineamiento de los resultados de secuenciación obtenidos para los plásmidos de las colonias 4P15XC1, 4P15XC2, 4P15XC4, el vector pJET1.2/blunt vacío y un constructo hipotético del vector pJET1.2/blunt y la secuencia codificante para Gt115. Las secuencias 4P15XC1, 4P15XC2 Y 4P15XC4 están destacadas con rojo, azul y morado, respectivamente. El inserto se muestra en naranja y el gen letal Eco47I/T7 está en verde. Se indica con una flecha roja donde comenzaría la secuenciación si el vector estuviera vacío, y con una flecha azul donde comenzaría si el vector tuviera el inserto.

Los resultados del alineamiento muestran que las secuencias obtenidas corresponden con un porcentaje de identidad del 99,89% para 4P15XC1, 99,69% para 4P15XC2 a la región del vector pJET1.2/blunt que contiene la segunda mitad del gen letal Eco47I/T7 y parte del origen de replicación plasmidial. Para 4P15XC4 se observa que el comienzo de la secuencia inicia aproximadamente 700 pb río abajo en comparación con las otras



colonias, cubriendo desde la sección previa al origen de replicación y cubriendo parte del gen de resistencia a la ampicilina teniendo un 94,69% de identidad.

## Discusión y proyecciones

Las familias de  $\alpha$ -glucuronidasas GH67 y GH115 presentan características comunes tales como el mecanismo catalítico de inversión y el dominio de barril TIM donde se encuentra el sitio activo, sin embargo, estas presentan diferencias fundamentales en cuanto a especificidad de sustrato, actividad y función biológica.

Por un lado, para la familia GH67 la especificidad de sustrato está dirigida de manera general a XOS sustituidos con GA/MeGA, aunque se ha reportado que existe una preferencia por los sustituidos con MeGA (Yagi y col., 2017). En algunos casos también se ha observado actividad contra residuos de ácido hexenurónico (Septiningrum y col., 2015). A pesar de ello, para la familia GH67 el mecanismo catalítico es el mismo para todas enzimas conocidas de la familia como también lo es su conservación en los residuos relevantes para la hidrólisis (Tabla 10 en el anexo). Las glucuronidasas de la familia GH67 son específicas por sustratos oligoméricos solubles provenientes de actividad xilanolíticas previas. Según la evidencia actual es necesario en un futuro enfocar los esfuerzos en aumentar la cantidad de enzimas caracterizadas de esta familia para tener resultados conclusivos sobre su especificidad de sustrato.

La familia GH115 se distingue por presentar enzimas capaces de hidrolizar residuos GA/MeGA que se encuentran en posiciones interiores de la cadena de xilano. A pesar de ese factor común, el conocimiento actual sobre esta familia indica que existe una gran diversidad a nivel estructural y funcional. La familia GH115 se puede subclasificar de

acuerdo con varios criterios, tales como la arquitectura de dominios, especificidad de sustrato y estructura cuaternaria. Tal diversidad ha llevado a la familia a ser clasificada como promiscua, sin embargo, a pesar de que los residuos catalíticos más importantes han sido identificados, tanto el mecanismo catalítico general de estas enzimas como los aminoácidos que determinan su especificidad de sustrato no están claros.

Para la determinación de la actividad y la especificidad de sustrato de la familia GH115, hay dos factores importantes a tomar en cuenta, el primero es la presencia del loop móvil donde se encuentra la base catalítica. Wang y col. (2016) reportaron que mutar el ácido catalítico (E216A) más lejano a la base catalítica (D335) en SdeAgu115A, provoca la pérdida total de la actividad de la enzima, mientras que mutando el ácido catalítico más cercano (D215A) el efecto no es total y se observa actividad remanente. Este hecho también se reporta en el estudio de Rogowski y col. (2014), donde una mutación equivalente a D215 de SdeAgu115A en BoAgu115A (D206A) reduce drásticamente la actividad sin hacerla desaparecer por completo.

En la Tabla 9 se explicitan las distancias entre los dos ácidos catalíticos de la diada con respecto a la base catalítica del loop móvil, AC1 es el primer ácido catalítico de la secuencia, correspondiendo en todos los casos a un ácido aspártico que se encuentra en el rango de distancia ideal de 10 Å con respecto a la base catalítica, lo que permite la ejecución del mecanismo de inversión según lo que establecen Davies y Henrissat (1995). AC2 por otro lado corresponde en todos los casos a un ácido glutámico a una distancia mayor a 10 Å de la base catalítica, lo que en un principio daría entender que es menos probable que AC2 cumpla la función de ejecutar el mecanismo de inversión

como donador de protones. A pesar de ello los resultados indican que la mutación de AC2 tiene una mayor influencia en la eficiencia catalítica de la enzima que AC1. Esto en suma sugeriría que el loop móvil al unir al sustrato adoptaría posiciones distintas a las observadas hasta ahora en los modelos experimentales, alterando la distancia que la base catalítica tiene con la diada de ácidos catalíticos, favoreciendo la interacción con AC2 (ahora más cercano) por sobre AC1 (ahora más lejano), pudiendo explicar por qué la pérdida de actividad en SdeAgu115A es total cuando se muta E216 (AC2) y parcial cuando se muta D215 (AC1), según lo que reportaron Wang y col. (2016).

El segundo criterio importante para la determinación de la actividad y especificidad de sustrato de la familia GH115 reside en la forma y composición del surco catalítico. Aalbers y col. (2015) han sugerido que una combinación específica de aminoácidos en el sitio de unión a sustrato podría diferenciar entre enzimas con preferencia a xilanos o galactanos, así como también el tamaño del bolsillo catalítico determinaría el tamaño del sustrato y la función biológica que cumplen estas enzimas. En suma, dado la diversidad encontrada en la familia GH115 se considera relevante enfocar los esfuerzos investigativos en conocer de manera más completa los factores determinantes en la catálisis y la especificidad de sustrato en orden de tener una base clara que permita su explotación como enzimas auxiliares en la hidrólisis de xilano tanto a nivel de investigación básica como aplicada.

Con respecto al modelo por homología de Gt115, obtenido y optimizado en este trabajo, se puede notar que aún existen residuos que no están en posiciones permitidas, particularmente un residuo catalíticamente relevante del dominio C+, llamado W708. Se

piensa que la función de este aminoácido es la unión de la cadena de xilano, ubicándose en la superficie del dominio C+ de la cadena opuesta a la que se encuentra el surco catalítico, como ya ha sido observado por Yan y col. (2021) para el residuo W680 de AxyAgu115A. Si bien W708 está en una posición no permitida en el gráfico de Ramachandran, no es descartable que esto sea porque cumple un rol en la unión a la cadena de xilano en el sitio activo, sin embargo, dado que el modelo de Gt115 es monomérico y fue optimizado y refinado sin tomar en cuenta la interacción que puedan tener las dos subunidades de su homodímero, se decidió no estudiar su rol en el sitio activo hasta que se pueda generar un modelo homodimérico en un futuro. A pesar de ello, el modelo refinado fue elegido en primera instancia por su Puntuación-Z de la distribución de Ramachandran (Rama-Z). Esta toma en cuenta la distribución total de ángulos  $\phi/\psi$  de la proteína comparándola con modelos de alta calidad, pudiendo observarse que el modelo de Gt115 está dentro del rango aceptable (Entre -2 y 2). Por lo tanto, la estructura modelada se acerca al que sería una proteína en un estado cercano al nativo (Sobolev y col., 2020).

Del análisis del modelo por homología a nivel general se observó que Gt115 tiene la estructura de una  $\alpha$ -glucuronidasa de la familia GH115 y que posee la inserción de un dominio C+, consistente con el hecho de que la mayoría de las enzimas fúngicas identificadas de esta familia presentan tal inserción según lo que reportan Wang y col. (2016).

A nivel del sitio catalítico se observó una sustitución de la segunda histidina por ácido glutámico (E294) en la triada HHE (Grupo 5 a 7 de la tabla 11) relacionada con la unión

a GA/MeGA. Este cambio no está presente en las  $\alpha$ -glucuronidasas caracterizadas anteriormente. Si bien tanto la histidina como el ácido glutámico son capaces de formar puentes de hidrogeno con el sustrato, la sustitución podría tener dos efectos en la catálisis. El primer efecto tiene que ver con la cercanía de 5 Å de E294 con la base catalítica D349, lo podría indicar un cambio del pKa de la base catalítica debido a un ambiente local más ácido comparado con las otras enzimas de la familia GH115, lo que se observaría como un pH óptimo más alto para Gt115 para mantener a la base catalítica en su estado desprotonado para el mecanismo de inversión.

El segundo efecto en la catálisis tiene que ver con un cambio en la afinidad que Gt115 tiene por el sustrato, la sustitución de Histidina a Ácido glutámico implica que, dependiendo del pH, se tendrá una cadena lateral con carga positiva (protonada) o neutra (desprotonada) en el caso de His, o una cadena lateral con carga negativa (desprotonada) o neutra (protonada) en el caso de Glu. Hasta la fecha no se ha podido obtener un modelo experimental de alguna enzima de la familia GH115 en complejo con Glucuronoxilano, Glucoronoarabinosilano o Ácido glucurónico/metil glucurónico, sin embargo, para la familia GH67 Nurizzo y col. (2002) describen que la interacción de GlcA67A con GA y Xilano (Figura 9) se basa principalmente en cadenas laterales con carga positiva (Arg, Lys y His) que reconocen de manera específica el grupo carboxilo de carga negativa del ácido glucurónico, mientras que cadenas laterales ácidas (Glu y Asp) con carga negativa o neutra establecen puentes de hidrogeno con los oxígenos y grupos hidroxilo de este y la cadena de xilano. Este precedente podría indicar que la sustitución en esa posición de una cadena lateral con carga positiva como es His por una

de carga negativa como es Glu disminuiría la afinidad que el sitio activo tiene por el grupo carboxilo del ácido glucurónico y por tanto su unión al sitio activo, lo que en suma disminuiría la afinidad que Gt115 podría tener con sustratos sustituidos con ácido glucurónico/metil glucurónico.

La importancia de la triada HHE/HEE se puede evidenciar en que si se muta la primera histidina de la triada (grupo 5 en la tabla 11 del anexo) por Alanina, conservada en Gt115 y entre todas las enzimas GH115 bacterianas mencionadas, disminuye la actividad catalítica 2600 veces en BoAgu115A (Rogowski y col., 2014), por ello si bien Gt115 podría conservar niveles de actividad similares a sus contrapartes bacterianas, su óptimo sería en condiciones distintas y su afinidad por sustratos como GAX, GX y GA/MeGA sería más baja.

A nivel del surco catalítico se encontraron varias ocurrencias relevantes, la primera es la separación de los residuos de Trp implicados en la unión de la cadena de xilano/galactano. Para Gt115 se observa que la separación entre estos dos residuos es casi el doble que la observada para AxyAgu115A (Figura 18), SdeAgu115A y BoAgu115A, y cercana a la observada en BtGH115A. La segunda ocurrencia corresponde al hecho que las rotaciones del ángulo  $\omega$  para los residuos de Trp de los sitios +1 y -1 en Gt115 son notablemente diferentes que todas las otras enzimas analizadas, incluyendo BtGH115A. Nishio (2011) concluyó que las interacciones entre aminoácidos aromáticos como Trp, Tyr y Phe con carbohidratos dependen de la distancia, ángulo y electronegatividad de los residuos aromáticos y como estos interactúan con las distintas distribuciones de electropositividad que poseen los enlaces

de distintos isómeros de carbohidratos. Esto implica que posiciones o rotaciones distintas de los residuos aromáticos podrían interferir o favorecer interacciones diferentes con varios tipos de mono y polisacáridos, por ello el hecho de que los residuos del sitio -1 y +1 de Gt115 presenten una rotación marcadamente distinta que sus análogos en otras enzimas bacterianas podrían implicar que la interacción Gt115 y posiblemente otras enzimas GH115 fúngicas presentarían interacciones y afinidades diferentes con xilano y arabinogalactano. La estabilidad de las torsiones  $\omega$  queda por comprobarse a través de resultados experimentales o simulaciones de dinámica molecular.

Tanto el largo del surco de unión del sustrato, como las rotaciones de Trp en los sitios +1 y -1 implican la posibilidad de que Gt115 prefiera sustratos más largos que sus contrapartes bacterianas, y que las interacciones con estos sustratos puedan condicionar una especificidad distinta. No obstante, se debe tomar en cuenta que el modelo de Gt115 se generó a base de todos los modelos de enzimas bacterianas GH115 actualmente disponibles, donde como máximo se alcanza un 40% de identidad para el caso de SdeAgu115A, por lo que características propias de una enzima GH115 fúngica podrían perderse en el proceso, y características bacterianas aparecer como artefactos del modelado de la misma manera. Se proyecta en el futuro lograr obtener una versión en forma de homodímero de esta enzima para poder investigar de manera más completa todos los residuos implicados en la unión al sustrato en el surco catalítico, particularmente W708 que se encuentra en el dominio C<sup>+</sup>.



Con los recursos computacionales actuales, las simulaciones de dinámica molecular y el docking permitirían dilucidar la dinámica del surco catalítico, y cómo este interactúa con distintos tipos de sustratos. En el largo plazo la obtención de estructuras cristalizadas y modelos 3D de enzimas GH115 fúngicas, idealmente unidas a ligandos, es un paso crítico para distinguir su rol biológico en la degradación de BML, sus posibles diferencias con sus contrapartes bacterianas y los factores determinantes de su especificidad de sustrato.

Con respecto al aislamiento e identificación de la secuencia codificante para Gt115, la primera etapa consistió en la cultivar el hongo *G. trabeum* bajo condiciones de inducción de la expresión de genes relacionados con la degradación de lignocelulosa. En hongos es ampliamente conocido que la presencia de monosacáridos u oligosacáridos simples, particularmente glucosa, inhibe la expresión de genes relacionados con la degradación de polisacáridos más complejos provenientes de tejido vegetal (Yan y col., 2020). Tomando este hecho en consideración se tomó como base un medio mínimo de hongos carente de fuente de carbono y se consideraron dos fuentes de carbono basadas en tejido vegetal. La primera fue aserrín proveniente de madera de la angiosperma *Aristotelia chilensis* obtenida de un bosque renoval de la X región de Chile. Esta elección fue en base a que la presencia de la secuencia de DNA codificante de Gt115 fue detectado cuando *G. trabeum* fue cultivado sobre madera de angiosperma proveniente de álamo (*Populus sp.*) según lo reportado por Presley y col. (2017). Por otro lado, la Alfalfa (*Medicago sativa*) es una leguminosa ampliamente usada y es comercialmente accesible, Se ha reportado la presencia de la secuencia de DNA codificante de Gt115 en

otras plantas herbáceas como la caña de azúcar o el sorgo (Persley y col., 2017; Valadares y col., 2019) lo que indicaría que *G. trabeum* sería capaz de crecer y posiblemente activar sus mecanismos de degradación de lignocelulosa si es crecida en alfalfa.

Se observó que *G. trabeum* pudo crecer en placas de agar con medio mínimo suplementado con alfalfa, pero no en aquellas suplementadas con aserrín de maqui, a pesar de contar con el mismo tamaño de partícula (<1 mm) y tratamiento previo a la elaboración del medio entre las dos fuentes de carbono. Este fenómeno puede ser explicado en primer lugar con el hecho de que los hongos de pudrición parda como *G. trabeum* prefieren maderas blandas y por ello tendría mayor dificultad en utilizar la fuente de carbono proveniente de una madera dura como es la de *A. chilensis*. Asimismo, el hecho de que el contenido de lignina sea más alto en plantas leñosas que en herbáceas (Ma y col., 2017) podría implicar un acceso y degradación más dificultosa de los carbohidratos de la pared de *A. chilensis*, una planta leñosa, mientras que, para la alfalfa, una planta herbácea, el acceso y subsecuente degradación de los carbohidratos se vería más favorecida en comparación. A nivel técnico, se pudo notar que los medios de agar donde se creció *G. trabeum* no eran completamente homogéneos, por ello la distribución de nutrientes accesibles es un factor que pudo haber afectado el crecimiento del hongo en los dos medios.

La evaluación de crecimiento de *G. trabeum* en distintas fuentes de carbono no fue contemplada en los alcances de este seminario, y por ello no se obtuvo evidencia suficiente que dé a conocer de manera completa los factores que afectan el crecimiento

de este hongo y su afinidad por distintos sustratos vegetales. En futuros trabajos se consideraría relevante evaluar como el tamaño de partícula, el pretratamiento y origen de la biomasa afectan el crecimiento de *G. trabeum* y otros hongos filamentosos descomponedores de madera en orden de observar la potencial actividad de sus enzimas implicadas en distintos tipos de tejido vegetal.

Durante el aislamiento de la secuencia de DNA codificante para Gt115 en una primera etapa se realizó PCR de gradiente al cDNA obtenido del RNA aislado, en orden de observar si es que se obtenía algún producto cercano a los 3 kb que corresponde al largo de esta secuencia de DNA codificante. El PCR de gradiente dio como resultado un producto cercano a las 3 kb más otros productos inespecíficos más pequeños (Figura 22), de este experimento se eligió una temperatura de annealing de 60°C, lo que resultó en bandas únicas y claras de aproximadamente 3 kb (Figura 23), las que se usaron para aislar el producto de PCR y ligarlo al vector pJET 1.2/Blunt. El producto de la ligación fue analizado a través de digestión con enzimas de restricción y secuenciación.

El vector pJET 1.2/Blunt tiene un solo sitio de corte de XmnI, ubicado en el gen de resistencia a la ampicilina (Figura 6), por ello solo debería linealizar el vector en condiciones normales, sin embargo, la presencia de un fragmento de 700 pb (Figura 24) en todas las digestiones implica la aparición de un nuevo sitio de corte de XmnI a una distancia aproximada de 700 pb en las dos direcciones desde el sitio de corte original, por lo que se encontraría dentro de la secuencia del gen letal Eco47I/T7 o en la sección final del gen de resistencia a la ampicilina.

La secuenciación con un partidador específico para el promotor T7 realizada a los vectores extraídos de cada colonia evaluada confirma que ninguno de los vectores contiene algún inserto. Para el caso de 4P15XC1 Y 4P15XC2 la lectura comienza en las cercanías del promotor T7, este punto fue elegido principalmente porque precede a la sección del vector que contendría el inserto en el caso de una ligación exitosa, sin embargo, los resultados de la secuenciación indican que el fragmento secuenciado corresponde a la segunda sección del gen letal y parte del origen de replicación, lo que evidencia que estas colonias contienen un vector que se recircularizó sin ningún inserto. El caso de 4P15XC4 es similar, sin embargo, la lectura de la secuenciación comienza cerca de 700 pb río abajo partiendo en la sección anterior al origen de replicación, extendiéndose hasta parte del gen de resistencia a la ampicilina y por ello se evidencia que tampoco posee inserto.

Para el caso del vector extraído de la colonia 4P15XC4, se puede argumentar una pérdida de material genético posiblemente derivada de una hidrólisis espontánea. Esta suposición se basa en los resultados la digestión con la enzima de restricción *XmnI*. Se observó que el vector extraído tenía un largo menor (700 pb aproximadamente) que los provenientes de otras colonias. Esto se confirma en los resultados de la secuenciación, donde el inicio de la lectura para el vector proveniente de 4P15XC4 inicia aproximadamente 700 pb río abajo de la posición original del promotor T7. De ello además se infiere que el fragmento perdido para 4P15XC4 fue la sección entre el promotor T7 y la sección previa al origen de replicación, la cual contenía la mitad del gen letal Eco47I/T7.

El fallo en la clonación puede ser atribuido a varias causas en diferentes etapas en el proceso de clonamiento. A nivel de la ligación se debe tomar en cuenta que los fragmentos a ligar poseían extremos romos, con una concentración en una relación de inserto:vector de 3:1 en las condiciones descritas por el fabricante de la enzima. En general se recomienda que las ligaciones de extremos romos se realicen con mayor concentración de DNA y de T4 DNA ligasa que las usadas en este trabajo, por lo que esta puede ser una razón para la falla en la reacción de ligación (Saunders, 2012). Debido a la baja concentración de DNA resultante de la extracción desde el gel, y a una poca cantidad de vector pJET 1.2/Blunt disponible, no se pudieron realizar reacciones de ligación con razones inserto:vector más altas por ello es posible que la cantidad de producto de ligación resultante haya sido muy pequeño.

Usar una mayor concentración de DNA sería clave para poder realizar ligaciones con razones más altas de inserto:vector, sumado a usar concentraciones más altas de PEG que las que contiene el buffer de reacción usado en la ligación (>25%) para favorecer el encuentro de los extremos del vector y el inserto y con ello la ligación.

A nivel del clonamiento se puede notar que la selección secundaria que debería asegurar que las colonias contengan vector recombinante, con inserto, no funcionó. El vector pJET 1.2/Blunt posee como marcador de selección secundaria un gen letal Eco47I/T7 que en situaciones ideales impediría el crecimiento de aquellas bacterias que contengan vector sin inserto, ya que necesitan de este para interrumpir la secuencia del marcador de selección. La supervivencia de la colonia 4P15XC4 en la placa de selección se podría explicar por la posible pérdida de una sección de su secuencia correspondiente al gen

mortal en el vector de clonación. Por lo tanto, la ausencia de inserto no activó el mecanismo de selección secundaria al tener un gen letal no funcional. Por otro lado, la supervivencia de las colonias 4P15XC1 y 4P15XC2 no puede ser explicada de manera tan directa, sin embargo, tomando el precedente del vector de 4P15XC4 y el hecho de que se detectó un sitio de restricción nuevo para todos los vectores, se teoriza que estos pudieron haber experimentado hidrólisis espontánea de pequeñas secciones de su secuencia que en suma hayan generado corrimientos del marco de lectura que dejaran inutilizado al gen letal Eco47I/T7 y que generaran sitios de restricción nuevos.

El fallo de la selección secundaria también puede atribuirse a células competentes insensibles a la expresión del gen letal Eco47I/T7, lo que explicaría también la aparición de colonias aparentemente recombinantes. En los experimentos de transformación no se incluyó un control positivo para pJET 1.2/Blunt cerrado sin inserto, por ello la sensibilidad de las células competentes usadas no pudo comprobarse. En el caso de seguir trabajando con estas células competentes se considera fundamental evaluar la sensibilidad de estas al mecanismo de selección secundaria en orden de evaluar un cambio de vector de clonamiento.

A pesar de que se aisló un fragmento de 3kb directamente desde el cDNA total de *G. trabeum*, la única evidencia de que corresponde a la secuencia de DNA codificante para Gt115 es que posee el mismo largo aproximado en pares de bases y que fue aislado por PCR con partidores específicos. Para comprobar de manera definitiva que el fragmento aislado corresponde a la secuencia de DNA codificante de Gt115 (XM\_007867777.1) se debería determinar la secuencia del fragmento.

En suma, aun tomando en cuenta que no hay evidencia concluyente de la identidad del fragmento aislado y que este no pudo ser clonado en *E. coli* DH5 $\alpha$ , se considera un buen indicio que este tenga aproximadamente el mismo largo que la secuencia de DNA codificante de Gt115 y que se haya aislado de micelio de *G. trabeum* crecido sobre el medio de alfalfa, lo que facilitará su obtención, secuenciamiento y clonamiento en futuros trabajos.

## Conclusiones

En la primera parte del presente estudio se realizó un estudio completo sobre las características de las enzimas  $\alpha$ -glucuronidasas de las familias GH67 y GH115. Los resultados reunidos en este trabajo permiten sugerir un mecanismo catalítico y los determinantes de la unión a sustrato en enzimas glucuronidasas de la familia GH67, no así para las de la familia GH115, cuya especificidad de sustrato no ha sido caracterizada a fondo, pero ha sido asociada con el movimiento de un loop de aminoácidos móvil del sitio activo y la composición de residuos del surco catalítico presente en su estructura.

Un modelo por homología de Gt115 fue generado exitosamente y refinado a nivel estructural hasta que fuera apto para el análisis. Del estudio de este modelo se identificó un surco catalítico largo, con dos residuos de triptófano en su interior, que poseen ángulos de rotación  $\omega$  distintos a los vistos en otras enzimas. Además de eso, se halló una sustitución de His por Glu en una triada de residuos relacionados con la unión a xilano. Estos dos rasgos observados sugieren que Gt115 podría acomodar sustratos más largos, una menor afinidad por residuos de ácido glucurónico, un pH óptimo más alto y afinidades distintas a varios tipos de azúcares que las vistas en  $\alpha$ -glucuronidasas bacterianas de la familia GH115.

El clonamiento de la secuencia codificante para Gt115 no fue exitoso, debido principalmente a problemas con el vector de clonación y la ligación. Se reporta haber aislado desde micelio de *G. trabeum* crecido en alfalfa un fragmento de DNA del mismo largo que el predicho para el gen maduro de Gt115, sin embargo, no se obtuvo evidencia conclusiva de que el fragmento corresponda a efectivamente a este gen.



Tomando en cuenta todo lo anteriormente dicho, durante este trabajo se logró cumplir la mayoría de los objetivos específicos, sin embargo, debido a que no se pudo clonar la secuencia de DNA codificante de Gt115 el objetivo general solo pudo cumplirse de manera parcial

## Referencias

1. A Roy, A Kucukural, Y Zhang 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*. **5**: 725-738
2. Aachary, A., Gobinath, D., Srinivasan, K., y Prapulla, S. 2015. Protective effect of xylooligosaccharides from corn cob on 1,2-dimethylhydrazine induced colon cancer in rats. *Bioactive Carbohydrates And Dietary Fibre*. **5**:146-152.
3. Aachary, A., y Prapulla, S. 2010. Xylooligosaccharides (XOS) as an Emerging Prebiotic: Microbial Synthesis, Utilization, Structural Characterization, Bioactive Properties, and Applications. *Comprehensive Reviews In Food Science And Food Safety*. **10**: 2-16.
4. Aalbers, F., Turkenburg, J. P., Davies, G. J., Dijkhuizen, L., y Lammerts van Bueren, A. 2015. Structural and Functional Characterization of a Novel Family GH115 4-O-Methyl- $\alpha$ -Glucuronidase con Specificity for Decorated Arabinogalactans. *Journal of Molecular Biology*. **427**: 3935–3946.
5. Abdo, A. A. A., Zhang, C., Lin, Y., Kaddour, B., Li, X., Fan, G., y Yang, R. 2021. Nutritive sweetener of short-chain xylooligosaccharides improved the foam properties of hen egg white protein via glycosylation. *Journal of Food Measurement and Characterization*. **15**: 1341-1348.
6. Alokika, y Singh, B. 2019. Production, characteristics, and biotechnological applications of microbial xylanases. *Applied Microbiology And Biotechnology*. **103**: 8763-8784.
7. Amorim, C., Silvério, S. C., y Rodrigues, L. R. 2019. One-step process for producing prebioticarabino-xylooligosaccharides from brewer's spent grain employing *Trichoderma* species. *Food Chemistry*. **270**: 86–94.
8. Asiegbu, F. O., y Kovalchuk, A. 2021. *Forest Microbiology*. En Langer, G. J., Bußkamp, J., Terhonen, E., & Blumenstein, K. *Fungi inhabiting woody tree tissues (1<sup>o</sup> ed.)*, pp 175-195. Academic press, Londres
9. Beldman, G., Schols, H. A., Pitson, S. M., Searle-van Leeuwen, M. J. F., y Voragen, A. G. J. 1997. Arabinans and arabinan degrading enzymes. *Advances in macromolecular carbohydrate research*. **1**:1-64.
10. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N, Weissig H., Shindyalov I.N., Bourne P.E. 2000. The Protein Data Bank *Nucleic Acids Research*. **28**: 235-242.
11. Bhardwaj, N. Kumar y B. Verma, P. 2019. A detailed overview of xylanases: an emerging biomolecule for current and future prospective. *Bioresources and Bioprocessing*. **6**: 40.
12. Boratyn, G. M., Schäffer, A. A., Agarwala, R., Altschul, S. F., Lipman, D. J., & Madden, T. L. 2012. Domain enhanced lookup time accelerated BLAST. *Biology direct*. **7**: 12.
13. Bornscheuer, U., Buchholz, K., & Seibel, J. 2014. Enzymatic Degradation of (Ligno)cellulose. *Angewandte Chemie International Edition*. **53**: 10876–10893.
14. Bragatto, J. Segato, F. Squina, F.M. 2013. Production of xylooligosaccharides (XOS) from delignified sugarcane bagasse by peroxide-HAC process using recombinant xylanase from *Bacillus subtilis*. *Industrial Crops and Products*. **51**:123–129.
15. Cardoso, B. B., Amorim, C., Silvério, S. C., & Rodrigues, L. R. 2021. Novel and emerging prebiotics: Advances and opportunities. *Advances in Food and Nutrition Research*. **95**: 41–95.
16. Collins, T., Gerday, C., y Feller, G. 2005. Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews*. **29**: 3-23.
17. Daniel, G. 2014. Fungal and Bacterial Biodegradation: White Rots, Brown Rots, Soft Rots, and Bacteria. *Deterioration and Protection of Sustainable Biomaterials (ACS Symposium Series)*. **1**: 23–58.
18. Davies, G., y Henrissat, B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure*. **3**: 853–859.
19. De Souza, P. M., y D Oliveira Magalhães, P. 2010. Application of microbial  $\alpha$ -amylase in industry - A review. *Brazilian journal of microbiology*. **41**:850–861.
20. Dodd, D., y Cann, I. 2009. Enzymatic deconstruction of xylan for biofuel production. *GCB Bioenergy*. **1**: 2-17.
21. Drula E., Garron M., Dogan S., Lombard V., Henrissat B. y Terrapon N. 2022. The carbohydrate-active enzyme database: functions and literature *Nucleic Acids*. **50**: 571–577.
22. Ejaz, U., Sohail, M., Y Ghanemi, A. 2021. Cellulases: From Bioactivity to a Variety of Industrial Applications. *Biomimetics*. **6**: 44.
23. Floudas, D. 2021. Evolution of lignin decomposition systems in fungi. *Wood Degradation and Ligninolytic Fungi*. **99**: 37–76.
24. Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R. A., Henrissat, B., Martínez, A. T., Otilar, R., Spatafora, J. W., Yadav, J. S., Aerts, A., Benoit, I., Boyd, A., Carlson, A., Copeland, A., Coutinho, P. M., de Vries, R. P., Ferreira, P., Findley, K., Foster, B., Hobbitt, D. S. 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science (New York, N.Y.)*. **336**: 1715–1719.
25. Fu, G., Hu, Y., Bian, J., Li, M., Peng, F., y Sun, R. 2019. Isolation, Purification, and Potential Applications of Xylan. *Production Of Materials From Sustainable Biomass Resources*, pp 3-35. Springer, Singapore
26. Golan, G., Shallom, D., Teplitsky, A., Zaide, G., Shulami, S., Baasov, T., Stojanoff, V., Thompson, A., Shoham, Y., y Shoham, G. 2004. Crystal Structures of *Geobacillus stearothermophilus*  $\alpha$ -Glucuronidase Complexed con Its Substrate and Products. *Journal of Biological Chemistry*. **279**: 3014–3024.
27. Grigoriev, I., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., Riley R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova T., Nordberg, H., Dubchak, I. y Shabalov, I. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research*. **42**: 699–704.
28. Habibi Y., L. Lucia A., Rojas O. J. 2010. Cellulose nanocrystals: chemistry, self-assembly, and applications, *Chem*. **110**: 3479–3500.
29. Heo, L., Park, H., y Seok, C. 2013. GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic acids research*. **41**: 384–388.
30. Houfani, Aicha Asma; Anders, Nico; Spiess, Antje C.; Baldrian, Petr; Benallaoua, Said 2020. Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars a review. *Biomass and Bioenergy*. **134**: 105481.
31. Imaizumi K., Nakatsu Y., Sato M., Sedarnawati Y., Sugano M. 1991. Effects of Xylooligosaccharides on Blood Glucose, Serum and Liver Lipids and Cecum Short-chain Fatty Acids in Diabetic Rats, *Agricultural and Biological Chemistry*. **55**: 199–205.
32. Jalak, J., Kurašin, M., Teugjas, H., & Väljamäe, P. 2012. Endo-exo synergism in cellulose hydrolysis revisited. *The Journal of biological chemistry*. **287**: 28802–28815.
33. Jia, X., Mi, S., Wang, J., Qiao, W., Peng, X., y Han, Y. 2014. Insight into Glycoside Hydrolases for Debranched Xylan Degradation from Extremely Thermophilic Bacterium *Caldicellulosiruptor lactoaceticus*. *PLoS ONE*. **9**: e106482.
34. Kaffenberger, J. T., y Schilling, J. S. 2013. Using a grass substrate to compare decay among two clades of brown rot fungi. *Applied microbiology and biotechnology*. **97**: 8831–8840.
35. Kim, Y. S., Funada, R., y Singh, A. P. 2016. *Secondary Xylem Biology (1a ed)*. pp 131-167 Elsevier *Gezondheidszorg*.
36. Krahl, F. S., Bässler, C., Heibl, C., Soghigian, J., Schaefer, H., y Hobbitt, D. S. 2018. Evolutionary dynamics of host specialization in wood-decay fungi. *BMC evolutionary biology*. **18**: 119.

37. Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., y Karplus, K. 2009. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins*. **9**: 114–122.
38. Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. 1993, PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**: 283-291.
39. Lee, G. R., Heo, L., y Seok, C. 2016. Effective protein model structure refinement by loop modeling and overall relaxation. *Proteins*. **84**: 293–301
40. Li, X., Dilokpimol, A., Kabel, M. A., & de Vries, R. P. 2022. Fungal xylanolytic enzymes: Diversity and applications. *Bioresource Technology*. **344**: 126290.
41. Ma, S., He, F., Tian, D., Zou, D., Yan, Z., Yang, Y., Zhou, T., Huang, K., Shen, H., y Fang, J. 2017. Variations and determinants of carbon content in plants: a global synthesis. *Biogeosciences*. **15**: 693-702.
42. Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., y Lopez, R. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*. **gkac240**.
43. Maehara, T., Fujimoto, Z., Ichinose, H., Michikawa, M., Harazono, K., y Kaneko, S. 2014. Crystal structure and characterization of the glycoside hydrolase family 62  $\alpha$ -L-arabinofuranosidase from *Streptomyces coelicolor*. *The Journal of biological chemistry*. **289**: 7962–7972.
44. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Bryant, S. H. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic acids research*. **45**: D200–D203.
45. Marcotuli, I., Hsieh, Y. S. Y., Lahnstein, J., Yap, K., Burton, R. A., Blanco, A., Fincher, G. B., y Gadaleta, A. 2016. Structural Variation and Content of Arabinoxylans in Endosperm and Bran of Durum Wheat (*Triticum turgidum* L.). *Journal of Agricultural and Food Chemistry*. **64**: 2883–2892.
46. McKee, L. S., Sunner, H., Anasontzis, G. E., Toriz, G., Gatenholm, P., Bulone, V., Vilaplana, F., & Olsson, L. 2016. A GH115  $\alpha$ -glucuronidase from *Schizophyllum commune* contributes to the synergistic enzymatic deconstruction of softwood glucuronoarabinoxylan. *Biotechnology for biofuels*. **9**: 2.
47. Morrell, J. and Zabel, R., 2020. *Wood Microbiology: Decay and Its Prevention*. 2nd ed, pp.198-213. Academic Press
48. Na, M. H., y Kim, W. K. 2007. Effects of xylooligosaccharide intake on fecal bifidobacteria, lactic acid and lipid metabolism in Korean young women. *Journal of Nutrition and Health*. **40**: 154-161.
49. Nagy, T., Emami, K., Fontes, C. M. G. A., Ferreira, L. M. A., Humphry, D. R., y Gilbert, H. J. 2002. The Membrane-Bound  $\alpha$ -Glucuronidase from *Pseudomonas cellulosa* Hydrolyzes 4-O-Methyl-d-Glucuronoxyloligosaccharides but Not 4-O-Methyl-d-Glucuronoxylan. *Journal of Bacteriology*. **184**: 4925–4929.
50. Nagy, T., Emami, K., Fontes, C. M. G. A., Ferreira, L. M. A., Humphry, D. R., & Gilbert, H. J. 2002. The Membrane-Bound  $\alpha$ -Glucuronidase from *Pseudomonas cellulosa* Hydrolyzes 4- O- Methyl- d -Glucuronoxyloligosaccharides but Not 4- O- Methyl- d -Glucuronoxylan. *Journal of Bacteriology*. **184**: 4925–4929.
51. New England Biolabs. 2022. PCR Protocol for Phusion High-Fidelity DNA Polymerase (M0530). *Protocols*.Io. <https://www.protocols.io/view/pcr-protocol-for-phusion-high-fidelity-dna-polymer-bdd3i28n>
52. Nishio, M. (2011). The CH/ $\pi$  hydrogen bond in chemistry. Conformation, supramolecules, optical resolution and interactions involving carbohydrates. *Physical Chemistry Chemical Physics*. **13**: 13873.
53. Nurizzo, D., Nagy, T., Gilbert, H. J., y Davies, G. J. 2002. The Structural Basis for Catalysis and Specificity of the *Pseudomonas cellulosa*  $\alpha$ -Glucuronidase, GlcA67A. *Structure*. **10**: 547–556.
54. Peralta, R. M., da Silva, B. P., Gomes C rrea, R. C., Kato, C. G., Vicente Seixas, F. A., y Bracht, A. 2017. Enzymes from Basidiomycetes—Peculiar and Efficient Tools for Biotechnology. *Biotechnology of Microbial Enzymes*. C. 5 pp 119–149. Academic Press
55. Pereira, G. F., de Bastiani, D., Gabardo, S., Squina, F., y Ayub, M. A. Z. 2018. Solidstate cultivation of recombinant *Aspergillus nidulans* to coproduce xylanase, arabinofuranosidase, and xylooligosaccharides from soybean fibre. *Biocatalysis and Agricultural Biotechnology*. **15**: 78-85.
56. Poletto, P., Pereira, G. N., Monteiro, C. R. M., Pereira, M. A. F., Bordignon, S. E., & de Oliveira, D. 2020. Xylooligosaccharides: Transforming the lignocellulosic biomasses into valuable 5-carbon sugar prebiotics. *Process Biochemistry*. **91**: 352–363.
57. Presley, G. N., y Schilling, J. S. 2017. Distinct Growth and Secretome Strategies for Two Taxonomically Divergent Brown Rot Fungi. *Applied and Environmental Microbiology*. **Rev 83**: e02987-16.
58. Presley, G. N., Zhang, J., Purvine, S. O., y Schilling, J. S. 2020. Functional Genomics, Transcriptomics, and Proteomics Reveal Distinct Combat Strategies Between Lineages of Wood-Degrading Fungi Con Redundant Wood Decay Mechanisms. *Frontiers in Microbiology*. **11**:1646.
59. Rafiei, V., V lez, H., y Tzelepis, G. 2021. The Role of Glycoside Hydrolases in Phytopathogenic Fungi and Oomycetes Virulence. *International journal of molecular sciences*. **22**: 9359.
60. Rajagopalan, G., Shanmugavelu, K., Yang, K. 2016. Production of xylooligosaccharides from hardwood xylan by using immobilized endoxylanase of *Clostridium* strain BOH3. *RSC Adv*. **6**: 81818–81825
61. Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. 1963. Stereochemistry of polypeptide chain configurations. *Journal of molecular biology*. **7**: 95–99.
62. Ricardo Soccol, C., Faraco, V., Karp, S., Vandenbergh, L. P., Thomaz-Soccol, V., Woiciechowski, A., y Pandey, A. 2011. Lignocellulosic Bioethanol. *Biofuels*. pp 101–122. Academic Press.
63. Rodrigues, J. P. G. L. M., Melquiond, A. S. J., Karaca, E., Trellet, M., Dijk, M., Zundert, G. C. P., Schmitz, C., Vries, S. J., Bordogna, A., Bonati, L., Kastiris, P. L., y Bonvin, A. M. J. J. 2013. Defining the limits of homology modeling in information-driven protein docking. *Proteins: Structure, Function, and Bioinformatics*. **Rev 81**: 2119–2128.
64. Rogowski, A., Basl , A., Farinas, C. S., Solovyova, A., Mortimer, J. C., Dupree, P., Gilbert, H. J., y Bolam, D. N. 2014. Evidence That GH115  $\alpha$ -Glucuronidase Activity, Which Is Required to Degrade Plant Biomass, Is Dependent on Conformational Flexibility. *Journal of Biological Chemistry*. **289**:53–64.
65. Rosa, L., Ravanal, M. C., Mardones, W., & Eyzaguirre, J. 2013. Characterization of a recombinant  $\alpha$ -glucuronidase from *Aspergillus fumigatus*. *Fungal Biology*. **117**: 380-387
66. Saini, J., Saini, R., y Tewari, L. 2014. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *Biotech. Rev* **5**: 337-353.
67. Sambrook, J., Russell, D. W., Irwin, N., Cold Spring Harbor Laboratory, Ralph Erskine Conrad Memorial Fund, & Janssen, K. A. 2001. *Molecular Cloning*. (3<sup>rd</sup> ed) Transformation of *E. coli* by Electroporation Protocol pp 119-122 Amsterdam University Press.
68. Santib niz, L., Henr quez, C., Corro-Tejada, R., Bernal, S., Armijo, B., y Salazar, O. 2021. Xylooligosaccharides from lignocellulosic biomass: A comprehensive review. *Carbohydrate Polymers*. **251**: 117118.
69. Saunders, V. A. 2012. *Microbial genetics applied to biotechnology : principles and techniques of gene transfer and manipulation (1<sup>a</sup> edici n)* pp. 66-70. Springer.
70. Scheller, H., y Ulvskov, P. (2010). Hemicelluloses. *Annual Review Of Plant Biology*. **61**: 263-289.
71. Schumann, U., Smith, N. A., y Wang, M. B. 2013. A fast and efficient method for preparation of high-quality RNA from fungal mycelia. *BMC Research notes*. **6**: 1-5.
72. Septiningrum, K., Ohi, H., Waeonukul, R., Pason, P., Tachaapaikoon, C., Ratanakhanokchai, K., Sermsathanaswadi, J., Deng, L., Prawitwong, P., y Kosugi, A. 2015. The GH67  $\alpha$ -glucuronidase of *Paenibacillus curdlanolyticus* B-6 removes hexenuronic acid groups and facilitates biodegradation of the model xylooligosaccharide hexenuronosyl xylotriose. *Enzyme and Microbial Technology*. **71**: 28–35.

73. Sethupathy, S., Morales, G. M., Li, Y., Wang, Y., Jiang, J., Sun, J., y Zhu, D. 2021. Harnessing microbial wealth for lignocellulose biomass valorization through secretomics: a review. *Biotechnology for biofuels*. Rev **14**: 154.
74. Silva, E. K., Arruda, H. S., Mekala, S., Pastore, G. M., Meireles, M. A. A., y Saldaña, M. D. 2022. Xylooligosaccharides and their chemical stability under high-pressure processing combined with heat treatment. *Food Hydrocolloids*. **124**: 107167.
75. Silva, E. K., Arruda, H. S., Pastore, G. M., Meireles, M. A. A., y Saldaña, M. D. 2020. Xylooligosaccharides chemical stability after high-intensity ultrasound processing of prebiotic orange juice. *Ultrasonics Sonochemistry*. **63**: 104942.
76. Singh, R. V., Sambyal, K., Negi, A., Sonwani, S., y Mahajan, R. 2021. Chitinases production: A robust enzyme and its industrial applications. *Biocatalysis and Biotransformation*. **39**: 161–189.
77. Sista Kameshwar, A. K., & Qin, W. (2017). Comparative study of genome-wide plant biomass-degrading CAZymes in white rot, brown rot and soft rot fungi. *Mycology*. **9**: 93–105.
78. Sizova, M. V., Izquierdo, J. A., Panikov, N. S., & Lynd, L. R. 2011. Cellulose- and xylan-degrading thermophilic anaerobic bacteria from biocompost. *Applied and environmental microbiology*. **77**: 2282–2291.
79. Sobolev, V.; Pavel V.; Moriarty, W., Hekkelman, L., Joosten, P. Perrakis, A. y Adams D. 2020. A Global Ramachandran Score Identifies Protein Structures with Unlikely Stereochemistry. *Structure*. Rev **28**: 1249–1258.
80. Stokland, J., Siitonen, J., Jonsson, B., 2012. Biodiversity in Dead Wood (Ecology, Biodiversity and Conservation). En Cambridge University Press, Cambridge, pp. 27–36
81. Sun, H. N., Yu, C. M., Fu, H. H., Wang, P., Fang, Z. G., Zhang, Y. Z., Chen, X. L., & Zhao, F. 2021. Diversity of Marine 1,3-Xylan-Utilizing Bacteria and Characters of Their Extracellular 1,3-Xylanases. *Frontiers in Microbiology*, **12**.
82. Terrone, C., Montesino de Freitas Nascimento, J. Fanchini C., Brienza, M., Carmona, E. 2020. Salt-tolerant  $\alpha$ -arabinofuranosidase from a new specie *Aspergillus hortai* CRM1919: Production in acid conditions, purification, characterization and application on xylan hydrolysis. *Biocatalysis and Agricultural Biotechnology*. **23**: 101460–.
83. Thermo Fisher. 2012. T4 DNA Ligase. Recuperado 5 de agosto de 2021, de [https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2Ft4dnaligase\\_1U\\_man.pdf](https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2Ft4dnaligase_1U_man.pdf)
84. Thermo Fisher. 2015a. User Guide: SuperScript IV Reverse Transcriptase. Thermo Fisher. Recuperado 5 de agosto de 2021, de [https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FSSIV\\_Reverse\\_Transcriptase\\_UG.pdf](https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FSSIV_Reverse_Transcriptase_UG.pdf)
85. Thermo Fisher. (2015b). User Guide: GeneJet Gel Extraction Kit. Recuperado 5 de agosto de 2021, de [https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0012661\\_GeneJET\\_Gel\\_Extraction\\_UG.pdf](https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FLSG%2Fmanuals%2FMAN0012661_GeneJET_Gel_Extraction_UG.pdf)
86. Thermo Fisher. (2018). CloneJET PCR Cloning Kit. Recuperado 5 de agosto de 2021, de <https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FBID%2FReference-Materials%2Fclonejet-pcr-cloning-kit-labaid.pdf>
87. Uffen, R. L. 1997. Xylan degradation: a glimpse at microbial diversity. *Journal of Industrial Microbiology and Biotechnology*. **19**: 1-6.
88. Valadares, F., Gonçalves, T. A., Damasio, A., Milagres, A. M., Squina, F. M., Segato, F., y Ferraz, A. 2019. The secretome of two representative lignocellulose-decay basidiomycetes growing on sugarcane bagasse solid-state cultures. *Enzyme and Microbial Technology*. **130**: 109370.
89. Wan A., Jahim, J.; Ismail, A., Fuzi, S. Rahman, R., Illias, R. 2016. High xylooligosaccharides (XOS) production from pretreated kenaf stem by enzyme mixture hydrolysis. *Industrial Crops and Products*. **81**: 11–19.
90. Wan Azelee, Nur Izyan; Jahim, Jamaliah Md; Ismail, Ahmad Fauzi; Fuzi, Siti Fatimah Zaharah Mohamad; Rahman, Roshanida A.; Md Illias, Rosli (2016). High xylooligosaccharides (XOS) production from pretreated kenaf stem by enzyme mixture hydrolysis. *Industrial Crops and Products*. **81**: 11–19.
91. Wang, B. T., Hu, S., Yu, X. Y., Jin, L., Zhu, Y. J., y Jin, F. J. 2020. Studies of Cellulose and Starch Utilization and the Regulatory Mechanisms of Related Enzymes in Fungi. *Polymers*. **12**: 530.
92. Wang, W., Yan, R., Nocek, B. P., Vuong, T. V., Di Leo, R., Xu, X., Cui, H., Gatenholm, P., Toriz, G., Tenkanen, M., Savchenko, A., y Master, E. R. (2016). Biochemical and Structural Characterization of a Five-domain GH115  $\alpha$ -Glucuronidase from the Marine Bacterium *Saccharophagus degradans* 2-40T. *Journal of Biological Chemistry*. **291**: 14120–14133.
93. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ 2009 Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. **25**: 1189-1191.
94. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., y Richardson, D. C. 2018. MolProbity: More and better reference data for improved all-atom structure validation. *Protein science : a publication of the Protein Society*. Rev **27**: 293–315.
95. Wlodawer A. 2017. Stereochemistry and Validation of Macromolecular Structures. *Methods in molecular biology*. Rev **1607**: 595–610.
96. Yagi, H., Maehara, T., Tanaka, T., Takehara, R., Teramoto, K., Yaoi, K., y Kaneko, S. 2017. 4-O-Methyl Modifications of Glucuronic Acids in Xylans Are Indispensable for Substrate Discrimination by GH67  $\alpha$ -Glucuronidase from *Bacillus halodurans* C-125. *Journal of Applied Glycoscience*. **64**: 115–121.
97. Yan, R., Wang, W., Vuong, T. V., Xiu, Y., Skarina, T., Di Leo, R., Gatenholm, P., Toriz, G., Tenkanen, M., Stogios, P. J., y Master, E. R. 2021. Structural characterization of the family GH115  $\alpha$ -glucuronidase from *Amphibacillus xylanus* yields insight into its coordinated action con  $\alpha$ -arabinofuranosidases. *New Biotechnology*. **62**: 49–56.
98. Yang J., Yan, R., Roy, A., Xu, D., Poisson, J., y Zhang 2015. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*. **12**: 7-8.
99. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden T 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. **13**:134
100. Yousuf, A., Pirozzi, D., y Sannino, F. 2020. Lignocellulosic Biomass to Liquid Biofuels (1ª ed). *Fundamentals of lignocellulosic biomass*. pp 1-15. Academic Press
101. Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*. **9**: 40.
102. Zhang, M., Chekan, J. R., Dodd, D., Hong, P. Y., Radlinski, L., Revindran, V., Nair, S. K., Mackie, R. I., & Cann, I. 2014. Xylan utilization in human gut commensal bacteria is orchestrated by unique modular organization of polysaccharide-degrading enzymes. *Proceedings of the National Academy of Sciences*. **111**: E3708-E3717
103. Zhou, T., Xue, Y., Ren, F., Dong, Y. 2018. Antioxidant activity of xylooligosaccharides prepared from *Thermotoga maritima* using recombinant enzyme cocktail of  $\beta$ -xylanase and  $\alpha$ -glucuronidase. *Journal of Carbohydrate Chemistry*. **37**: 210-224
104. Zhou, X., Li, W., Mabon, R., y Broadbelt, L. J. 2016. A Critical Review on Hemicellulose Pyrolysis. *Energy Technology*. **5**: 52–79.

## Anexo

### 1. Alineamientos de secuencias para las enzimas de las familias GH67 y GH115

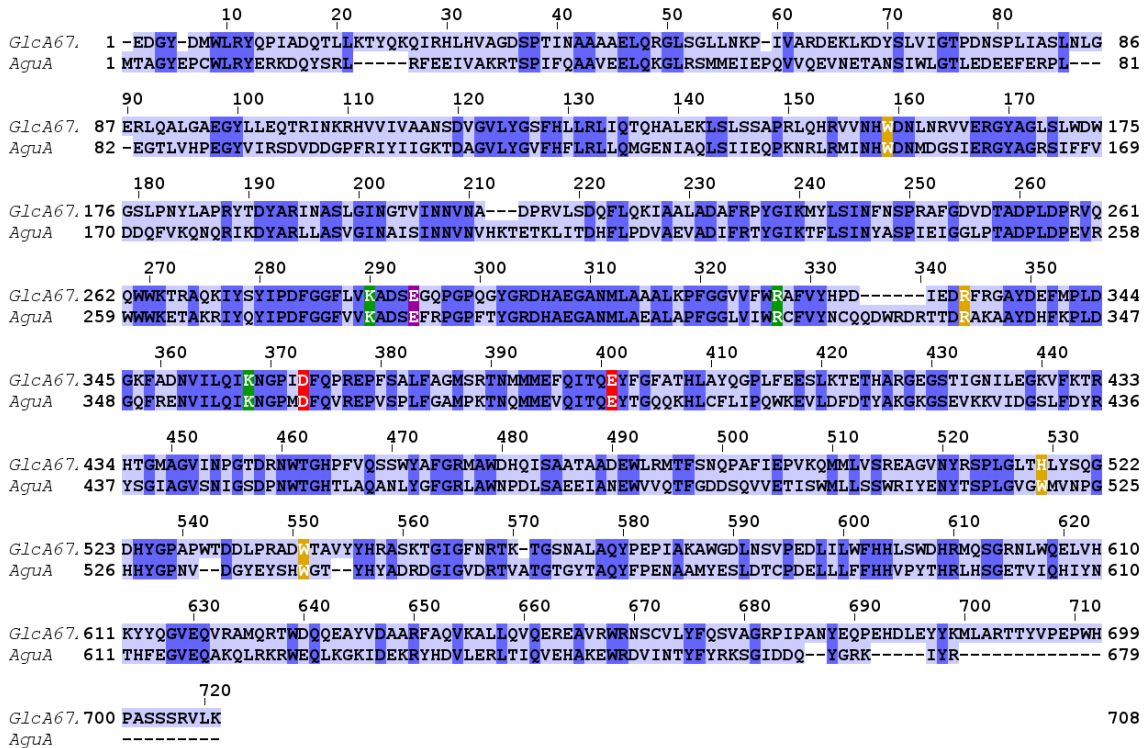


Figura 26. Alineamiento por Clustal  $\Omega$  de las secuencias de GlcA67 y AguA. Los residuos conservados más importantes están marcados con colores, el ácido catalítico está marcado en morado, las bases catalíticas están marcadas en rojo, los residuos de unión a xilano en amarillo y los de unión a ácido glucurónico en verde. El porcentaje de identidad está representado por una gradiente ascendente de azul.

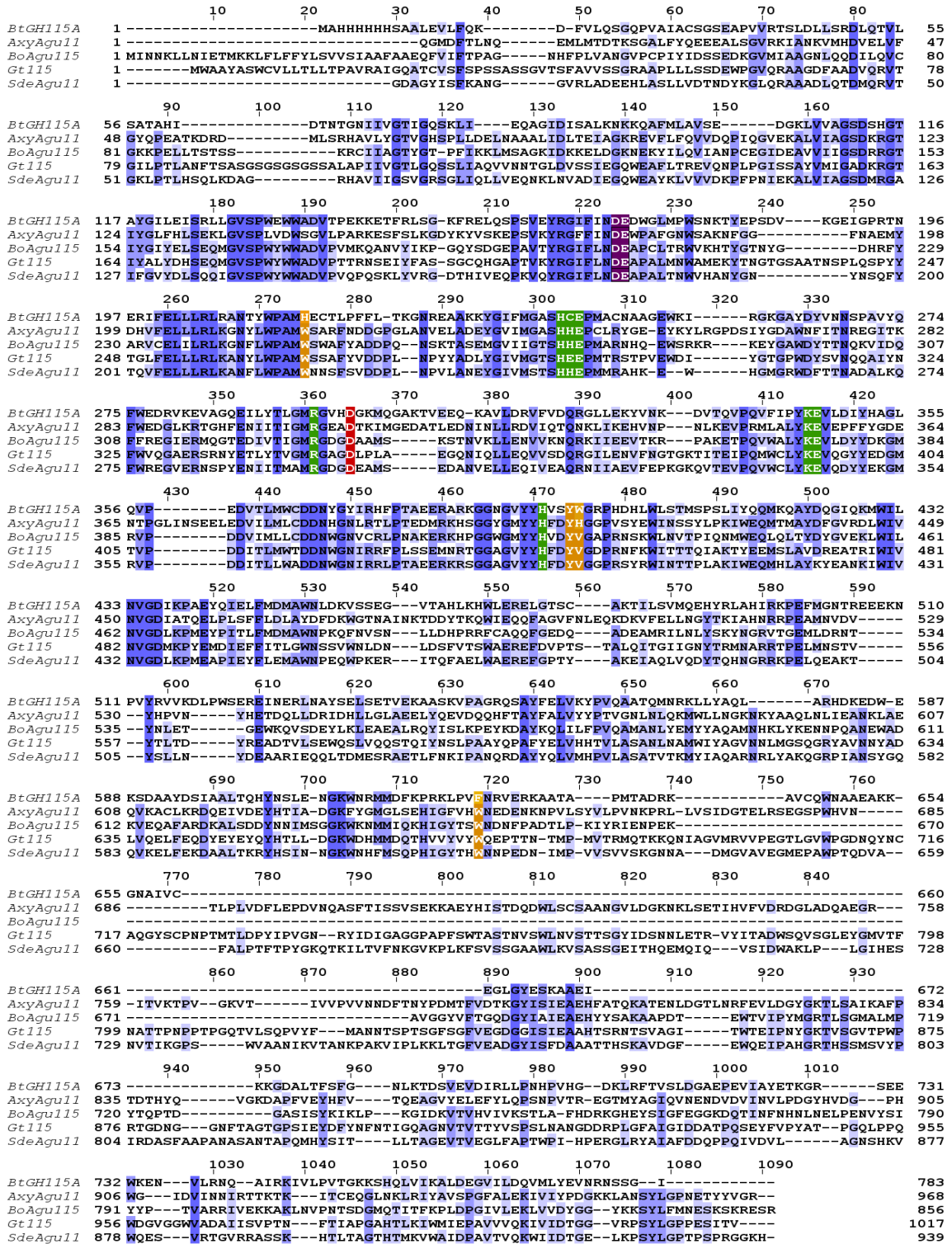


Figura 27. Alineamiento por Clustal  $\Omega$  de las secuencias de BtGH115, AxyAgu115A, BoAgu115A, Gt115 y SdeAgu115A. Los residuos conservados más importantes están marcados con colores, la diada de ácidos catalíticos está marcada en morado, la base catalítica marcada en rojo, los residuos de unión a

xilano en amarillo y los de unión a ácido glucurónico en verde. El porcentaje de identidad está representado por una gradiente ascendente de azul.

## 2. Datos derivados del análisis bioinformático

Tabla 10. Grupos de residuos conservados catalíticamente relevantes en posición y/o función para las enzimas de la familia GH67 de las que actualmente se dispone un modelo 3D obtenido experimentalmente. Los residuos están indicados por su código de una letra, y su posición indicada con un número relativo a la secuencia de cada una de las enzimas. (UX) se refiere a la función de unión a xilano, (UG) a la función de unión a ácido glucurónico o metil glucurónico, (AC) se refiere a la función del residuo como ácido catalítico, (BC) a la función del residuo como base catalítica y (S/E) para indicar que no hay residuo equivalente para ese grupo. El dominio donde se encuentra cada grupo se anota en números romanos del I al III.

Grupo	Residuos <i>P. cellulosa</i>	Residuos <i>G.stearothermophilus</i>	Función del residuo	Dominio
1	W160	W150	UX	I
2	K288	K281	UG	II
3	E292	E285	AC	II
4	R325	R318	UG	II
5	R336	R335	UX	II
6	K360	K359	UG	II
7	D365	D364	BC	II
8	E393	E392	BC	II
9	H521	W520	UX	III
10	W543	W540	UX	III

Tabla 11. Grupos de residuos conservados catalíticamente relevantes en posición y/o función para las enzimas de la familia GH115 de las que actualmente se dispone un modelo 3D obtenido experimentalmente. Los residuos están indicados por su código de una letra, y su posición indicada con un número relativo a la secuencia de cada una de las enzimas. (UX) se refiere a la función de unión a xilano, (UG) a la función de unión a ácido glucurónico o metil glucurónico, (AC) se refiere a la función del residuo como ácido catalítico, (BC) a la función del residuo como base catalítica y (S/E) para indicar que no hay residuo equivalente para ese grupo. El dominio donde se encuentra cada grupo se anota según la convención de la literatura A, B, C, C<sup>+</sup> o D.

Grupo	Residuos <i>B. ovatus</i>	Residuos <i>S. degradans</i>	Residuos <i>A. xylanus</i>	Residuos <i>B. thetaiotaomicron</i>	Residuos <i>G.trabeum</i>	Función	Dominio
1	D206	D215	D175	D169	D216	AC	B
2	E207	E216	E176	E170	E217	AC	B
3	S/E	S/E	W177	S/E	S/E	UX	B
4	W249	W256	W216	H216	W267	UX	B
5	H275	H282	H244	H243	H293	UG	B
6	H276	H283	H245	C244	E294	UG	B

7	E277	E284	E246	E245	E295	UG	B
9	R328	R331	R301	R295	R345	UG	B
10	D332	D335	D305	D299	D349	BC	B
11	K374	K380	K352	K345	K394	UG	B
12	E375	E381	E353	E346	E395	UG	B
13	H422	H428	H408	H399	H442	UG	B
14	Y425	Y431	Y411	Y402	Y445	UX	B
15	V426	V432	H412	W397	V446	UX	B
16	S/E	S/E	S/E	E507	N/A	UX	C
17	S649	H655	H642	S/E	Y671	UX	C
18	W650	W656	W643	F625	W672	UX	C
19 <sup>a</sup>	S/E	W689	W680	S/E	W708	UX	C+
20 <sup>a</sup>	N/A	F696	H681	N/A	Y714	UX	C+
21 <sup>a</sup>	N/A	W773	N/A	N/A	Y720	UX	C+

a. Grupos que poseen residuos que comparte una función común, pero difieren en la posición de la estructura tridimensional

Tabla 12. Análisis de MolProbity de residuos catalíticamente relevantes de la proteína Gt115 (XP\_007865968). Se describen los parámetros más relevantes de la geometría como son el solapamiento estérico, los ángulos  $\phi/\psi$  del gráfico de Ramachandran, los rotámeros y las desviaciones del carbono  $\beta$ . El análisis completo puede ser hallado [aquí](#).

Residuo	Solapamiento (>0,4 Å)	Ramachandran	Rotámero	Desviación carbono $\beta$
D216	-	Favorecido (8.84%) General / -154.8,179.7	Favorecido (26.7%) <i>p0</i> Ángulo chi: 59.7,59.9	0.16 Å
E217	-	Favorecido (51.43%) General / -68.0,-10.2	Favorecido (32.1%) <i>pm20</i> Ángulo chi: 69.7,274.8,9.5	0.06 Å
W267	0.43 Å HZ3 con A 222 MET SD	Favorecido (84.58%) General / -60.3,-47.8	Favorecido (77%) <i>t60</i> Ángulo chi: 182.8,81.3	0.14 Å
H293	0.54 Å HD2 con A 345 ARG HD3	Favorecido (63.32%) General / -58.0,-27.3	Favorecido (48.9%) <i>p-80</i> Ángulo chi: 57.3,283.4	0.09 Å
E294	-	Favorecido (39.88%) General / -83.8,-17.8	Favorecido (98.2%) <i>mt-10</i> Ángulo chi: 292.8,177.2,2	0.11 Å
E295	0.53 Å HG3 con A 299 ARG O	Favorecido (32.28%) Pre-Pro / -109.1,98.7	Favorecido (82.1%) <i>tt0</i> Ángulo chi: 182.3,177.9,14.2	0.21 Å
R345	0.54 Å HD3 con A 293 HIS HD2	Favorecido (17.23%) General / -123.8,165.5	Favorecido (61.8%) <i>mtp180</i> Ángulo chi: 301.4,194.5,69.2,187	0.15 Å
D349	-	Favorecido (11.65%) General / 56.6,27.0	Favorecido (60.5%) <i>m-30</i> Ángulo chi: 295.1,305.8	0.10 Å
K394	-	Favorecido (20.49%) General / -50.2,135.8	Favorecido (61.1%) <i>pttt</i> Ángulo chi: 65.5,182.6,181.9,180.3	0.11 Å
E395	-	Permitido (0.49%) General / 78.3,-53.0	Favorecido (57.9%) <i>mt-10</i> Ángulo chi: 290.8,180.9,298.9	0.30 Å
H442	-	Favorecido (52.15%) General / -	Favorecido (57.2%) <i>m-70</i> Ángulo chi: 298.5,259.4	0.11 Å



		125.5,134.2		
Y445	-	Favorecido (3.29%) General / - 165.4,134.2	Favorecido (59%) <i>t80</i> Ángulo chi: 186.5,70.4	0.10Å
V446	-	Favorecido (57.81%) Ile or Val / - 108.8,119.2	Favorecido (84.2%) <i>t</i> Ángulo chi: 176.3	0.06Å
Y671	-	Favorecido (2.74%) General / -100.8,- 174.9	Favorecido (86.5%) <i>m-80</i> Ángulo chi: 293.4,282.8	0.09Å
W672	-	Favorecido (69.91%) General / -62.6,-50.8	Favorecido (58.2%) <i>m-10</i> Ángulo chi: 295.1,350.3	0.13Å
W708	0.48Å HB3 con A 709 PRO HD3	No permitido (0.02%) Pre-Pro / -152.4,-72.7	Favorecido (55.8%) <i>t-100</i> Ángulo chi: 192.2,255.9	0.29Å
Y714	0.92Å HD1 con A 722 CYS HG	Favorecido (58.55%) General / -75.2,-9.4	Favorecido (27.7%) <i>t80</i> Ángulo chi: 193.1,65.7	0.29Å
Y720	-	Favorecido (65.49%) General / -73.5,-37.7	Favorecido (48.1%) <i>m-80</i> Ángulo chi: 283.2,282.2	0.10Å

### 3. Secuencias de DNA y alineamientos para las colonias 4P15XC1 y 4P15XC2

>4P15XC1

```
gctgccatggaaaatcgatggttcttcttttattctctcaagattttcaggctgtatatta
aaacttatattaagaactatgctaaccacctcatcaggaaccggttgtaggtggcgtgggt
tttcttggcaatcgactctcatgaaaactacgagctaaatattcaatatgttccctcttga
ccaactttattctgcatTTTTTTTTgaacgaggttttagagcaagcttcaggaaactgagac
aggaatTTTattaaaaatttaaattttgaagaaagttcagggTTaatagcatccatTTTT
tgctttgcaagttcctcagcattcttaacaaaagacgtctcttttgacatgTTTaaagtt
taaacctcctgtgtgaaattattatccgctcataattccacacattatacgagccggaag
cataaagtgtaaagcctggggtgcctaataagtgagtgagctaactcacattaattgCGTtgcg
ctcactgccaattgctttccagtcgggaaacctgctcgtgccagctgcattaatgaatcgg
ccaacgcgcggggagagggcggTTTgCGtattgggCGctcttccgcttcctcgctcactga
ctcgtgCGctcggtcgttcggctgCGgcgagcggtatcagctcactcaaaggcggtaat
acggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagca
aaaggccaggaaccgtaaaaaggcCGcgttgctggCGTTTTTccataggctccgCCCCC
tgacgagcatcacaAAAatcgacgctcaagtcagaggtggcgaaaccgacaggactata
aagataccaggcgtttccccctggaagctccctcgtgCGctctcctgTTccgacctgCC
gcttaccggataacctgtccgcTTTTctcccttcgggaagcgtgCGctttctcatagctca
cgctgtaggtatctcagttcgggtgtagtcgTTcGctccaagctgggctgtgtgcacgaac
ccccgttcagcccgaccgc
```

>4P15XC2

```
gatggctcgagTTTTTcagcaagaatctttctagaagatctcctacaatattctcagctg
ccatggaaaatcgatggttcttcttttattctctcaagatTTTcaggctgtatattaAAC
ttatattaagaactatgctaaccacctcatcaggaaccggttgtaggtggcgtgggttttc
ttggcaatcgactctcatgaaaactacgagctaaatattcaatatgttccctcttgaccaa
```

ctttattctgcatttttttgaacgaggttttagagcaagcttcaggaaactgagacagga  
atthttattaaaaatthaaatthttgaagaaagttcagggttaatagcatccattttttgct  
ttgcaagttcctcagcattcttaacaaaagacgtctcttttgacatgtttaaagtttaa  
cctcctgtgtgaaattattatccgctcataattccacacattatacgagccggaagcata  
aagtgtaaagcctggggtgcctaataagtgagtaactcacattaattgcggtgcgctca  
ctgccaattgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaa  
cgcgcggggagagggcggtttgcgtattgggcgctcttccgcttcctcgtcactgactcg  
ctgcgctcggtcggtcggctgcggcgagcgggtatcagctcactcaaaggcggtaatacgg  
ttatccacagaatcaggggataacgcaggaagaacatgtgagcaaaaggccagcaaaag  
gccaggaaccgtaaaaaggccgcggttgctggcggtttttccataggctccgccccctgac  
gagcatcaaaaaatcgacgctcaagtcagaggtggcgaaccgcagaggactataaagat  
accaggcgtttccccctggaagctccctcgtgcgctctcctggtccgaccctgcccgtta  
ccggataacctgtccgcctttctcccttcggg

>4P15XC4

gtttttcagcaagatthtatccacagaatcaggggataacgcaggaagaacatgtgagcaa  
aaggccagcaaaaggccaggaaccgtaaaaaggccgcggttgctggcggtttttccataggct  
ccgccccctgacgagcatcaaaaaatcgacgctcaagtcagaggtggcgaaaaccgcaca  
ggactataaagataaccaggcgtttccccctggaagctccctcgtgcgctctcctggtccga  
ccctgccgcttaccggataacctgtccgcctttctcccttcgggaagcgtggcgctttctca  
tagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaagctgggctgtgtg  
cacgaacccccggttcagcccgaaccgctgcgccttatccggtaactatcgtcttgagtcca  
accggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagc  
gaggtatgtaggcgggtgctacagagttcttgaagtggcctaactacggctacactaga  
agaacagtatthggatctgcgctctgctgaagccagttaccttcgaaaaagagttggta  
gctcttgatccggcaaaaaaccaccgctggtagcgggtgggtttttttggttgcaagcagca  
gattacgcgcagaaaaaaaggatctcaagaagatcccttgatctttttctacggggtctgac  
gctcagtggaacgaaaactcacgttaagggatthttggatcatgagattatcaaaaaggatct  
tcacctagatcctthtaattaaaaatgaagthtttaaatcaatctaaagtatatatgagta  
aacttggtctgacagttaccaatgcttaatacagtgaggcacctatctcagcgatctgtcta  
thtcggtcatccatagttgcctgactccccgctcgtgtagataactacgatacgggagggt  
taccatctggccccagtgctgcaatgataaccgcgagaccacgctcacggctccagattta  
tcagcaataaaccagccagccgagcgcagcagaagtggtctgcaactt