**UNIVERSIDAD COMPLUTENSE DE MADRID**
FACULTAD DE FARMACIA


# TESIS DOCTORAL

# Estudio funcional de las Beta Glucosidasas del hongo Talaromyces amestolkiae: aplicaciones biotecnológicas y diseño racional de catalizadores

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Juan Antonio Méndez Líter

DIRECTOR

María Jesús Martínez Hernández
Laura Isabel de Eugenio Martínez

*A mi padre, Juan, descansa en paz*

*A mi madre Ana, y mi hermano, Hugo*

*“The individual has always had to struggle to keep from being overwhelmed by the tribe. To be your own man is a hard business. If you try it, you'll be lonely often, and sometimes frightened. But no price is too high to pay for the privilege of owning yourself”*

- Rudyard Kipling.

# Agradecimientos

*La presente tesis doctoral se ha desarrollado en el Centro de Investigaciones Biológicas, del Consejo Superior de Investigaciones Científicas. Ha sido financiada gracias a los proyectos BIO2015-68387-R del MINECO y RETO-PROSOT S2013/MAE-2972 de la Comunidad de Madrid. Además, la tesis doctoral se realizó gracias a una beca predoctoral concedida por la Fundación Tatiana Pérez de Guzmán el Bueno, en el área de medioambiente. Me gustaría agradecer especialmente la beca otorgada por dicha fundación. En una época en la que la financiación pública no pasa por su mejor momento, el hecho de que existan iniciativas privadas como las que desempeña dicha Fundación, resulta absolutamente fundamental para la supervivencia de la ciencia en España.*

Antes de comenzar con los agradecimientos, me gustaría dedicar un par de líneas a hacer una reflexión sobre lo que significa la realización de una tesis doctoral. Y tras estos 4 años experimentándolo, la mejor definición que se me ocurre es que es un viaje personal incomparable, y una lucha constante con uno mismo, en la cual se aprende, se crece, se lidia con todo tipo de problemas y complicaciones, y se termina siendo una persona bastante diferente a la que empezó todo el proceso. La experiencia es dura, a la par que enriquecedora, y lo que está claro es que no es un reto que se pueda abordar solo, sino que se necesita la ayuda y colaboración de decenas de personas que, en mayor o menor medida, contribuyen durante este tiempo a que la tesis doctoral se haga realidad.

Y la primera de esas personas, como no puede ser de otra manera, es la directora de esta tesis doctoral, la doctora María Jesús Martínez. Gracias por acogerme desde enero de 2014 cuando era un pipiolo que venía a hacer el trabajo fin de grado, pasando por el trabajo fin de máster, y hasta hoy, siempre con un enorme cariño, dedicación y esfuerzo, a la par que una gran capacidad de enseñarme cualquier cosa relacionada con la biomasa lignocelulósica y sus intrincados mecanismos. Y te querría agradecer con especial atención estos últimos meses, duros para todos pero sobre todo para ti, en los que has demostrado no sólo ser maestra, sino una de las personas más fuertes que conozco. Gracias, María Jesús.

Infinitas gracias también a mi codirectora de tesis (y de TFG y de TFM, parece que llevas toda la vida corrigiéndome cosas), la doctora Laura de Eugenio, que como ya se ha dicho alguna vez hace no tanto, es la mejor codirectora que haya existido jamás. Estructural en el CIB, en el grupo de Biotecnología para la Biomasa Lignocelulósica, y en cualquier tesis que pase por el laboratorio 249. Para la posteridad quedarán nuestras

discusiones de ciencia en el HPLC, molestando a la gente que se sienta por ahí cerca, pero que tanto he disfrutado. Gracias Lidem.

Por último, una mención en esta parte de los agradecimientos a la doctora Alicia Prieto. Co-jefa del grupo, y si no codirectora de esta tesis, casi casi. Muchas gracias por tus consejos, por dedicarnos los minutos que hagan falta siempre que lo necesitamos. Y por supuesto, por tu buen humor, y por las desinteresadas invitaciones a cañas y a tu piscina. Nuestro grupo no sería lo mismo sin ti.

Gracias también a todos los miembros del grupo de Biotecnología para la Biomasa Lignocelulósica porque, quien más, quien menos, todos habéis aportado vuestro granito de arena a que esta tesis se haga realidad. Empezando por el laboratorio 249, en primer lugar, gracias a Manu. ¿También se podría decir aquí esto de casi codirector? Probablemente. Gracias por toda tu ayuda y consejos, por las eternas conversaciones de ciencia que tanto estoy echando de menos desde que los vikingos te capturaron, y gracias también por cosas de fuera del labo, por esos planes con los Manolos, por tu incomparable sentido del humor, por las tan interesantes conversaciones culturales… La lista es inagotable, y desde luego para mi eres un amigo para siempre. Gracias también a Fran. Seguramente la mejor persona (y la más popular) que ha estado (¿y estará de nuevo?) nunca por el CIB. Super trabajador, supusiste un soplo de aire fresco al laboratorio con tus medios, placas, bufferes… Y fuera del labo, integrándote perfectamente en los Manolos (como suplente), y haciéndote poco a poco indispensable en nuestra vida. Ah, se me olvidaba, ¡a casa ya Fran! Gracias a Jezú. Otro manolo, madridista empedernido, y maestro (te guste o no, lo dice el título que te dieron) de las maravillosas e inigualables β-glucosidasas. Gracias porque también me ayudaste un montón en el TFG y TFM, aunque tu partida hacia tierras barbateñas fuera más rápida de lo que me habría gustado. Me salto el orden para agradecer aquí también a Rosita y Marco (recuerda siempre que fui el primer Manolo en conocerte, queda escrito), el nuevo y encantador fichaje. Gracias, como no, al doctor Jorge Barriuso. Siempre he pensado que nuestro grupo no sería ni la mitad de divertido sin ti, y cada vez estoy más convencido de ello. Gracias por tus consejos, por transmitir siempre la forma apasionada con la que vives la ciencia, y por estar siempre dispuesto a unas cañas o un partido de pádel. Gracias a María Molina, reina de las inmovilizaciones. Gracias por siempre haber sido tan amable conmigo, por transmitir buen rollo a todo el grupo. Lo petarás seguro en la industria farmacéutica, pero no sin antes celebrar tu futura tesis como dios manda. Gracias a Rosa, que fue coetánea mía durante 3 años en el 249, por ser tan buena persona, y ser capaz de lidiar con la DyP sin perder las fuerzas. Ojalá te vaya muy bien en tu doctorado en la maravillosa Austria. Gracias a Mariu, sabia de las lipasas, capaz de

organizar el labo como nadie, pero de sumarse al buen rollo cuando era necesario. Gracias a Neumara, tan trabajadora, conseguiste aportar ese toque de aplicación industrial que siempre nos falta en la investigación básica. Ojalá te vaya muy bien. Gracias también a Anna y Carlos, ex-TFMs, y nuevas incorporaciones, sois los herederos de todo esto. Ojalá podáis tener una experiencia tan enriquecedora en el laboratorio como la he tenido yo. Por último, mencionar a algunos de los becarios, que en los meses que han pasado por el laboratorio, han dejado huella: Iñigo, que lástima que no saliese esa beca, te la merecías. Albertonero, eres un tío encantador. María Mer, acabé muy orgulloso de tu 9,8 en el TFG. Lara, capaz de renunciar a un mes de veraneo por venir con nosotros, espero que sigas disfrutando de la ciencia en tu futuro TFG aquí. Jaime, inigualable e incomparable, sólo tú eres capaz de poner freno a tu infinito talento. Albertini, casi veterano del laboratorio, y dignísimo aprendiz del intrincado quorum sensing, eres una de las personas más divertidas que conozco, ojalá te vaya estupendamente en el futuro. Gracias por estos meses en los que conseguías que el 249 estuviese un poco menos vacío. También gracias a Marta, otra excelente aprendiz del quorum sensing, por sumarte cada vez que me metía con Albertini, y hacer tan divertido el labo estos meses. Sé que faltáis muchos (Iván, recién llegado, Malina, Amaia, Cynthia, los Joses, Lucía…), pero no quiero que esto se eternice, así que gracias a todos.

Cambio por fin del laboratorio, porque el grupo de Biotecnología para la biomasa lignocelulósica no es sólo un único laboratorio… De hecho, somos 5, así que aún queda mucha gente a la que agradecer. Gracias a Ángel, super jefazo, porque eres un ejemplo científico en el que mirarse con prácticamente cualquier cosa que haces. Gracias a Susana, por tus innumerables consejos, simpatía, y por haber sido siempre tan amable conmigo cada vez que he necesitado tu ayuda. Gracias a Javi, por su conocimiento científico, las charlas sobre las LPMOs, y su enorme simpatía y disponibilidad. Gracias a Marta, por ser siempre un ejemplo de afecto que sube el estado de ánimo de todo el grupo. Gracias a Lola, por su enorme amabilidad, simpatía, y estar siempre dispuesta a echarte una mano. Gracias a Ana, aragonesa hasta la sepultura, con un enorme corazón (aunque dicen por ahí que un poco autoritaria), y siempre la primera a la hora de tomarse unas cañas donde sea. Gracias a Juan, tan divertido e inteligente, eres un ejemplo de que en la vida la actitud correcta es saber reírse de uno mismo. Gracias a Iván, por ser un ejemplo a seguir, por esas conversaciones sobre videojuegos o manga, y por tener tanta paciencia con la lata que te damos Lidem y yo en nuestros eternos intercambios de opinión. Gracias a Felipe, por ser un encanto de persona, tan divertido y creativo. Estas charlas sobre burocracia UCM y demás cosas de las tesis durante estos últimos meses han sido muy liberadoras. Gracias a Marisa, por aguantar el marrón de presentarme en el seminario de departamento, y

por ese humor maligno que tiene. Tienes un gran futuro por delante. Gracias a Pablo, tan trabajador e inteligente. Espero que algún día puedas perdonar lo del contagio de las paperas, poco nos faltó para iniciar una epidemia. Gracias a David, capaz de alegrar cualquier reunión del grupo. Gracias a Gonzalo, por esa forma de ver el mundo tan oscura que tanta gracia me hace. Gracias a Rash, por ser tan majete. Seguro que tu camino inverso doctorando – masterando acaba teniendo recompensa. Seguid trabajando duro, chicos del 200, lo petaréis seguro. Gracias a Vero, ahora exploradora del páramo sueco, por ser un ejemplo de candidez y amabilidad. Gracias a Isa Pardo, postdoc modélica del 245, ojalá lo estés petando bastante en Estados Unidos. Gracias a Isa Vicente, mitad CIB mitad ICP, por ser siempre tan divertida y preocuparse por cómo nos va la tesis. Gracias a Elena, por tus consejos y por enseñarnos lo que es la eficiencia personificada. Y por supuesto, muchas gracias a toda la demás gente que ha pasado por uno de estos laboratorios a lo largo de estos años: Cris Coscolín, Bea Balcells, Rocío, Ander, Mario, Julia, Chantal…

He acabado (creo) con el grupo de Biotecnología para la Biomasa Lignocelulósica, pero, como es normal, una de las claves de la ciencia a día de hoy es colaborar con otros grupos. Y sin salir del CIB, los primeros que se merecen mi agradecimiento son el grupo de RMN y reconocimiento molecular. Gracias al doctor Javier Cañada, jefe de dicho grupo, por ser siempre tan amable y buena persona, por sus acertados consejos, y prestarnos su tiempo siempre que lo hemos necesitado. Y gracias a Bea, seguramente la persona más eficiente del CIB, por tantas horas de trabajo para resolver los glicósidos que te mandamos sin descanso. Tendrás un futuro brillante, en la ciencia o dónde sea. Esta tesis no sería igual sin vuestro curro. Especial agradecimiento también a dos servicios del centro, el de Cromatografía de Gases, y el de Genómica y Proteómica.

Moviéndonos entre otros centros del CSIC, hay que hacer una parada en el IQOG. Muchas gracias a Andrés. Sin tu ayuda no habríamos hecho ni la mitad en nuestros proyectos transglicosiladores, ni nos lo habríamos pasado tan bien en el congreso de Lisboa.

Otra parada debe hacerse en el ICP. Berni, Javi, Noa, Paloma, Diana… sea en unas cañas o en congresos, gracias por el apoyo y la diversión. Agradecimientos especiales para Kiko Plou, por prestarnos los diferentes compuestos fenólicos para nuestros ensayos de glicosilación, sus consejos, y su enorme amabilidad.

Dentro de las colaboraciones de la UCM, no me puedo olvidar del Grupo de Señalización por Cannabinoides, de la Universidad Complutense de Madrid. En especial, de la parte del grupo centrada en la investigación del cáncer de mama. Gracias a Cristina Sánchez, por organizar el trabajo, y

aceptar la colaboración para probar las actividades biológicas de los glucósidos. Gracias a Isa, por el excelente y riguroso trabajo que hiciste hasta que conseguiste poner a punto todo el ensayo. Ojalá podamos volver a colaborar con vosotros en el futuro, ha sido un placer.

Creo que existe un cierto consenso entre los becarios predoctorales en que el período de estancia internacional es de los que más se disfruta de la tesis. En mi caso no ha sido menos. Los tres meses que pasé en Viena fueron maravillosos, tuve resultados, y conocí a gente genial. Así, creo que merecen un agradecimiento especial en esta sección:

*Thanks to Dr. Roland Ludwig, from The University of Natural Resources and Life Sciences, Vienna (BOKU), for welcoming me into his laboratory, and helping me to improve my knowledge about the complicated (really complicated) LPMOs. Thank you for your kindness, sympathy, and for having such a nice group of people. Thanks to Florian, for being my main guide during the first weeks there. Your vast knowledge in fermentations was a great help. Thanks to Daniel, an example to be followed, for letting me help you with your CDHs and mutant LPMOs while waiting for my genes to arrive. Thanks to Erik and Christophe, experts in LPMOs, for your help and advice. Thanks to Stefan and Lena, for welcoming me so well the first few days, and always being willing to joke or have some beers. Thank you all because I learned so much there and it was an unparalleled experience.*

Por supuesto, también debe haber un espacio para la gente que no está tan relacionada con el trabajo, porque han sido el apoyo fundamental para que estos años se hayan hecho mucho más llevaderos. En primer lugar, muy relacionado con el CIB, destacar al grupo de los Manolos. Manu, Jesús y Fran ya habéis tenido vuestro hueco en esta sección de la tesis, así que paso a los otros miembros. Gracias a Emi. Fundador y administrador del grupo, tu agorerismo y quejas sólo están al nivel de tu incomparable inteligencia y extensa cultura. Gracias por los momentos en los viajes, cenas, y películas en las que te quedabas dormido, pero venías a ver por estar con nosotros. Y por supuesto, por ayudarme con los experimentos de RT-qPCR, un pedacito del capítulo 3 de la tesis es gracias a ti. Gracias a David, el figure, nuevo líder de los Manolos. Gracias por ser una persona increíble, con uno de los corazones más grandes que he conocido. Gracias por hacer entretenidos los ratos más aburridos. Tu elocuencia, a veces tan criticada, nunca dejará de ser indispensable. Trabajador incansable y dedicado, estoy seguro de que más pronto que tarde encontrarás el ansiado contrato indefinido, pero si mientras tenemos el gusto de asistir a la tesis del futuro doctor Almendral, mejor que mejor.

Agradecimiento destacado a todos los miembros del grupo de Zahara. Gracias a Carlos, por ser el mejor amigo que nadie pueda tener. Gracias por tantos ratos desde 3º de la ESO. Gracias a Marco, coétaneo de cumpleaños, y en mi vida desde 1º de primaria hasta hoy (26 años, se dice pronto…). Gracias a Diego, recién bautizado como Iron Troll, una de las personas con las que más me divierto (y discuto). Gracias a Utri, por ser tan buena persona y compañero de debates futbolísticos. Gracias a Alfred, mentor y amigo, porque gracias a ti despertó en mi ese interés por el baloncesto que ha sido primordial en mi vida.

Gracias también a todos mis compañeros en equipos de baloncesto. Gracias a todos los miembros de Glastonbury/Black Label, equipo con el que más he disfrutado jugando. Gracias también a la gente de Botika Classic, con los que por fin he podido ganar una liga municipal. A pesar de que la edad no perdona, he seguido disfrutando del basket gracias a vosotros.

Gracias a los biotecnólogos UEM. Con especial atención a Álvaro e Isa (Musasis, por cierto Isa vuelves a salir), con los que he vivido momentos muy importantes en mi vida. Agradecimiento especial para Noelia, una de las principales responsables de que en tercero de carrera la llama de la ciencia prendiese en mí. Nombrar, por supuesto, a mucha más gente como Jaime Rojas, Jaime Domínguez, David, Óscar, Alberto, Nacho, Esteban, o Charly. Todos sois geniales, ojalá tuviese más espacio para agradeceros como merecéis.

Gracias a los compañeros del máster de microbiología de la UCM. En especial a Adri y Gema. Y claro, Carol, que te mereces un punto aparte para ti sola. Mi mejor amiga, desde que te conocí has sabido animar mi vida hasta límites insospechados, y hemos vivido juntos (por skype) esta etapa predoc, compartiendo sinsabores, y celebrando los éxitos cuando los hemos tenido. Eres la persona que es capaz de hacerme sentir más cómodo en mi vida. Parece poco, pero es mucho decir. Una lástima que en estos 4 años tu marcha a la República Checa haya hecho que tengamos que restringir nuestras quedadas a verano y navidades. Pero bueno, todo acaba, y ya te queda poquita tesis.

Gracias a los compañeros de esa etapa fallida, pero nunca por vosotros, en la facultad de Periodismo de la UCM. Gracias a Javi, Candela, Rober, Cris, Sergio, Elena, y Laura. Ojalá os esté yendo muy bien.

Por último, pero en absoluto menos importante, la familia. Gracias a mis tíos Carmen y Luis, y mi primo David. Gracias por vuestro incansable esfuerzo en mantenernos unidos siempre, por vuestros consejos y apoyo incondicional. Gracias a mi tía Nana, mi tío Antonio (que en paz descanse) y mis primos Margarita, Mariví, Juan y Dara. Gracias por vuestros ánimos y esas comidas cumpleañeras o navideñas que tanto disfruto. Gracias a mis

abuelos Carmen y Roberto, que me enseñaron que el cariño y la bondad son indispensables en la vida. También contribuyó a ello mi tía abuela Pilar. Todos habéis sido un modelo de cómo llegar a altas edades de la vida sin nada que reprochar.

Gracias a mi Padre. A pesar de no haber podido disfrutar muchos años de ti, fuiste esencial en inculcar en nosotros tu pasión por la cultura y el estímulo intelectual. Gracias por tanto cariño y esfuerzo. Por último, gracias a mi hermano Hugo, y mi madre Ana. Gracias a Hugo por ser el mejor hermano y amigo que uno pueda tener, por preocuparse tanto por mí, y por estar siempre dispuesto a pasar tiempo conmigo. Estoy muy orgulloso de cómo te va la vida, pero nunca te conformes. Gracias a mi madre por ser de forma objetiva e indiscutible la mejor persona que existe en el planeta. Gracias por no hundirte en 2004, y seguir adelante con una fuerza titánica cuidando de nosotros hasta hoy (aunque ya nos haga menos falta). Gracias a ambos por mantener mi mundo estructurado cuando más falta hace. Sin vosotros esta tesis habría sido mucho más dura.

# ÍNDICE

# ABREVIATURAS

## Bases nitrogenadas

| | |
|---|---|
| A | Adenosina |
| C | Citosina |
| G | Guanina |
| T | Timina |

## Aminoácidos

| | | | |
|---|---|---|---|
| A/Ala | Alanina | M/Met | Metionina |
| C/Cys | Cisteína | N/Asn | Asparagina |
| D/Asp | Ácido Aspartico | P/Pro | Prolina |
| E/Glu | Ácido Glutamico | Q/Gln | Glutamina |
| F/Phe | Fenilalanina | I/Ile | Isoleucina |
| G/Gly | Glicina | R/Arg | Arginina |
| H/His | Histidina | T/Thr | Treonina |
| I/Ile | Isoleucina | V/Val | Valina |
| K/Lys | Lisina | W/Trp | Triptófano |
| L/Leu | Leucina | Y/Tyr | Tirosina |

## Enzimas

| | |
|---|---|
| AA | Enzimas de actividad auxiliar |
| BGL-1 | β-glucosidasa 1 de *T. amestolkiae* |
| BGL-2 | β-glucosidasa 2 de *T. amestolkiae* |
| BGL-3 | β-glucosidasa 3 de *T. amestolkiae* |
| BGLs | β-glucosidasas |
| BXL | β-xilosidasas |
| CBD | Dominios de unión a celulosa |
| CBHs | Celobiohidrolasas |
| CBMs | Dominio de unión a carbohidratos |
| EGs | Endoglucanasas |
| GH | Glicosil hidrolasa |
| LPMO | Monooxigenasas degradadoras de polisacáridos |

## Resonancia magnética nuclear

| | |
|---|---|
| $^{1}$H-$^{13}$C-NMR | Two-dimensional proton carbon-13 nuclear magnetic resonance |
| $^{1}$H-NMR | Resonancia magnética nuclear del protón |
| HSQC | Heteronuclear single quantum coherence |
| NMR | Resonancia magnética nuclear |

## Otras

| | |
|---|---|
| 1G | Primera generación |
| 2G | Segunda generación |
| 3G | Tercera generación |
| Å | Angstrom |
| ACN | Acetonitrilo |
| ADN | Ácido desoxirribonucleico |
| ADNc | ADN complementario |
| ARN | Ácido ribonucleico |
| BLAST | Basic Local Alignment Search Tool |
| BSA | Albúmina de suero bovino |
| C2 | Celobiosa |
| C3 | Celotriosa |
| C4 | Celotetraosa |
| C5 | Celopentaosa |
| C6 | Celohexaosa |
| CAZy | Carbohydrate Active Enzymes database |
| CMC | Carboximetilcelulosa |
| CSS | Maíz macerado sólido (Corn Steep Solid) |
| EGCG | Epigallocatechin gallate |
| ESI | Electrospray ionization |
| FPLC | Fast protein liquid chromatography |
| HPLC | Cromatografía líquida de alta eficacia (High Performance Liquid Chromatography) |
| $k_{cat}$ | Constante catalítica |
| $k_{cat}/K_m$ | Eficiencia catalítica |
| kDa | kiloDalton |
| *K*i | Constante de inhibición |
| $K_m$ | Constante de Michaelis-Menten |
| KOG | euKaryotic Orthologous Groups |
| LC-MS/MS | Cromatografía líquida acoplada a espectrómetro de masas |
| MALDI-TOF | Desorción/ionización mediante láser asistida por matriz con un analizador de tiempo de vuelo (Matrix-assisted laser desorption/ionization-time of flight) |
| NCBInr | National Center for Biotechnology Information (non-redundant) |
| *o*- | orto-/2- |
| ORFs | Marcos de lectura abiertos |
| PCR | Reacción en cadena de la polimerasa (Polymerase Chain Reaction) |
| PDB | Protein Data Bank |
| pI | Punto isoeléctrico |
| *p*- | para-/4- |
| *p*NP | *p*-nitrofenol |
| *p*NPG | *p*-Nitrofenil-β-D-glucopiranósido |
| *p*NPgal | *p*-Nitrofenil-β-D-galactopiranosido |

| | |
|---|---|
| *p*NPX | *p*-Nitrofenil-β-D-xilopiranosido |
| PSM | Coincidencia del espectro peptídico |
| SDS-PAGE | Electroforesis en gel de poliacrilamida con dodecilsulfato sódico (Dodecyl Sulfate PolyAcrilamide Gel Electrophoresis) |
| TLC | Thin layer chromatography |
| $V_{max}$ | Velocidad máxima de la reacción |

# RESUMEN/SUMMARY

# RESUMEN


## Introducción

La biomasa vegetal representa una importante fuente renovable de materia prima, ya que sus paredes celulares constituyen la mayor parte del carbono fijado por fotosíntesis. Estas paredes celulares están compuestas principalmente por celulosa, hemicelulosa y lignina. La degradación de esta última es clave para utilizar los polisacáridos de la pared celular vegetal, pero también es clave encontrar complejos enzimáticos eficaces para la hidrólisis completa de la celulosa y hemicelulosa, y así poder utilizar todos los carbohidratos que éstas contienen en diferentes aplicaciones biotecnológicas. En lo referente a la celulosa, es el componente mayoritario de la pared celular vegetal (30-50%, dependiendo del tipo de planta). Industrialmente, interesa tanto su sacarificación en monómeros de glucosa, con vistas a la obtención de biocombustibles, como su conversión en productos de alto valor añadido. Existen en la naturaleza numerosos organismos capaces de aprovechar con éxito la celulosa, secretando enzimas que pueden ser usadas en diferentes aplicaciones biotecnológicas. Actualmente, los hongos filamentosos son la principal fuente de estas enzimas para uso comercial, y suelen caracterizarse por su alta productividad y eficacia catalítica, con respecto a las de otros microorganismos.

Existen tres tipos de celulasas, que actúan de forma coordinada para degradar la celulosa hasta moléculas de glucosa: endoglucanasas, celobiohidrolasas y β-glucosidasas. Aunque todas son necesarias en los cócteles enzimáticos, los mayores esfuerzos en los últimos años se han centrado en descubrir nuevas β-glucosidasas, ya que representan el paso clave para la sacarificación de la celulosa y la mayor parte de los cócteles enzimáticos comerciales requieren ser suplementados con β-glucosidasas robustas, especialmente para la producción de bioetanol de segunda generación.

## Objetivos

El objetivo del trabajo fue profundizar en el conocimiento de las β-glucosidasas producidas por el hongo *T. amestolkiae*, enzimas robustas que poseen un papel esencial en la transformación de la celulosa, y son capaces de actuar tanto en reacciones de hidrólisis, como en la síntesis de glucósidos con potencial interés en diferentes sectores tecnológicos. Para conseguir estos objetivos, el plan de trabajo fue:

**1-** Análisis del genoma de *T. amestolkiae* y de su secretoma, con especial interés en las β-glucosidasas producidas por este hongo.

**2-** Expresión heteróloga de las principales β-glucosidasas de *T. amestolkiae* en *P. pastoris*, su purificación y caracterización, y su comparación con las enzimas nativas del hongo.

**3-** Estudio de la sacarificación de residuos lignocelulósicos utilizando cócteles comerciales suplementados con β-glucosidasas de *T. amestolkiae.*

**4-** Obtención de glucósidos de interés mediante reacciones de transglicosilación catalizadas por las β-glucosidasas de *T. amestolkiae*.

**5-** Obtención de variantes de estas enzimas para mejorar los rendimientos en la síntesis de glicósidos de interés.

## Resultados

### Estudio del genoma y proteoma del hongo *T. amestolkiae*

Los análisis genómicos y proteómicos son herramientas potentes para la caracterización metabólica de microorganismos. El hongo *Talaromyces amestolkiae* demostró ser un excelente productor de celulasas, secretando altos niveles de β-glucosidasas, β-1,4-endoglucanasas, y celobiohidrolasas. La secuenciación y el análisis del genoma de este hongo revelaron la existencia de múltiples genes que codifican β-glucosidasas. Además, en los análisis de los secretomas del hongo, se identificaron tres β-glucosidasas, de las cuales dos fueron mayoritarias. La primera se produjo en cultivos inducidos en presencia de celulosa o lignocelulosa, mientras la otra se produjo en todos los medios estudiados, cuando el carbono se agotaba en el cultivo.

### Caracterización de BGL-2, la primera β-1,4-glucosidasa con un dominio de unión a celulosa

El descubrimiento de nuevas β-glucosidasas altamente eficientes sigue siendo uno de los principales cuellos de botella para la degradación de la celulosa. La BGL-2 es la principal β-glucosidasa secretada por este hongo *T. amestolkiae*, en presencia de inductores celulósicos. Es una glicosil hidrolasa de la familia GH3 y posee un CBD (dominio de unión a celulosa), una característica inusual en este tipo de proteínas.

Cuando se analizó la expresión de esta enzima, se vió que el hongo produjo dos formas de RNAm para el gen *bgl-2*. Ambas formas se retrotranscribieron a cDNA y se clonaron en *Pichia pastoris*, obteniéndose una proteína con el CBD (BGL-2*) y su forma truncada, que carece de CBD (BGL-2T*). Las dos versiones de la enzima BGL-2 fueron perfectamente funcionales, y altamente eficientes utilizando celobiosa y otros oligosacáridos. La sacarificación de bagazo de cerveza con un coctel

comercial basal y estas enzimas, frente a β-glucosidasas comerciales, puso de manifiesto que con las enzimas de *T. amestolkiae* se producen altos rendimientos en la sacarificación de este material lignocelulósico, comparable a los obtenidos al suplementar los cócteles con β-glucosidasas comerciales.

**Caracterización de BGL-3, una β-glucosidasa versátil e inducible cuando se agota la fuente de carbono**

La β-glucosidasa BGL-3, fue purificada, caracterizada y expresada heterólogamente. La síntesis de esta enzima en el hongo *T. amestolkiae* no requirió la presencia de una fuente de carbono específica debido a que se produce bajo estímulos de agotamiento de carbono. La enzima mostró una alta estabilidad térmica y una muy alta eficiencia sobre sustratos como *p*NPG, celobiosa y otros celooligosacáridos. Sorprendentemente, también mostró una capacidad notable para hidrolizar laminarina, un β-1,3-glucano presente en las algas. La eficacia de la enzima se examinó en procesos de sacarificación de paja de trigo, en los que BGL-3 funcionó mejor que un cóctel comercial de β-glucosidasas. Además, fue capaz de hidrolizar laminarina más eficientemente que una laminarinasa comercial.

**Caracterización de BGL-1, una β-glucosidasa glucotolerante**

Una de las características más deseadas para estas enzimas es la tolerancia a la glucosa, que les permite actuar de manera óptima en procesos que requieran concentraciones altas de este monosacárido. La β-glucosidasa BGL-1 del hongo *T. amestolkiae* fue una enzima muy minoritaria en todos los culivos estudiados. Esta enzima, que pertenece a la familia GH1, fue expresada heterólogamente en *P. pastoris*, purificada y caracterizada. La producción de BGL-1 en la levadura fue muy alta, alcanzando 75 U/mL, y se purificó en un solo paso cromatográfico, con un rendimiento del 80%. La característica más sobresaliente de BGL-1 fue su glucotolerancia, presentando un $K_i$ de 3.87 mM, uno de los más altos detectados hasta la fecha. Aunque esta enzima resultó no ser tan eficiente como las BGL-2 y BGL-3 hidrolizando celobiosa y celooligosacarido, presentó una alta eficacia catalítica sobre soforosa, el dímero de glucosa con enlaces β-1,2.

**Síntesis de glicósidos de interés usando las β-glucosidasa de *T. amestolkiae***

La transglicosilación enzimática es una de las técnicas más prometedoras para obtener nuevos glicoconjugados. Diferentes experimentos de transglicosilación se llevaron a cabo utilizando las β-glucosidasas BGL-2 y BGL-3 del hongo *T. amestolkiae*. Después de un primer análisis, en el que se puso de manifiesto que podían utilizar una

amplia variedad de aceptores potenciales para llevar a cabo su transglicosilación, se identificaron derivados mono-glucosilados de hidroxitirosol, alcohol vainillínico, alcohol 4-hidroxibencílico e hidroquinona. El hidroxitirosol y el alcohol vainillínico se seleccionaron para continuar estos estudios por sus posibles aplicaciones. Tras obtener sus respectivos glucósidos y purificarlos, se evaluó su efecto biológico usando modelos de células de cáncer de mama. Ambos compuestos mostraron un incremento de la capacidad antiproliferativa, respecto a los compuestos no glicosilados. Mediante ingeniería de proteínas (diseño racional), se convirtió la BGL-1 en una glicosintasa. BGL-1-E521G, que fue capaz de utilizar glucosa fluorada como donador en reacciones de transglicosilación, en las que se sintetizaron de manera regioselectiva derivados glicosilados de diferentes *p*-nitrofenoles, así como de algunos compuestos fenólicos de gran interés como el galato de epigalocatequina, un antioxidante presente en el té verde con interesantes propiedades para la salud.

**Conclusiones**

El trabajo realizado en la presente tesis doctoral demuestra que el genoma del hongo *T. amestolkiae* contiene varios genes que codifican β-glucosidasas y secreta altos niveles de esta actividad, independientemente de la disponibilidad de la fuente de carbono. Se han identificado dos principales β-glucosidasas diferentes a partir del análisis proteómico. Una de ellas se produjo en todas las condiciones estudiadas, mientras que la otra se indujo únicamente en sustratos celulósicos.

Se han purificado y caracterizado tres nuevas β-glucosidasas que pueden ser aplicadas en procesos de sacarificación de residuos lignocelulósicos, y en la obtención de otros productos de valor añadido. BGL-2 y BGL-3 se utilizaron exitosamente en la sacarificación de la celulosa de bagazo de cerveza y paja de trigo pretratada. BGL-1 mostró una enorme glucotolerancia, lo cual puede resultar ventajoso en su utilización en procesos donde se produzcan elevadas concentraciones de glucosa. Además, todas las enzimas fueron capaces de catalizar reacciones de transglicosilación. BGL-2 puede utilizar un amplio rango de posibles moléculas aceptoras de transglicosilación, mostrando altos rendimientos y regioselectividad. El aprovechamiento de estas características permitió la síntesis enzimática de dos glucósidos, de hidroxitirosol y vainillina, con propiedades bioactivas. Para profundizar en la capacidad de transglicosilación de estas enzimas, BGL-1 se transformó en su variante glicosintasa mediante diseño racional, siendo capaz de transferir moléculas de glucosa a diferentes compuestos de interés.

# SUMMARY


## Introduction

Plant biomass represents an important renewable source of raw materials since most of the carbon fixed by photosynthesis is contained in the cell walls. These cell walls are mainly composed of cellulose, hemicellulose and lignin. The degradation of the latter is the key to have access to polysaccharides. It is also essential to find specific enzymes for the complete hydrolysis of cellulose and hemicellulose, and then be able to use all their components for different biotechnological applications. Regarding cellulose, this polysaccharide is the major constituent of the plant cell wall (30-50%, depending on the plant type). By its conversion into glucose monomers it could be transformed into biofuels, or other high added value products. There are many organisms able to successfully degrade cellulose, secreting enzymes that can be used for different biotechnological applications. Currently, filamentous fungi are the main source of this type of enzymes for commercial use, which are characterized by their high productivity and high catalytic efficiency with respect to the ones secreted by other microorganisms.

There are three types of cellulases, which act in a coordinated way to degrade cellulose to glucose: endoglucanases, cellobiohydrolases and β-glucosidases. So far, the greatest efforts have been focused on discovering new β-glucosidases, which represent the key step in cellulose degradation.

## Aims

The aim of the work was to deepen in the knowledge of the β-glucosidases produced by the fungus *T. amestolkiae*, robust enzymes that have an essential role in the transformation of cellulose, and are capable of acting both in hydrolysis reactions, or in the synthesis of glycosides with potential interest in different technological sectors. To achieve these objectives, the work plan was:

**1-** Analysis of the genome and secretome of *T. amestolkiae*, showing special interest in the β-glucosidases produced by this fungus.

**2-** Heterologous expression of the main β-glucosidases of *T. amestolkiae* in *P. pastoris*, its purification, characterization and comparison with the native enzymes of the fungus.

**3-** Study of the saccharification of lignocellulosic residues using commercial enzymatic cocktails supplemented with *T. amestolkiae* β-glucosidases.

**4-** Synthesis of glycosides of interest with transglycosylation reactions catalyzed by *T. amestolkiae* β-glucosidases.

**5-** Obtaining variants of these enzymes to improve yields in the synthesis of glycosides of interest.

## Results

### Study of *T. amestolkiae* genome and proteome

Genomic and proteomic analysis are potent tools for metabolic characterization of microorganisms. *T. amestolkiae* was described as an excellent cellulase-producer, secreting high levels of β-1,4-endoglucanase, cellobiohydrolase and β-glucosidase activities. Its genome contains several genes encoding β-glucosidases and the fungus secretes high levels of this activity, regardless of the carbon source available. Two main different β-glucosidases have been identified from proteomic shotgun analysis. One of them is produced under different carbon sources, while the other is induced in cellulosic substrates and is a good supplement to Celluclast in saccharification of pretreated wheat straw.

### BGL-2 characterization: the first β-1,4-glucosidase described with a cellulose binding domain

The discovery of novel, highly-efficient β-glucosidases remains as one of the major bottlenecks for cellulose degradation. BGL-2 is the major β-glucosidase secreted by this fungus in the presence of cellulosic inductors. This enzyme possesses a Cellulose Binding Domain (CBD), an unusual feature among this type of proteins. Besides, when growing on cellulose, the fungus produced two different *bgl-2* mRNAs that were retrotranscribed into cDNA, cloned and expressed in *Pichia pastoris*. A complete recombinant protein (BGL-2*) and its truncated form, lacking CBD (BGL-2T*) were produced, being both completely functional proteins. All tested BGL-2 forms were highly efficient using cellobiose and other short oligosaccharides as substrates. As one of its potential biotechnological applications, the recombinant *T. amestolkiae* enzymes were studied in saccharification of brewers' spent grain, reaching a yield comparable to that of commercial β-glucosidase cocktails.

### Characterization of BGL-3, a versatile β-glucosidase

BGL-3 has been purified, characterized, and heterologously produced. The synthesis of this β-glucosidase (BGL-3) was not induced by cellulose, and the presence of a specific carbon source is not required for its secretion, since its synthesis is triggered by carbon starvation stimuli. The enzyme showed high thermal stability, and very high efficiency on *p*NPG,

cellobiose and other cellooligosaccharides. Surprisingly, it also showed remarkable ability to hydrolyze laminarin, a β-1,3-glucan with β-1,6 branches present in algae. The enzyme's efficiency was examined in wheat straw saccharification, in which BGL-3 worked better supplementing Celluclast 1.5L than the commercial cellulase cocktail N-50010. Besides, BGL-3 hydrolyzed laminarin more efficiently than a commercial laminarinase.

**Characterization of BGL-1, a glucotolerant β-glucosidase**

One of the most desired features for these enzymes is glucose tolerance, which allows them to act optimally under elevated glucose concentrations. A glucotolerant β-glucosidase, named BGL-1, has been heterologously expressed in *P. pastoris*, purified, and characterized. The production of the enzyme in the yeast was very high, reaching 75 U/mL, and it was purified in just one step with a yield of 80%. The most outstanding feature of BGL-1 was its glucotolerance, with a $K_i$ of 3.87 mM, one of the highest reported up to date. Interestingly, when examining the substrate specificity of BGL-1, it was more active over sophorose, the β-1,2 disaccharide of glucose, than over cellobiose.

**Synthesis of glycosides using the β-glucosidases of *T. amestolkiae***

Transglycosylation represents one of the most promising approaches for obtaining novel glycosides. Transglycosylation experiments were carried out using the recombinant BGL-2 and BGL-3 β-glucosidases from the fungus *T. amestolkiae*. After a first screening with a wide variety of potential transglycosylation acceptors, mono-glucosylated derivatives of hydroxytyrosol, vanillin alcohol, 4-hydroxybenzyl alcohol, and hydroquinone were detected. Hydroxytyrosol and vanillyl alcohol were selected as the best options for transglycosylation. The evaluation of the biological effect of these glucosides using models of breast cancer cells, showed an enhancement in the anti-proliferative capacity of the vanillin derivative, and an improved safety profile of both glucosides.

Besides, the transglycosylation profile of BGL-1 was also examined, and, for expanding its synthesis capacities, it was converted into a glycosynthase. The mutant enzyme, named BGL-1-E521G, was able to use α-D-glucosyl-fluoride as donor in glycosylation reactions, and synthesized in a regioselective manner glycosylated derivatives of different *p*NP-sugars, but also of some phenolic compounds of industrial interest as epigallocatechin gallate.

**Conclusions**

The work carried out in this Doctoral Thesis has disclosed that the genome of the fungus *T. amestolkiae* contains several genes that encode β-

glucosidases and secretes high levels of this activity, regardless of the availability of the carbon source. Two main different β-glucosidases have been identified from proteomic analysis. One of them was produced under several carbon sources, while the other was induced only in cellulosic substrates.

In addition, three new β-glucosidases, enzymes that can be applied in saccharification processes of lignocellulosic residues, and in obtaining other value-added products, have been isolated and characterized. BGL-2 and BGL-3 were successfully used in the saccharification of pre-treated wheat straw and brewers spent grain. BGL-1 showed an enormous glucotolerance, which can position it to be used in industrial processes that require high glucose concentrations. In addition, all enzymes showed transglycosylation capacity. BGL-2 had a wide range of potential acceptors, altogether with high yields in the synthesis, and is regioselective. The use of these characteristics allowed the enzymatic production of two glycosides, from hydroxytyrosol and vanillin, with improved bioactive properties. To raise the transglycosylation capacity of these enzymes, BGL-1 was converted into a glycosynthase, which was able to transfer glucose molecules to different compounds of interest.

# ESTRUCTURA DE LA TESIS

La presente Tesis Doctoral está estructurada de la siguiente forma: Los principales resultados están recogidos en cinco capítulos, que corresponden a cuatro artículos científicos ya publicados en revistas indexadas y a uno en preparación. Los cinco capítulos están relacionados a través de la introducción y la discusión general. Los capítulos han sido maquetados para mantener una misma estructura a lo largo de toda la tesis, pero el contenido de los trabajos publicados se ha mantenido íntegro. El material suplementario ha sido incluido al final de cada capítulo. El listado de capítulos contenidos en la tesis se describe a continuación:

**Capítulo 1:** de Eugenio, L.I., Méndez-Líter, J.A., Nieto-Domínguez, M., Alonso, L., Gil-Muñoz, J., Barriuso, J. *et al*., 2017. Differential β-glucosidase expression as a function of carbon source availability in *Talaromyces amestolkiae*: a genomic and proteomic approach. Biotechnol. Biofuels. 10, 161-161. doi: 10.1186/s13068-017-0844-7

**Capítulo 2:** Méndez-Líter, J.A., Gil-Muñoz, J., Nieto-Domínguez, M., Barriuso, J., de Eugenio, L.I., Martínez, M.J., 2017. A novel, highly efficient β-glucosidase with a cellulose-binding domain: characterization and properties of native and recombinant proteins. Biotechnol. Biofuels. 10, 256. doi: 10.1186/s13068-017-0946-2

**Capítulo 3:** Méndez-Líter, J.A., de Eugenio, L.I., Prieto, A., Martínez, M.J., 2018. The β-glucosidase secreted by *Talaromyces amestolkiae* under carbon starvation: a versatile catalyst for biofuel production from plant and algal biomass. Biotechnol. Biofuels. 11, 123-123. doi: 10.1186/s13068-018-1125-9

**Capítulo 4:** Méndez-Líter, J.A., Tundidor, I., Nieto-Domínguez, M., de Toro, B.F., González Santana, A., de Eugenio, L.I. *et al*., 2019. Transglycosylation products generated by *Talaromyces amestolkiae* GH3 β-glucosidases: effect of hydroxytyrosol, vanillin and its glucosides on breast cancer cells. Microb. Cell Fact. 18, 97. doi: 10.1186/s12934-019-1147-4

**Capítulo 5:** Méndez-Líter, J.A., Nieto-Domínguez, M., de Toro, B.F., González Santana, A., Prieto, A., *et al*., 2019. A glucotolerant β-glucosidase from the fungus *Talaromyces amestolkiae* and its conversion into a glucosynthase capable to glycosylate epigallocatechin gallate. (En preparación).

# INTRODUCCIÓN GENERAL

# INTRODUCCIÓN

## 1. Carbohidratos y enzimas activas sobre carbohidratos

Las cuatro clases principales de moléculas orgánicas en los sistemas vivos son las proteínas, los lípidos, los ácidos nucleicos y los carbohidratos. Éstos últimos representan el grupo de moléculas orgánicas más abundante en la naturaleza, y casi todos los organismos son capaces de sintetizarlos y metabolizarlos (Reeves, 1996). Un carbohidrato es una biomolécula compuesta por átomos de carbono, hidrógeno, y oxígeno. La composición de la mayoría de estas moléculas responde a la fórmula empírica $C_nH_{2n}O_n$, dónde "n" representa un número mayor o igual a 3 (Butterworth, 2005). Estas proporciones pueden ser modificadas por la adición de algunos grupos funcionales, como metilos o acetilos, u otras especies químicas como átomos de nitrógeno, azufre o fósforo.

Los carbohidratos desempeñan numerosos papeles en los organismos vivos. Por ejemplo, pueden servir como sustancias de reserva energética (almidón en plantas o glucógeno en animales), actuar como componentes estructurales (celulosa en plantas y quitina en artrópodos), o formar parte de los ácidos nucleicos (la ribosa y la desoxirribosa son componentes esenciales del ARN y el ADN, respectivamente) (Bailey y Ollis, 1986; Kamerling, 2007). En general, los carbohidratos se pueden definir como polihidroxi-compuestos que poseen un grupo carbonilo potencialmente activo, que se denomina carbono anomérico, el cual puede ser un aldehído o una cetona. Se pueden clasificar, en base al número de repeticiones de monómeros presentes en su estructura, en monosacáridos, oligosacáridos, y polisacáridos.

Los monosacáridos son los carbohidratos más pequeños, ya que no pueden descomponerse en azúcares más simples, y contienen de tres a nueve átomos de carbono. Estos bloques básicos pueden formar moléculas más complejas cuando se combinan entre sí a través de enlaces. Cuando la unión es de varios residuos de monosacáridos se denominan oligosacáridos y, si poseen más de 20 unidades, se consideran polisacáridos. Estos a su vez se dividen en homopolisacáridos, cuando están formado por un único tipo de monómero, y heteropolisacáridos, formados por dos o más monosacáridos distintos (Butterworth, 2005). Oligo y polisacáridos se forman cuando dos o más monosacáridos se unen covalentemente mediante un enlace *O*-glicosídico, que implica la reacción de condensación entre cualquiera de los grupos hidroxilo de un monosacárido y el hidroxilo del carbono anomérico de otro.

Es importante destacar que todos los monosacáridos y, en general, la mayoría de los carbohidratos, pueden actuar como agentes reductores, ya que siempre existe un carbono anomérico libre susceptible de ser oxidado.

Sólo hay algunas excepciones, como es el caso de la sacarosa (α-D-glucopiranosil-(1→2)-β-D-fructofuranósido), que no tiene poder reductor debido a la implicación de los carbonos anoméricos de los dos monómeros en el establecimiento del enlace glicosídico. Además de ser esencial para comprender la reactividad de los carbohidratos, el poder reductor de los carbohidratos es clave en multitud de métodos utilizados para evaluar la hidrólisis de los polisacáridos (Nelson, 1944; Miller, 1959). Debido a que al menos un carbono anomérico está involucrado en la creación de un enlace glicosídico, la hidrólisis de un polisacárido siempre conducirá a un aumento del poder reductor de los productos.

Además de los enlaces *O*-glicosídicos entre monómeros, los hidratos de carbono pueden formar glicoconjugados al unirse a otras moléculas como lípidos o proteínas. En el caso de las uniones a proteínas o a los ácidos nucleicos, la unión suele formarse mediante el establecimiento de un enlace *N*-glicosídico, aunque con las proteínas también pueden existir enlaces *O*-glicosídicos (Roth, 2002). Además, aunque mucho menos comunes en la naturaleza, se han identificado carbohidratos unidos a moléculas por enlaces *S*-glicosídicos (Jahn y Withers, 2003) y *C*-glicosídicos (Nicotra, 2008). En todas estas uniones también es necesaria la participación del carbono anomérico del carbohidrato, que en este caso interactúa con átomos de nitrógeno, azufre o carbono (figura 1).

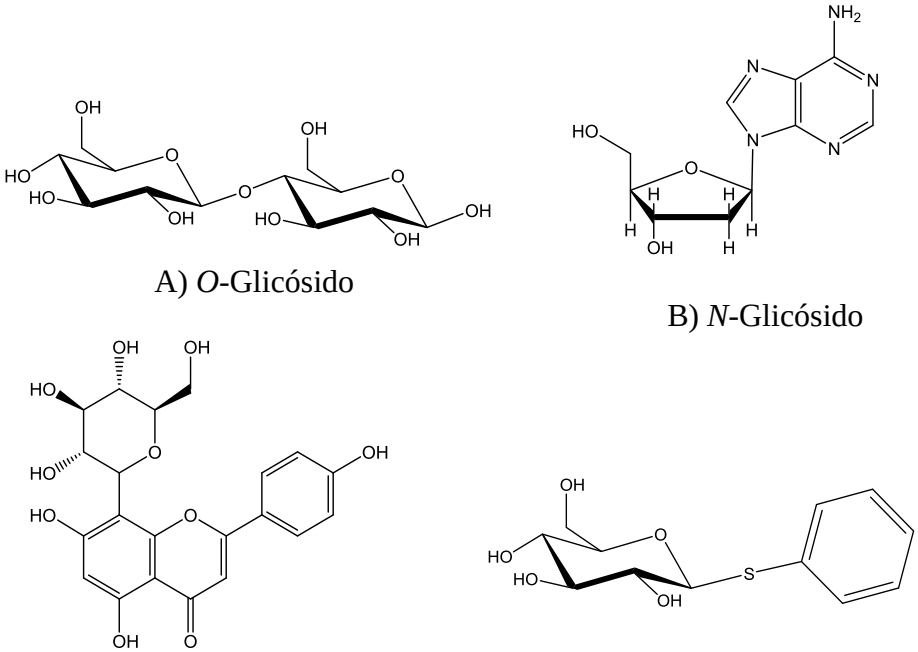



**Figura 1.** Principales tipos de glicósidos según el enlace glicosídico. A) Celobiosa, un ejemplo de monosacáridos unidos por enlaces *O*-glicosídicos. B) Adenosina, un ejemplo de enlace *N*-glicosídico, en el cual una molécula de ribosa se une a la base nitrogenada adenina. C) Vitexina, un ejemplo de monosacárido unido a otra molécula, en este caso un flavonoide, por enlace *C*-glicosídico. D) Tiofenol glucósido, un ejemplo de monosacárido unido a un benceno, mediante un enlace *S*-glicosídico.

## 1.1. Glicobiología

La glicobiología es la ciencia encargada del estudio de los polisacáridos libres o presentes en glicoconjugados a nivel de su estructura y función biológica. Aunque históricamente su repercusión es menor que la de otras disciplinas como la genómica, la proteómica o la biología de sistemas, cada vez se considera más importante poseer un mayor conocimiento de los azúcares y su composición para comprender la biología a una escala superior, debido a todos los procesos en los que intervienen (Merry y Merry, 2005).

La extrema variedad de estructuras de monosacáridos, la diversidad de enlaces existentes, y el hecho de que, como se ha mencionado anteriormente, casi todos los tipos de biomoléculas pueden estar glicosiladas, hace que el conjunto de enzimas que actúan sobre glicoconjugados, oligosacáridos y polisacáridos sea uno de los más variados del planeta (Davies y Henrissat, 1995). Este extenso grupo de enzimas se denomina "enzimas activas sobre carbohidratos", o CAZymes (Carbohydrate Active Enzymes) y se caracteriza por catalizar la síntesis y degradación de carbohidratos y/o glicoconjugados.

En la actualidad, la base de datos CAZY (www.cazy.org) incluye todas las enzimas activas sobre carbohidratos identificadas, y las clasifica de acuerdo a su homología de secuencia (Lombard *et al*., 2014). Así, actualmente se distinguen cinco grandes grupos:

1. **Glicosil hidrolasas (GHs):** Compuesta por 156 familias de enzimas diferentes, que se encargan de llevar a cabo reacciones de hidrólisis o transglicosilación (apartado 1.1.3) de los enlaces glicosídicos (Henrissat y Davies, 1997). Los genes que codifican GHs son muy abundantes en los genomas de los seres vivos. Actualmente cerca de la mitad de las enzimas existentes en CAZY pertenecen a este grupo.

2. **Glicosil transferasas:** Estas enzimas catalizan la síntesis de un enlace glicosídico por transferencia de una molécula glucídica desde un carbohidrato activado con fósforo a distintos aceptores (Campbell *et al*., 1997). Se han identificado hasta el momento 106 familias diferentes de glicosil transferasas.

3. **Polisacárido liasas:** Son un grupo de enzimas que escinden las cadenas de polisacáridos que contienen ácido urónico, a través de un mecanismo de β-eliminación para generar un residuo de ácido hexenurónico y un nuevo extremo reductor. Se han descrito 29 familias diferentes hasta la fecha (Lombard *et al*., 2010).

4. **Carbohidrato esterasas:** Son las encargadas de eliminar modificaciones de tipo éster en mono-, oligo- y polisacáridos, con el objetivo final de mejorar la accesibilidad de las GHs para degradar estos sustratos. En la actualidad, tan sólo se han identificado 16 familias de estas enzimas (Lombard *et al*., 2010).

5. **Enzimas con actividades auxiliares (AAs):** Este grupo incluye enzimas, habitualmente oxidorreductasas, que actúan facilitando la acción de otras CAZymes. Este grupo incluye 16 familias distintas (Levasseur *et al*., 2013).

Además, varias de las CAZymes pueden tener **módulos de unión a carbohidratos (CBM)** que aumentan su eficiencia. Los CBM están formados por una secuencia de aminoácidos con capacidad para unirse a carbohidratos (Boraston *et al*., 2004). Actualmente, se ha demostrado que estas regiones pueden desempeñar las siguientes funciones: 1) mantener la enzima cerca del sustrato, aumentando la concentración de glicosidasas en el entorno de los carbohidratos y, de esta manera favorecer su degradación (Teeri *et al*., 1998); 2) dirigir la enzima a una región específica del sustrato en la que va a ejercer su acción catalítica, como por ejemplo los extremos reductores o no reductores (Teeri *et al*., 1998); y 3) disgregar el sustrato, como ocurre en algunas celulasas cuyos CBM disminuyen la cristalinidad de la celulosa en las regiones de la interacción enzima-sustrato (Arantes y Saddler, 2010). Estos dominios pueden adoptar una amplia variedad de estructuras (Gilbert *et al*., 2013), y tienen su propia clasificación en la base de datos CAZy, en la que en la actualidad hay descritas 84 familias de CBM.

De todos los grupos enzimáticos explicados, las glicosil hidrolasas son probablemente el más importante, no solo porque representan el mayor número de secuencias depositadas, sino también por su amplia variedad de sustratos (Yamamoto *et al*., 2000), sus versatilidad catalítica (Mao *et al*., 2010), y por las numerosas aplicaciones en las que se utilizan (Liu *et al*., 2007; Rempel y Withers, 2008).

## 1.2. Glicosil hidrolasas: hidrólisis

Las glicosil hidrolasas, como se ha mencionado anteriormente, son enzimas que catalizan la hidrólisis del enlace glicosídico entre dos carbohidratos, o un carbohidrato y un aglicón (parte no glucídica) (Henrissat, 1991).

El enlace glicosídico es una de las uniones más estables de la naturaleza. Los experimentos realizados por Wolfenden y colaboradores (1998) demostraron que estos enlaces poseen una vida media de cerca de 5

millones de años, y que la capacidad de aceleración de la reacción de las glicosil hidrolasas es del orden de $10^{17}$ órdenes de magnitud, con lo que representan uno de los biocatalizadores más eficientes de la naturaleza (Rye y Withers, 2000).

Los números EC (*Enzyme Comission number*) son códigos que ayudan a organizar los diferentes tipos de enzimas existentes, basándose en las reacciones químicas que catalizan. Cada número de EC está asociado con un nombre recomendado para la enzima respectiva, pero hay que destacar que estos números no indican enzimas específicas, sino las reacciones catalizadas por éstas. Si diferentes enzimas catalizan la misma reacción, entonces reciben el mismo número de EC. De acuerdo con esta clasificación, las GHs están incluidas en el grupo EC 3.2.1.-, comprendiendo desde el 3.2.1.1 hasta el 3.2.1.208, es decir, catalizan 208 tipos de reacciones distintas.

La existencia de diferentes criterios de clasificación, como el que organiza las GHs por secuencias utilizado en la base de datos CAZY o el basado en las reacciones que catalizan (EC), provoca que enzimas de una misma familia puedan encontrarse en varios números EC, y al mismo tiempo, que enzimas de diferentes familias tengan mismo número EC por desempeñar la misma función.

La hidrólisis enzimática del enlace glicosídico llevado a cabo por las GHs se realiza mediante un proceso de catálisis ácida general, que requiere la participación de dos residuos catalíticos: un donante de protones, denominado aminoácido de la catálisis ácido-base, y un aminoácido nucleófilo. Éstos suelen estar representados en casi todas las familias por ácidos aspárticos y ácidos glutámicos (Davies *et al*., 1997). Independientemente del tipo de catálisis y su mecanismo, las glicosidasas comparten una característica básica, que es la presencia de subsitios (Davies *et al*., 1997). Éstos se definen como regiones contiguas a los aminoácidos catalíticos de la enzima que interactúan de manera no covalente con las unidades de sustrato que se va a degradar para facilitar su reconocimiento. En el sitio activo de las glicosil hidrolasas, los subsitios se indican numéricamente a partir del enlace hidrolizado, teniendo en cuenta la posición de los extremos no reductores y reductores del carbohidrato. Por lo tanto, los subsitios se definen en función de la unidad de sustrato que alojan, y se denominan (+1), (+2), (+3) ... hacia el extremo con poder reductor, y (-1), (-2) (-3) ... en la dirección opuesta. Por lo tanto, la hidrólisis siempre ocurrirá entre los subsitios (+1) y (-1).

De forma general, se puede decir que las GHs se pueden clasificar también de acuerdo a otros dos criterios:

1. En función de la región del sustrato en la que actúen, distinguimos entre GHs con actividad exo-, si hidrolizan el sustrato desde los extremos (tanto en el reductor, como en el no reductor) o con actividad endo-, si atacan puntos internos de la cadena glucídica.

2. Según los mecanismos de hidrólisis del enlace glicosídico: Basándose en la configuración del carbono anomérico del producto de la reacción de hidrólisis, se han descrito dos mecanismos distintos, GHs de retención de la configuración, y GHs de inversión de la configuración, que explicamos a continuación.

Estos dos mecanismos fueron postulados por Koshland (1953), y se puede observar su representación esquemática en la figura 2. En las GHs de **retención de la configuración**, la hidrólisis se logra a través de un mecanismo de doble desplazamiento en dos etapas que involucra la formación de un intermediario enzima-sustrato. La reacción también la realizan los dos aminoácidos catalíticos, que en las enzimas que usan este mecanismo se encuentran separados por aproximadamente 5.5 Å (McCarter y Withers, 1994). En el primer paso, el aminoácido nucleófilo ataca el carbono anomérico del sustrato, liberándose una parte del carbohidrato con la ayuda del aminoácido que realiza la catálisis ácido-base, que cede un protón, y termina con la formación de un intermediario enzima-sustrato a través del aminoácido nucleófilo. En la segunda fase, una molécula nucleófila externa, como el agua, realiza un segundo ataque en el carbono anomérico del sustrato, separando el intermediario anterior. Este nuevo ataque también es asistido por el segundo residuo de la catálisis ácido-base, que recibe un protón del agua para activarla y permitir este ataque. De acuerdo con este mecanismo, el carbohidrato liberado mantiene la configuración anomérica del inicial (Rempel y Withers, 2008).

En las GHs de **inversión de la configuración**, la hidrólisis se realiza en un solo paso. La reacción se lleva a cabo por la acción de los dos aminoácidos catalíticos, que en las enzimas que usan este mecanismo se encuentran separados por aproximadamente 10 Å (McCarter y Withers, 1994). Cuando una molécula nucleófila, generalmente agua, entra en el centro activo de la enzima, uno de los dos aminoácidos actúa como base y la desprotona, lo que le permite atacar el carbono anomérico del carbohidrato. Este ataque nucleofílico es asistido por el otro aminoácido catalítico, que actuando como ácido transfiere un protón al azúcar/aglicón liberado. Finalmente, se libera un nuevo carbohidrato, cuya configuración anomérica es opuesta a la del sustrato inicial.

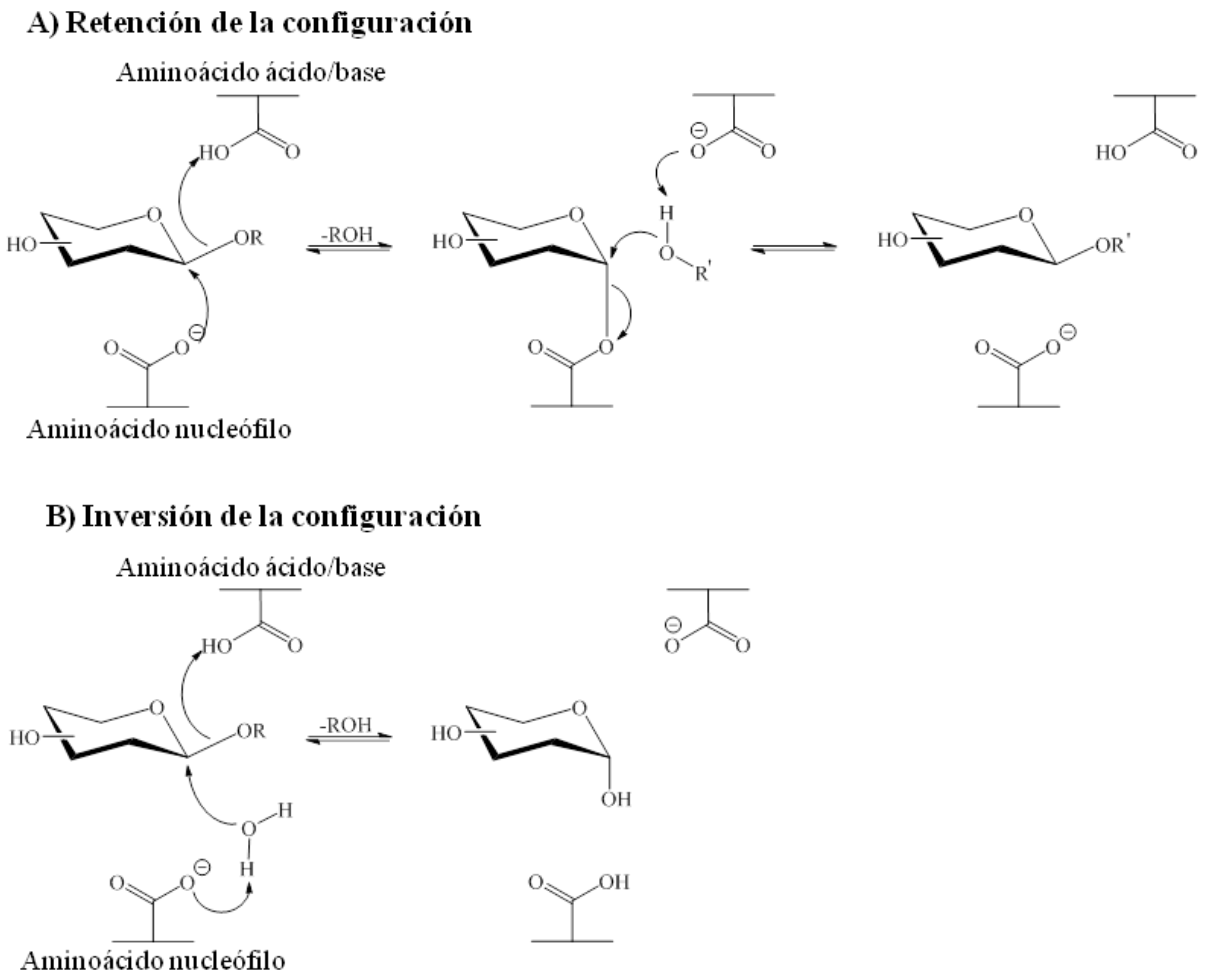



**Figura 2.** Mecanismos de acción de las glicosil hidrolasas de retención de la configuración (A), y de inversión de la configuración (B) (figura adaptada de Nieto-Domínguez, 2017).

Aunque la gran mayoría de las GHs conocidas siguen uno de estos dos mecanismos, en los últimos años se han encontrado otras alternativas catalíticas:

- Las GHs de las familias 18, 20, 25, 56, 84 y 85 no tienen un aminoácido catalítico nucleófilo, y necesitan la participación de otros aminoácidos del centro activo, denominados "*neighboring groups*" (grupos vecinos) para estabilizar la reacción y poder llevar a cabo la hidrólisis (Terwisscha van Scheltinga *et al*., 1995; Mark *et al*., 2001).

- Las mirosinasas, un tipo de GHs pertenecientes a la familia GH1, catalizan la hidrólisis de tioglicósidos presentes en plantas. En estas enzimas, el ácido glutámico responsable de la catálisis ácido/base es reemplazado por una glutamina, lo cual reduce la repulsión de cargas

presentes en el centro activo de la enzima, permitiendo así la entrada de los tioglicósidos (Burmeister *et al*., 2000). Las sialidasas y trans-sialidasas de las familias GH33 y GH34, tienen la particularidad de utilizar una tirosina como aminoácido nucleófilo. Una posible explicación para esta diferencia puede ser la reducción de la repulsión de cargas negativas presentes en el centro activo de las GH tradicionales (Amaya *et al*., 2004; Ndeh *et al*., 2017).

- Las GHs de las familias 4 y 109 usan un particular mecanismo que requiere la participación de nicotinamida adenina dinucleótido (NAD) como cofactor. El mecanismo de reacción avanza a través de estados de transición que implican procesos de oxidorreducción, alejándose bastante del de las GHs tradicionales (Yip *et al*., 2004).

Con respecto a la estructura y forma de plegamiento de las GHs, no se pueden establecer motivos comunes debido a la enorme variedad de enzimas de diferentes familias que componen el grupo. En cambio, la topología general de los sitios activos sí se puede dividir en tres clases principales (Lynd *et al*., 2002; Juturu y Wu, 2014), incluso independientemente de si la enzima es de inversión o de retención de la configuración:

**1.** Sitio activo con forma de "bolsillo". Esta topología es óptima para el reconocimiento de los extremos no reductores de carbohidratos, y se ha identificado en varias enzimas, tales como β-galactosidasas, β-glucosidasas, glucoamilasas y β-amilasas.

**2.** Sitio activo con forma de "surco". Es una estructura abierta que permite la unión de varias unidades de azúcares en sustratos poliméricos, y se ha descrito en enzimas que degradan polisacáridos, como endocelulasas, quitinasas, α-amilasas, o xilanasas.

**3.** Sitio activo con forma de "túnel". Esta topología es una variación de la anterior, desarrollándose cubiertas a partir del surco. Es mucho menos común en GHs, habiendo sido identificada solo en celobiohidrolasas. Esta topología permite que estas enzimas liberen el producto mientras permanecen firmemente unidas a la cadena de polisacáridos, la cual atraviesa el túnel, manteniéndose así constantes las condiciones para la degradación del sustrato (Payne *et al*., 2015).

## 1.3. Glicosil hidrolasas: transglicosilación

Tradicionalmente, el estudio de las GHs se ha centrado en su capacidad para hidrolizar enlaces glicosídicos. Pero, además de hidrolizarlos, en determinadas condiciones pueden sintetizarlos mediante transglicosilación (Edelman, 1956). Hasta la fecha, este tipo de reacciones se ha descrito solo en las GHs de retención de la configuración. Como hemos comentado, cuando se genera el intermediario enzima–sustrato, el complejo es separado por el ataque de una molécula nucleófilica. Si se trata de una molécula de agua, tiene lugar un proceso de hidrólisis normal. En cambio, si es realizado por otro compuesto suficientemente nucleófilo para separar el intermediario enzima–sustrato, se produce la transglicosilación (Bissaro *et al*., 2015). La molécula que cede el carbohidrato se denomina donador, y la molécula a la cual éste va a quedar unido, aceptor.

Dentro de las diferentes familias de GHs, hay algunas que, en vez de realizar el proceso de hidrólisis de forma normal, catalizan de forma preferente la transferencia de azúcares a ciertas moléculas aceptoras, diferentes del agua, y reciben el nombre de transglicosilasas. El mecanismo de éstas enzimas suele basarse en primero cortar algunos tipos de carbohidratos para, a continuación, transglicosilarlos con una regioselectividad distinta al compuesto del que proceden. Las familias de GHs en las que se han descrito enzimas que pueden ser consideradas transglicosilasas incluyen: la GH2, la GH13, GH16, GH31, GH70, GH77. Estas enzimas catalizan reacciones tan diversas como la conversión de lactosa en alolactosa, la de amilosa en ciclodextrinas, la reorganización de las estructuras de los xiloglucanos de la pared celular vegetal de algunas plantas, o la síntesis de polisacáridos de alto peso molecular a partir de sacarosa (Uitdehaag *et al*., 1999; van Hijum *et al*., 2006; Eklöf y Brumer, 2010; Juers *et al*., 2012; Larsbrink *et al*., 2012).

La síntesis enzimática, tanto de oligosacáridos como de glicósidos de otras moléculas, a partir de distintos carbohidratos, es interesante industrialmente como alternativa a la síntesis química para producir compuestos de interés biotecnológico. La síntesis química es compleja y suele conllevar varios pasos, empleando catalizadores que pueden dañar el medio ambiente, y la variada estereoquímica de los oligosacáridos y sus derivados dificulta la obtención de moléculas enantioméricamente puras. El uso de enzimas para la obtención de estos compuestos presenta grandes ventajas debido a que suelen ser regio- y estereoespecíficas. Además, hay que considerar que la biocatálisis se considera un proceso medioambientalmente limpio, que no genera tantos contaminantes como los procesos de síntesis química tradicionales (Danby y Withers, 2016).

A pesar de las ventajas mencionadas, los procesos de

transglicosilación catalizados por GHs también presentan algunos inconvenientes, principalmente relacionados con el bajo rendimiento del producto que se quiere obtener. Esto puede deberse a la autohidrólisis del nuevo glicósido provocada por la desaparición de aceptores nucleófilos en el medio de reacción. En la actualidad, se están desarrollando algunas aproximaciones enfocadas a superar este problema. Entre ellas, las más relevantes consisten bien en disminuir el contenido de agua en la reacción, para desplazar hacia la síntesis el equilibrio hidrólisis-transglicosilación, o bien en obtener, mediante ingeniería de proteínas, versiones de las GHs con los aminoácidos implicados en la catálisis mutados, lo que aumenta su capacidad de síntesis con respecto a las enzimas silvestres (Danby y Withers, 2016).

- La disminución de la concentración de agua en las reacciones minimiza la competencia entre hidrólisis y transglicosilación. Para ello, pueden realizarse ensayos en entornos orgánicos, aunque en estas condiciones la estabilidad de las GHs suele ser baja (Mori *et al*., 1997). Para resolver este problema podrían utilizarse líquidos iónicos, que proporcionarían un entorno polar en el que la enzima pueda ser estable, evitando la presencia de agua (Gorke *et al*., 2010). Esta alternativa, aunque prometedora, no es rentable en la actualidad debido al alto precio de estos solventes.

- Las aproximaciones más exitosas para mejorar los rendimientos de transglicosilación han sido las obtenidas con versiones mutadas de las enzimas originales. Este enfoque ha conducido a la obtención de dos clases principales de mutantes: 1) las glicosintasas (Perugino *et al*., 2004) y 2) las tioglicoligasas (Jahn *et al*., 2003). Ambas variantes se basan en la sustitución de uno de los aminoácidos catalíticos del centro activo por uno neutro, habitualmente alanina, serina, glicina o glutamina. En el caso de las glicosintasas, el residuo reemplazado es el responsable del ataque nucleofílico, por lo tanto, el mutante no es capaz de catalizar por sí mismo el primer paso del mecanismo de doble desplazamiento. Por esta razón, las glicosintasas requieren el uso de un donador activado con flúor, que sí consigue generar el intermediario enzima-sustrato. Una vez realizado el primer paso, se produce un ataque nucleófilo convencional por parte de una molécula aceptora, teniendo lugar una reacción de transglicosilación. En estas variantes, la capacidad hidrólítica se elimina por completo, con lo que las glicosintasas no son capaces de degradar los productos de reacción que se generan, obteniéndose así rendimientos de transglicosilación que pueden llegar al 100% de conversión. Con respecto a las tioglicoligasas, se obtienen reemplazando el

aminoácido de la catálisis ácido–base por uno neutro. Sin la presencia de ese residuo, los mutantes necesitan un donador con un buen grupo saliente, como un azúcar unido a una molécula de nitrofenol, para tener éxito en la formación del intermedio enzima-sustrato, ya que deja de ser posible formar este intermediario con carbohidratos naturales. Luego, el segundo ataque nucleófilo debe ser llevado a cabo por un aceptor mucho más nucleófilo que los que se necesitan en los procesos normales de hidrólisis o transglicosilación. Estos mutantes han sido utilizados con éxito en la síntesis de enlaces *S*-glicosídicos, transfiriéndose un carbohidrato a moléculas con un grupo tiol (Jahn et al., 2003).

Los mecanismos con los que funcionan ambos tipos de mutantes, se pueden observar en la figura 3.

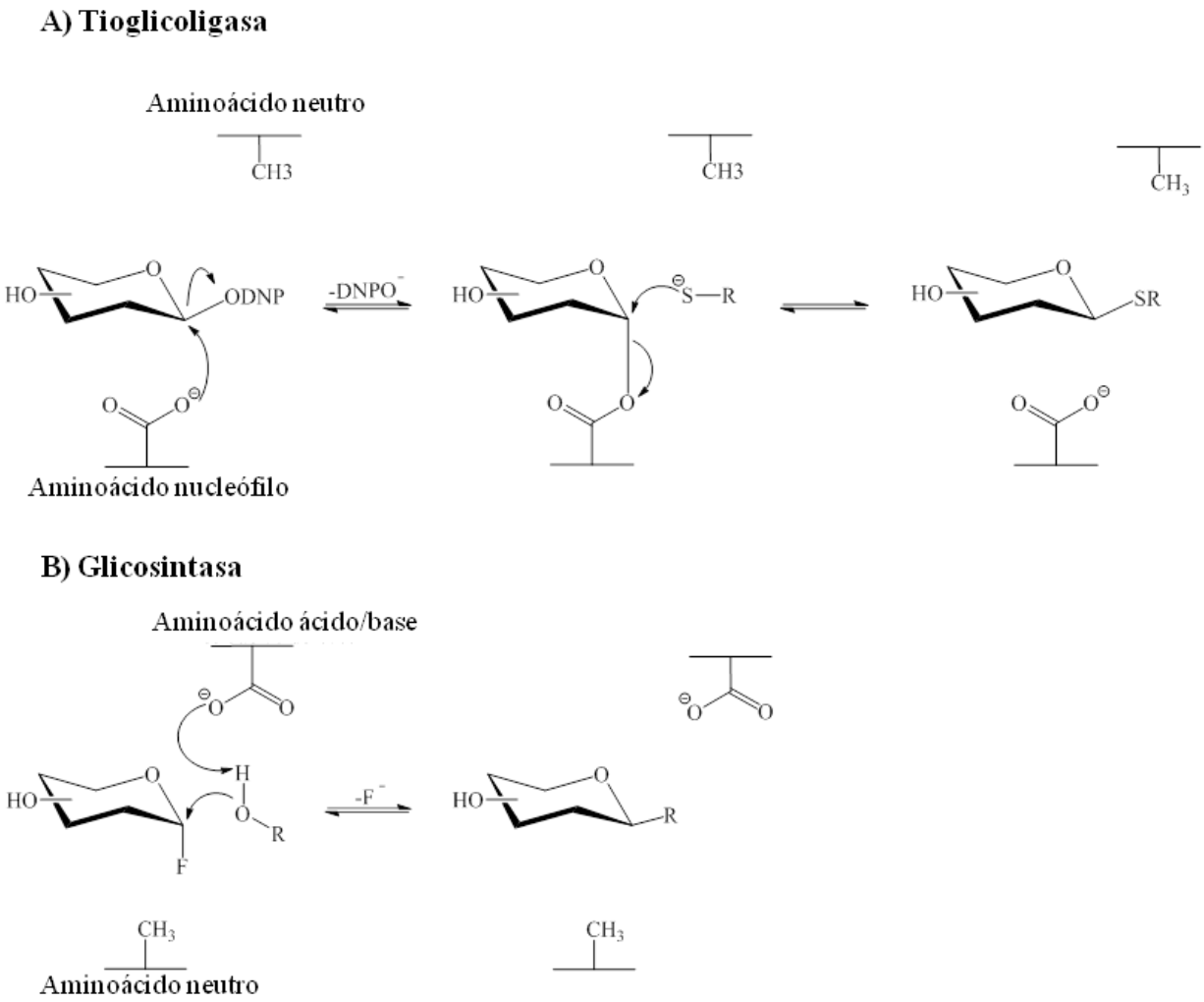



**Figura 3.** Mecanismos de acción de las principales variantes mutantes de glicosil hidrolasas: A) Tioglicoligasas, y B) Glicosintasas. (DNPO: Di-nitrofenil. Figura adaptada de Nieto-Domínguez, 2017).

## 1.4. Aplicaciones de las glicosil hidrolasas

Teniendo en cuenta la enorme variedad de sustratos diferentes sobre los que son activas las GHs, la diversidad de mecanismos catalíticos descritos, y el gran número de enzimas descubiertas hasta la fecha, no es sorprendente que sean uno de los catalizadores enzimáticos más producidos a nivel industrial, solo precedidos por proteasas y lipasas. A continuación, se resumen algunas de las aplicaciones potenciales más notables descubiertas hasta la fecha:

1. Industria textil. Esta industria es una de las principales consumidoras de celulasas, con el fin de conseguir superficies modificadas de tejidos, y mejorar el tacto y la suavidad de los mismos. En esta aplicación, cabe destacar el uso de las β-1,4-endoglucanasas para obtener el efecto de "lavado a la piedra" en telas. El papel de estas enzimas en el proceso es degradar fibras de celulosa, proporcionando regiones donde el tinte se puede quitar fácilmente, generando el aspecto envejecido deseado (Pazarlioglu *et al*., 2005; Kuhad *et al*., 2011).

2. Industria alimentaria. Se han desarrollado principalmente tres aplicaciones en las que se usan GHs (Kuhad *et al*., 2011):
    - Clarificación de zumos. El uso de pectinasas y otras GHs consiguen reducir la turbidez de los zumos industriales, consiguiendo un aspecto más atractivo para los consumidores.
    - Fermentación de masas en fabricación de panes y bollos. Las α-amilasas pueden añadirse para incrementar la hidrólisis del almidón de las harinas, mejorando así los procesos de fermentación y la calidad final de las masas.
    - Producción de prebióticos. Éstos se definen como ingredientes alimentarios, no digeribles por los humanos, que tienen un efecto beneficioso a través de su metabolismo en la microbiota del tracto intestinal (Gibson *et al*., 2004). La mayoría de los prebióticos conocidos se obtienen a partir de la hidrólisis controlada de los polisacáridos, como es el caso de los fructooligosacáridos derivados de la inulina, o los xilooligosacáridos derivados del xilano, o por glicosilación, como la producción de galactooligosacáridos usando lactosa y galactosa para el proceso de síntesis (Pandey *et al*., 2015). Para obtener estos productos, pueden utilizarse diferentes GHs como endoxilanasas, inulinasas o β-galactosidasas.

3. Industria energética. Quizá la aplicación más importante de las GHs es la relacionada con su uso como catalizadores en la producción de bioetanol de primera (1G) y segunda (2G) generación. En el primer caso, la fuente de azúcares fermentables es el almidón o la sacarosa, presentes en diversas materias primas como granos de cereales, remolacha o caña de azúcar. Por el contrario, el de segunda generación se basa en la degradación de la biomasa lignocelulósica, donde las fuentes de azúcares fermentables son la celulosa y la hemicelulosa de la pared celular de las células vegetales presentes, por ejemplo, en paja de cereal, bagazo de maíz, o madera (Gray *et al*., 2006). En este proceso, los polisacáridos de la pared celular o de los materiales de reserva de la planta se hidrolizan en sus correspondientes monosacáridos, esencialmente glucosa y xilosa, que luego se fermentan para obtener etanol. Una sacarificación eficiente de los polisacáridos requiere la acción concertada de diferentes tipos de GHs, dependiendo del tipo de etanol que se esté produciendo, principalmente α-1,4-amilasas (bioetanol 1G), β-1,4-endoglucanasas, celobiohidrolasas, β-1,4-glucosidasas, endo 1,4-β-xilanasas, y β-1,4-xilosidasas (bioetanol 2G) (Martínez *et al*., 2009). En ocasiones, estas enzimas son ayudadas por otras que desempeñan una actividad auxiliar que incrementa los rendimientos del proceso.

4. Industria biomédica. El campo médico constituye un caso particular, ya que las aplicaciones derivadas no están tan relacionadas con las capacidades catalíticas de las GHs si no con su papel en el desarrollo de algunas enfermedades. Así, se han ensayado estrategias en las que se busca inhibir diferentes GHs implicadas en procesos de patogénesis causados por virus, bacterias u hongos, o también en otras enfermedades como diabetes o algunos tipos de cáncer.
   Por otro lado, el uso de la capacidad de transglicosilación de las GHs puede utilizarse para la mejora de propiedades de moléculas bioactivas. Gracias a la adición de una o más unidades de carbohidratos a estas moléculas, se han reportado efectos beneficiosos en comparación a sus precursores no glicosilados, en términos de mayor solubilidad, bioseguridad y estabilidad. (Kometani *et al*., 1994; Prodanovic *et al*., 2005; Torres *et al*., 2011; Woo *et al*., 2012)

## 2. Biomasa lignocelulósica

El $CO_2$ atmosférico es fijado por los tejidos vegetales a través del complejo fotosintético, constituyendo el paso más importante en la incorporación de carbono inorgánico a los ecosistemas (Leschine, 1995). Así es como se produce la biomasa lignocelulósica, que es la materia prima orgánica más abundante en el planeta tierra.

La biomasa lignocelulósica está formada por tres componentes mayoritarios: celulosa, hemicelulosa, y lignina (figura 4). La celulosa y la hemicelulosa son polisacáridos, mientras que la lignina es un polímero de unidades fenólicas (Doherty *et al*., 2011). El porcentaje de estos componentes varía en función de la especie vegetal, del tejido seleccionado y de su etapa de desarrollo (Liu *et al*., 2016). Las fibras de celulosa, formadas por cadenas poliméricas ordenadas que contienen regiones cristalinas muy compactas, se incrustan dentro de una matriz formada por hemicelulosa y lignina.

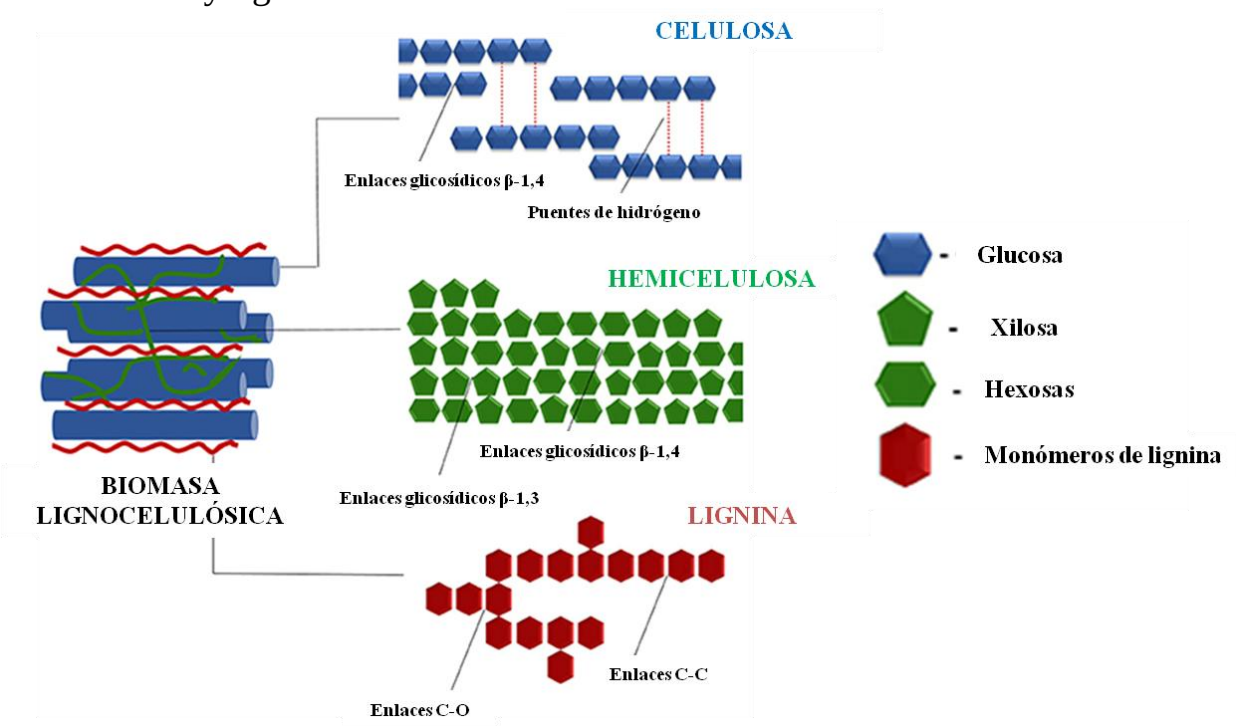



**Figura 4.** Composición de la biomasa lignocelulósica presente en la pared celular vegetal (Adaptada de Baruah *et al*., 2018).

### 2.1. Lignina

La principal función de este polímero es dar soporte estructural a los tejidos vegetales. Debido a su carácter recalcitrante, también proporciona una gran resistencia frente al ataque de patógenos. La lignina se forma mediante la polimerización de tres tipos de alcoholes cinamílicos: el *p*-

cumarílico, coniferílico, y sinapílico (Doherty *et al*., 2011). La figura 5 muestra un esquema representativo de la lignina. Su composición es muy variable en función de los tipos de plantas, y su desestructuración constituye un paso clave para permitir el acceso a la celulosa y hemicelulosa, debido a que es el componente más difícil de degradar del material lignocelulósico. Por ejemplo, en relación a esto, se han realizado estudios que demuestran que las plantas con un mayor contenido de lignina son más plásticas, elásticas, y resistentes a la luz solar directa y las heladas (Farahi *et al*., 2017). Los tres tipos de monómeros de la lignina comparten la misma estructura química básica, difiriendo en el grado de sustituciones del anillo fenólico. El porcentaje de entrecruzamiento, y por lo tanto la rigidez de la estructura, depende de la proporción de los monómeros que tenga cada tipo de material. Hasta la fecha, la lignina se ha considerado más como un obstáculo a retirar para mejorar el acceso a los polisacáridos que como una materia prima de la que obtener productos de alto valor añadido. Sin embargo, trabajos recientes muestran el potencial de este polímero en diferentes aplicaciones, como la producción de nuevos combustibles (Azadi *et al*., 2013), su uso como "building blocks" para diferentes procesos químicos (Azadi *et al*., 2013), o la producción de compuestos con actividad antimicrobiana, antifúngica y antioxidante (Dong *et al*., 2011; Guo *et al*., 2018).

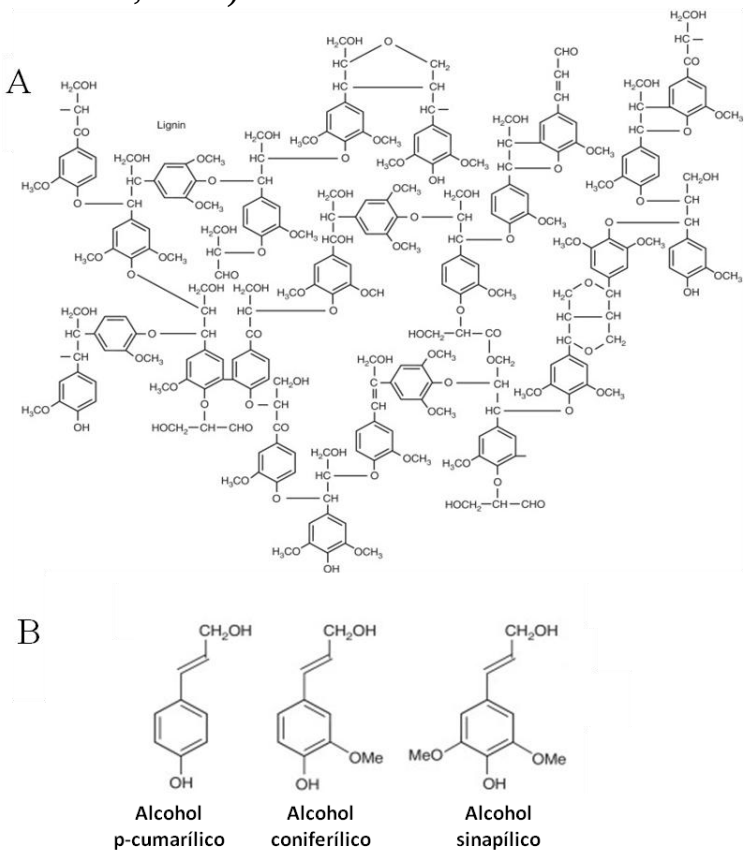



**Figura 5.** Representación esquemática de la estructura de la lignina (A), y estructura molecular de los diferentes monómeros que la componen (B). (Adaptado de Watkins *et al*., 2015).

## 2.2. Hemicelulosa

El término hemicelulosa engloba una serie de polisacáridos de estructuras heterogéneas que forman parte de las paredes celulares de las plantas junto a la celulosa. Las hemicelulosas se clasifican generalmente en cuatro tipos: xilanos, mananos, xiloglucanos, y β-glucanos de enlace mixto, que difieren entre sí en su estructura principal, sus ramificaciones, y el tipo y distribución de los enlaces glicosídicos (Scheller y Ulvskov, 2010). En su composición de carbohidratos participan pentosas (β-D-xilosa y β-L-arabinosa), hexosas (β-D-glucosa, β-D-manosa y β-D-galactosa), desoxihexosas (α-L-ramnosa y α-L-fucosa) y ácidos urónicos (β-D-glucurónico, y α-D-galacturónico) (Quiroz Castañeda y Folch-Mallol, 2011).

- Los **xilanos** son heteropolisacáridos ramificados formados por una cadena principal de moléculas de xilosa, unidas por enlace glucosídicos β-1,4, y diferentes cadenas laterales (figura 6) (Ebringerová, 2005). Si las cadenas laterales tienen residuos de ácido glucurónico, el polisacárido será denominado glucuronoxilano. Si, además contiene arabinosa, se llama glucuronoarabinoxilano, o arabinoxilano si las proporciones de arabinosa son mayoritarias. Una particularidad de los arabinoxilanos, es que pueden tener unidades fenólicas, principalmente ácido ferúlico, unidas al carbono 5 de las arabinosas a través de un enlace éster.

- Los **mananos** son polímeros ramificados, en los cuales la cadena principal está formada por manosas unidas por enlaces glucosídicos β-1,4. En la posición *O*-6 de la cadena principal de manano tienen ramificaciones de galactosa. Otras variedades de mananos, denominadas glucomananos, incluyen residuos de β-1,4-glucopiranosa en la cadena principal (Scheller y Ulvskov, 2010).

- Los **xiloglucanos** tienen una cadena principal compuesta por moléculas de glucosa unidas por enlaces β-1,4, a la que se añaden ramificaciones de xilosa unidas al O-6 de las glucosas de la cadena principal (Scheller y Ulvskov, 2010).

- Por último, los **β-glucanos de enlace mixto** tienen una cadena lineal compuesta por bloques de glucosa unidos por enlaces β-1,4 o β-1,3, en diferente proporción según el organismo de origen (Scheller y Ulvskov, 2010).

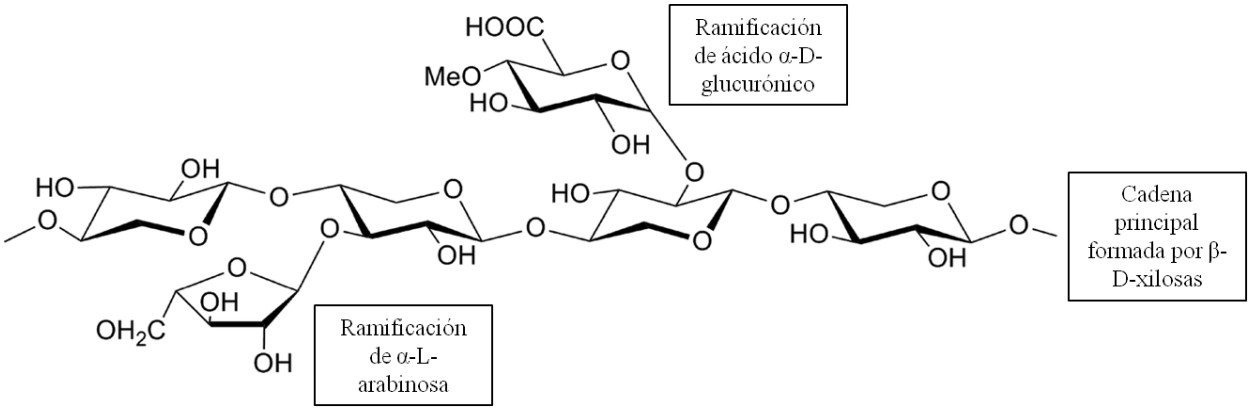



**Figura 6.** Estructura del xilano como ejemplo de hemicelulosas. Se puede observar la cadena principal de xilosa, de la que parten ramificaciones de ácido β-D-glucurónico, y de β-L-arabinosa (Adaptado de Chen *et al*., 2014).

## 2.3. Celulosa

La celulosa, que destaca por ser el polímero más abundante de la Tierra, es un homopolisacárido formado por largas cadenas lineales de glucosa unidas por enlaces glucosídicos β-1,4. Estas cadenas se unen entre sí por puentes de hidrógeno y fuerzas de Van der Waals para formar las microfibrillas de celulosa. Éstas son la base de la celulosa microcristalina, también conocida como celulosa I, un polímero insoluble, altamente recalcitrante, y muy difícil de degradar (Himmel *et al*., 2007). Las regiones cristalinas están altamente organizadas y forman un tipo de celulosa insoluble en la mayoría de solventes, dificultando de esta manera que sea atacada por glicosil hidrolasas (Hildén y Johansson, 2004). En la naturaleza, esta estructura se combina con otras regiones de celulosa amorfa, que son regiones no cristalinas situadas en el espacio interfibrilar, las cuales permiten un mayor acceso a moléculas de agua, mejorando la capacidad de acción de las enzimas (Klemm *et al*., 2005). Dada su abundancia en la biosfera y su tremenda utilidad como precursor de combustibles y otros químicos, el estudio de la celulosa ha sido vasto y diverso (Payne *et al*., 2015).

A pesar de su aparente homogeneidad, la celulosa puede presentar diferentes conformaciones. Los organismos que son capaces de sintetizarla, producen siempre celulosa I, que se caracteriza por tener las cadenas de glucosa organizadas de forma paralela (Gardner y Blackwell, 1974). A su vez, la celulosa I se divide entre Iα y Iβ, que se diferencian en los patrones que adoptan los enlaces de hidrógeno, y en la disposición de las cadenas (Finkenstadt y Millane, 1998). Aparte, ciertos tratamientos químicos pueden convertir la celulosa I en otras formas cristalinas. Por ejemplo, tratar el polisacárido con disolventes como el hidróxido de sodio, puede

convertir las cadenas paralelas de la celulosa nativa en antiparalelas, produciendo así la denominada celulosa II (Langan *et al*., 2001). Tratamientos químicos menos severos, como por ejemplo la incubación en amoníaco, son capaces de convertir tanto la celulosa I como la celulosa II en celulosa III, que presenta mayor separación entre las diferentes cadenas de celulosa, favoreciendo la entrada de enzimas (Wada *et al*., 2001). La gran ventaja que muestran poseer estas formas de celulosa, comparadas con la celulosa I, es que exhiben una mayor digestibilidad por parte de las celulasas, poniendo en valor que el uso de un proceso de pretratamiento químico previo al uso de glicosil hidrolasas puede mejorar la eficiencia del proceso (Chundawat *et al*., 2011).

En la actualidad se han desarrollado diferentes tipos de celulosas comerciales que pueden ser utilizados como sustrato de las enzimas celulolíticas:

- Celulosa de elevada cristalinidad: en este grupo se encuentan la celulosa microcristalina bacteriana, el Avicel y la celulosa procedente de algas.
- Celulosa amorfa. Puede tratarse de celulosa regenerada, carboximetil celulosa o celulosa cristalina tratada con ácido fosfórico.

El grado de polimerización y de cristalinidad, o el número de microfibrillas de cada uno de estos sustratos puede ser significativamente diferente, lo que influye en la efectividad de las enzimas que llevan a cabo su degradación. Por ello, la utilización de distintos sustratos modelo, aun siendo todos ellos celulosas, a menudo conduce a unas variaciones significativas de los resultados y puede dificultar la comparación entre distintas enzimas.

## 2.4. Biomasa lignocelulósica para la producción de bioetanol de segunda generación

La biomasa lignocelulósica tiene un enorme potencial para cubrir las demandas mundiales de materias primas en diferentes industrias. Solo en los Estados Unidos, se ha estimado que podrían sustituirse el 30% de los productos derivados del petróleo mediante la conversión de biomasa vegetal en combustible. La creciente demanda de energía a nivel mundial y la preocupación suscitada por el uso continuo de los combustibles fósiles y su influencia en el cambio climático han convertido las aplicaciones derivadas de la utilización de biomasa lignocelulósica en una herramienta clave para la sostenibilidad del planeta (Amorim *et al*., 2011).

Debido a su bajo coste, abundancia, y diversidad química y estructural, su potencial biotecnológico se considera enorme, y se ha

descrito como una fuente casi inagotable de energía y precursores para diferentes procesos químicos (Cherubini y Strømman, 2011). En los últimos años, la mayoría de las investigaciones se han centrado en la producción de bioetanol 2G, procedente de la degradación de la ligocelulosa, como alternativa a los combustibles fósiles y a la producción de bioetanol 1G, para el que se utilizan materias primas compatibles con el consumo humano. Sin embargo, la síntesis de bioetanol 2G presenta otros inconvenientes. Su principal limitación es la presencia de lignina en la lignocelulosa, que hay que eliminar o transformar para utilizar los polisacáridos de la pared celular de las plantas (Naik *et al*., 2010). Al mismo tiempo, la sacarificación de estos hidratos de carbono requiere cócteles enzimáticos más complejos que los utilizados para la producción del bioetanol 1G, que se realiza a partir de carbohidratos fácilmente hidrolizables, como la sacarosa o el almidón. Las mejoras tecnológicas del proceso, junto con la necesidad de elaborar bioetanol 2G de una forma rentable, han promovido el concepto de biorrefinería, que se puede definir como la infraestructura donde se combinan procesos de aprovechamiento de biomasa para producir múltiples productos (Martínez *et al*., 2009). Así, además del bioetanol, los residuos vegetales se pueden utilizar para obtener productos químicos, proteínas, biomateriales (Witzler *et al*., 2018) y otras energías como biohidrógeno o biogas (Mulat *et al*., 2018), con el fin de hacer más rentable el proceso (figura 7).

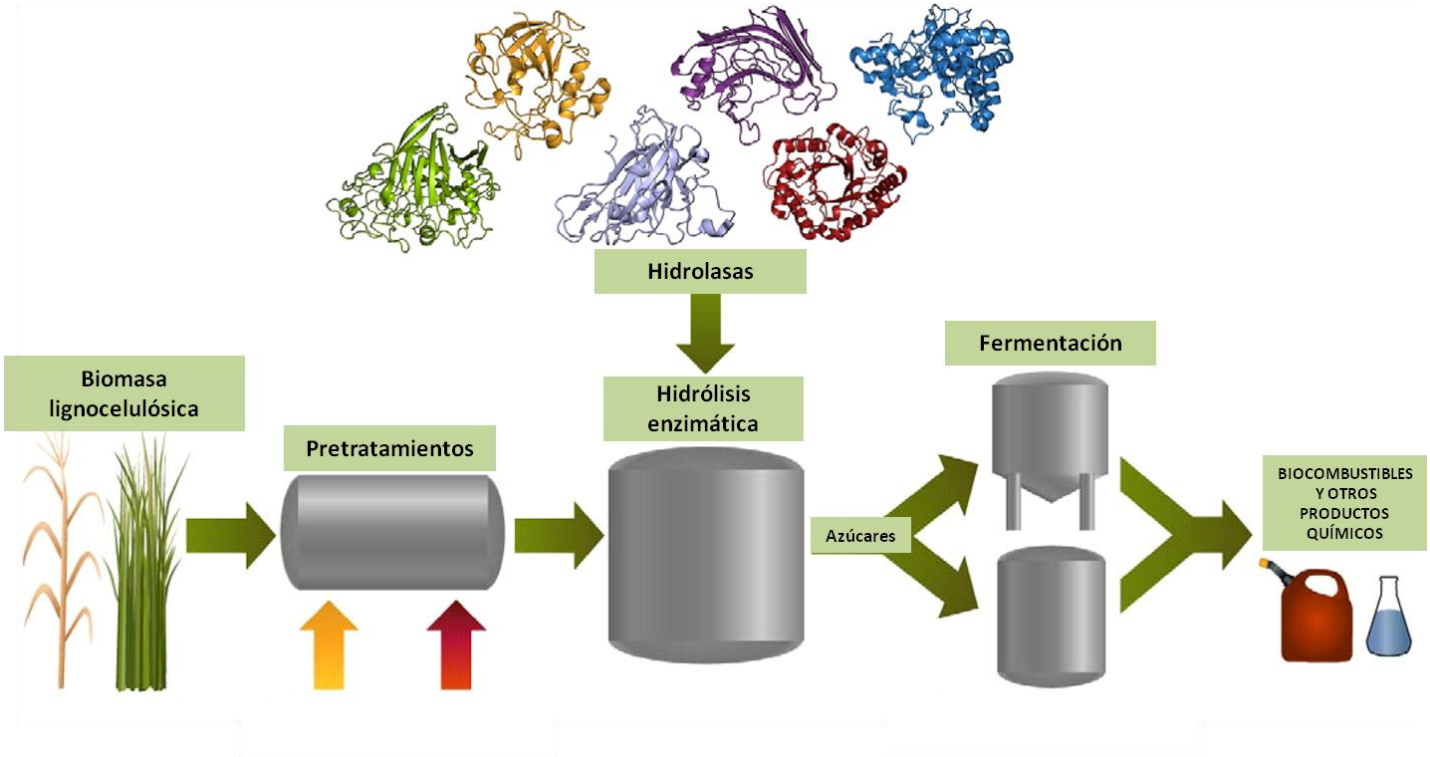



**Figura 7.** Producción de biocombustibles y otros productos químicos a partir de biomasa lignocelulósica (Adaptado de Payne *et al*., 2015).

## 2.5. Pretratamientos de la biomasa lignocelulósica

En la mayoría de casos, para conseguir un mayor rendimiento en los procesos de aprovechamiento de biomasa lignocelulósica, es necesario realizar un pretratamiento de los residuos vegetales (Alvira *et al*., 2010), que permita retirar o alterar la lignina y aumentar la disponibilidad de la celulosa y hemicelulosa.

Los procesos de pretratamiento más eficaces se describen a continuación:

- **Físicos.** La reducción del tamaño mediante métodos de estrés mecánico, como la trituración, el astillado o la molienda, constituye el pretratamiento más básico de la biomasa. Así se puede mejorar la eficiencia de procesos posteriores, pero el coste energético de conseguir partículas muy pequeñas puede ser muy elevado (Hendriks y Zeeman, 2009). Por regla general, a menor tamaño de partícula y mayor humedad del material de partida, se requiere un mayor gasto de energía. Otro tratamiento físico de gran importancia es la extrusión. En ésta, los materiales se someten a un calentamiento, y cizallamiento, que provoca modificaciones físicas y químicas (Chinnadurai *et al*., 2008). Aunque los tratamientos físicos suelen ser necesarios en cualquier proceso de degradación de biomasa lignocelulósica, suelen realizarse en combinación con otros tipos de pretratamientos.

- **Químicos.** Existen distintos tipos de pretratamientos químicos para mejorar la digestibilidad de la biomasa lignocelulósica. Una de las opciones más eficientes es incubar la biomasa en soluciones alcalinas (Kumar y Wyman, 2009) para incrementar la digestibilidad de la celulosa y hemicelulosa, ya que a pHs básicos la lignina se solubiliza. Son bastante eficientes porque el porcentaje de celulosa y hemicelulosa que se pierde es pequeño. Otro tipo de tratamientos se basan en la aplicación de ácidos (Wyman 1996). El principal objetivo de estos pretratamientos es solubilizar la fracción hemicelulósica de la biomasa y hacer que la celulosa sea más accesible. Estos pretratamientos pueden realizarse con ácido, generalmente diluido, aunque pueden formarse compuestos secundarios que pueden actuar como inhibidores del proceso. Otro tipo de tratamiento químico se basa en el uso de ozono, un potente oxidante que muestra una alta eficiencia en la deslignificación de residuos vegetales (Sun y Cheng, 2002). En este caso, el pretratamiento no suele producir compuestos inhibidores de la

posterior hidrólisis de los polisacáridos. Otros dos procesos que también han sido utilizados como pretratamientos, aunque con menor eficacia, son el denominado "organosolv", o uso de solventes orgánicos para solubilizar la lignina y separarla de los polisacáridos (Papatheofanous *et al*., 1995), y la aplicación de líquidos iónicos para conseguir disolver componentes de la biomasa lingocelulósica (Li *et al*., 2009).

- **Físico-químicos.** Se basan en la combinación de procesos físicos y químicos, y en general son los que más eficiencia han demostrado. Dentro de este grupo, el más utilizado es la explosión con vapor (o "*steam explosion*") (Alfani *et al*., 2000). Consiste en tratar la biomasa lignocelulósica con vapor de agua a altas presiones, seguido de una descompresión rápida. De forma general, la biomasa se expone a 210-290 ºC y 20-50 bares. En ocasiones, se puede añadir ácido sulfúrico diluido, aumentando en general la eficiencia del tratamiento. El proceso de explosión con vapor presenta varias características atractivas en comparación con otras técnicas, como son su menor impacto ambiental y coste económico, las condiciones de reacción menos peligrosas, y una buena recuperación del material inicial, que simplemente sufre una modificación (Avellar y Glasser, 1998). A pesar de todo ello, también posee alguna desventaja ya que, aunque se solubiliza la hemicelulosa en presencia de ácido y sus azúcares podrían fermentarse, en el proceso se generan algunos compuestos inhibidores producto de las modificaciones que sufre la lignina y los polisacáridos. Por ello, se han desarrollado técnicas físico-químicas alternativas. Una de ellas es la denominada "*liquid hot water*", que es un tratamiento hidrotérmico que no requiere una rápida descompresión y en el que no se emplea ningún producto como catalizador (Mosier *et al*., 2005; Pérez *et al*., 2008). Se basa en aplicar presión para mantener el agua en estado líquido a temperaturas elevadas (entre 160 y 240 °C) y provocar alteraciones en la estructura de la lignocelulosa. El objetivo del agua caliente líquida es solubilizar principalmente la hemicelulosa, haciendo que la celulosa sea más accesible y con un proceso que evita la formación de inhibidores, pero el rendimiento global del proceso es menor que el de la "*steam explosion"*. Otra alternativa es la explosión con amonio, también conocida como "*Ammonia fiber explosion*" o por sus siglas, "AFEX" (Laureano-Perez *et al*., 2005). En este proceso, la biomasa se trata con amoníaco líquido a temperaturas entre 60 y 100 ° C y altas presiones durante un período de tiempo variable. Luego se disminuye la presión rápidamente, lo que da como resultado una rápida expansión del gas amoníaco que

provoca la rotura de las fibras de biomasa y una reducción parcial de la cristalinidad de la celulosa (Galbe y Zacchi, 2007). Como principales ventajas de este método, está que la producción de inhibidores para los procesos biológicos posteriores es mucho más baja que con otros métodos, aunque algunos derivados fenólicos de la lignina pueden permanecer en la superficie de la celulosa.
Aparte de estas tres técnicas, hay otros nuevos métodos de tratamientos físico-químicos que han demostrado buenos resultados, como por ejemplo el uso de microondas, ultrasonidos, o explosión de dióxido de carbono (Alvira *et al*., 2010).

- **Biológicos.** En este tipo de pretratamiento, los materiales lignocelulósicos son modificados por la acción de determinados microorganismos, entre los que destacan hongos basidiomicetos (Sánchez, 2009; Salvachúa *et al*., 2011) , que poseen gran variedad de lacasas y peroxidasas para la degradación de la lignina. Aunque es un proceso respetuoso con el medio ambiente y no genera inhibidores indeseables, el uso de hongos para hacer pretratamientos de biomasa lignocelulósica presenta dos inconvenientes principales (Sun y Cheng, 2002): 1) Muchos de estos hongos no se limitan a degradar la lignina, sino que también consumen los polisacáridos, con lo que se disminuye el rendimiento de la fermentación. 2) Los rendimientos de los tratamientos biológicos son muy inferiores, y además requieren de tiempos largos para conseguir una buena deslignificación. En cualquier caso, para llevar a cabo con éxito pretratamientos biológicos se han usado principalmente hongos de la podredumbre blanca (Shi *et al*., 2008; Alvira *et al*., 2010) como *Phanerochaete chrysosporium*, *Ceriporia lacerata*, *Cyathus stercolerus*, *Ceriporiopsis subvermispora*, *Pycnoporus cinnarbarinus* y *Pleurotus ostreaus*. Cabe destacar también entre estos basidiomicetos al hongo *Irpex lacteus*, que en presencia de lignocelulosa secreta peroxidasas que actuarían sobre la lignina, facilitando la deconstrucción de la pared celular, tal y como se ha evidenciado mediante el estudio de su secretoma (Salvachúa *et al*., 2013). Además, este hongo posee una baja cantidad de celulasas y xilanasas y, por tanto, apenas consume carbohidratos, dejándolos disponibles para la producción de etanol 2G.

## 3. Degradación enzimática de la biomasa lignocelulósica

Como hemos visto, la pared celular vegetal está formada por varios componentes de naturaleza recalcitrante, por lo que solo determinados microorganismos son capaces de degradarla. Entre ellos, se encuentran varias especies de hongos filamentosos, denominados hongos de la podredumbre de la madera (Blanchette, 1991)**.** Éstos desempeñan un papel central en la ecología de los sistemas forestales, pues son elementos clave para el desarrollo del ciclo del carbono. En función de los patrones de degradación que producen sobre la biomasa lignocelulósica, se clasifican en:

- **Hongos de podredumbre blanca**. La mayoría son hongos basidiomicetos que son los únicos organismos capaces de degradar completamente la lignina (Kirk y Farrell, 1987). Comienzan su ataque degradando lignina y hemicelulosa, y tras retirar éstas, finalmente atacan la celulosa a un ritmo reducido (Martínez *et al*., 2005).

- **Hongos de podredumbre parda**. También son hongos basidiomicetos, pero principalmente degradan celulosa y hemicelulosa, modificando la lignina, pero sin degradarla (Martínez *et al*., 2011).

- **Hongos de podredumbre blanda**. Este grupo está compuesto por hongos ascomicetos (Levy 1966). Funcionan de manera parecida a los de podredumbre parda, alterando ligeramente la lignina para poder acceder a los polisacáridos que utilizarán como fuente de carbono.

Para degradar la biomasa lignocelulósica los hongos filamentosos hacen uso de una gran variedad de enzimas, entre ellas una amplia batería de glicosil hidrolasas. En primer lugar se suele llevar a cabo el ataque a la lignina, en el que intervienen oxidorreductasas extracelulares (lacasas, peroxidasas y oxidasas productoras de $H_2O_2$), intracelulares, metabolitos, y especies activas de oxígeno (Kirk y Farrell, 1987; Guillén *et al*., 1990; Valmaseda *et al*., 1991).

Una vez que la lignina se ha alterado o degradado, quedan más disponibles en la pared celular vegetal la celulosa y la hemicelulosa, polisacáridos que tras su hidrólisis pueden ser fermentados a etanol, pero también utilizados en distintas aplicaciones biotecnológicas que ayuden a la valorización del proceso y su comercialización. β-xilosidasas y endo-β-1,4-

xilanasas, aparte de otras enzimas como α-arabinofuranosidasas y α-glucuronidasas, son las encargadas de la hidrólisis de hemicelulosas.
En el caso de la celulosa, las enzimas implicadas se denominan celulasas y, dado que son el objetivo fundamental de este trabajo, se tratarán en el siguiente apartado.

### 3.1. Hidrólisis enzimática de la celulosa

La celulosa es el polisacárido más abundante del planeta, y puede ser utilizada para desarrollar un gran número de aplicaciones. Uno de los usos más explotados es su conversión en azúcares fermentables. La celulosa experimenta una degradación por procesos de hidrólisis a través de la colaboración coordinada de un complejo de múltiples enzimas, las celulasas. Las enzimas celulolíticas están ampliamente distribuidas en la naturaleza y se organizan en muchas familias y subfamilias. Existen principalmente tres clases (Tabla 1), típicamente extracelulares, que se han estudiado ampliamente a lo largo de las últimas cinco décadas (Bhat y Bhat, 1997; Castro y Pereira Jr, 2009):

- Las **β-1,4-endoglucanasas** (EGs) (EC 3.2.1.4) (Payne *et al*., 2015), que actúan sobre los enlaces β-1,4 del interior de la cadena de celulosa, rompiéndola en unidades más pequeñas, y proporcionando así más moléculas susceptibles de ser degradadas por las enzimas que atacan los extremos de las fibras de celulosa. La mayoría de EGs estudiadas hasta la fecha se encuentran en las siguientes familias de GHs: 5, 6, 7, 8, 9, 10, 12, 26, 44, 45, 48, 51, 74, 124, y 148.

- **β-1,4-Exoglucanasas** (EC 3.2.1.91) (Payne *et al*., 2015), también conocidas como celobiohidrolasas (CBHs), son enzimas que actúan en los extremos de las cadenas de celulosa, tanto reductores como no reductores, liberando moléculas de celobiosa y celooligosacáridos. Estas enzimas pertenecen principalmente a las familias 5, 6, 7, 9 y 48 de GHs.

- Las **β-1,4-Glucosidasas** (BGLs) (EC 3.2.1.21) (Singhania *et al*., 2013; Sørensen *et al*., 2013), que llevan a cabo el último paso de la hidrólisis de la celulosa, hidrolizando la celobiosa y los oligosacáridos de cadena corta en unidades monoméricas de glucosa. Enzimas con esta actividad han sido descubiertas en las familias de GHs 1, 2, 3, 5, 9, 16, 30, 39 y 116.

En general, se puede afirmar que todas estas enzimas exhiben un comportamiento sinérgico, debido a que trabajan de forma coordinada,

ayudándose unas a otras, y mejorando la eficiencia de la hidrólisis de la celulosa (Lynd *et al*., 2002). El primer paso es llevado a cabo por las EGs, que rompen la cadena de celulosa lineal al actuar sobre los enlaces glicosídicos internos en las regiones amorfas. Posteriormente, las CBHs actúan en los nuevos extremos de las cadenas generados por las EGs, liberando celooligosacáridos y celobiosa, que finalmente son convertidos en glucosa por las BGLs (Van Dyk y Pletschke, 2012).

**Tabla 1.** Características de las principales familias de enzimas celulasas.

| Actividad | Familia GH | pH óptimo | Temperatura óptima (ºC) | Tamaño molecular (kDa) | Sustrato |
|---|---|---|---|---|---|
| **BGLs** | 1 | 3.5-6.3 | 40-55 | 52-94 | Celobiosa y otros |
| | 3 | 3.5-8.0 | 37-72 | 74-145 | celooligosacáridos |
| **CBHs** | 6 | 4.8-9.0 | 40-50 | 40-60 | Extremos de los |
| | 7 | 5.0-6.0 | 35-65 | 47-90 | polisacáridos de celulosa |
| **EGs** | 5 | 3.5-8.5 | 40-75 | 35-56 | Partes amorfas del |
| | 7 | 4.0-5.5 | 45-57 | 46-56 | polisacárido de celulosa |
| | 12 | 2.0-5.0 | 55-70 | 25-32 | |
| | 45 | 5.0-7.0 | 30-65 | 20-47 | |

A lo largo de la última década, se han descubierto otro tipo de enzimas, a las que inicialmente se les atribuyó actividad hidrolítica, que degradan la celulosa por reacciones de oxidorreducción (Martínez, 2016). Los primeros estudios sobre la degradación de la celulosa indicaban la posibilidad de que existiera un sistema de dos componentes (Reese *et al*., 1950): 1) En primer lugar actuaría un primer componente, no identificado, que ayudaría a reducir la cristalinidad de la celulosa, y 2) un segundo componente, que correspondería a las glicosil hidrolasas explicadas en el párrafo anterior. Pasaron sesenta años entre estos estudios y la descripción de las monooxigenasas degradadoras de polisacáridos (LPMOs) a las que se ha atribuido esta actividad (Harris *et al*., 2010; Vaaje-Kolstad *et al*., 2010). Estas nuevas enzimas se han incluido en la base de datos CAZY en la sección de actividades auxiliares, y actualmente pertenecen a las familias AA9, 10, 11, 13, 14, 15 y 16 (Levasseur *et al*., 2013), aunque muchas de ellas sólo tienen uno o dos miembros caracterizados hasta la fecha. La diversidad de LPMOs parece muy grande, pero todavía ha sido poco explorada, y el número de enzimas identificadas es bajo. A pesar de ello, el

número de genes de potenciales LPMOs descubiertas en genomas fúngicos ha aumentado exponencialmente en los últimos años, lo que demuestra que son enzimas muy antiguas y ofrecen un nuevo y vasto campo a estudiar en la degradación de la celulosa. La figura 8 muestra una representación del funcionamiento sinérgico de las diferentes enzimas celulasas.

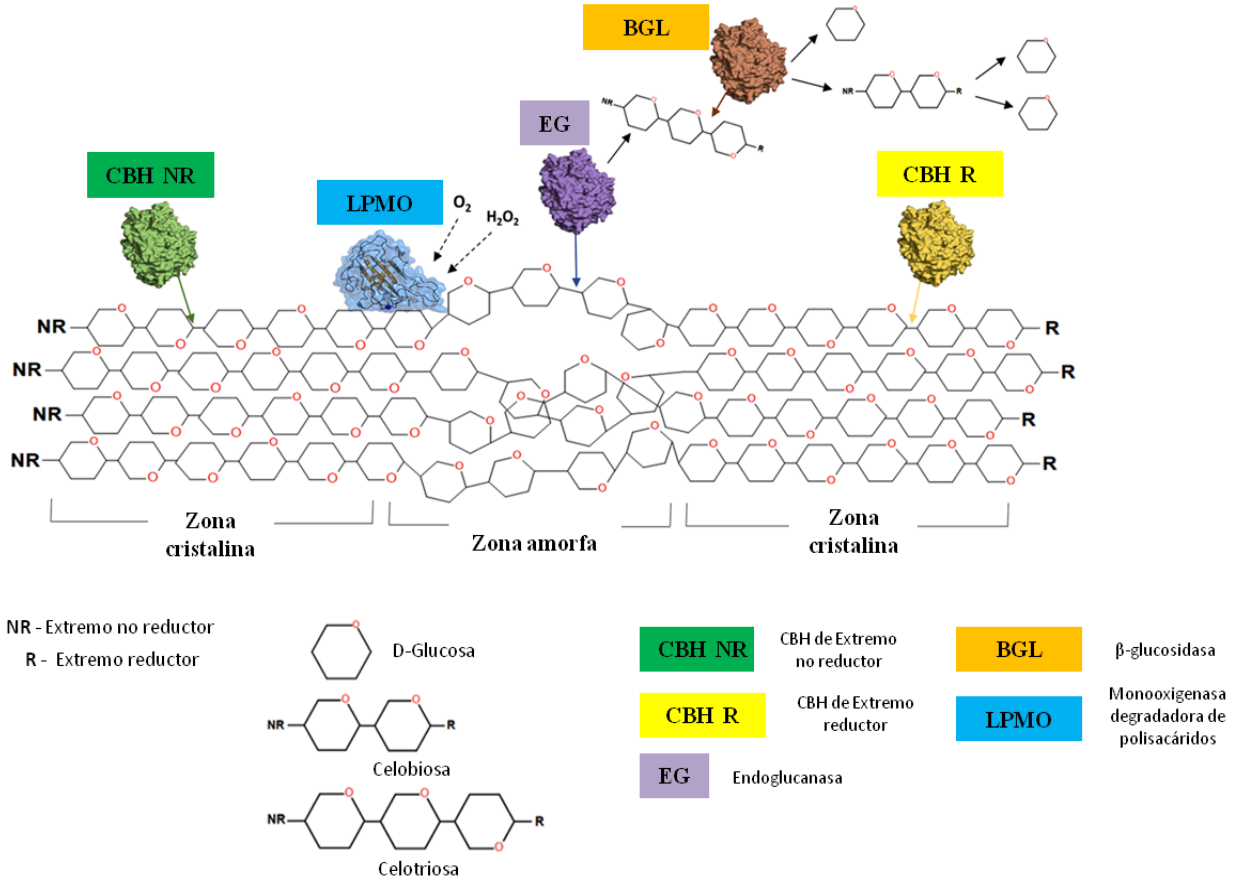



**Figura 8.** Representación de la degradación enzimática de la celulosa por la acción sinérgica de los 4 tipos de enzimas celulasas (adaptado de Andlar *et al*., 2018).

De forma general, las enzimas que degradan la biomasa lignocelulósica trabajan en una interfaz sólido-líquido y, en estas situaciones, la concentración de unidades catalíticas en la superficie del sustrato está directamente relacionada con la eficacia de la degradación. Debido a esto, la mayoría de celulasas son multimodulares con uno o varios dominios catalíticos acoplados a uno o más CBMs (figura 9) (Linder y Teeri, 1997; Schülein, 1997).

El dominio catalítico constituye la mayor parte de la secuencia de la proteína y contiene el sitio activo. La especificidad de sustrato de una determinada enzima se puede predecir con frecuencia en base a la estructura del sitio activo: la forma de surco y de bolsillo son características de enzimas con actividad exo, mientras que la forma de túnel se da

típicamente en enzimas de tipo endo (Lynd *et al*., 2002). A menudo este dominio catalítico se puede subdividir en varios dominios estructurales, como ocurre en algunas BGLs y β-xilosidasas de la familia GH3.

El dominio catalítico y los CBM se unen a través de una secuencia denominada "*linker*" (figura 9), que actúa como conector flexible, lo que permite que ambos dominios actúen de forma independiente (Srisodsuk *et al*., 1993). Esta secuencia varía de una enzima a otra aunque su composición suele ser muy rica en treoninas y serinas, que son dianas habituales de glicosilaciones, lo que proporciona a los linkers más rigidez y protección frente a procesos proteolíticos (Langsford *et al*., 1987).

Hasta la fecha, se han descrito dominios de unión a celulosa en CBHs y EGs, pero no se ha descubierto ninguno en BGLs. Esto seguramente esté relacionado con la naturaleza de los sustratos de éstas, que suelen ser solubles, y por tanto, más accesibles para las enzimas que deben degradarlos.

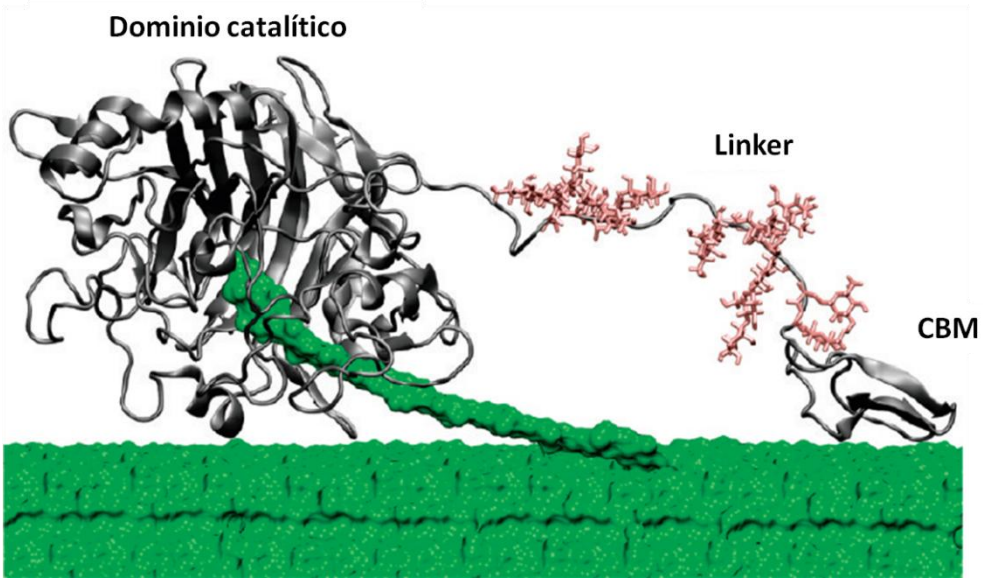



**Figura 9.** Ejemplo de representación de la estructura de una enzima celulolítica, tomando como modelo la CBH Cel7A de *Trichoderma reesei*. Las cadenas de celulosa se representan en color verde. Las glicosilaciones del linker, de color rosa (adaptado de Beckham *et al*., 2010).

## 3.2. β-glucosidasas: enzimas clave en la degradación de la celulosa

La celulosa es degradada hasta glucosa por la acción sinérgica de LPMOs, EGS, CBHs y BGLs. Las EGs y CBHs actúan en primer lugar, hidrolizando el polímero de celulosa hasta liberar celobiosa y otros celooligosacáridos, que son finalmente convertidos por las BGLs en glucosa (Singhania *et al*., 2013; Sørensen *et al*., 2013). Este paso es esencial ya que las CBHs y EGs a menudo se inhiben por la celobiosa

(Murphy *et al*., 2013), por lo que la acción de las BGLs disminuye la inhibición por producto de éstas enzimas. De esta manera, podrían mantenerse rendimientos de hidrólisis de celulosa constantes a lo largo del tiempo. Sin embargo, las BGLs pueden ser también inhibidas por producto (Shewale, 1982; Xiao *et al*., 2004), y esta inhibición es difícilmente evitable ya que el monosacárido es el producto último de la reacción de hidrólisis de la celulosa. Por ello, se considera que las BGLs son los componentes que limitan la eficiencia del proceso. El mantenimiento de una alta tasa de hidrólisis de celulosa, por tanto, demanda BGLs muy eficientes que sean capaces de tolerar altas concentraciones de glucosa.
En general, se puede considerar que las características clave para evaluar la capacidad de una BGL son: 1) su tasa de hidrólisis, 2) su inhibición por glucosa, y 3) su estabilidad en las condiciones de la reacción.

La estabilidad de una enzima viene definida generalmente por su comportamiento frente a pH y temperatura. En general, las BGLs presentan un pH óptimo de funcionamiento en torno a 4-5 (Bhatia *et al*., 2002; Eyzaguirre *et al*., 2005) disminuyendo significativamente la actividad a pH mucho más bajos o altos. Por lo tanto, en la mayoría de los casos, las reacciones enzimáticas deben ajustarse en torno a estos valores de pH. Con respecto a la temperatura, cuando se utilizan valores muy altos, la estabilidad de las proteínas se ve afectada, lo que puede producir desnaturalización y, por tanto, inactivación irreversible de la enzima. La mayoría de BGLs producidas por hongos mesófilos que normalmente crecen a 28 ºC, producen BGLs con temperaturas óptimas de alrededor de 50 ºC, pero solo se observan aumentos destacables de la estabilidad térmica en las enzimas derivadas de los hongos termófilos (Yeoman *et al*., 2010). Debido a que los procesos de hidrólisis de biomasa suelen durar muchas horas o incluso días, la estabilidad de la enzima al pH y la temperatura del proceso es fundamental.

Otro problema que afecta a la eficacia de las BGLs es la disminución en el rendimiento de la hidrólisis de la celulosa, causada por eventos de transglicosilación. La transglicosilación es, obviamente, un evento no deseado en la hidrólisis de la biomasa, y tiene lugar con bastante facilidad en varias BGLs descritas hasta la fecha (Bhatia *et al*., 2002), especialmente a altas concentraciones de sustrato (Bohlin *et al*., 2013). La mutagénesis dirigida enfocada a modificar los aminoácidos esenciales implicados en la transglicosilación podría reducir potencialmente este inconveniente (Frutuoso, 2013).

## 3.3. Clasificación y estructura de la BGLs

Las BGLs, como otras glicosil hidrolasas, se clasifican habitualmente en función de su secuencia, pero existe otra clasificación basada en su

especificidad de sustrato: así, pueden ser 1) celobiasas (alta especificidad hacia la celobiosa y otros celooligosacáridos), 2) aril-β-glucosidasas (alta especificidad hacia los sustratos sintéticos como *p*-nitrofenil-beta-D-glucopiranosido (*p*NPG), o 3) BGLs con actividad similar sobre ambos tipos de sustratos (Eyzaguirre *et al*., 2005). La mayoría se encuentran en la última categoría.

A pesar de que se ha descrito actividad BGL en numerosas familias de GHs, las BGLs más estudiadas hasta ahora pertenecen a las familias GH1 y GH3, que son glicosil hidrolasas con un mecanismo de retención de la configuración.

Estructuralmente, las BGLs de la familia GH1 presentan un solo dominio catalítico en forma de barril TIM $(\alpha/\beta)_8$, donde se encuentra el centro activo. En él, están los dos aminoácidos catalíticos, que en esta familia son siempre dos ácidos glutámicos, muy conservados entre las diferentes proteínas caracterizadas (Jenkins *et al*., 1995).

En lo que respecta a la familia GH3, poseen un dominio principal, subdividido en tres dominios estructuralmente diferentes: un dominio N-terminal, que se pliega en forma de barril TIM $(\alpha/\beta)_8$, seguido del dominio C-terminal, que lo hace en α/β sándwich, y por último, un dominio fibronectina tipo III (Karkehabadi *et al*., 2014). La función del dominio FnIII es aún incierta, aunque algunos trabajos han descrito que podría dotar a las enzimas de una mayor termoestabilidad (Suzuki *et al*., 2013). Otros trabajos postulan que las enzimas de la familia GH3, al degradar biomasa lignocelulósica, se unen a la lignina de forma productiva (Gao *et al*., 2014) a través de este dominio FnIII (Lima *et al*., 2013), consiguiendo así no perder actividad hidrolítca a pesar de quedar atrapadas por el polímero. En cuanto a los aminoácidos catalíticos del centro activo, en esta familia son un ácido aspártico y un ácido glutámico, también muy conservados entre miembros GH3.

En general, la mayoría de las BGLs utilizadas para la degradación industrial de la celulosa pertenecen a la familia GH3. Son enzimas con eficacias catalíticas muy altas que, sin embargo, sufren inhibición por glucosa a concentraciones mucho más bajas que varios representantes de la familia GH1. Por lo tanto, uno de los principales objetivos para lograr una BGL eficiente y glucotolerante sería combinar la tolerancia a la glucosa de las GH1 con los valores de eficiencia catalítica de las GH3 (Cao *et al*., 2015). Diferentes técnicas de bioinformática e ingeniería genética pueden ser utilizadas para identificar y cambiar los aminoácidos esenciales implicados en cada funcionalidad, y así conseguir desarrollar BGLs diseñadas a la carta, que posean una alta eficiencia catalítica pero también tengan aumentada su glucotolerancia.

## 4. Hongos celulolíticos

Existen en la naturaleza numerosos organismos capaces de aprovechar con éxito los polisacáridos de la pared celular vegetal como fuente de carbono, secretando enzimas que pueden ser utilizadas en el desarrollo de diferentes aplicaciones biotecnológicas (Martínez *et al*., 2005). Actualmente, los hongos filamentosos son la principal fuente de enzimas de este tipo para uso comercial, debido a que producen una mezcla de enzimas degradadoras de celulosa y hemicelulosa que se caracterizan por su alta productividad y su gran eficacia catalítica respecto a las procedentes de otros organismos (Singh and Singh, 2014). Entre los hongos filamentosos, el *phylum* Ascomycota es considerado como el principal productor de enzimas involucradas en la degradación de la celulosa y la hemicelulosa. En particular, especies de los géneros *Trichoderma* y *Aspergillus* han sido estudiadas exhaustivamente para esta aplicación, y sus enzimas se usan a nivel industrial. Recientemente especies del género *Penicillium* han surgido como fuentes prometedoras de glicosil hidrolasas con gran potencial biotecnológico (Chávez *et al*., 2006).

La demanda mundial de celulasas está aumentando debido a sus múltiples aplicaciones. Actualmente, ya existen un buen número de empresas que se dedican a la producción de cócteles enzimáticos para la degradación de polisacáridos de biomasa vegetal. Entre ellas se encuentran algunas como Novozyme, Genencore, DSM, Dyadic, y las principales preparaciones enzimáticas de cada una de ellas están recogidas en la Tabla 2. Sin embargo, al surgir nuevas aplicaciones, como la producción de bioetanol 2G, en las que las enzimas deben trabajar en soluciones muy densas y habitualmente en presencia de inhibidores derivados de la lignina, se requiere la búsqueda de enzimas más robustas y eficaces para mejorar cada proceso.

### 4.1. Género *Trichoderma*

La mayoría de cepas de *Trichoderma* descritas hasta la fecha tiene probada capacidad celulolítica, como *T. reesei*, *T. lignorum*, *T. harzianum*, *T. longibrachiatrum*, *T. virens*, y *T. pseudokoningii*. Muchas de ellas han sido mejoradas genéticamente en los últimos 50 años (Payne *et al*., 2015). De todas ellas, el descubrimiento más relevante ha sido el del ascomiceto *T. reesei* (anamorfo de *Hypocrea jecorina*), hongo utilizado como modelo en la degradación de polisacáridos lignocelulósicos. El trabajo en la obtención de mutantes hiperproductores de celulasas de *T. reesei* ha sido constante desde la década de los 70 hasta la obtención de la cepa RUT-C30, que produce de 15 a 20 veces más celulasas que la cepa parental,

además de evitar la represión catabólica por glucosa (Peterson and Nevalainen, 2012).

Sin embargo, *T. reesei* posee pocos genes que codifiquen enzimas celulasas. De los 200 genes identificados que pertenecen a distintas familias de GHs, solo dos de ellos son CBHs, y siete son EGs (Foreman *et al*., 2003; Martinez *et al*., 2008). En cuanto a BGLs, a pesar de que se han descrito hasta 15 genes que las codifican en el genoma de *T. reesei*, la gran mayoría de ellos son intracelulares, con lo que generar mezclas enzimáticas ricas en BGLs a partir de este hongo es muy complicado. En cuanto a la detección de proteínas en los sobrenadantes de sus cultivos, la producción es aún más limitada: sólo se han detectado dos celobiohidrolasas (CBH I y CBH II), y dos endoglucanasas (EG1 y EG2), en una proporción aproximada de 60:20:10:10, mientras que las siete BGLs (BGLI-BGLVII) secretadas por este hongo constituyen menos del 1% del secretoma (Singhania *et al*., 2010; Gusakov, 2011). Debido a esto, las preparaciones del género *Trichoderma* deben ser suplementadas con BGLs procedentes de otros microorganismos (Gusakov, 2011).

### 4.2. Género *Aspergillus*

Las enzimas de los hongos de este género han sido objeto de numerosas aplicaciones biotecnológicas, entre ellas, varias relacionadas con el aprovechamiento de la celulosa (de Vries and Visser, 2001). Además, *Aspergillus* es un buen hospedador heterólogo de proteínas debido a su gran capacidad de secreción. Se han identificado una gran variedad de enzimas celulasas en el género *Aspergillus*, destacando los trabajos desarrollados durante décadas en especies como *A. niger* (Cairns *et al*., 2018), *A. nidulans* (Culleton *et al*., 2013), *A. terreus* (Ivanova *et al*., 1983) y *A. ficuum* (Hayashida *et al*., 1988).

Además, en algunas de las enzimas secretadas por *Aspergillus*, se han descrito actividades de transglicosilación (Becker *et al*., 1991), usando moléculas como celobiosa, celotriosa, metil-glucósido y metilglucósido como sustratos donadores (Yan *et al*., 1998).

Dentro de las principales familias de GHs detectadas en *Aspergillus*, las EGs se asignan principalmente a las familias 5 y 12, las CBHs a las familias 6 y 7, y la gran mayoría de las BGLs de éste género se han asignado a la familia GH3 (de Vries and Visser, 2001).

La gran ventaja que presenta el género *Aspergillus* con respecto a *Trichoderma*, es que es capaz de producir buenos niveles de actividad BGL. Generalmente, se utilizan cepas de *A. niger* para producir cócteles enzimáticos ricos en actividad BGL, que son usados para complementar los generados a partir de *T. reesei* (Sukumaran *et al*., 2005).

## 4.3. Géneros *Penicillium* y *Talaromyces*

Además de *Aspergillus*, algunas cepas de *Penicillium* han sido estudiadas por su gran capacidad para producir BGLs (Vaishnav *et al*., 2018) Industrialmente, las cepas de este género han sido ampliamente utilizadas en la producción de antibióticos, aunque en trabajos más recientes se ha demostrado que también son una gran fuente de celulasas. Los estudios que se han realizado con los hongos de esta familia han demostrado que tienen los ratios más equilibrados de producción de BGLs, EGs y CBHs, con lo que representan la especie capaz de degradar la celulosa más eficientemente (Maeda *et al*., 2011; Martins *et al*., 2008). Además, las enzimas lignocelulósicas generadas por el género *Penicillium* parecen tener menos tendencia a ser atrapadas por la lignina, y ser menos susceptibles a la inhibición por compuestos derivados de ella (Berlin *et al*., 2006). Se han descrito numerosas especies de *Penicillium* que tienen capacidad para la biodegradación de celulosa, entre las que se destacan: *P. brasilianum* (Jørgensen *et al*., 2003), *P. chrysogenum* (Hou *et al*., 2007), *P. funiculosum* (de Albuquerque de Carvalho *et al*., 2014), *P. purpurogenum* (Suto *et al*., 1991) y *P. verruculosum* (Solov'eva *et al*., 2005).

Una de las especies más interesantes de este género es *P. oxalicum*, anteriormente conocido como *P. decumbens* (Liu *et al*., 2013), que ya es utilizado por la industria china para la producción a gran escala de celulasas, y se considera que supera a *T. reesei* en la eficiencia de sus preparaciones enzimáticas.

**Tabla 2.** Principales preparaciones enzimáticas de celulasas comerciales.

| Preparación | Empresa | Fuente |
|---|---|---|
| N-50010 | Novozymes (Dinamarca) | *A. niger* |
| N-188 | Novozymes | *A. niger* |
| Celluclast 1.5L | Novozymes | *T. reesei* |
| Celtec2 | Novozymes | *T. reesei* |
| Celtec3 | Novozymes | *T. reesei* |
| Biocellulase A | Ques Intl. (EE.UU) | *A. niger* |
| Biocellulose Tri | Ques Intl. | *A. niger* |
| GC 220 | Genecor (EE.UU) | *T. longibrachiatum/T. reesei* |
| GC 440 | Genecor | *T. longibrachiatum/T. reesei* |
| GC 880 | Genecor | *T. longibrachiatum/T. reesei* |
| Accelerase 1500 | Genecor | *T. reesei* |
| Rovabio | Adisseo (Fracia) | *P. funiculosum* |
| Cellulase AP30K | Amano enzyme (Japón) | *A. niger* |
| Cellulase TAP106 | Amano enzyme | *T. viride* |

Recientemente, diferentes especies asociadas tradicionalmente a *Penicillium*, han sido agrupadas en un género taxonómicamente muy cercano, denominado *Talaromyces* (Samson *et al*., 2011), y su capacidad de degradación de la lignocelulosa ha sido analizada. Diversos miembros del género como *Talaromyces stipitatus* (Vafiadi *et al*., 2006), *Talaromyces cellulolyticus* (Fujii *et al*., 2014) y *Talaromyces verruculosus* (Goyari *et al*., 2015) han mostrado un gran potencial biotecnológico para la producción de enzimas celulasas y hemicelulasas. En la presente tesis doctoral se estudiará en detalle el potencial celulolítico del hongo *Talaromyces amestolkiae*. La cepa de *T. amestolkiae* utilizada fue aislada a partir de residuos de cereales por el grupo de la Dra. Covadonga Vázquez (Departamento de Microbiología, Facultad de Biología, Universidad Complutense de Madrid), y fue identificada en nuestro grupo de investigación e incluida en la colección del Instituto Jaime Ferrán de Microbiología (IJFM) del Centro de Investigaciones Biológicas, con el número A795. Las primeras investigaciones mostraron que el hongo era capaz de producir muy altos niveles de BGLs, lo que indicaba que podría interesante a nivel industrial (Gil Muñoz, 2015).

En el cóctel enzimático de este hongo también se han caracterizado dos EGs (de Eugenio *et al*., 2018), una β-xilosidasa (Nieto-Domínguez *et al*., 2015), y una endoxilanasa (Nieto-Domínguez *et al*., 2017). Todas ellas, tras ser purificadas y caracterizadas, demostraron ser enzimas altamente eficientes en comparación a otras enzimas fúngicas caracterizadas previamente. En la actualidad se continúa el estudio de las enzimas de este hongo implicadas en la degradación de los polisacáridos de la pared celular vegetal, mediante la expresión heteróloga de las enzimas previamente caracterizadas y la búsqueda de nuevas enzimas encontradas en su genoma y su secretoma con el fin de valorizar la biomasa vegetal, especialmente la que procede de los residuos lignocelulósicos.

Además, en cuanto a procesos de síntesis, tanto las BGLs como la β-xilosidasa han probado ser capaces de catalizar reacciones de transglicosilación, lo cual diversifica sus potenciales aplicaciones biotecnológicas (Gil Muñoz, 2015; Nieto-Domínguez, 2017).

## 5. Genómica y proteómica funcional: herramientas para el estudio de nuevas enzimas fúngicas

La genómica y proteómica funcional es un campo de la biología molecular que intenta describir y asignar funciones a los genes, tránscritos de ARN, y proteínas identificados a partir de los datos proporcionados por estudios genómicos y transcriptómicos (Bader *et al*., 2003; de Oliveira and de Graaff, 2011).

Los hongos forman un gran reino eucariótico de más de un millón y medio de especies, siendo la mayoría de ellos hongos filamentosos. Debido a esta enorme diversidad, han sido capaces de extenderse y colonizar diversos recursos naturales, existiendo en numerosos hábitats en los que viven de forma saprófita o parasitaria, degradando diferentes fuentes de material orgánico (Ronald de Vries *et al*., 2018). Son probablemente el grupo de microoorganismos más útil biotecnológicamente, y se utilizan para sintetizar una amplia gama de compuestos de importancia económica, como enzimas, metabolitos secundarios, y otros productos farmacéuticos (Hawksworth, 1991).

La disponibilidad de las tecnologías genómicas ha tenido un gran impacto en el campo de la investigación fúngica. El primer genoma fúngico que se publicó fue el de la levadura *Saccharomyces cerevisiae* (Goffeau *et al*., 1996) y años más tarde se anotó el del hongo ascomiceto filamentoso *Neurospora crassa* (Galagan *et al*., 2003). Un año después se secuenció el primer genoma de un basidiomiceto, *Phanerochaete chrysosporium* (Martinez *et al*., 2004), un hongo de la podredumbre blanca, capaz de despolimerizar de manera eficiente la lignina. En los últimos años, gracias a los nuevos y potentes programas de secuenciación masiva, se han obtenido más de 1000 genomas de hongos. La diversidad genómica en este reino se encuentra a todos los niveles, incluso dentro del mismo género, como se ha descrito recientemente entre diferentes especies de *Aspergillus* (de Vries *et al*., 2017).

El estudio genómico y proteómico en profundidad de hongos filamentosos ha puesto de manifiesto la presencia de genes que pueden tener interesantes aplicaciones industriales. Aplicado a la biotecnología para la degradación de biomasa lignocelulósica, estos genomas suelen analizarse enfrentando los datos a la base de datos CAZY, lo que es una manera sencilla de identificar el número de estas enzimas presentes en un genoma fúngico. El análisis de los datos suele revelar perfiles específicos de CAZymes para cada especie. Los distintos perfiles de CAZymes entre los hongos que degradan la pared celular de las plantas se han relacionado con los variados modos nutricionales de cada especie, y suelen implicar diferentes estrategias para atacar la lignocelulosa a nivel enzimático (Zhao *et al*., 2013).

Aunque estudios recientes han puesto de manifiesto que, como regla general, los hongos patógenos de plantas contienen más CAZymes que los patógenos de animales (Zhao *et al*., 2013), aún no se han realizado suficientes investigaciones como para tener un análisis comparativo completo y sistemático de las CAZymes presentes en todo el reino Fungi, ni de la eficacia de cada una de ellas. Además, en cuanto a análisis proteómicos, aún se conoce poco acerca de la expresión de los diferentes genes que codifican CAZYmes en presencia de sustratos complejos (Doyle,

2011). Por tanto, este es un campo aún en rápida expansión, y un mayor conocimiento puede permitir la optimización en la producción de enzimas con interés industrial.

# OBJETIVOS

# OBJETIVOS

Teniendo en cuenta lo anteriormente expuesto, el objetivo del presente trabajo fue profundizar en el estudio de las β-glucosidasas producidas por el hongo *T. amestolkiae*. El empleo de estas enzimas podría aumentar el rendimiento de sacarificación residuos lignocelulósicos destinados a la elaboración de etanol de segunda generación, contribuyendo a su revalorización. Además, mediante reacciones de transglicosilación, también catalizadas por estas β-glucosidasas, podrían mejorarse las propiedades de moléculas con interesantes propiedades, o encontrar nuevos glucósidos con interés biotecnológico.

Para conseguir estos objetivos, se planearon las siguientes tareas:

**1-** Análisis del genoma de *T. amestolkiae* y de su secretoma, con especial interés en las β-glucosidasas producidas por este hongo.

**2-** Expresión heteróloga de las principales β-glucosidasas de *T. amestolkiae* en *P. pastoris*. Purificación, caracterización y comparación con las enzimas nativas del hongo.

**3-** Estudio de cócteles comerciales suplementados con β-glucosidasas de *T. amestolkiae* para la sacarificación de residuos lignocelulósicos.

**4-** Obtención de glucósidos de interés mediante reacciones de transglicosilación catalizadas por las β-glucosidasas de *T. amestolkiae*.

**5-** Obtención de variantes de estas enzimas para mejorar los rendimientos en la síntesis de glicósidos de interés.

# CHAPTER 1

**Differential β-glucosidase expression as a function of carbon source availability in *Talaromyces amestolkiae*: a genomic and proteomic approach**

De Eugenio, L.I., **Méndez-Líter, J.A.**, Nieto-Domínguez, M., Alonso, L., Gil-Muñoz, J., Barriuso, J., Prieto, A., Martínez, M.J., 2017. Differential β-glucosidase expression as a function of carbon source availability in *Talaromyces amestolkiae*: a genomic and proteomic approach. Biotechnol. Biofuels. 10, 161-161. doi: 10.1186/s13068-017-0844-7

Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

## ABSTRACT

**Background.** Genomic and proteomic analysis are potent tools for metabolic characterization of microorganisms. Although cellulose usually triggers cellulase production in cellulolytic fungi, the secretion of the different enzymes involved in polymer conversion is subjected to different factors, depending on growth conditions. These enzymes are key factors in biomass exploitation for second generation bioethanol production. Although highly effective commercial cocktails are available, they are usually deficient for β-glucosidase activity, and genera like *Penicillium* and *Talaromyces* are being explored for its production.

**Results.** This article presents the description of *Talaromyces amestolkiae* as a cellulase-producer fungus that secretes high levels of β-glucosidase. β-1,4-endoglucanase, exoglucanase and β-glucosidase activities were quantified in the presence of different carbon sources. Although the two first activities were only induced with cellulosic substrates, β-glucosidase levels were similar in all carbon sources tested. Sequencing and analysis of the genome of this fungus revealed multiple genes encoding β-glucosidases. Extracellular proteome analysis showed different induction patterns. In all conditions assayed, glycosyl hydrolases were the most abundant proteins in the supernatants, albeit the ratio of the diverse enzymes from this family depended on the carbon source. At least two different β-glucosidases have been identified in this work: one is induced by cellulose and the other one is carbon source-independent. The crudes induced by Avicel and glucose were independently used as supplements for saccharification of slurry from acid-catalyzed steam-exploded wheat straw, obtaining the highest yields of fermentable glucose using crudes induced by cellulose.

**Conclusions.** The genome of *T. amestolkiae* contains several genes encoding β-glucosidases and the fungus secretes high levels of this activity, regardless of the carbon source availability, although its production is repressed by glucose. Two main different β-glucosidases have been identified from proteomic shotgun analysis. One of them is produced with different carbon sources, while the other is induced in cellulosic substrates and is a good supplement to Celluclast in saccharification of pretreated wheat straw


**Keywords:** Fungi; Ascomycete; Glycosidase; Lignocellulosic biomass; Saccharification.

# BACKGROUND

Second generation bioethanol represents an efficient alternative to conventional energy supply, involving the exploitation of renewable sources, usually disposable wastes from other industrial or agricultural activities. In this process, polysaccharides from lignocellulosic biomass are hydrolyzed to glucose and xylose units that are further converted to ethanol (Naik *et al*., 2010). When raw materials with hardly available polysaccharides are used, a pretreatment is needed in order to facilitate its accessibility. Steam-explosion is maybe the most commonly applied method to perform this pretreatment, generating a biomass slurry that usually contains by-products which adversely affect downstream steps like enzymatic hydrolysis or ethanol fermentation (Jurado *et al*., 2009; Alvira *et al*., 2010). Although all stages are relevant for the process, the improvement of the composition and dosage of the hydrolytic enzymatic cocktails remains a subject of great interest to industry. An efficient hydrolysis of carbohydrates, especially cellulose, the major polysaccharide from plant biomass, is indispensable to recover fermentable sugars from feedstock. Cellulose is a homopolysaccharide of β-1,4-linked D-glucose residues whose enzymatic conversion into free glucose requires the coordinated work of the cellulase system, consisting of three different kinds of hydrolases: i) cellobiohydrolases (CBH) named also as exoglucanases, which cleave cellobiose units from the chain ends, ii) β-1,4-endoglucanases (EG), which hydrolyze the polymers internally, reducing the degree of polymerization and iii) β-glucosidases (BGL), which convert cellobiose and soluble cellodextrins into glucose (Zhang and Lynd, 2004).

The main industrial producers of cellulases are *Trichoderma reesei* and *Aspergillus niger* (Duff and Murray, 1996; Martinez *et al*., 2008), but the enzymatic crudes released by these microorganisms, in particular those from *Trichoderma* sp., are deficient in β-glucosidase (Maeda *et al*., 2011), and the saccharification cocktails must be supplemented with this activity to increase the efficiency of cellulose degradation.

The current approach for production of lignocellulosic bioethanol relies on combining different cellulase cocktails, which is very effective but involves a significant increase of the production costs. Then, many efforts are being devoted to discover microorganisms that secrete high amounts of glycosidases and to search for novel robust enzymes, which efficiently transform cell wall polysaccharides, even in the presence of the undesirable by-products found in the process streams. In this sense, *Penicillium* sp. and its perfect states (*Talaromyces* or *Eupenicillium*) have attracted much attention in the last years for their high cellulase activity (Bhiri *et al*., 2008; Szijarto *et al*., 2004). In addition, these fungi could have special interest for cellulose and hemicelluloses transformation in the context of bioethanol

production or other industrial applications, since they also produce high levels of xylanases (Chávez *et al*., 2006; Adsul *et al*., 2007; Ustinov *et al*., 2008).

This work reports the characterization of *T. amestolkiae* as cellulase producer, and the sequencing and assembly of its genome. The cellulolytic activities secreted upon fungal growth in several carbon sources were quantified, and the pool of extracellular proteins in each one of the secretomes was analyzed by massive peptide analysis. The data obtained were comprehensively related to the number of genes codifying for members of different glycosyl hydrolases families (GH) in *T. amestolkiae* genome, showing that β-glucosidases are highly represented from DNA to proteins secreted in this fungus. The role of these enzymes in saccharification of wheat straw acid slurry is discussed.

## RESULTS AND DISCUSSION

### Cellulase production by *T. amestolkiae* CIB

Different carbon sources, namely Avicel (microcrystalline cellulose), beechwood xylan, and wheat straw slurry, were tested as inducers of the cellulase activities. In addition, a culture with glucose, which has been described as cellulases' inhibitor (Hanif *et al*., 2004; Chen, 2014), was grown under the same conditions. According to the results shown in Figure 1, Avicel seemed to be the best substrate for cellulase production. Cellobiohydrolases and β-1,4-endoglucanase activities (Figs. 1A and 1B) were mainly detected in media containing Avicel or acid wheat straw slurry. Nevertheless, as this last substrate is a mixture of lignocellulose components, the production of cellulolytic enzymes was slightly lower. β-1,4-endoglucanase activity was strongly induced by Avicel (over 10 U/mL after 8 days), being also measurable in cultures containing slurry. BGL activity reached similar values regardless of the carbon source used (1.4-1.8 U/mL), even with glucose (figure 1C). It is well known that, in general, Avicelase and β-1,4-endoglucanase production is repressed in the presence of glucose (Amore *et al*., 2013), but little is known about BGL production in the presence of easily available carbon sources.

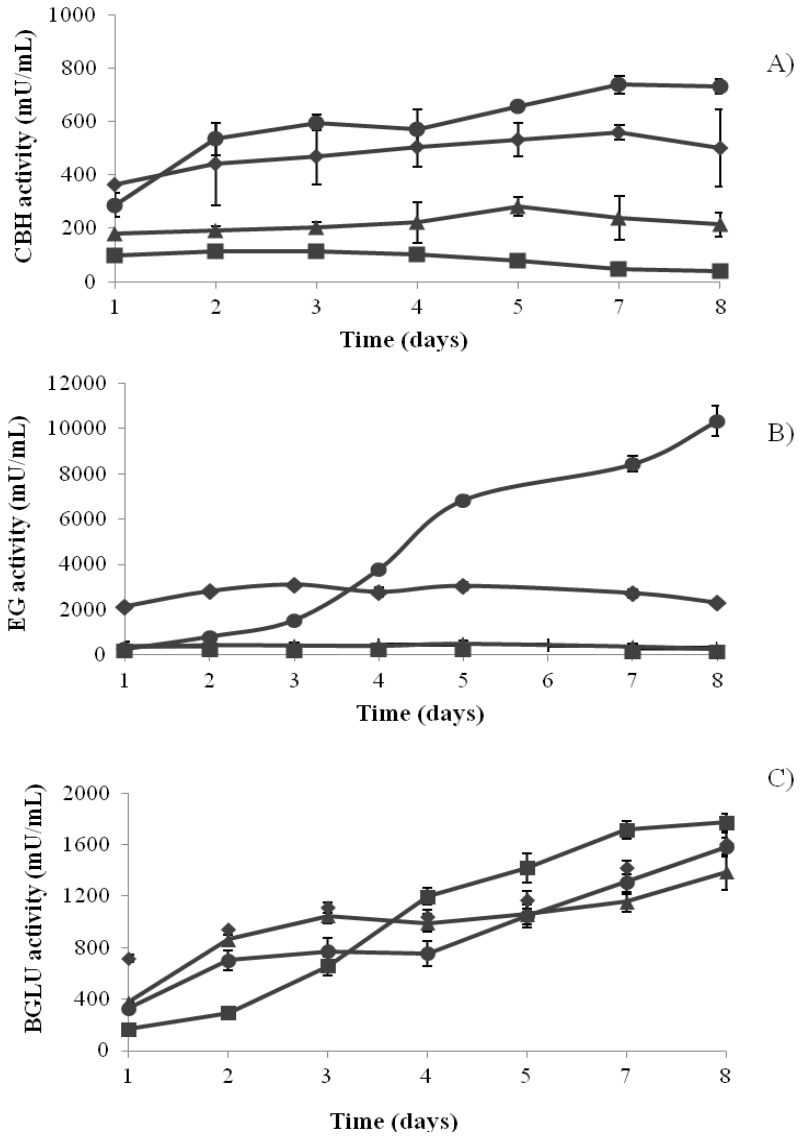



**Figure 1.** Extracellular cellulase production by *Talaromyces amestolkiae* grown on Mandels medium with different carbon sources. *Circles* Avicel (1%); *squares* glucose (1%); *triangles* xylan (2%); *diamonds* slurry (1%). Samples were taken each 24 h, and the different enzymatic activities were quantified in the supernatants: **A)** exocellulase, **B)** endo β-(1,4) glucanase and **C)** β-glucosidase

Non-inducible BGL activity has been described in some fungal species (Shewale, 1982; Lee *et al*., 1996) but, in most cases, the secretion of this enzyme is induced by cellulose (Shewale, 1982). In addition, the regulation and expression of cellulolytic enzymes can be affected by the presence of inducers, derived from cellulose metabolism, as oligosaccharides or transglycosylation products, or experience catabolic repression by glucose, as reported in *T. reesei* (Ilmen *et al*., 1997). Then, the effect of small molecules as cellobiose, glycerol, fructose and xylose on BGL production was also assayed (figure 2A). In virtually all cases, the released activity was very similar (around 1.4 U/mL) and comparable to

that obtained before for polymeric cellulosic substrates and glucose, getting lower values for cellobiose, the only disaccharide evaluated (1 U/mL), which needs to be transformed in glucose before being consumed by the fungus. These data demonstrate that this *T. amestolkiae* strain does not require specific inducers for BGL production.

Monitoring glucose consumption in cultures containing this monosaccharide as carbon source revealed the complete depletion of the sugar after 24 h (figure 2B), which may indicate that, in this case, carbon starvation is triggering BGL production. Daily addition of 0.5 or 1% glucose to cultures in Mandels medium containing either 1% glucose or Avicel (figure 2C) caused only basal production of BGL, confirming that the production of this enzyme is repressed by glucose. Thus, when easily assimilable carbon sources are exhausted, carbon starvation might induce BGL production in *T. amestolkiae,* while the production of CBH and EG is induced only when cellulosic substrates are present.

However, these data raised the question of whether the BGLU secreted under carbon starvation and in the presence of cellulose are the same enzyme or not. Preliminary results from zimograms after isoelectric focusing suggested that they could be different proteins (not shown). The production of basal levels of cellulases, usually Avicelase and β-1,4-endoglucanase, has been described and interpreted as a fungal strategy intended to take advantage of any cellulosic material present in the surrounding medium. On the contrary, the production of BGL in a carbon source-independent way could be due to its role in the regulation of cellulase production. In this sense, BGL from *T. reesei*, can convert cellobiose into sophorose, a strong cellulase inducer (Fowler and Brown Jr, 1992).

In this work, the analysis of cellulase diversity in *T. amestolkiae* and its differential production in specific conditions was tackled according to several strategies, as reported above.

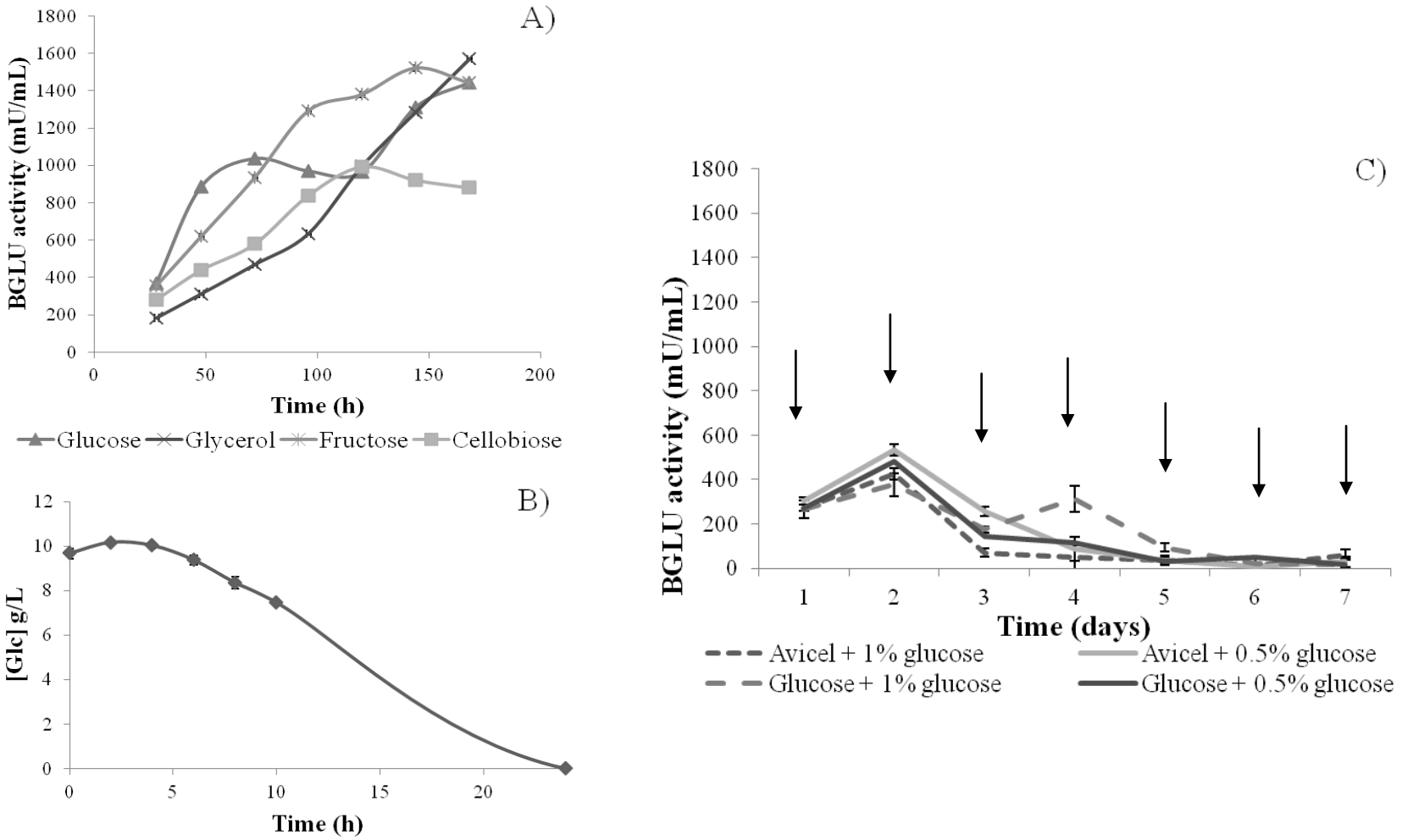



**Figure 2.** Effect of the carbon source and glucose addition on BGL production. **A)** BGL production in different soluble carbon sources. The fungus was grown in Mandels with 1% glucose (*triangles*), glycerol (*crosses*), fructose (*asterisks*), or cellobiose (*squares*) for 7 days. Samples were taken daily and BGL activity was measured in the supernatants. **B)** Glucose consumption measured in Mandels + 1% glucose cultures. Samples were taken daily and glucose concentration of supernatants was measured by HPLC as described in "Methods". **C)** Effect of glucose addition on BGL production. To study the influence of glucose, 0.5 or 1% glucose pulse feeds were daily added to 1% glucose or Avicel cultures. Arrows indicate the addition of the monosaccharide. Samples were taken daily and BGL activity measured in the supernatants

## General features of *T. amestolkiae* CIB genome

In a first approach, the cellulolytic potential of the fungus was analyzed through the sequencing and annotation of its genome. The draft genome sequence of *T. amestolkiae* CIB was based on high-throughput sequencing system (~65-fold coverage) and *de novo* assembling. The 24,330,860 pair-ended reads were incorporated into 215 scaffolds, among which 132 comprised more than 1 kb. Genome size was determined to be 33.7 Mb and N50 and L50 statistics displayed values of 1,486,010 and 9 respectively. The assembled genome resulted in the prediction of 10,408 ORFs (open reading frames) (Additional file 2). As will be seen below, 342 of these ORFs were supported by the secretome data. The prediction of GH enzymes was performed by submitting the putative ORFs to the dbCAN server and filtering the results with the selected cutoff. By this way 325 glycosyl hydrolases were identified, a number close to those described for *Fusarium verticillioides* (332), *Aspergillus flavus* (334) and *Aspergillus oryzae* (317) (Zhao *et al*., 2013).

## β-glucosidases are highly represented in the genome of *T. amestolkiae*

Glycosyl hydrolases involved in hydrolysis of lignocellulose are widely extended across the CAZy families (Zhao *et al*., 2013). Figure 3 depicts the profile of the main families of CAZymes implicated in plant cell wall-degradation encoded by the genomes of *T. amestolkiae* and six microorganisms of remarkable lignocellulolytic interest. The selection of these species was made attending to the criteria of relevance and availability of accurately annotated genomic data. *T. reesei* is the main industrial source of the cellulases and hemicellulases added to commercial cocktails for lignocellulose saccharification (Martinez *et al*., 2008; Häkkinen *et al*., 2012). Other fungi, as *Aspergillus niger* and *Penicillium oxalicum* (previously classified as *Penicillum decumbens*) are also industrial producers of cellulolytic enzymes (Pel *et al*., 2007; Liu *et al*., 2013) . *Thielavia terrestris* secretes thermostable hydrolases (Berka *et al*., 2011), and *Clostridium thermocellum* is a well-known Firmicutes capable of directly convert cellulose into ethanol and other value-added products (Stevenson and Weimer, 2005; Roberts *et al*., 2010).

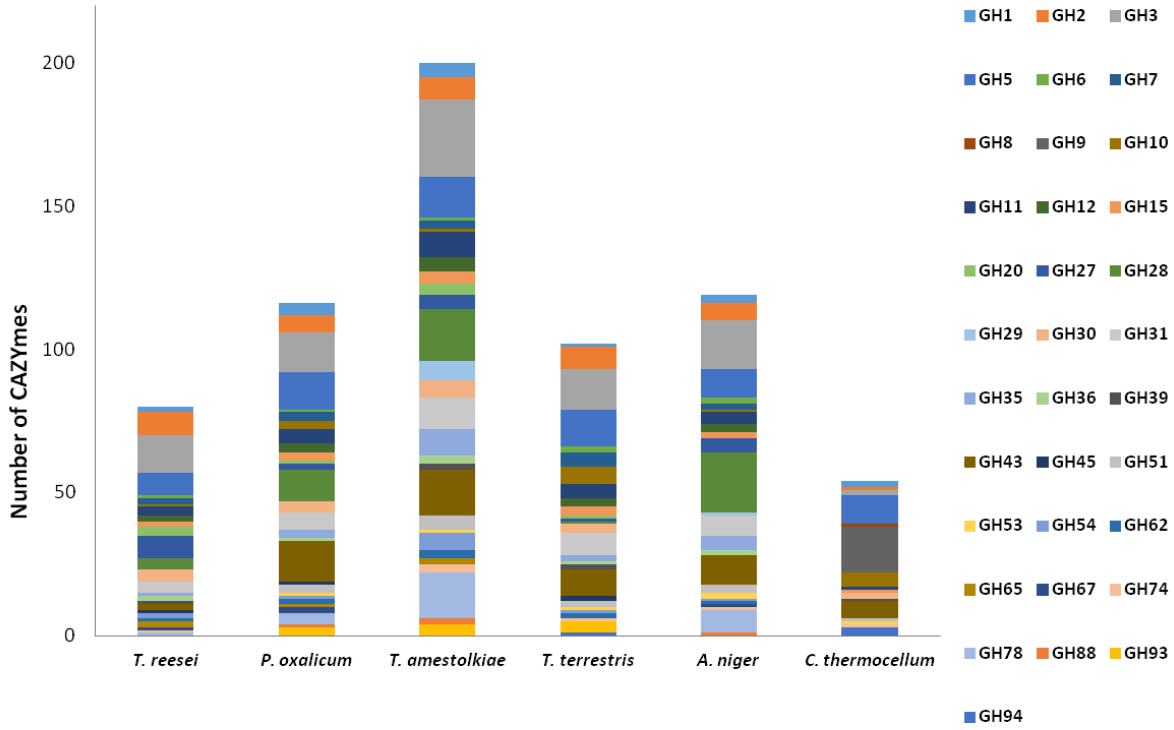



**Figure 3.** Comparative analysis of the number of CAZyme proteins predicted from genomes of different cellulolytic organisms. Data from GH families that include cell wall-degrading enzymes have been selected for this representation. From left to right, putative GHs from the genomes of *T. reesei*, *P. oxalicum* 114-2, *T. amestolkiae* CIB, *T. terrestris* NRRL 8126, *A. niger* CBS 513.88 and *C. thermocellum* ATCC 27405.

Data analysis revealed specific CAZyme profiles for each species although, as expected, the bacterium displayed the biggest differences. Distinctive CAZyme profiles among plant cell wall-degrading fungi have been linked to nutritional mode adaptations and may imply different strategies for attacking lignocellulose at the enzymatic level (Martinez *et al*., 2009; Zhao *et al*., 2013; Znameroski and Glass, 2013). Regarding *T. amestolkiae*, the number of encoded CAZymes was significantly higher than in the other organisms. In particular, the high amount of genes encoding enzymes from the GH3 family, which comprises many of the reported fungal β-glucosidases, was intriguing and it could be related to the findings of BGL production reported above. The presence of this activity in the culture supernatants of this organism, regardless of the carbon source, may suggest a central role of these enzymes in the metabolic strategy adopted by *T. amestolkiae* along evolution.

To assess this possibility, we carried out a deeper comparison focusing on GH1, GH3, GH5 and GH30, the main GH families encoding fungal β-glucosidases (figure 4A). The total number of genes for these enzymes in the genome of *T. amestolkiae* was annotated by running

BLASTP against the characterized GHs from the CAZy database, and compared with those reported for each of the cellulolytic species used as reference (figure 4B).

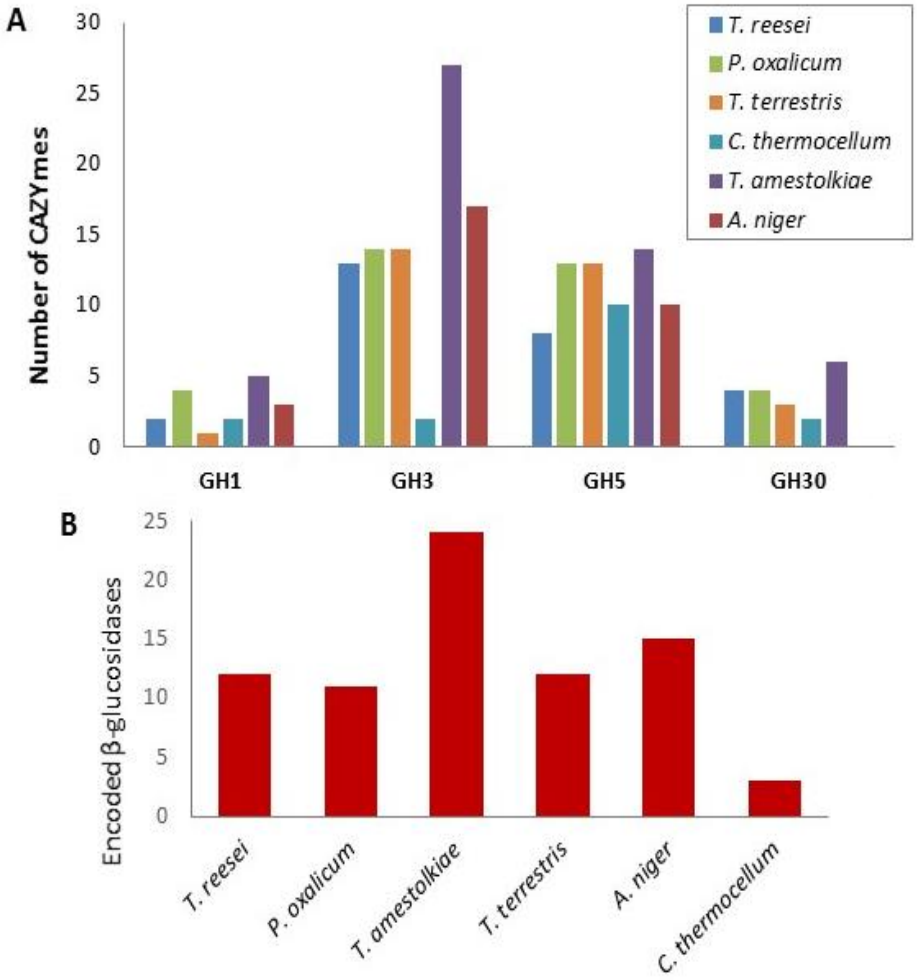



**Figure 4.** Comparison of the BGLs predicted for *T. amestolkiae* and the referenced cellulolytic species. **A)** Number of CAZyme proteins from the main families encoding BGL (GH1, GH3, GH5 and GH30). These proteins were manually annotated by running a BLASTP against the GHs from CAZy database, and those similar to previously characterized BGL were selected, and represented in **B)**.

According to our BLASTP analysis, 5 BGLs from GH1 family, 18 from GH3 and 1 from GH5 are codified in the genome of this isolate (Table 1). The CAZymes types and abundance represented in a given fungal genome may condition a certain strategy for hydrolyzing lignocellulose (Zhao *et al*., 2013).

**Table 1.** Putative BGLs identified from analysis of *T. amestolkiae* genome. The presence of signal peptide was analyzed using the Phobius software.

| ID | CAZY Family | MM (kDa) | Signal peptide |
|---|---|---|---|
| g8961 | GH1 | 53.9 | No |
| g5848 | GH1 | 56.1 | No |
| g7413 | GH1 | 55.5 | No |
| g4650 | GH1 | 55.5 | No |
| g8384 | GH1 | 243.6 | Yes |
| g2731 | GH3 | 66.7 | No |
| g1119 | GH3 | 230.5 | No |
| g377 | GH3 | 88.7 | Yes |
| g3618 | GH3 | 84.3 | Yes |
| g3821 | GH3 | 78.9 | Yes |
| g7527 | GH3 | 92.2 | No |
| g7273 | GH3 | 92.9 | No |
| g3126 | GH3 | 86.7 | No |
| g3139 | GH3 | 93.6 | Yes |
| g6753 | GH3 | 81.8 | Yes |
| g4056 | GH3 | 97.4 | No |
| g6367 | GH3 | 89.7 | No |
| g7815 | GH3 | 84.3 | Yes |
| g9082 | GH3 | 181.2 | No |
| g9183 | GH3 | 87.6 | No |
| g9861 | GH3 | 150.1 | No |
| g6857 | GH3 | 109.0 | Yes |
| g9398 | GH3 | 81.4 | Yes |
| g9150 | GH3 | 86.5 | Yes |

The high number of β-glucosidases annotated in *T. amestolkiae* pointed them as key factors during lignocellulose degradation by this fungus. This is especially relevant considering that commercial mixtures based on *T. reesei* crudes (the main industrial source of this type of enzymes) are deficient in β-glucosidase activity (Kumar *et al*., 2008). In this sense, this unexpectedly high number of cellulases and hemicellulases in *T. amestolkiae* opens the possibility of improving lignocellulose saccharification through the synergistic addition of several CAZymes.
The expression of β-glucosidases by *T. amestolkiae* in the presence of different lignocellulosic substrates was validated, beyond the genomic level, by proteomic analysis of the extracellular proteins secreted by this fungus.

## Differential shotgun analysis of proteins secreted by *T. amestolkiae*

The enzymatic pools released under the four assayed conditions were compared by sequential tryptic digestion and LC-MS/MS of the whole

peptide mixtures produced (additional file 3). The number of proteins identified in each secretome differed according to the complexity of the carbon source used for fungal growth. The maximum number of identifications came from samples produced in the two heterogenic substrates, xylan (184) and slurry (144), while only 119 and 104 proteins were recognized in the secretomes from glucose and Avicel cultures, respectively. Xylan and slurry are structurally complex materials that contain heterogeneous and branched polymers. Hence, it is expectable that the amount and variety of extracellular enzymes required to accomplish their degradation is higher than those needed to metabolize simpler substrates. Table 2 lists the ten extracellular proteins identified with maximal confidence (with the highest scores) in each condition. Mainly CBH, followed by xylanases and BGLs were produced when the fungus grew in the medium with Avicel, and the same distribution was observed in cultures with wheat straw slurry. The presence of xylanases in the secretome of *P. chrysosporium* grown in cellulose has already been reported (Hori *et al*., 2011). β-xylosidases were detected only in xylan cultures, as previously described (Nieto-Domínguez *et al*., 2015). Surprisingly, a glutaminase was expressed in all conditions. This protein might be released upon induction of protein catabolism by early carbon starvation (van Munster *et al*., 2014) and it was produced in *Thermomyces lanuginosus* SSBP cultures in corncob as has been previously described (Winger *et al*., 2014). Other extracellular proteins with different functionalities were also detected in glucose cultures. Some of them are probably involved in the metabolism of fungal carbohydrates, as an hexosaminidase (Gooday *et al*., 1992). For example, cellulase regulation in *H. jecorina* proceeds through deglycosylation catalyzed by glucosaminidases, which can modulate enzyme activity by this mechanism (Stals *et al*., 2012). Under starvation stress, the increased production of other enzymes involved in fungal cell wall degradation, as β-1,3-glucanases, has also been observed. The heterogeneity of the proteins identified in the extracellular pool of proteins from glucose cultures is in agreement with the high variability detected by KOG functional analysis (see below). The PSM values of two independent experiments were taken as low precision, semi-quantitative records for further evaluation of the shotgun analysis. These values account for the number of scans identified for each protein, which are roughly related to the amount of a given protein in the sample (Roberts *et al*., 2015), being analyzed from different perspectives: functional analysis (KOG attribution), GHs diversity, and relative abundance of cellulases and hemicellulases.

**Table 2.** Summary of the ten extracellular proteins identified with maximal scores from the shotgun analysis. UP: unique peptides.

| | ID | Predicted protein function | GH Family | Score | UP | MM [kDa] | pI |
|---|---|---|---|---|---|---|---|
| **Avicel** | g2234 | cellobiohydrolase | GH7 | 49,485.86 | 14 | 55.8 | 4.81 |
| | g5707 | endoglucanase | GH6 | 15,441.61 | 8 | 48.3 | 4.79 |
| | g9427 | endoxylanase | GH10 | 4,256.38 | 12 | 43.5 | 5.25 |
| | g3995 | glutaminase | | 3,534.39 | 10 | 76.4 | 4.53 |
| | g4058 | swollenin | | 3,179.49 | 4 | 41.6 | 4.54 |
| | g3821 | beta-glucosidase | GH3 | 2,764.60 | 14 | 78.9 | 5.22 |
| | g2018 | endoglucanase | GH74 | 2,689.52 | 10 | 78.2 | 4.94 |
| | g1296 | endoxylanase | GH11 | 2,684.90 | 3 | 22.8 | 5.39 |
| | g2158 | glucoamylase | GH15 | 2,148.19 | 9 | 65.2 | 4.56 |
| **Glucose** | g377 | beta-glucosidase | GH3 | 20,095.77 | 24 | 88.7 | 4.91 |
| | g2140 | glucoamylase | GH15 | 18,940.98 | 18 | 67.7 | 4.92 |
| | g3995 | glutaminase | | 10,898.10 | 14 | 76.4 | 4.53 |
| | g2158 | glucoamylase | GH15 | 9,928.19 | 12 | 65.2 | 4.56 |
| | g9324 | exo-beta-1.3-glucanase | GH55 | 8,297.36 | 12 | 84.3 | 4.79 |
| | g8295 | alpha-glucosidase | GH31 | 7,988.51 | 19 | 98.6 | 4.77 |
| | g3279 | alpha-amylase | GH13 | 6,080.68 | 11 | 66.2 | 4.51 |
| | g1839 | endo-1.3(4)-beta-glucanase | GH16 | 6,050.23 | 5 | 30.4 | 4.82 |
| | g2234 | cellobiohydrolase; | GH7 | 5,462.33 | 9 | 55.8 | 4.81 |
| | g40761 | hexosaminidase | GH20 | 4,795.60 | 18 | 67.9 | 5.19 |
| **Slurry** | g2234 | cellobiohydrolase | GH7 | 37,368.64 | 14 | 55.8 | 4.81 |
| | g9427 | endoxylanase | GH10 | 7,497.84 | 15 | 43.5 | 5.25 |
| | g6537 | endoglucanase | GH5 | 5,378.70 | 11 | 62.0 | 4.81 |
| | g3995 | glutaminase | | 5,355.00 | 13 | 76.4 | 4.53 |
| | g5707 | endoglucanase | GH6 | 5,154.87 | 5 | 48.3 | 4.79 |
| | g2140 | glucoamylase | GH15 | 5,064.97 | 13 | 67.7 | 4.92 |
| | g377 | beta-glucosidase | GH3 | 4,539.30 | 19 | 88.7 | 4.91 |
| | g8295 | alpha-glucosidase | GH31 | 3,690.24 | 18 | 98.6 | 4.77 |
| | g5915 | hypothetical protein | GH55 | 3,061.95 | 13 | 68.8 | 4.61 |
| | g3707 | arabinofuranosidase | GH62 | 2,450.28 | 6 | 41.0 | 4.81 |
| **Xylan** | g2234 | cellobiohydrolase | GH7 | 49,601.87 | 14 | 55.8 | 4.81 |
| | g2140 | glucoamylase | GH15 | 37,213.51 | 18 | 67.7 | 4.92 |
| | g377 | beta-glucosidase | GH3 | 31,858.88 | 25 | 88.7 | 4.91 |
| | g3995 | glutaminase | | 26,983.15 | 14 | 76.4 | 4.53 |
| | g2158 | glucoamylase | GH15 | 22,043.67 | 12 | 65.2 | 4.56 |
| | g8295 | alpha-glucosidase | GH31 | 17,081.31 | 19 | 98.6 | 4.77 |
| | g9427 | xylanase | GH10 | 15,562.62 | 15 | 43.5 | 5.25 |
| | g9324 | exo-beta-1.3-glucanase | GH55 | 11,780.55 | 12 | 84.3 | 4.79 |
| | g8981 | beta-D-xylosidase | GH3 | 9,394.38 | 23 | 83.3 | 4.84 |
| | g4068 | arabinofuranosidase | GH51 | 9,345.18 | 15 | 60.3 | 5.07 |

**Functional analysis of the proteins secreted in the presence of different carbon sources**

Functional analysis based on KOG categories (Table 3) showed that most proteins detected in all conditions tested were implicated in carbohydrate metabolism and transport (70.5-91.8%). The second most abundant group clustered enzymes involved in amino acid metabolism and transport. The rest of proteins from each one of the secretomes studied belonged to a different number of KOG categories. For each condition, the categories that group more than 2% of the total proteins identified were considered to be the most representative. These outlines describe perfectly the results deduced for Avicel cultures, in which more than 95% of the extracellular proteins are categorized in the two main groups. The enzymes in the other three secretomes studied share significant amounts of proteins involved in signal transduction and transport, energy production and conversion. The extracellular proteins from glucose cultures had the lowest representation of enzymes involved in carbohydrate metabolism and a significantly high amount of proteins related to cell wall biogenesis, as compared to the other samples. Similarly, the notable representation of enzymes engaged in post-translational modification, protein turnover, and with chaperone functions in the proteins from glucose, xylan and slurry supernatants merits especial attention. As commented before, the number of proteins detected in the shotgun analysis of the last two secretomes is considerably high due to the need of using a battery of enzymes to metabolize these complex and heterogeneous lignocellulosic materials, and this fact would also justify the over-representation of enzymes involved in the synthesis and secretion of the proteins required to accomplish this task. Although these are mainly intracellular proteins, they can be detected in the culture supernatant due to fungal autolysis. In fact, when the sequences of the proteins identified were examined for the presence of signal peptide, it turned out that more than 30% of proteins in glucose, xylan and slurry supernatants were intracellular, which contrasts with the 14% of intracellular proteins in Avicel supernatants. These data coincide with the results from semi-quantitative analysis (%PSM) that indicated that more than 90% of the total protein amount from Avicel supernatants corresponds to extracellular enzymes, against 74-79% of extracellular proteins in xylan, glucose and slurry. Similarly, the classification of the carbohydrolases identified into GH families reflected 88% of extracellular enzymes in the medium with Avicel (95% PSM), and 74%, 72% and 82% (89, 78 and 89 %PSM) in the supernatants from glucose, xylan and slurry respectively. More than 5% of the proteins detected in slurry supernatants were categorized as energy production and conversion enzymes, and may have been secreted due to the presence of lignin in the culture medium, a heterogeneous substrate that comes from plant cell wall and is catabolized

by oxidases (Liu *et al*., 2013). The number of different proteins identified in the extracellular medium recovered upon fungal growth in the four carbon sources was analyzed by a Venn representation (figure 5). According to these data, 37 proteins were produced in all conditions evaluated, and 21, 5, 44 and 15 enzymes were exclusive from Avicel, glucose, xylan and slurry supernatants, respectively. These data agree with those accounted for total proteins, since the maximum number of total proteins was also found in xylan cultures.

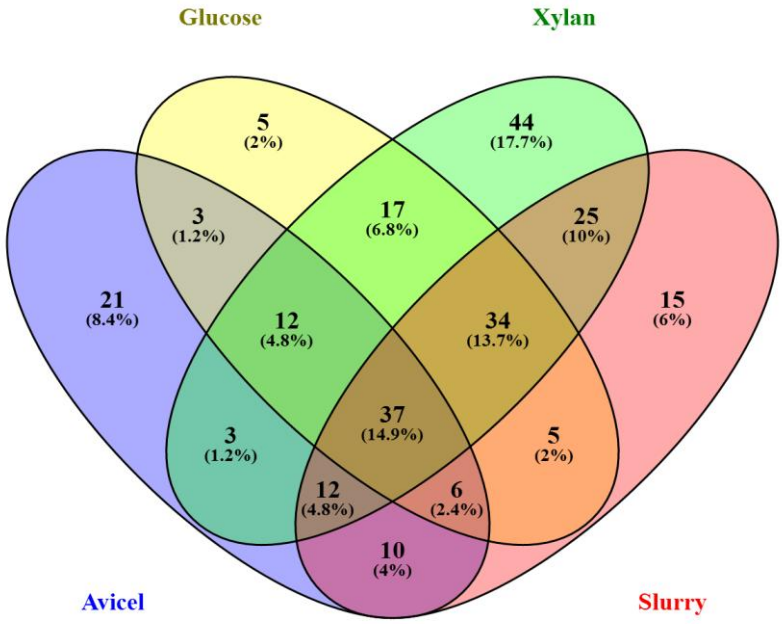



**Figure 5.** Venn representation of the number of proteins found in the different growth conditions. Numbers in brackets account for protein percentages.

In the present work, the proteomic analysis was carried out in samples from 7-day-old cultures. As already mentioned, glucose was depleted after 24 h of growth, which corresponds to severe carbon starvation in these cultures. On the other hand, the scarce solubility of xylan and slurry components may lead to an overall pseudo-starvation, as oligo- and monosaccharides are presumably slowly released and quickly metabolized by the fungus. Similar cellulase induction has been reported after depletion of the carbon source and during early exposure to cellulose in *Neurospora crassa* and *A. niger* cultures (van Munster *et al*., 2014). Our data are consistent with those obtained for the extracellular proteins detected in carbon-starved cultures of the ascomycete *Aspergillus niger* that also cluster into multiple KOG categories (Nitsche *et al*., 2012).

**Table 3.** Functional classification of the proteins identified in the extracellular pool of proteins of *T. amestolkiae* cultured with several carbon sources. The categories containing over 2% of the total EPP are marked in bold italics.

| | % PSM | | | |
|---|---|---|---|---|
| | Avicel | Glucose | Slurry | Xylan |
| A– RNA processing and modification | 0.03 | 0.22 | 0.00 | 0.11 |
| C- Energy production and conversion | 1.48 | ***4.90*** | ***5.44*** | 0.11 |
| E- Amino Acid metabolism and transport | ***3.53*** | ***10.38*** | ***7.47*** | ***9.92*** |
| F- Nucleotide metabolism and transport | 0.00 | 1.42 | 0.14 | 1.07 |
| G- Carbohydrate metabolism and transport | ***91.84*** | ***65.16*** | ***71.73*** | ***70.52*** |
| I -Lipid metabolism | 0.62 | 0.17 | 2.02 | 1.53 |
| M- Cell wall/membrane/envelop biogenesis | 0.00 | ***3.61*** | 0.89 | 1.13 |
| O- Post-translational modification, protein turnover, chaperone functions | 0.51 | 1.71 | 2.15 | 1.64 |
| Q- Secondary Structure | 0.06 | 0.91 | 0.48 | 1.12 |
| R- General Functional Prediction only | 0.63 | 1.89 | ***2.56*** | 1.67 |
| S- Function Unknown | 0.37 | ***3.71*** | ***3.16*** | 2.31 |
| T- Signal Transduction | 0.40 | ***4.86*** | ***3.42*** | 4.03 |

**GHs are the main extracellular enzymes secreted by *T. amestolkiae***

The major enzymes secreted in all conditions tested are GHs, which constitute a very complex group of carbohydrate-active enzymes. Their current classification into families is entirely based on the similarity of proteins' sequence, and two enzymes with the same activity can be included in different families.

As explained before, low precision quantitative data for GHs were also based on the number of PSMs (mean value from the two biological replicates). Total GH abundance was first calculated as the sum of PSMs from all GHs identified in each individual sample, and then the percentage of each enzyme was obtained. The GHs identified in the proteomic analysis were categorized into 31 families (Table 4). Proteins from family GH3, associated to degradation of plant cell wall polysaccharides (β-glucosidases, β-xylosidases, arabinofuranosidases and exo-1,3-1,4-glucanases), are produced in all carbon sources tested. In terms of relative abundance (PSM%), these enzymes are the main extracellular GHs found in the xylan-containing medium (>25%), representing 16.3% and 13.7% in glucose and slurry media and less than 10% in the Avicel crude. GH7 proteins, which include endo-β-1,4-glucanases and reducing end-acting cellobiohydrolases, were profusely produced in the two cellulose-containing media, Avicel (47.1% of the total GHs) and slurry (29.5%), while GH6 endoglucanases and non-reducing end cellobiohydrolases were especially abundant in the secretome from the Avicel medium (15.1%).

**Table 4.** Glycosyl hydrolase families identified in secretomes of *T. amestolkiae* grown in different carbon sources. The most represented families, containing over 2% of the total extracellular pool of proteins are marked in bold italics. Percentage values refer to % of all GHs.

| GH Family | % PSM | | | |
|---|---|---|---|---|
| | Avicel | Glucose | Xylan | Slurry |
| **GH1** | 0.2 | ***2.2*** | 0.4 | 0.5 |
| **GH2** | 0.5 | 1.7 | 1.9 | 1.5 |
| **GH3** | ***6.3*** | ***16.3*** | ***25.1*** | ***13.7*** |
| **GH5** | ***4.2*** | 0.5 | 1.7 | ***4.5*** |
| **GH6** | ***15.1*** | 0.3 | 1.5 | ***3.3*** |
| **GH7** | ***47.1*** | ***2.6*** | ***5.6*** | ***29.5*** |
| **GH10** | ***3.5*** | 0.0 | ***3.8*** | ***5.3*** |
| **GH11** | 1.2 | 0.0 | 0.0 | 0.0 |
| **GH13** | 0.2 | ***4.3*** | 1.8 | 0.8 |
| **GH15** | ***3.6*** | ***28.4*** | ***12.9*** | ***7.8*** |
| **GH16** | 0.4 | ***3.5*** | 0.0 | 1.5 |
| **GH17** | 0.8 | 1.7 | 1.7 | 0.7 |
| **GH18** | 0.1 | 0.6 | ***2.1*** | ***2.0*** |
| **GH20** | 0.3 | ***4.8*** | ***2.8*** | 1.3 |
| **GH27** | 1.8 | 1.7 | ***2.7*** | 2.0 |
| **GH30** | 1.0 | 0.5 | ***2.1*** | 1.3 |
| **GH31** | 1.7 | ***11.8*** | ***9.8*** | 4.0 |
| **GH35** | 0.7 | 1.8 | ***3.6*** | 0.4 |
| **GH43** | 0.2 | 0.1 | 0.8 | 0.4 |
| **GH47** | 0.4 | 1.1 | 0.9 | 0.2 |
| **GH51** | 0.0 | 0.0 | ***3.3*** | 0.9 |
| **GH54** | 1.2 | 0.0 | 0.0 | ***3.2*** |
| **GH55** | 1.0 | ***3.9*** | ***2.5*** | ***2.4*** |
| **GH62** | 1.5 | 0.3 | 1.1 | ***2.7*** |
| **GH71** | 1.0 | 0.1 | 0.2 | 1.0 |
| **GH72** | 0.2 | 1.3 | 1.5 | 1.1 |
| **GH74** | 1.6 | 0.1 | 0.5 | 1.1 |
| **GH92** | 0.4 | 1.1 | 1.2 | 0.4 |
| **GH95** | 0.5 | 1.2 | 1.1 | 2.4 |
| **GH125** | 0.8 | 1.7 | 1.3 | 0.2 |
| **GH127** | 0.5 | ***3.2*** | ***3.4*** | ***2.1*** |

Proteins from these families were induced by xylan to a lesser extent (1.5% for GH6 and 5.6% for GH7). Family GH10 groups enzymes with endo-xylanase activity and is well represented in Avicel, slurry and xylan crudes. In addition, α-carbohydrolases belonging to GH15 and GH31 families were identified in the samples analyzed. The huge amount of GH15 glucoamylases/glucodextranases detected in glucose cultures (28%)

can be related to fungal autolysis, since α-(1-3) and α-(1-4) glucans have been reported as cell-wall components of several *Talaromyces* species (Prieto *et al*., 1995). On the other hand, GH31 proteins act on α-linked xylo-, manno-, galacto- or gluco-oligosaccharides contributing to the final breakdown of oligomers released from plant hemicellulose or fungal polysaccharides.

The percentages of the different types of cellulases and hemicellulases secreted in each medium are summarized in Figure 6.

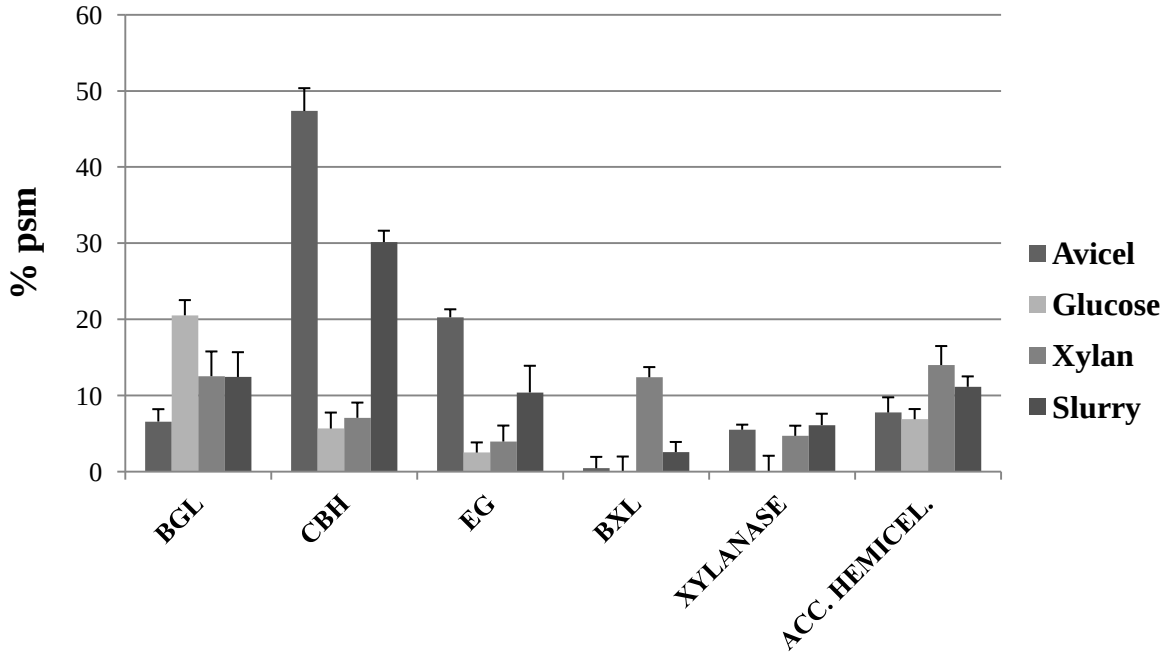



**Figure 6.** Relative abundance of cellulases and hemicellulases as a function of the carbon source. The sum of the number of PSMs of all GH proteins was taken as 100%. The putative proteins were annotated searching against entries in the CAZy database using BLASTP. Numbers account for the arithmetic mean of two biological replicates. BXL: β-xylosidases. ACC. HEMICEL: accessory hemicellulases.

## β-glucosidases produced in different culture conditions

Further analysis of the proteomic data allowed the evaluation of the β-glucosidases induced in the carbon sources tested (figure 7).

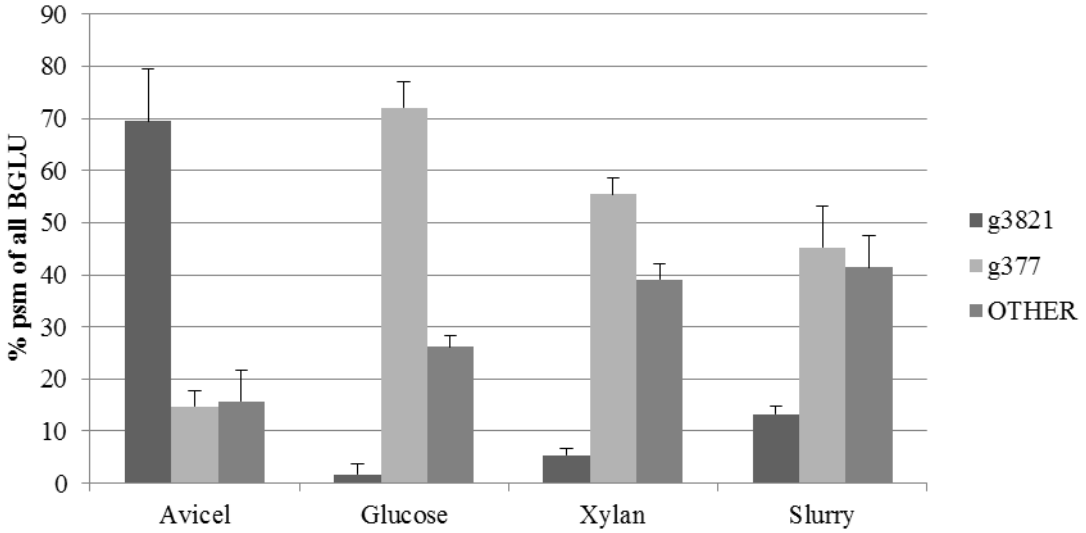



**Figure 7.** Production and percentages of different BGLs as a function of the carbon source. The putative proteins were annotated searching against entries in the CAZy database using BLASTP. The sum of the number of PSMs of all BGLs was taken as 100%.

An approximation to the relative abundance of each one of them, calculated as the percentage of total β-glucosidases, was analyzed. The number of hypothetical BGLs detected varied from 6 different enzymes in Avicel and glucose media to 8 and 10 for slurry and xylan cultures, respectively. Protein g377 (see Additional file 1 for sequence) was the most abundant in media containing slurry, glucose, or xylan, representing 72%, 55% and 45% of the total BGLs, respectively, while in Avicel the main BGL was protein g3821 (70%). However, low amounts of this enzyme were also produced in media with glucose (<2%), xylan (5%), or slurry (13%). These findings suggest that BGL g3821 is strongly induced by cellulose but constitutively expressed in basal levels in the other conditions studied. On the contrary, the production of BGL g377 seems to be carbon source-independent, although its synthesis and secretion may be triggered upon starvation. Preliminary RT-PCR experiments indicated that g3821 is overexpressed in Avicel cultures and repressed when glucose was used as the carbon source, while g377 is overexpressed in both media, which is compatible with a constitutive expression, independent of the presence of cellulose in the medium (data not shown). The enzyme g377 could be responsible for the BGL activity detected when *T. amestolkiae* was cultivated using non-cellulosic materials as carbon sources (mono and disaccharides). Similarly to this, one of the two different β-glucosidases

produced by *Stachybotrys art* is non-inducible (De Gussem *et al*., 1978). Its role could also be related to cell wall remodeling, as seen in some plants (Ketudat Cairns *et al*., 2015).

### *T. amestolkiae* crude supernatants as BGL sources for saccharification

Preliminary saccharification assays were performed using Celluclast 1.5 L FG or *T. amestolkiae* crudes from Avicel as the main sources of cellulases. Released glucose was 33% higher when Celluclast 1.5 L FG was used, probably for its high content in CBH and EG compared to *T. amestolkiae* crudes, rich in BGL.

For that reason, enzymatic hydrolysis of wheat straw slurry was assayed using Celluclast 1.5 L FG as the main source of cellulases. As BGL supplements for saccharification, 0.5 units of this activity were incorporated to the saccharification mixture using *T. amestolkiae* crudes from glucose (TAM377) or Avicel (TAM3821) culture media, or the commercial preparation N50010 from Novozymes (figure 8).

The treatment of wheat straw slurry with only Celluclast released 50% glucose after 96 h. In the same time period, glucose release enhanced significantly when Celluclast was supplemented with TAM3821 (80%), N50010 (73%) or TAM377 (61%).

These data show that these enzymatic crudes from *T. amestolkiae* are very efficient as supplements for lignocellulosic biomass saccharification, especially TAM3821 that provided superior glucose solubilization than N50010. This could be due to a high affinity of BGL g3821 for cellulosic substrates, since this protein was only induced by cellulose. Nevertheless, the existence of synergistic catalysis between this BGL and the EGs and CBHs induced by cellulosic substrates cannot be ruled out. Further purification and characterization of both BGL will be carried out in order to ascertain their different catalytic constants, substrate specificities and efficiencies.

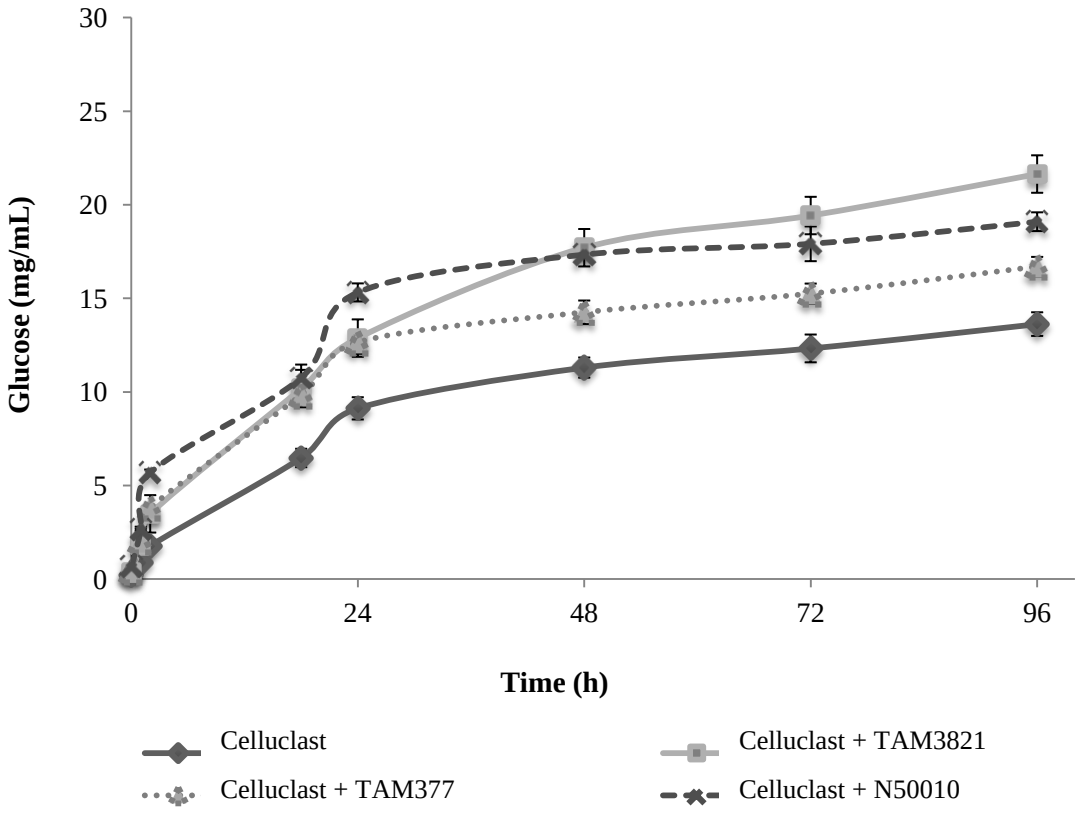



**Figure 8.** Saccharification of pretreated wheat straw with different enzymatic cocktails. Celluclast was used as the base enzymatic cocktail to perform supplementation studies adding N500510 (commercial Novozymes preparation rich in BGL), TAM377 or TAM3821 (*T. amestolkiae* enzymatic supernatants from cultures grown in glucose or Avicel, respectively). To measure the saccharification efficiency, samples were taken periodically, and glucose was measured in supernatants as described in materials and methods.

## CONCLUSIONS

A strong extracellular β-glucosidase activity was detected in *T. amestolkiae* cultures grown in four carbon sources of different composition and complexity. Genome sequencing and annotation disclosed a high number of genes encoding BGLs, which suggests the relevance of these enzymes for this organism. Differential semi-quantitative proteomic analysis of the secretomes recovered from each culture revealed the production of two main proteins with BGL activity. One of them was constitutively released regardless the carbon source used, although was repressed by glucose, and showed to be less efficient than the commercial

BGL preparation used for comparison in saccharification experiments. The other was strongly induced by cellulose and revealed to be an excellent supplement to Celluclast for saccharification of pretreated wheat straw.

## METHODS

### Fungal strain and culture media

The fungus was isolated from cereal wastes and deposited in the IJFM culture collection at "Centro de Investigaciones Biológicas" (Madrid, Spain), with the reference A795. This isolate was identified on the basis of molecular and phenotypic analyses (see Additional file 1), and hereinafter it will be denominated as *T. amestolkiae* CIB. Fungal cultures were grown on 2% agar-malt Petri dishes at 28 ºC and incubated for 7 days to obtain spore suspensions. One agar plug of about 1 $cm^2$ was cut from actively growing mycelium and deposited in 15 mL Falcon tubes with 5 mL of 1% NaCl with 0.1% Tween 80, shaken, and 200 μL withdrawn to inoculate 250 mL Erlenmeyer flasks with 50 mL of CSS medium (28 ºC, 250 rpm for 7 days). CSS medium (pH 5.6) contained ($L^{-1}$): 40 g glucose, 0.4 g $FeSO_4$ $7H_2O$, 9 g $(NH_4)_2SO_4$, 4 g $K_2HPO_4$, 26.3 g corn steep solids, 7 g $CaCO_3$ and 2.8 mL soybean oil. Two mL of these cultures were used as inocula for the production of cellulase activities in Mandels medium (Mandels and Weber 1969). Its components are ($L^{-1}$): 2.0 g $KH_2PO_4$, 1.3 g $(NH_4)_2SO_4$, 0.3 g urea, 0.3 g $MgSO_4 \cdot 7H_2O$, 0.3 g $CaCl_2$, 5 mg $FeSO_4 \cdot 7H_2O$, 1.6 mg $MnSO_4 \cdot H_2O$, 1.4 mg $ZnSO_4 \cdot 7H_2O$, 1 g Bacto Peptone. The medium was supplemented with 1% Avicel, glucose, acid slurry from wheat straw (acid catalyzed steam exploded wheat straw provided by Abengoa), maltose, fructose, glycerol, xylose or cellobiose, or 2% beechwood xylan (reported as the best concentration for xylanases' production (Mandels and Weber 1969)) as carbon sources. In all cases the cultures were carried out in 250 mL Erlenmeyer flasks with 50 mL of culture medium and incubated at 28 ºC and 250 rpm. Samples were taken periodically from three replicate flasks and the mycelium separated from the culture liquid by centrifugation at 13,000 *g* and 4 °C for 5 min.

### Enzyme and protein assays and other determinations

Proteins were quantified according to the Bradford method using the Bio-Rad reagent and bovine serum albumin as the standard. Avicelase (total microcrystalline-cellulose hydrolyzing activity, as indicative of exocellulase activity) and β-1,4-endoglucanase activities were measured by determining the release of reducing sugars by the Somogyi-Nelson method (Somogyi, 1945). The standard enzymatic assays were performed in 50 mM sodium citrate buffer, pH 5.0, containing appropriately diluted crudes and 1% Avicel (Merck) or 2% low viscosity carboxymethylcellulose

(CMC, Sigma), as substrates for Avicelase and β-1,4-endoglucanase activities, respectively. Linearity of the enzymatic assays was checked by performing the reaction at two different incubation times (5 and 10 min). β-glucosidase activity was assayed spectrophotometrically following *p*-nitrophenol (*p*NP) release ($\varepsilon_{410}$ = 15,200 $M^{-1}$ $cm^{-1}$) from *p*-nitrophenyl-β-D-glucopyranoside (Sigma-Aldrich), in 100 mM sodium citrate buffer pH 5, using 1.4% sodium carbonate to stop the reaction. One unit of activity was defined as the amount of enzyme releasing 1 μmol of reducing sugars or *p*NP per minute under the above conditions. Direct quantification of glucose and cellobiose was accomplished by high-performance liquid chromatography (HPLC) on an Agilent 1200 series system equipped with a refractive index detector. Aliquots of 100 μL were loaded onto a SUPELCOGEL C-G610H column (Sigma) equilibrated with 5 mM $H_2SO_4$. The column was previously calibrated by injecting 100 μL of glucose or cellobiose in a concentration range from 0.5 to 60 mM. From the area under the peaks, a calibration curve was calculated for each compound. Peaks were identified from their retention times, by comparison with those of the commercial standards, and their concentrations were calculated from the calibration curves.

**Genome sequencing and assembly**

DNA was extracted with the DNeasy Plant minikit (Qiagen). Sequencing was performed on an Illumina HiSeqTM 2000 system. 90 bp-long reads were obtained by paired-end sequencing of a DNA library composed of 537±396 bp inserts. The Illumina GA Pipeline version 1.5 was used for a first removal of adapter sequences, contamination and low-quality reads from raw reads. Then A5-miseq pipeline (Tritt *et al*., 2012) was used in order to carry out a subsequent filtering of the reads, the assembly and the scaffolding. Assembly was performed *de novo* and evaluated in terms of N50 and L50. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MIKG00000000. The version described in this paper is version MIKG01000000.

**Gene prediction, general function annotation and CAZyME prediction**

Gene models were predicted by AUGUSTUS trained with *Aspergillus fumigatus* (Hoff and Stanke, 2013). The obtained open reading frames (ORFs) were used as query to interrogate the KEGG and KOG databases by using KAAS (Moriya *et al*., 2007) and WebMGA (Wu *et al*., 2011) servers, respectively, in order to assign general protein functions profiles. The specific annotation of CAZy families (Cantarel *et al*., 2008; Lombard *et al*., 2014) was accomplished by submitting the predicted genes to dbCAN and filtering the results by applying an E-value < $1 \cdot 10^{-20}$ as

cutoff. Then, the assignment of β-glucosidase function was carried out by doing a BLASTP of the predicted glycosyl hydrolases from *T. amestolkiae* against the characterized GHs from the CAZy database and annotating the function of the best hit with an E-value < $1 \cdot 10^{-20}$. The presence or absence of signal peptide in all the predicted proteins was analyzed using Phobius, a combined transmembrane topology and signal peptide predictor (Käll *et al*., 2004)

**Shotgun analysis of the secretomes**

For differential proteomic analysis, samples of the extracellular pool of proteins from culture supernatants of *T. amestolkiae* grown in the different substrates were independently analyzed. Two mL of the liquid supernatants from 7-day-old cultures were centrifuged at 14,000 × *g* for 5 min to remove insoluble material and freeze-dried. Then, samples containing 5 µg of proteins were first dissolved in sample buffer, denatured, loaded in a 12% SDS-gel and allowed to run for 10 min in order to remove non-protein compounds before the proteomic analysis. The protein band was horizontally divided into two similar portions and each one was excised in small pieces and destained with 50 mM ammonium bicarbonate/50% ACN (acetonitrile), dehydrated with ACN and dried. Samples were reduced by adding DTT to a final concentration of 10 mM and alkylated with iodoacetamide to a final concentration of 50 mM. Then, gel pieces were dried, rehydrated with 12.5 ng/µL trypsin in 50 mM ammonium bicarbonate and incubated overnight at 30 °C. Peptides were extracted at 37 °C using 100% ACN and 0.5% TFA, respectively. The peptide pool was dried and then cleaned using ZipTip with 0.6 µL C18 resin (Millipore) and reconstituted in 5 µL 0.1% formic acid/2% ACN (v/v) prior to MS analysis.

All peptide separations were carried out on a NanoEasy HPLC (Proxeon Biosystems) coupled to a nanoelectrospray ion source (Proxeon Biosystems). For each analysis, the sample peptides were loaded onto a C18-A1 ASY-Column 2 cm precolumn (Thermo Scientific) and then eluted onto a Biosphere C18 column (C18, inner diameter 75 µm, 15 cm long, 3 µm particle size, Nano Separations). The mobile phase flow rate was 250 nL/min using 0.1% formic acid, 2% ACN in water (solvent A) and 0.1% formic acid and 100% ACN (solvent B). The gradient profile was set as follows: 0–35% solvent B for 120 min, 35%-45% solvent B for 20 min, 45%-95% solvent B for 9 min and 5 min isocratically at 95%). Six microliters of each sample were injected. Full-scan MS spectra (m/z 300-1800) were acquired in the LTQ-Orbitrap Velos in the positive ion mode with a target value of 1,000,000 at a resolution of 60,000 and the 15 most intense ions were selected for collision induced dissociation (CID)

fragmentation in the LTQ with a target value of 10,000 and normalized collision energy of 35%.

Mass spectra *.raw files were searched against an in –house specific database of the *T. amestolkiae* genome (10408 sequences, 5662098 residues) using the SEQUEST search engine through Proteome Discoverer (version 1.4.1.14, Thermo). Precursor and fragments mass tolerance were set to 10 ppm and 0.5 Da, respectively. Search parameters included a maximum of two missed cleavages allowed, carbamidomethylation of cysteines as a fixed modification and oxidation of methionine as a variable modification. Peptides were validated through the algorithm Percolator (Käll *et al*., 2007) and only those with high confidence were admitted (FDR 0.01). Unless otherwise specified, protein identifications were accepted if they contained at least two identified peptides. Results were inferred from data obtained from two technical replicates from two different biological samples. Relative low precision quantification of the proteins in the samples analyzed was estimated from the sum of the number of peptide spectrum matches (PSMs) corresponding to each protein (Liu *et al*., 2004; Schulze and Usadel, 2010; Salvachúa *et al*., 2016): These values were used to calculate the percentage of individual proteins or categories (% of GHs, % of BGLs) in the whole protein pool. The data provided in tables and figures correspond to the mean value of the %PSMs from the biological replicates.

**Wheat straw saccharification**

Celluclast 1.5 L FG (Novozymes) was used as commercial cocktail for saccharification. An amount of Celluclast containing 0.5 U of BGL activity was added to 66.6 mg/mL of acid slurry from wheat straw (40% content in cellulose) in 1.5 mL sodium acetate 100 mM, pH 4, at 50 ºC. Samples were taken periodically and the glucose released was quantified using the Glucose TR kit from Spinreact.

Celluclast was supplemented with 0.5 U of BGL activity from N50010 (Novozyme), TAM377 (*T. amestolkiae* enzymatic crude produced using glucose as carbon source) or TAM3821 (*T. amestolkiae* enzymatic crude produced using Avicel as carbon source).

**Declarations**

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Availability of supporting data**

This Whole Genome Shotgun project has been deposited at DDBJ/ ENA/ GenBank under the accession MIKG00000000. The version described in this paper is MIKG01000000.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This work has been funded by projects RTC-2014-1777-3 and BIO2015-68387-R from MINECO and S2013/MAE-2907 from Comunidad de Madrid. M. Nieto-Domínguez and J.A. Méndez-Líter thank their fellowships, FPU of the MINECO and Tatiana Pérez de Guzmán el Bueno Foundation, respectively.

**Authors' contributions**

LE, JM, JG, JB and MN contributed to design the research and participate in the experiments. LA and MN analyzed and assembled fungal genome. LE and MN drafted the manuscript. MJM and AP co-coordinated the research, and reviewed and edited the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

The authors thank the Proteomics and Genomics facility at CIB and also Ph.D. Alexander Holm Viborg for kindly providing the bioinformatics tool (http://research.ahv.dk/cazy) used for characterized GHs sequences extraction from CAZy database.

**Authors' information**

Not applicable

## REFERENCES

Adsul, M.G., Bastawde, K.B., Varma, A.J., Gokhale, D.V., 2007. Strain improvement of *Penicillium janthinellum* NCIM 1171 for increased cellulase production. Bioresour. Technol. 98, 1467-1473.

Alvira, P., Tomás-Pejó, E., Ballesteros, M., Negro, M.J., 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. Bioresour. Technol. 101, 4851-4861.

Amore, A., Giacobbe, S., Faraco, V., 2013. Regulation of cellulase and hemicellulase gene expression in fungi. Curr. Genomics. 14, 230-249.

Berka, R.M., Grigoriev, I.V., Otillar, R., Salamov, A., Grimwood, J., Reid, I. *et al*., 2011. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. Nat. Biotechnol. 29, 922.

Bhiri, F., Chaabouni, S.E., Limam, F., Ghrir, R., Marzouki, N., 2008. Purification and biochemical characterization of extracellular β−glucosidases from the hypercellulolytic pol6 mutant of *Penicillium occitanis*. Appl. Biochem. Biotechnol. 149, 169-182.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2008. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 37, D233-D238.

Chávez, R., Bull, P., Eyzaguirre, J., 2006. The xylanolytic enzyme system from the genus *Penicillium*. J. Biotechnol. 123, 413-433.

Chen, H. 2014. Biological fundamentals for the biotechnology of lignocellulose, in: Biotechnology of Lignocellulose. Springer, pp. 73-141.

De Gussem, R.L., Aerts, G.M., Claeyssens, M., De Bruyne, C.K., 1978. Purification and properties of an induced β−D-glucosidase from *Stachybotrys atra*. Biochim. Biophys. Acta. 525, 142-153.

Duff, S.J.B., Murray, W.D., 1996. Bioconversion of forest products industry waste cellulosics to fuel ethanol: A review. Bioresour. Technol 55, 1-33.

Fowler, T., Brown Jr, R.D., 1992. The *bgI1* gene encoding extracellular β−glucosidase from *Trichoderma reesei* is required for rapid induction of the cellulase complex. Mol. Microbiol. 6, 3225-3235.

Gooday, G.W., Wei-Yun, Z., O'Donnell, R.W., 1992. What are the roles of chitinases in the growing fungus? FEMS Microbiol. Lett. 100, 387-391.

Häkkinen, M., Arvas, M., Oja, M., Aro, N., Penttilä, M., Saloheimo, M. *et al*., 2012. Re-annotation of the CAZy genes of *Trichoderma reesei* and transcription in the presence of lignocellulosic substrates. Microb. Cell. Fact. 11, 134.

Hanif, A., Yasmeen, A., Rajoka, M.I., 2004. Induction, production, repression, and de-repression of exoglucanase synthesis in *Aspergillus niger*. Bioresour. Technol. 94, 311-319.

Hoff, K.J., Stanke, M., 2013. WebAUGUSTUS--a web service for training AUGUSTUS and predicting genes in eukaryotes. Nucleic Acids Res. 2013/05/21, W123-W128.

Hori, C., Igarashi, K., Katayama, A., Samejima, M., 2011. Effects of xylan and starch on secretome of the basidiomycete *Phanerochaete chrysosporium* grown on cellulose. FEMS Microbiol. Lett. 321, 14-23.

Ilmen, M., Saloheimo, A., Onnela, M.L., Penttilä, M.E., 1997. Regulation of cellulase gene expression in the filamentous fungus *Trichoderma reesei*. Appl. Environ. Microbiol. 63, 1298-1306.

Jurado, M., Prieto, A., Martínez-Alcalá, Á., Martínez, Á.T., Martínez, M.J., 2009. Laccase detoxification of steam-exploded wheat straw for second generation bioethanol. Bioresour. Technol. 100, 6378-6384.

Käll, L., Canterbury, J.D., Weston, J., Noble, W.S., MacCoss, M.J., 2007. Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat. Methods. 4, 923.

Käll, L., Krogh, A., Sonnhammer, E.L.L., 2004. A combined transmembrane topology and signal peptide prediction method. J. Mol. Biol. 338, 1027-1036.

Ketudat Cairns, J.R., Mahong, B., Baiya, S., Jeon, J.S., 2015. β–Glucosidases: Multitasking, moonlighting or simply misunderstood? Plant Sci. 241, 246-259.

Kumar, R., Singh, S., Singh, O.V., 2008. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. J. Ind. Microbiol. Biotechnol. 35, 377-391.

Lee, J., Kwon, K.S., Hah, Y.C., 1996. Regulation of β–glucosidase biosynthesis in *Aspergillus nidulans*. FEMS Microbiol. Lett. 135, 79-84.

Liu, G., Zhang, L., Wei, X., Zou, G., Qin, Y., Ma, L. *et al*., 2013. Genomic and secretomic analyses reveal unique features of the lignocellulolytic enzyme system of *Penicillium decumbens*. PLoS ONE 8, e55185.

Liu, H., Sadygov, R.G., Yates, J.R., 2004. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal. Chem. 76, 4193-4201.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2013/11/21, D490-D495.

Maeda, R.N., Serpa, V.I., Rocha, V.A.L., Mesquita, R.A.A., Anna, L.M.M.S., de Castro, A.M. *et al*., 2011. Enzymatic hydrolysis of pretreated sugar cane bagasse using Penicillium funiculosum and *Trichoderma harzianum* cellulases. Process Biochem. 46, 1196-1201.

Mandels M, Weber J. 1969. Production of cellulases. Adv. Chem. Ser. 391-414.

Martinez, D., Berka, R.M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S.E. *et al*., 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat. Biotechnol. 26, 553-560.

Martinez, D., Challacombe, J., Morgenstern, I., Hibbett, D., Schmoll, M., Kubicek, C.P. *et al*., 2009. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc. Natl. Acad. Sci. 1954-1959.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., Kanehisa, M., 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 2007/05/25, W182-W185.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A comprehensive review. Renew. Sustain. Energy Rev. 14, 578-597.

Nieto-Domínguez, M., de Eugenio, L.I., Barriuso, J., Prieto, A., Fernández de Toro, B., Canales-Mayordomo, A. *et al*., 2015. Novel pH-stable glycoside hydrolase family 3 β–xylosidase from *Talaromyces amestolkiae*: an enzyme displaying regioselective transxylosylation. Appl. Environ. Microbiol. 6380-6392.

Nitsche, B.M., Jørgensen, T.R., Akeroyd, M., Meyer, V., Ram, A.F., 2012. The carbon starvation response of *Aspergillus niger* during submerged cultivation: insights from the transcriptome and secretome. BMC Genomics. 13, 380.

Pel, H.J., De Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J. *et al*., 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat. Biotechnol. 25, 221.

Prieto, A., Bernabé, M., Leal, J.A., 1995. Isolation, purification and chemical characterization of alkali-extractable polysaccharides from the cell walls of *Talaromyces* species. Mycol. Res. 99, 69-75.

Roberts, S.B., Gowen, C.M., Brooks, J.P., Fong, S.S., 2010. Genome-scale metabolic analysis of *Clostridium thermocellum* for bioethanol production. BMC Systems Biol. 4, 31.

Roberts, T.C., Johansson, H.J., McClorey, G., Godfrey, C., Blomberg, K.E., Coursindel, T. *et al*., 2015. Multi-level omics analysis in a murine model of dystrophin loss and therapeutic restoration. Hum. Mol. Genet. 24, 6756-6768.

Salvachúa, D., Katahira, R., Cleveland, N.S., Khanna, P., Resch, M.G., Black, B.A. *et al*., 2016. Lignin depolymerization by fungal secretomes and a microbial sink. Green Chem. 18, 6046-6062.

Schulze, W.X., Usadel, B., 2010. Quantitation in mass-spectrometry-based proteomics. Annu. Rev. Plant Biol. 61, 491-516.

Shewale, J.G., 1982. β–Glucosidase: Its role in cellulase synthesis and hydrolysis of cellulose. In. J. Biochem. 14, 435-443.

Somogyi, M., 1945. A new reagent for the determination of sugars. J. Biol. Chem. 160, 61-68.

Stals, I., Karkehabadi, S., Kim, S., Ward, M., Van Landschoot, A., Devreese, B. *et al*., 2012. High resolution crystal structure of the endo-N-Acetyl-β–D-glucosaminidase responsible for the deglycosylation of *Hypocrea jecorina* cellulases. PLoS ONE e40854-e40854.

Stevenson, D.M., Weimer, P.J., 2005. Expression of 17 genes in *Clostridium thermocellum* ATCC 27405 during fermentation of cellulose or cellobiose in continuous culture. Appl. Environ. Microbiol. 71, 4672-4678.

Szijarto N, Szengyel Z, Liden G, Reczey K, 2004. Dynamics of cellulase production by glucose grown cultures of *Trichoderma reesei* Rut-C30 as a response to addition of cellulose. Appl. Biochem. Biotechnol.115-124.

Tritt, A., Eisen, J.A., Facciotti, M.T., Darling, A.E., 2012. An integrated pipeline for de novo assembly of microbial genomes. PLoS ONE 7, e42304.

Ustinov, B.B., Gusakov, A.V., Antonov, A.I., Sinitsyn, A.P., 2008. Comparison of properties and mode of action of six secreted xylanases from *Chrysosporium lucknowense*. Enz. Microb. Technol. 43, 56-65.

van Munster, J.M., Daly, P., Delmas, S.p., Pullan, S.T., Blythe, M.J., Malla, S. *et al*., 2014. The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in *Aspergillus niger*. Fungal Genet. Biol. 72, 34-47.

Winger, A.M., Heazlewood, J.L., Chan, L.J.G., Petzold, C.J., Permaul, K., Singh, S., 2014. Secretome analysis of the thermophilic xylanase hyper-producer *Thermomyces lanuginosus* SSBP cultivated on corn cobs. J. Ind. Microbiol. Biotechnol. 41, 1687-1696.

Wu, S., Zhu, Z., Fu, L., Niu, B., Li, W., 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics. 12, 444.

Zhang, Y.H.P., Lynd, L.R., 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. Biotechnol. Bioeng. 88, 797-824.

Zhao, Z., Liu, H., Wang, C., Xu, J.R., 2013. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. BMC Genomics. 14, 274.

Znameroski, E.A., Glass, N.L., 2013. Using a model filamentous fungus to unravel mechanisms of lignocellulose deconstruction. Biotechnol. Biofuels. 6, 6.

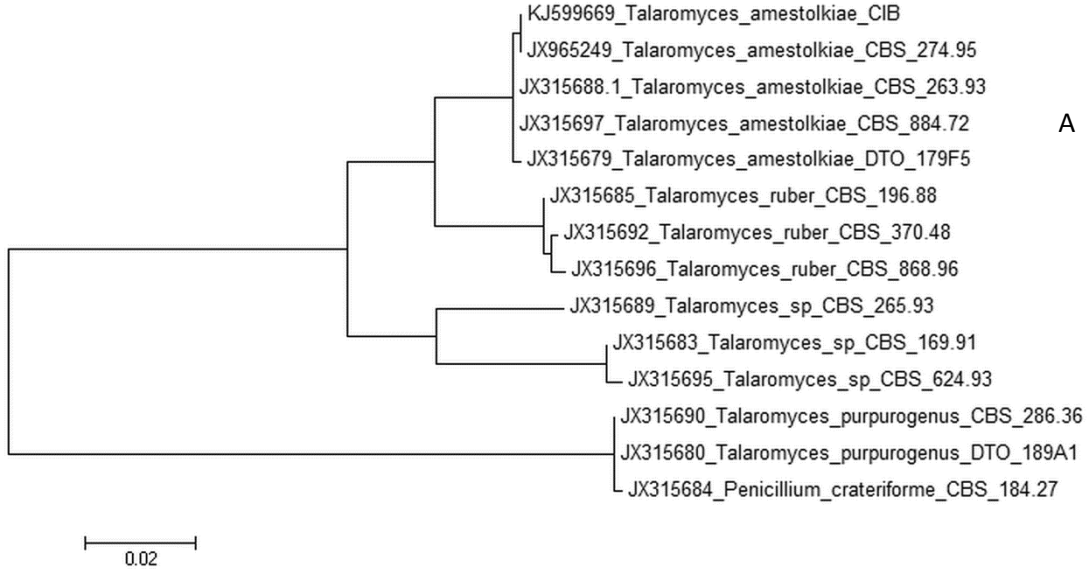
## SUPPLEMENTARY INFORMATION

In the original version of the article, there were three additional files. The version which is part of this thesis includes only additional file 1. Additional files 2 and 3 can be found in the digital version of the document (endorsed as a CD) or in the online version of the article due to their excessive sizes.

**Additional file 1.** Identification of the fungal isolate. This file contains three figures, the materials and methods associated to the information presented, and a brief discussion of the data presented. Figure S1. Maximum likelihood phylogenetic analysis of RPB1 (A), ITS (B) and BT2 (C) regions from different *Talaromyces* strains. Figure S2. Agar colonies of *T. amestolkiae, Penicillium purpurogenum var. rubrisclerotium* and *Penicillium rubrum*. Figure S3. SEM micrography of conidiophores and hyphae from the three fungal species.

**Identification of the fungal isolate**.

The molecular identification of the cellulolytic fungus isolated was initially carried out from the sequences of the ribosomal DNA including ITS1 and ITS2 regions, which have been described as suitable targets for analysis of fungal phylogeny (Romero *et al*., 2007). This isolate showed 100% identity with *T. amestolkiae* (accession number JX965247.1). To corroborate the identification of this new cellulolytic strain, β-tubulin (BT2) and RPB1 genes were also sequenced and showed to be 100% identical to those of previously analyzed *T. amestolkiae* strains (Yilmaz *et al*., 2012) (Fig S1):



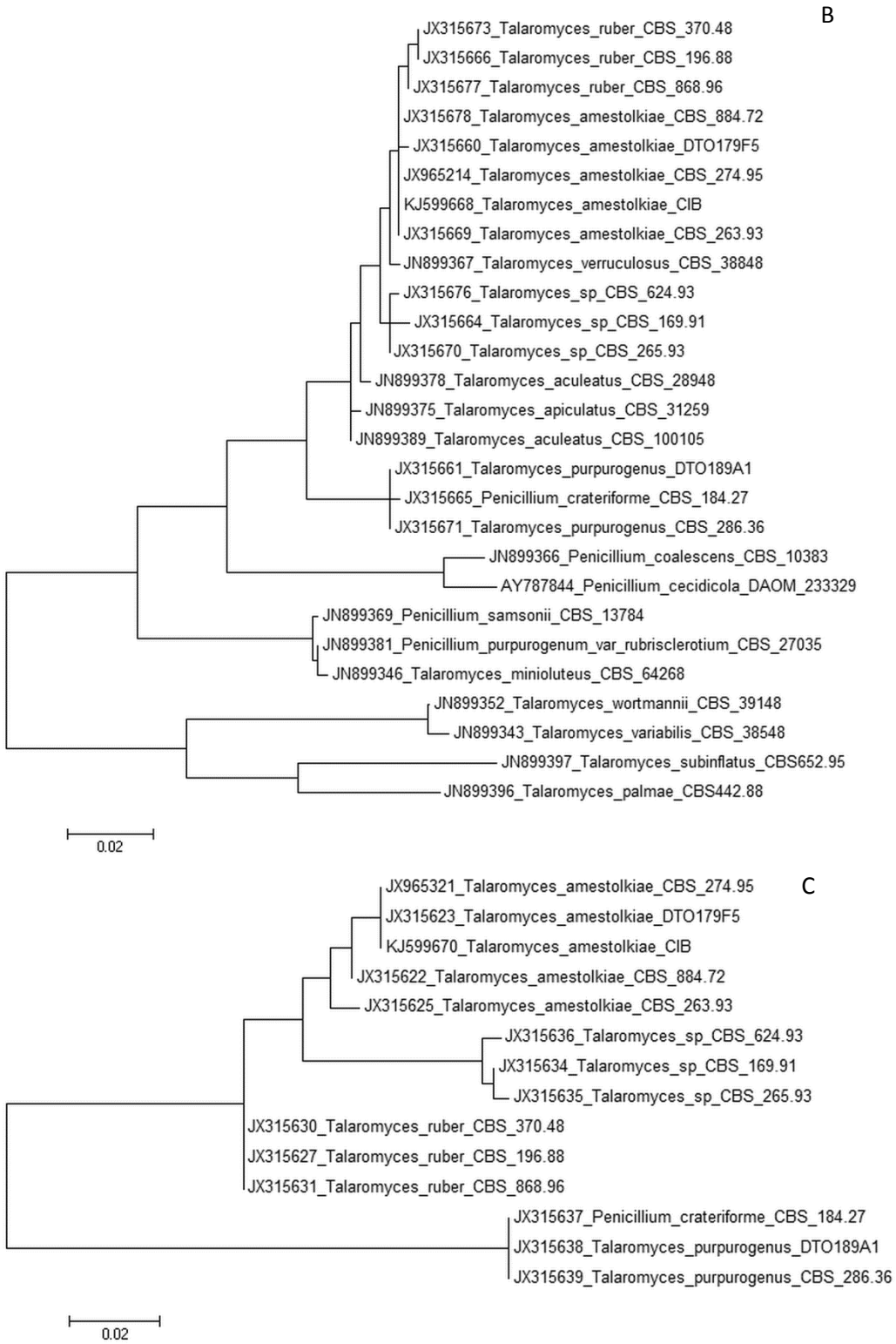



**Figure S1.** Maximum likelihood phylogenetic analysis of RPB1 (A), ITS (B) and BT2 (C) regions from different *Talaromyces* strains. The unrooted neighbor-joining phylogenetic tree with 1000 bootstrap replications was generated based on a previous alignment using MEGA6 software. Sequences accession numbers are shown preceding strain names.

The morphological traits of the novel strain, grown in agar plates on several carbon sources, were also examined and compared to those reported by Yilmaz *et al*. (2012). Cultures of *P. purpurogenum* var. *rubrisclerotium* and *P. rubrum* from the CBS fungal culture collection, which have recently been reassigned as *T. amestolkiae*, were grown in parallel under the same conditions. Growth on CYA and YES after 7 days was similar for the three strains. All of them produced a red pigment in MEA plates, although it was more abundant in our *T. amestolkiae* isolate. In spite of being the same species, small differences in pigment production on YES medium were observed between CBS 263.93 and CBS 274.95 strains (Fig. S2).

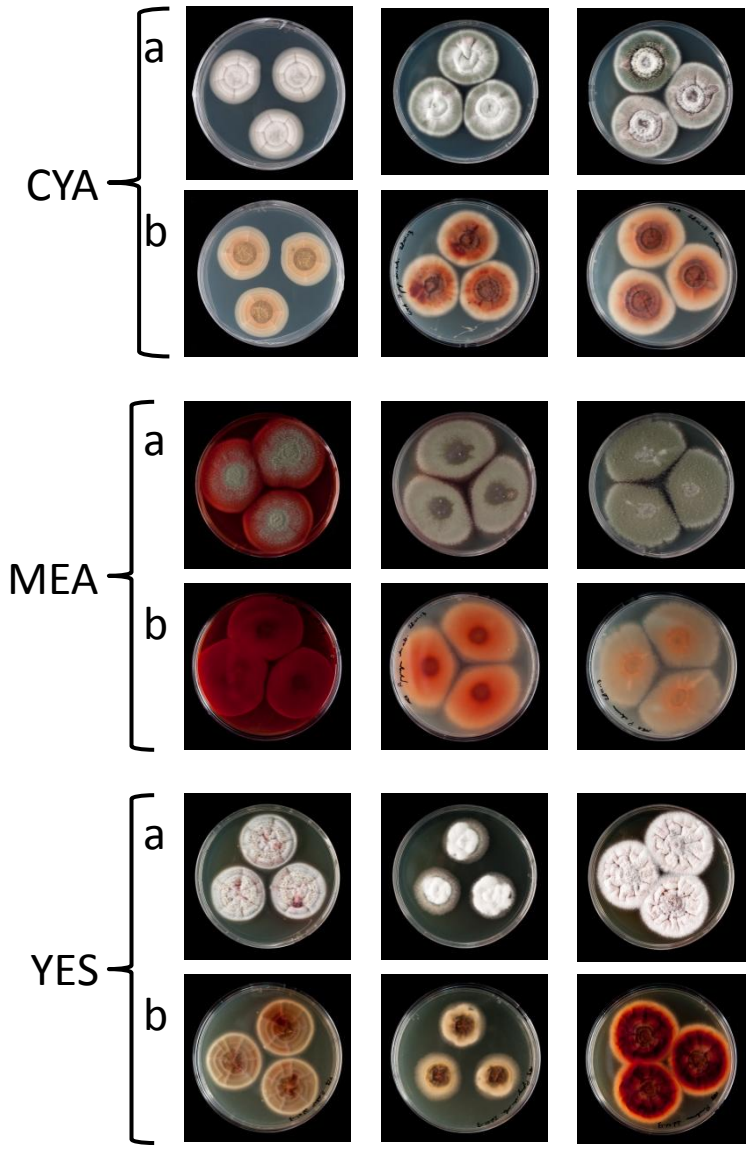



**Figure S2.** Agar colonies of *T. amestolkiae* (left column), *Penicillium purpurogenum var. rubrisclerotium* (middle column) and *Penicillium rubrum* (right column). Pictures of the different fungal strains in CYA, MEA and YES media: obverse (a) and reverse (b) plates incubated at 25 °C for 7 days.

The microscopic observation of the three strains also showed similar features, with biverticillate, vesiculate conidiophores, and rough conidia (Fig. S3).

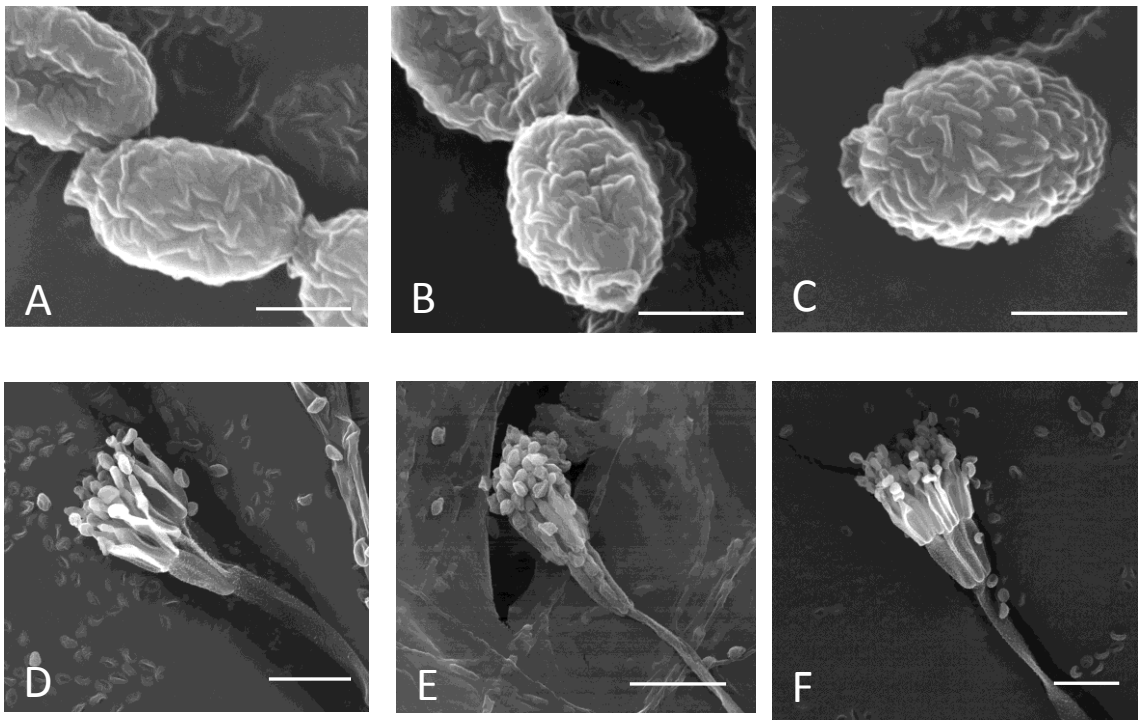



**Figure S3.** SEM micrography of conidiophores and hyphae from: A and D) *T. amestolkiae*; B and E) *P. purpurogenum var rubrisclerotium*; C and F) *P. rubrum*. All strains were grown on wheat straw and samples were treated as described in Materials and Methods. White bars represent 1 mm (upper row) or 10 mm (lower row).

## Materials and methods

Genomic DNA of the selected isolates was obtained using the protocol described in Romero *et al*. (2007). The extracted DNA was used as the template in a PCR to amplify the ITS1 and ITS2 regions and the 5.8S rRNA gene. The primers used for the amplification were its1 (5'-TCCGTAGGTGAACCTGCGG-3') and its4 (5'-TCCTCCGCTTATTGATATGC-3') for ITS1 and ITS2 amplification, Bt2a (5´- GGTAACCAAATCGGTGCTGCTTTC-3´) and Bt2b (5´-ACCCTCAGTGTAGTGACCCTTGGC-3´) for β-tubulin amplification, and RPB2-F (5´- ATTTYGAYGGTGAYGARATGAAC-3´) and RPB2-R (GRACRGTDCCRTCATAYTTRACC-3´) for RPB2 amplification.

For PCR amplifications, 1.2 U of Taq DNA polymerase from Applied Biosystems were used, according to the manufacturer instructions; nucleotide primers were added at 1 μM (final concentration) in a final reaction volume of 50 μL. After template concentration optimization, 0.2 μg DNA were used. The PCR program was composed of 35 cycles: 94 ºC, 1 min; 52 ºC, 40 s; 72 ºC, 1 min. A final extension of 10 min at 72 ºC was added to finish all reactions. In all experiments, control reactions without

template were carried out. For detection of the PCR products 1% agarose gel containing GelRed (Biotium) were run in TAE 1× buffer (Tris–acetate 40 mM and EDTA 1.0 mM). Gel bands were excised and purified using the QIAquick PCR purification kit (Qiagen). DNA sequencing was carried out using an automated ABI Prism 3730 DNA sequencer (Applied Biosystems), and the nucleotide sequences were compared by BLAST search (Altschul *et al*., 1997) to those of National Centre for Biotechnology Information (NCBI) databases.

Macroscopic characters were studied on Czapek yeast extract agar (CYA), yeast extract sucrose agar (YES) and malt extract agar (MEA). The strains were inoculated at three points on 90-mm Petri dishes and incubated for 7 days at 25 °C in darkness. All media were prepared as described by Yilmaz *et al.* (2012). *P. purpurogenum* var. *rubrisclerotium* CBS 274.95 and *P. rubrum* CBS 263.93 strains used for morphological comparison purposes were obtained from The Centraalbureau voor Schimmelcultures (CBS, Fungal Biodiversity Centre, Royal Netherlands Academy of Arts and Sciences, Utrecht)

For electron microscopy studies, strains were grown in sterilized wheat straw for 7 days and fixed in 5 mM glutaraldehyde for 12 h, 4 ºC. The aldehyde excess was removed by washing in distilled water and subsequent dehydration with increasing concentrations of ethanol (up to 100%). Samples were prepared onto small aluminum stubs, dried under low vacuum conditions, covered with sputtered gold in a Bio-Rad SC510 sputter coater and observed under high vacuum conditions using a FEI Inspect-S scanning electron microscope (SEM).

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.

Romero, E., Speranza, M., García-Guinea, J., Martínez, A.T., Martínez, M.J., Garcia-Guinea, J., Martinez, A.T., Martinez, M.J. 2007. An anamorph of the white-rot fungus *Bjerkandera adusta* capable of colonizing and degrading compact disc components. FEMS Microbiol. Lett. 275:122–9.

Yilmaz, N., Houbraken, J., Hoekstra, E., Frisvad, J., Visagie, C., Samson, R. 2012. Delimitation and characterisation of *Talaromyces purpurogenus* and related species. Persoonia. 29:39–54.

**Additional file 2.** Additional files 2 can be found in the digital version of the document (endorsed as a CD) or in the online version of the article due to its excessive sizes. Gene prediction from *T. amestolkiae* genome. The

file contains the amino acid sequences of the proteins encoded by the genome of *T. amestolkiae.*

**Additional file 3.** Additional file 3 can be found in the digital version of the document (endorsed as a CD) or in the online version of the article due to its excessive size. Proteins in secretomes of *T. amestolkiae* growing with different carbon sources: 3A) Avicel; 3B) Glucose; 3C) Slurry; 3D) Xylan. The file contains four tables with the list of proteins identified in each condition.

# CHAPTER 2

**A novel, highly efficient β-glucosidase with a cellulose-binding domain: characterization and properties of native and recombinant proteins.**

**Méndez-Líter, J.A.**, Gil-Muñoz, J., Nieto-Domínguez, M., Barriuso, J., de Eugenio, L.I., Martínez, M.J., 2017. A novel, highly efficient β-glucosidase with a cellulose-binding domain: characterization and properties of native and recombinant proteins. Biotechnol. Biofuels. 10, 256. doi: 10.1186/s13068-017-0946-2

Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

## ABSTRACT

**Background:** Cellulose, the most abundant biopolymer on earth, is an alternative for fossil fuels as a renewable feedstock for the production of second-generation (2G) biofuels and other chemicals. The discovery of novel, highly efficient β-glucosidases, remains as one of the major bottlenecks for cellulose degradation. In this context, the ascomycete *Talaromyces amestolkiae,* isolated from cereal samples, has been studied as a promising source for these enzymes.
**Results:** BGL-2 is the major β-glucosidase secreted by this fungus in the presence of cellulosic inductors. This enzyme possesses a CBD (Cellulose Binding Domain), an unusual feature among this type of proteins. Besides, when growing on cellulose, the fungus produced two different *bgl-2* mRNAs that were cloned and expressed in *Pichia pastoris*. A complete recombinant protein (BGL-2*) and its truncated form, lacking CBD (BGL-2T*), have been purified, characterized and compared with the native enzyme (BGL-2). The three BGL-2 forms presented are highly stable in a wide pH range, but BGL-2T* showed an improved thermal stability at 50 $^{o}$C after 72 h. Using *p*-nitrophenyl-β-D-glucopyranoside as a substrate, the steady-state kinetic characterization of the three proteins showed a similar $K_m$ and $k_{cat}$ for BGL-2 and BGL-2*, while the truncated protein displayed a threefold higher value for $k_{cat}$. All tested BGL-2 enzymes were as well very efficient using cellobiose and other short oligosaccharides as substrates. As biotechnological application, the recombinant *T. amestolkiae* enzymes were studied in saccharification of brewers' spent grain, being comparable to commercial β-glucosidase cocktails.
**Conclusion:** A new β-glucosidase from *T. amestolkiae* has been studied. The enzyme, containing a functional CBD, has been expressed in *P. pastoris*. The comparative analyses of the native protein and its recombinant forms, with and without CBD, suggest that they could be suitable tools for valorization of lignocellulosic biomass.

**Keywords:** Fungi; Glycosyl hydrolases; *Pichia pastoris*; Carbohydrate binding modules; Saccharification; Brewers spent grain.

# BACKGROUND

The current need of finding sustainable and renewable energy sources is making the degradation of lignocellulosic biomass to produce second-generation biofuels a blooming subject, since they may represent a viable alternative to fossil fuels (Naik *et al*., 2010; Amorim *et al*., 2011). Lignocellulosic biomass degradation requires different steps: i) pretreatment to improve polysaccharide accessibility, ii) enzymatic hydrolysis of cellulose and hemicellulose and iii) fermentation of free monosaccharides to ethanol (Lin and Tanaka, 2006). Steam explosion is the most common pretreatment method used to disrupt plant cell wall components. Although very effective, the use of high pressures and temperatures generates undesirable compounds from sugars and lignin, which produce negative effects on enzymatic hydrolysis and fermentation (Alvira *et al*., 2010).

The plant cell wall consists of three major components, whose proportions depend on the source of the lignocellulosic biomass: cellulose, hemicellulose, and lignin (Martínez *et al*., 2009). Nowadays, commercial enzymatic cocktails used for hydrolysis of lignocellulosic biomass contain cellulases, hemicellulases, and other complementary enzymes that facilitate the complete degradation of plant cell wall. Cellulose, the most abundant polymer on earth, is a polysaccharide composed of long linear chains of D-glucose linked by β-(1,4) glycosidic bonds. Besides being used for biofuels production, this polysaccharide is basic to a multitude of industrial processes like paper pulp and chemicals production (Martínez *et al*., 2009). Cellulose is enzymatically hydrolyzed by the combined action of endoglucanases (EGs) (EC 3.2.1.4), cellobiohydrolases (CBHs) (EC 3.2.1.91), and β-glucosidases (BGLs) (EC 3.2.1.21) (Payne *et al*., 2015). The effect of CBHs and EGs on cellulose generates short soluble oligosaccharides, which are converted into glucose by β-glucosidases. These enzymes are members of the glycoside hydrolases (GHs) family and attack polysaccharides like cellulose in a relatively inefficient way, as their glycosidic bonds are often inaccessible to the active site of the enzymes. This is due to the low solubility of these polysaccharides and/or to their crystalline structures. A wide variety of glycoside hydrolases that degrade such insoluble substrates have structural domains known as carbohydrate binding modules (CBMs), for example cellulose binding domains (CBDs), which are useful for the recognition and attachment of GHs to their substrates (Boraston *et al*., 2004).

Although the synergistic action of all the cellulolytic activities is required to fully degrade cellulose, β-glucosidases are considered as the key enzymes for this process as they are indispensible for releasing free glucose. β-glucosidases are generally found in little proportion in

commercial preparations, usually produced from *Trichoderma* and *Aspergillus* species (Singhania *et al*., 2010). Hence, many studies are focusing on finding robust β-glucosidases, since enzyme cocktails must be supplemented with this activity to increase the efficiency of cellulose saccharification (Singhania *et al*., 2010).

Commercial β-glucosidases are typically obtained from filamentous fungi because fungal enzymes usually have a higher catalytic productivity/efficiency than other microorganisms. In the last years, the high cellulolytic potential of various *Penicillium* species and its perfect states, *Talaromyces* or *Eupenicillium*, has been envisioned (Maeda *et al*., 2011; Singhania *et al*., 2014). In this sense, a recent work revealed that *T. amestolkiae* produces several β-glucosidases in the presence of different carbon sources (de Eugenio *et al*., 2017). Here we report the purification, heterologous expression and biochemical characterization of a β-glucosidase produced by this fungus in the presence of cellulosic substrates. In addition, we evaluated its role in saccharification of brewers spent grain.

## RESULTS AND DISCUSSION

### Production, purification and properties of native BGL-2

*Talaromyces amestolkiae* has been recently described as a fungus producing a wide variety of cellulases. Different types of these enzymes were secreted as a function of the available carbon source. On cellulosic substrates, a β-glucosidase was abundantly produced. In this work, this novel β-glucosidase has been purified from 8-days old *T. amestolkiae* cultures growing in Mandels medium containing 1% Avicel, when maximal BGL activity levels were reached. For purification, three consecutive chromatographic steps were needed. The filtered, concentrated and dialyzed crude was first loaded on a HiTrap Capto Adhere cartridge, a strong anion exchanger with multimodal functionality. Two peaks containing BGL activity were recovered indicating that, in these conditions, the fungus produced at least two β-glucosidases. This finding agrees with the results from a proteomic analysis of the secretome released in Avicel cultures of this fungus, which disclosed the production of different β-glucosidases (de Eugenio *et al*., 2017). Proteins from peak 2 (containing around 79% of initial BGL activity), were concentrated and loaded in a Mono Q column. After this chromatographic stage, again, two peaks with BGL activity were separated. The first peak, containing 80% of total BGL activity of this step, was concentrated and subjected to size exclusion chromatography in Superdex 75 HR 10/30 to complete the protein

purification. This protein was denominated BGL-2, and purified around 4-fold with a final yield of 6.3% (Table 1).

**Table 1.** Purification yield of native and recombinant BGL-2 isoforms.

| BGL-2 | | | | |
|---|---|---|---|---|
| **Step** | **Total protein (mg)** | **Total activity (U)** | **Specific activity (U/mg)** | **Yield (%)** |
| Crude extracts | 53.4 | 1,039.7 | 19.5 | 100.0 |
| HiTrap Capto Adhere | 15.8 | 795.6 | 50.4 | 76.5 |
| Mono Q 5/50 | 7.6 | 596.4 | 78.5 | 57.3 |
| Superdex 75 HR 10/30 | 0.8 | 66.1 | 82.6 | 6.3 |

| BGL-2* | | | | |
|---|---|---|---|---|
| **Step** | **Total protein (mg)** | **Total activity (U)** | **Specific activity (U/mg)** | **Yield (%)** |
| Crude extracts | 25.1 | 300.1 | 11.9 | 100.0 |
| HiTrap Capto Adhere | 7.6 | 126.5 | 16.5 | 42.2 |

| BGL-2T* | | | | |
|---|---|---|---|---|
| **Step** | **Total protein (mg)** | **Total activity (U)** | **Specific activity (U/mg)** | **Yield (%)** |
| Crude extracts | 19.9 | 436.1 | 21.8 | 100 |
| HiTrap Capto Adhere | 3.6 | 128.8 | 35.4 | 29.5 |

The molecular mass of BGL-2, estimated by SDS-PAGE, was around 100 kDa (figure 1A). Besides, although enzymatic *N*- and *O*-deglycosylation of BGL-2 did not alter its molecular mass, PAS staining showed a purple BGL-2 band, indicating its glycoprotein nature (figure 1B). This result suggests that BGL-2 may be *O*-glycosylated since enzymatic *O*-deglycosylation is not very efficient for fungal glycoproteins due to the heterogeneity of this post-translational modification, oppositely to *N*-deglycosylases that are usually very efficient because *N*-glycosylation is evolutionary conserved (Maruyama *et al.*, 1998).

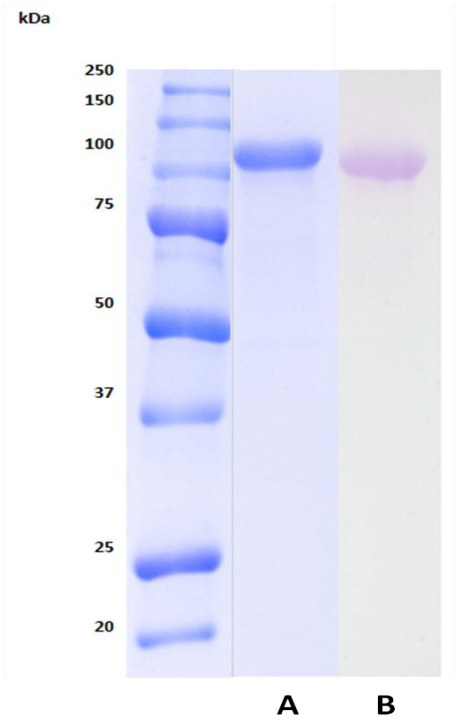



**Figure 1.** SDS-PAGE (A) and PAS staining on SDS-PAGE (B) of BGL-2 from *T. amestolkiae*.

## Sequence analysis and homology modeling of BGL-2

The peptide mass fingerprinting of this protein evidenced internal peptides that showed high similarities with a hypothetical β-glucosidase (XP_002149046) from *Talaromyces marneffei* ATCC 18224. Using TBLASTN, the sequence of these peptides was used to identify the *bgl2* gene in the genome of *T. amestolkiae*, deposited in the NCBI database with accession number of MIKG00000000.

Analysis of *bgl-2* gene indicated that this protein belongs to the GH3 family. Surprisingly, a sequence compatible with the presence of a cellulose binding domain (CBD) was found. This region is common in cellobiohydrolases and endoglucanases (Gilkes *et al*., 1991), but very rare in a β-glucosidase. The CBD was linked to the catalytic domain through a serine-threonine-rich region that would act as a flexible connector between both regions, as reported for *Trichoderma reesei* cellobiohydrolase I (Srisodsuk *et al*., 1993). The ability to bind cellulose has only been described for the β-glucosidase of *Phanerochaete chrysosporium* (Lymar *et al*., 1995), which has a CBD in its N-terminal domain, but this enzyme was finally identified as a glucan-1,3-β-glucosidase (Igarashi *et al*., 2003), with much more activity on laminaribiose and laminarin than on cellobiose. Hence, *T. amestolkiae* BGL-2 is the first 1,4-β-glucosidase where a CBD has been characterized.

Intron and exon identification was performed by comparison with transcriptomic data from the most similar sequences (XM_002485083 and XM_002149010) using BlastN. It was concluded that *bgl2* gene is

interrupted by three introns, with a sequence of 2460 bp, coding for 819 amino acids (GenBank accession number: KM393203). Intron prediction was confirmed after RNA extraction and cDNA production.

The information gathered in the databases on the crystal structures of fungal β-glucosidases from the GH3 family is very limited. Only four GH3 crystal structures are available in Protein Data Bank (PDB), three from *Hypocrea jecorina* (synonym of *Trichoderma reesei*) (4I8D, 3ZZ1 and 3ZYZ) (Karkehabadi *et al*., 2014) and one from *Aspergillus aculeatus* (4IIB) (Suzuki *et al*., 2013). As in BGL-2, the catalytic domain of these proteins is divided into three subdomains: a N-terminal domain with TIM (α/β) 8 barrel form, a C-terminal domain with α/β sandwich, and a fibronectin III-like domain, with unknown role and possible implications in thermal stability (Pozzo *et al*., 2010; Suzuki *et al*., 2013). The model structure of *T. amestolkiae* BGL-2 was obtained using structures 3ZYZ and 4I8D from PDB as templates, which showed an identity of 64.2%. QMean value in both cases was -3, indicating that the model was well adjusted to the experimental structures.

The CBD of BGL-2 was modeled separately since the fungal β-glucosidases of *H. jecorina* lack this domain. The model that best adjusted (QMean -0.51) came from a CBH of this fungus (PDB: 1CBH). Finally, the modeling of the linker between the catalytic domain and CBD was performed manually because there is no such crystallized structure deposited, probably due to its high flexibility (figure 2).

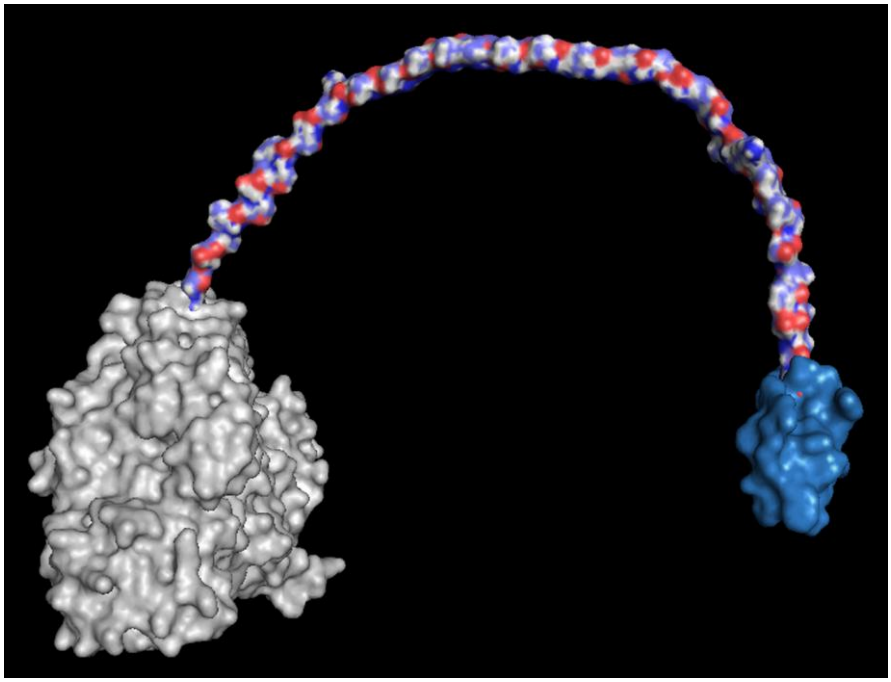


**Figure 2.** Molecular surface model of BGL-2 from *T. amestolkiae*. The catalytic domain (grey), the connector (blue, red and white) and the CBD (blue) are depicted.

The BGL-2 CBD sequence was compared with CBMs reported for other cellulases by using the protein-protein Blast database (BlastP), indicating that it belongs to family 1 (CBM1) since it had most of the conserved amino acids characteristic of this family (figure 3). With some of the best matches, a sequence alignment using the T-coffee program was performed.

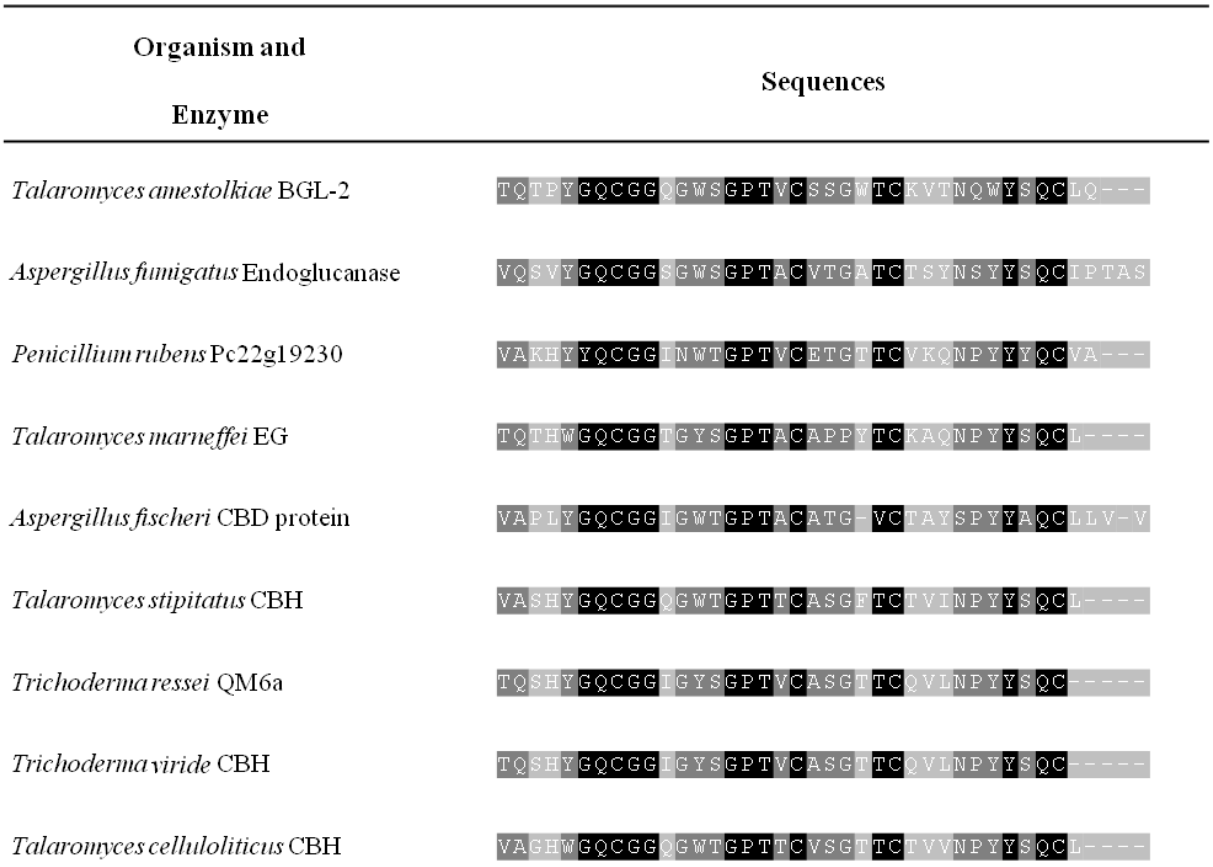



**Figure 3.** Sequence alignment of the CBD region of *T. amestolkiae* BGL-2 with other CBDs reported in different fungal cellulases using the T-coffee alignment program. The accession numbers of the compared enzymes are: XP_748707.1 (*A. fumigatus*), XP_002565826.1 (*P. rubens)*, XP_002152969.1 (*T. marneffei*), XP_001262453.1 (*A. fischeri*), and XP_002484839.1 (*T. stipitatus*), XP_006969224.1 (*T. reesei*), AAQ76092.1 (*T. viride*), and GAM33347.1 (*T. cellulolíticus*). Black color shows the conserved residues in all cases. Dark grey shows the amino acids with 1 or 2 variations. Light grey indicates the residues with 3 or more variations between the sequences.

## Heterologous expression of *bgl2*

PCR fragments containing *bgl2* cDNA obtained by reverse transcription were purified and cloned into pPIC9. Unexpectedly, two different sequences were obtained: one with the expected size for the fungal gene sequence after excision of the three predicted introns (2460 bp length) and another 65 bp longer (2525 bp), which had retained the third intron, maybe due to an alternative or wrong RNA processing. Clones with

this sequence will produce a truncated form of BGL-2 (without CBD) since this intron contains a stop codon upstream the CBD coding region. Thus, two constructs were isolated: I) a fragment that will produce a truncated form of *bgl2* without CBD, named BGL-2T*, with a coding sequence of 2202 bp and 734 amino acids; and II), the fragment that will generate the complete form of *bgl2*, including the CBD, that will be named BGL-2*, with a coding sequence of 2460 bp and 820 amino acids (Additional file 1: Figure S1). This misprocessing of mRNA in *bgl2* gene could be related to an alternative splicing phenomenon. In fact, alternative splicing mechanisms have been described in some cellobiohydrolases from *P. chrysosporium*, in which selective transcription of CBD occurs depending on the carbon source (Birch *et al*., 1995). As other filamentous fungi, *T. amestolkiae* exhibits the property of expressing different isoforms of cellulases depending on the culture conditions or carbon sources.

The pPIC9 vectors containing both cDNA fragments encoding BGL-2 and its truncated form BGL-2T* were transformed into *P. pastoris* KM71 strain. Transformants selected in YNB-His medium were plated in YPM in order to detect β-glucosidase activity after MUG-agar incubation, as described in material and methods section. The transformants that produced higher fluorescence were chosen as potential enzyme producers because they were able to hydrolyze MUG efficiently.

**Production of BGL-2* and BGL-2T* in *Pichia pastoris* and purification of the recombinant enzymes**

Figure 4 shows that recombinant *P. pastoris* yeast strains, which produce either the complete, or the truncated form of BGL-2, secreted higher β-glucosidase activity levels than *T. amestolkiae* growing in Mandels medium with Avicel as carbon source (5 and 2-fold, respectively). Although the time needed for optimal secretion was similar, the higher production levels and the few extracellular proteins secreted by *P. pastoris* strains facilitated its further purification.

Both recombinant proteins were purified in a single chromatographic step using a HiTrap Capto Adhere cartridge, with a yield of 42% and 29% for BGL-2* and BGL-2T*, respectively, which represent an increased yield of 7-fold and 5-fold compared with the native protein.

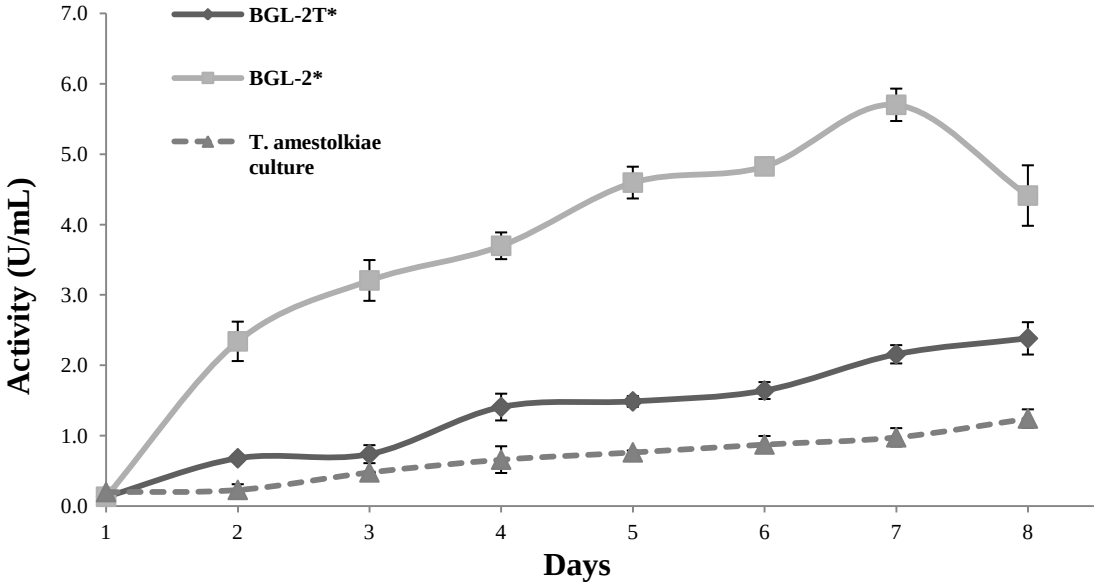



**Figure 4.** β-glucosidase activity secreted by *T. amestolkiae* in the presence of Avicel, and BGL-2* and BGL-2T* production in *P. pastoris*.

## Comparative properties of native and recombinant BGL-2 proteins

Molecular mass of BGL-2, BGL-2* and BGL-2T* were studied by MALDI-TOF MS and size exclusion chromatography. All BGL-2 forms are monomeric enzymes, with an accurate molecular mass of 97.2 kDa, 102.1 kDa and 80.8 kDa for BGL-2, BGL-2* and BGL-2T*, respectively (Additional file 1: Figure S2). The differences found between theoretical (83.1 kDa for the complete enzymes and 74.4 kDa for the truncated BGL-2) and empirical molecular masses can be attributed to protein glycosylation: 14.7%, for BGL-2, 18.7% for BGL-2* and 8% for BGL-2T*. It is noticeable that the truncated form without CBD is fairly less glycosylated than the proteins with this domain. This could be explained in terms of protein composition, since BGL-2T* also lacks the linker region, where 39 possible glycosylation sites have been identified (Additional file 1: Figure S3). Linkers are usually strongly glycosylated, preventing proteolysis of this region (Beckham *et al*., 2010; Guan *et al*., 2015).

Isoelectrofocusing indicated that the isoelectric point of purified β-glucosidases were 5.6, 7.4 and 5.2 for BGL-2, BGL2* and BGL-2T*, respectively. The major difference was found between BGL-2 and its recombinant form. The theoretical IEF value predicted from the BGL-2T* sequence was 4.89, similar to the one obtained experimentally. However, for BGL-2* the theoretical value was 5.06. These differences between theoretical and experimental isoelectric points of the complete forms of the native and the recombinant enzyme, could be attributed to their different glycosylation patterns, that can affect the isoelectric point of a given protein (Marsh *et al*., 1977).

Temperature and pH are crucial factors for the enzymatic hydrolysis of lignocellulosic biomass degradation. All BGL-2 isoforms studied were stable up to 40 ºC after 72 h. Remarkably, the lack of CBD increased the temperature stability of BGL-2T* to 50 ºC after 3 days (figure 5A).

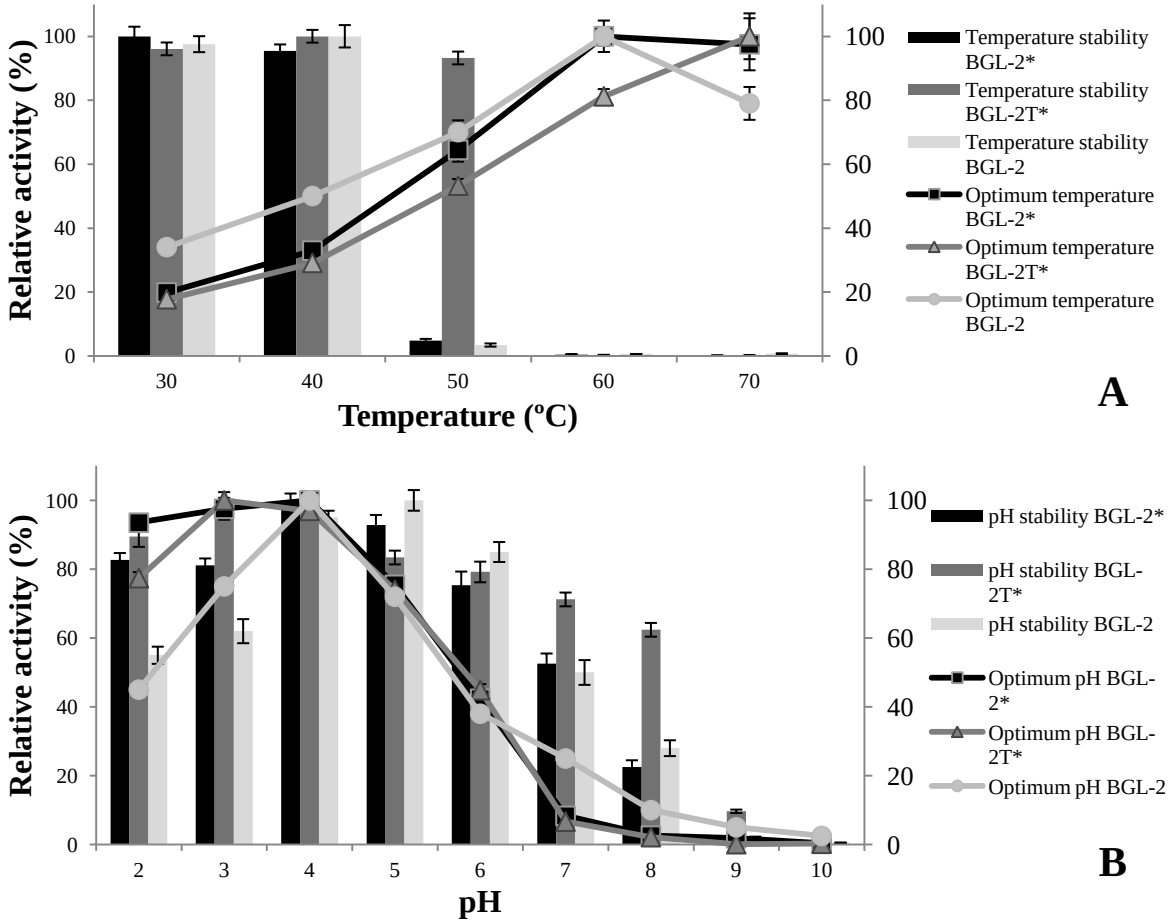



**Figure 5.** Effect of temperature **(A)** and pH **(B)** on BGL-2 enzyme activity. Lines indicate the optimum temperature or pH for enzyme activity; bars show its 72 h-stability in a range of temperatures from 30 °C to 70 °C or pH from 2 to 10.

BGL-2T* also was more stable in the most acidic and basic pH assayed (figure 5B). Optimal activity of BGL-2T* was obtained between pH 3-4 and 70 °C, while both, BGL-2 and BGL-2* showed their highest activity at a similar pH but at 60 °C. These results suggest that all BGL-2 forms studied could be useful for biotechnological applications such as 2G ethanol production from lignocellulosic wastes, especially those acid-pretreated (Sun and Cheng, 2002). Comparing with the data obtained for other native β-glucosidases, all BGL-2 forms had superior optimal temperatures (Bhatia *et al*., 2002) and were stable at broader pH ranges (Wallecha and Mishra, 2003; Toonkool *et al*., 2006). The recombinant β-glucosidases from *Thermoascus aurantiacus* (Hong *et al*., 2007), *Aspergillus fumigatus* (Liu *et al*., 2012), *Myceliophthora thermophila*

(Karnaouri *et al*., 2013), *Neosartorya fischeri* (Yang *et al*., 2014), *Penicillium funiculosum* (Ramani *et al*., 2015) and *Neurospora crassa* (Pei *et al*., 2016), expressed in *P. pastoris*, had optimum temperatures similar to BGL-2*, between 60 °C and 70 °C, with the exception of *N. crassa*, whose optimal temperature is 80 ºC. The optimal pH for these proteins was between 5 and 6, whereas the optimum pH of the recombinant BGL-2 forms is more acidic.

The circular dichroism analysis of native and recombinant proteins revealed that all BGL-2 forms have a typical spectrum of α+β folded structures. The native and recombinant full-length proteins had very similar spectra, confirming their common secondary structure. However, the spectrum of BGL-2T* lacked the minimum at 208 nm typical of proteins with CBD (Additional file 1: Figure S4). This study evidenced that the two forms of the protein do not have the same secondary structure, suggesting that BGL-2T* could have a slightly different folding. This may be one of the reasons for its increased stability to temperature and pH. A recent report has shown the negative effect of excessive *O*-glycosylation on the pH stability of a recombinant β-glucosidase from *Talaromyces leycettanus* (Xia *et al*., 2016), which is in good agreement with the lower glycosylation degree and higher stability determined for BGL-2T* against those of the full length proteins. However, the improved stability of BGL-2T* could be also explained in terms of protein composition: an endo-β-1,4-glucanase from *Bacillus subtilis* JA18 showed increased thermal stability and catalytic efficiency after CBD depletion (Wang *et al*., 2009). These authors postulated that the increased stability could be related to an enhanced refolding after thermal stress in the absence of CBD. In this sense, the deletion of CBD and its linker could result in a more compact protein, with a larger fractional polar surface, increasing hydrogen bonding density to water (Vogt *et al*., 1997).

**Cellulose binding assay**

In order to determine whether the CBD had the ability to bind cellulose, Avicel adsorption tests were performed with native and recombinant BGL-2, with and without CBD.

The results showed that native BGL-2 and BGL-2* quickly bind to Avicel, decreasing β-glucosidase activity in the supernatants, while no activity changes were observed in the supernatants of the assays with the truncated protein (figure 6). In addition, the interaction between Avicel and the proteins' CBD was strong and stable after 24 h, which suggests that this domain is functional, and could play an important role in binding to natural cellulosic substrates. Finally we have examined the CBD binding ability to other polysaccharides like xylan and chitin. The results showed that the

CBD of BGL-2 binds specifically to cellulose (Additional file 1: Figure S5).

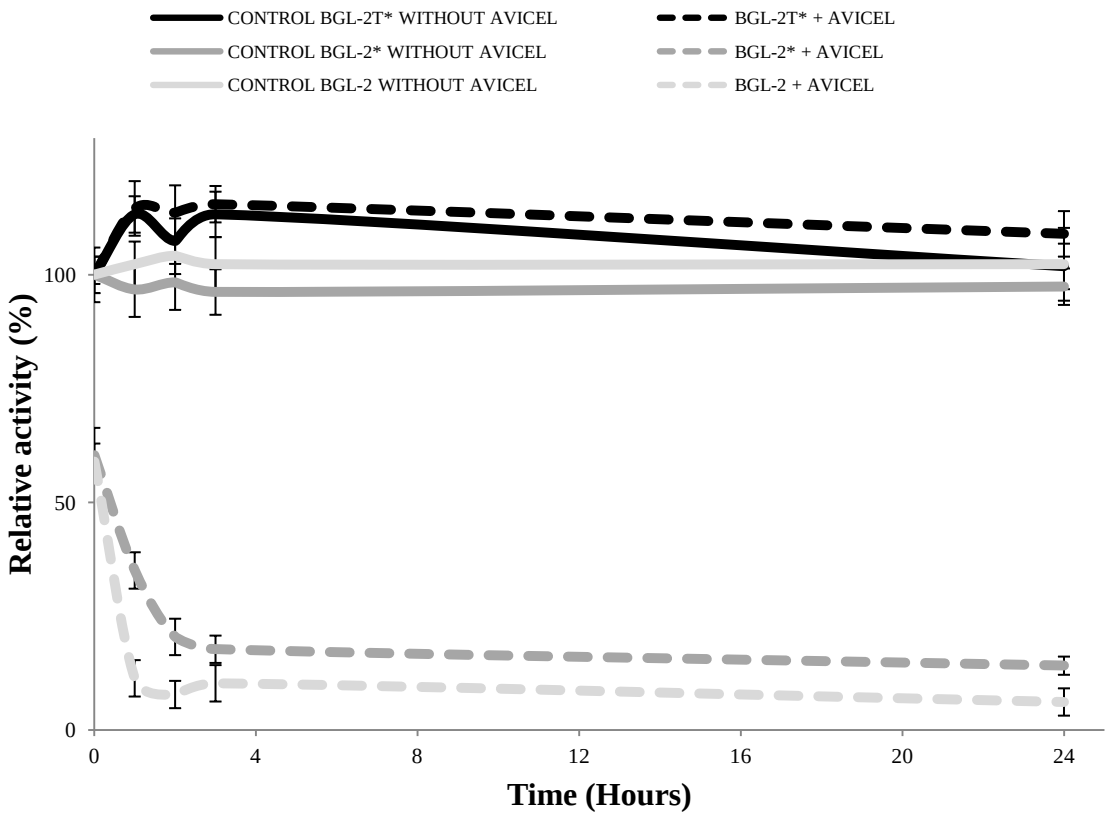



**Figure 6.** Cellulose-binding assays. The three BGL-2 forms were incubated with Avicel for 24 h at 4 ºC and residual β-glucosidase activity was measured in the supernatants.

## Substrate specificity of the native and recombinant BGL-2 forms

The three forms of BGL-2 hydrolyzed *p*NPG, *o*NPG and cellooligosaccharides from cellobiose to cellohexaose with different efficiency. As a general rule, the recombinant BGL-2 enzymes were more active and showed higher affinity towards all substrates tested. More specifically, BGL-2T* presented a superior catalytic efficiency towards all substrates but cellopentaose and cellohexaose (Table 2). In detail, it can be seen that the two recombinant proteins showed a better affinity for the longer cellooligosaccharides. The different catalytic efficiency of BGL-2* and BGL-2T* on cellooligosaccharides of different lengths could be related to the presence or absence of the CBD, since the role of CBMs in binding to cellotetraose or cellohexaose has already been demonstrated (Notenboom *et al.*, 2001; Viegas *et al.*, 2008). The higher catalytic efficiency of BGL-2T* on the shorter substrates could also be explained considering that the CBD could interfere with the access of small oligosaccharides to the catalytic site. On the other hand, the importance of the CBD to bind the enzyme to longer substrates, like cellopentaose and

cellohexaose, might favor their hydrolysis rate, being responsible for the improved catalytic efficiency of these molecules by BGL-2*. The lower effect of the native BGL-2 on all the cellooligosaccharides assayed could be due to some negative effect on the integrity of the purified enzyme across its multistep purification process.

**Table 2.** Kinetic parameters of all BGL-2 isoforms against different substrates.

| Substrate | $K_m$ (mM) | | | $k_{cat}$ ($s^{-1}$) | | | $k_{cat}/K_m$ ($mM^{-1}.s^{-1}$) | | |
|---|---|---|---|---|---|---|---|---|---|
| | BGL-2 | BGL-2* | BGL-2T* | BGL-2 | BGL-2* | BGL-2T* | BGL-2 | BGL-2* | BGL-2T* |
| *p*NPG | 0.41±0.02 | 0.19±0.01 | 0.34±0.02 | 485 | 444 | 874 | 1167 | 2243 | 2563 |
| *o*NPG | 0.86±0.04 | 1.48±0.02 | 0.67±0.04 | 174 | 59 | 217 | 201 | 39 | 323 |
| Cellobiose | 1.21±0.04 | 1.11±0.02 | 0.91±0.05 | 303 | 630 | 569 | 249 | 567 | 619 |
| Cellotriose | 1.37±0.05 | 1.87±0.13 | 1.42±0.06 | 272 | 617 | 716 | 198 | 329 | 502 |
| Cellotetraose | 1.68±0.04 | 0.92±0.01 | 0.49±0.03 | 313 | 580 | 402 | 185 | 629 | 813 |
| Cellopentaose | 1.28±0.02 | 0.71±0.03 | 0.68±0.01 | 309 | 671 | 482 | 239 | 936 | 701 |
| Cellohexaose | 1.01±0.04 | 0.51±0.01 | 0.49±0.03 | 313 | 405 | 359 | 307 | 794 | 720 |

The kinetic parameters determined for all BGL-2 isoforms, and especially the recombinant enzymes, show that these enzymes are among the most efficient BGLs discovered up to now (Additional file 1: Table S1). However, very few studies describe high BGL activities on natural substrates as cellobiose or cellooligosaccharides, despite these are their typical substrates in saccharification processes (Singhania *et al*., 2013). All the BGL-2 isoforms analyzed hydrolyzed cellooligosaccharides C2 to C6 more efficiently than all BGL studied so far (Table 3). These results put these enzymes in the forefront of the known β-glucosidases for the saccharification of these natural substrates, which is highly relevant from a biotechnological perspective. A set of aryl glycosides and disaccharides were also assayed to test the enzymes' specificity. Residual activity levels were detected towards 4-nitrophenyl β-D-xylopyranoside and 4-nitrophenyl α-D-glucopyranoside, and no activity was observed against 4-nitrophenyl β-D-galactopyranoside, 4-nitrophenyl α-D-galactopyranoside, 4-nitrophenyl α-L-rhamnopyranoside, 4-nitrophenyl β-D-fucopyranoside, maltose, lactose or sucrose. Remarkably, the three enzymes were quite active towards laminaribiose and gentiobiose (data not shown).

**Table 3.** Comparison of catalytic efficiencies against cellooligosaccharides (C2-C6) of fungal β-glucosidases.

| Organism and reference | Enzyme | Cellobiose | Cellotriose | | Cellotetraose | | Cellopentaose | | Cellohexaose | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $K_{cat}/K_m$ | $K_m$ | $V_{max}$ | $K_m$ | $V_{max}$ | $K_m$ | $V_{max}$ | $K_m$ | $V_{max}$ |
| **Native BGLs** | | | | | | | | | | |
| *Trichoderma reesei* (Guo *et al*., 2016) | Cel3A | 37.3 | 0.2 | 38.0 | | 41.0 | | | | |
| | Cel3B | 21.6 | 0.3 | 36.0 | | 36.0 | | | | |
| | Cel1A | 70.6 | 1.1 | 18.0 | | 2.6 | | | | |
| Metagenome from compost (Uchiyama *et al*., 2013) | Td2f2 | 1.6 | 3.1 | 10.1 | 1.53 | 7.8 | 8.4 | 7.9 | | |
| *T. thermosaccharolyticum* (Pei *et al*., 2012) | | 13.3 | | | | | | | | |
| *Penicillium purpurogenum* (Jeya *et al*., 2010) | | | | 111.0 | | 96.0 | | 68.4 | | |
| **BGLs expressed in *P. pastoris*** | | | | | | | | | | |
| *Aspergillus fumigatus* (Liu *et al*., 2012) | rBgl3 | 52.1 | | | | | | | | |
| *Penicillium funiculosum* (Ramani *et al*., 2015) | rBgl4 | 3610.4 | | 59.4 | | 39.3 | | 32.7 | | 23.7 |
| *Neurospora crassa* (Pei *et al*., 2016) | BGL2 | | | 12.3 | | 9.4 | | | | |
| *Myceliophthora thermophila* (Karnaouri *et al*., 2013) | MtBgl3a | 17.4 | | | | | | | | |
| **BGLs of this work** | | | | | | | | | | |
| ***T. amestolkiae*** | **BGL-2** | **249.1** | **1.37** | **167.8** | **1.69** | **193.1** | **1.2** | **190.7** | **1.0** | **193.2** |
| | **BGL-2*** | **567.1** | **1.88** | **363.2** | **0.92** | **341.5** | **0.7** | **394.8** | **0.5** | **238.7** |
| | **BGL-2T** | **619.1** | **1.43** | **532.2** | **0.50** | **298.9** | **0.6** | **358.1** | **0.5** | **266.9** |

In addition, all BGL-2 isoforms were capable of releasing reducing sugars from Avicel, carboxymethyl cellulose (CMC), and beechwood xylan (Table 4).

**Table 4.** Specific activity of BGL-2, BGL-2* and BGL-2T* against different polysaccharides.

| **Substrate** | **Specific activity (U/mg)** | | | | |
|---|---|---|---|---|---|
| | BGL-2 | BGL-2* | BGL-2T* | Celluclast 1.5L | Inactivated BGL-2* ** |
| Avicel | 11.2±0.5 | 15.8±0.4 | 6.5±0.3 | 4.4±0.2 | 0.51±0.1 |
| CMC | 9.1±0.2 | 10.0±0.4 | 5.4±0.4 | 11.3±0.4 | 0.88±0.12 |
| Xylan | 21.6±0.2 | 18.3±0.3 | 14.2±0.2 | 3.5±0.3 | 0.74±0.12 |

** boiled for 20 min

This is quite unusual since polysaccharides are typically not degraded by these enzymes. It was noticeable that, unlike for small carbohydrates, the native enzyme and the recombinant BGL-2 had similar catalytic activity. The full-length BGL-2 showed slightly more specific activity against these polymers than the truncated enzyme. Hence, although the CBD is not strictly needed for BGL-2 to hydrolyze small soluble substrates, its presence increases polysaccharide conversion, probably by improving the enzyme binding. The activity over Avicel of a fungal β-glucosidase from *A. fumigatus*, expressed in *P. pastoris*, was significantly lower (1.7 U/mg) (Liu *et al*., 2012) than that detected for BGL-2 variants from *T. amestolkiae*. On the contrary, a β-glucosidase from *P. purpurogenum* (Jeya *et al*., 2010) showed higher specific activity, but it was inactive against other polysaccharides like xylan or CMC. In general, most investigations report that BGLs have no activity on polysaccharides. The ability of BGL-2 to hydrolyze these substrates shows its versatility to be used in industrial applications.

## Brewers spent grain saccharification

The enzymatic conversion of cellulose to glucose is a crucial step in the production of bioethanol from lignocellulosic biomass. Since commercial cellulolytic cocktails based on enzymes from *T. reesei*, like Celluclast, are usually deficient in β-glucosidase activity, they are supplemented with this enzymatic activity for lignocellulosic biomass treatment. β-glucosidases from *Penicillium* sp. or *Talaromyces* sp. have successfully been applied for saccharification of cellulosic materials, being

more efficient than *Trichoderma* sp. enzyme preparations, since they have higher β-glucosidase levels (Gusakov, 2011). For saccharification of brewers spent grain, the synergistic effect of Celluclast 1.5L and a β-glucosidase-rich supplement (the native or recombinant isoforms of *T. amestolkiae* BGL-2 or the commercial preparation NS-50010) were evaluated.

The results shown in Figure 7 indicate that supplementation with any of the BGL-2 forms enhanced the saccharification of brewers spent grain, more than using only Celluclast (8% with BGL-2, 32% with BGL-2* and 24% for BGL-2T). The minor effect of the native *T. amestolkiae* protein could be related to their kinetic constants, slightly worse than those determined for the recombinant proteins against the substrates assayed. The differences found when Celluclast was supplemented with BGL-2T*, could indicate that the CBD could have a relevant role in binding cellulose in natural substrates increasing its degradation efficiency. Interestingly, the supplementation with BGL-2* showed a similar yield than with NS-50010. This is remarkable since BGL-2* is a purified protein, whereas NS-50010 is an enzyme cocktail, containing other cellulolytic and hemicellulolytic activities which could promote the yield of the saccharification process. Hence, this results suggest that both recombinant *T. amestolkiae* β-glucosidases, especially BGL-2*, could be suitable proteins in the valorization process of lignocellulosic biomass.

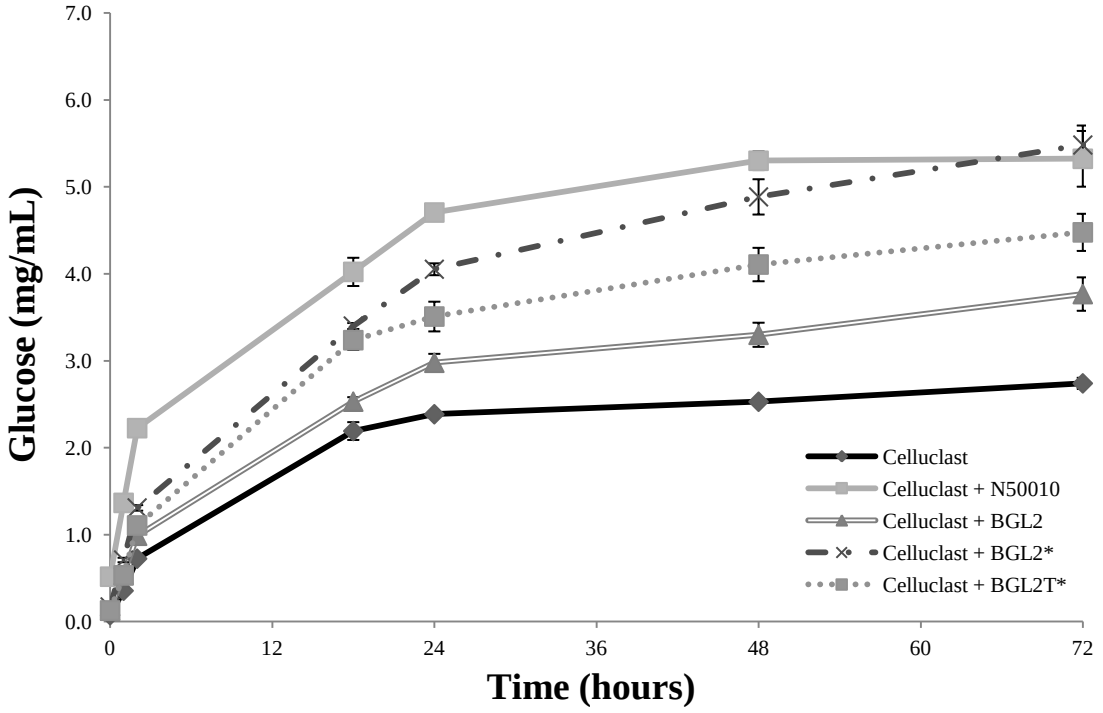



**Figure 7.** Glucose yield from saccharification of brewers spent grain. Addition of one of the three BGL-2 enzymes or NS-50010 to Celluclast 1.5 L improved substrate degradation.

## CONCLUSIONS

β-glucosidases are very versatile enzymes that play an essential role in the enzymatic hydrolysis of plant biomass for the production of 2G biofuels. In the present work, BGL-2 from *T. amestolkiae* was discovered, purified and characterized, standing out for being the first 1,4-β-glucosidase with a functional cellulose binding domain, similar to others found in CBHs and EGs. Besides, an isoform without CBD was also isolated and characterized. The work has revealed the high potential of the native BGL-2 and its recombinant forms, with and without CBD, to be applied for the saccharification of plant biomass. The complete characterization of the cellulolytic system of this fungus is currently being carried out, with special interest in the purification of other β-glucosidases produced by this strain, to evaluate their properties and features to be used in different biotechnological applications.

## MATERIALS AND METHODS

### Microorganism and culture conditions

The fungus used in this work, isolated from cereal samples by the group of Dr. Covadonga Vazquez (Department of Microbiology, Faculty of Biology, Complutense University of Madrid), was identified and included in the Collection of the Institute Jaime Ferrán of Microbiology (IJFM) with the number A795. The fungus was grown in potato dextrose agar (PDA) for 5 days at 28 ºC. Spore suspensions were obtained by placing 1 $cm^2$ agar slants in 5 mL of a 1% NaCl solution, and 0.1% of Tween 80. The mixture was shaken and 200 µL were used to inoculate 250 mL flasks with 50 mL of CSS medium (40 g/L glucose, 0.4 g/L $FeSO_4 \cdot 7H_2O$, 9 g/L $(NH_4)_2SO_4$, 4 g/L $K_2HPO_4$, 26.3 g/L corn steep solid, 7 g/L $CaCO_3$, and 2.8 mL/L soybean oil). The culture was incubated at 28 °C and 250 rpm for 5 days.

In order to grow *T. amestolkiae* for enzyme production, 2 mL from a CSS culture were inoculated in Mandels medium (Mandels and Weber, 1969), composed by: 2.0 g/L $KH_2PO_4$, 1.3 g/L $(NH_4)_2SO_4$, 0.3 g/L urea, 0.3 g/L $MgSO_4 \cdot 7H_2O$, 0.3 g/L $CaCl_2$, 5 mg/L $FeSO_4 \cdot 7H_2O$, 1.6 mg/L $MnSO_4 \cdot H_2O$, 1.4 mg/L $ZnSO_4 \cdot 7H_2O$, and 1 g/L Bacto peptone. The pH was adjusted to 4.5. This medium was supplemented with 1% Avicel (microcrystalline cellulose) as carbon source and inducer of cellulolytic activities. The inoculated flasks were incubated at 28 °C and 250 rpm for 8 days, taking daily culture samples for analytical determinations.

*Escherichia coli* DH5α (Invitrogen) was used for plasmid propagation. It was grown in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, and 15 g/L agar), at 37 ºC, overnight. For the growth and selection of

ampicillin resistant transformants, this antibiotic was included in the LB medium, as its sodium salt, in a final concentration of 100 mg/L, and bacteria were grown overnight at 37 ºC.

The heterologous expression of β-glucosidases was performed using *P. pastoris* KM71 strain (Invitrogen) as the host, which was preserved in YPD medium plates (10 g/L Yeast extract, 20 g/L peptone, 20 g/L glucose and 10 g/L of agar). The transformants were screened in a selective medium, YNB-His (20 g/L glucose, 6,7 g/L YNB, 1,92 g/L Yeast synthetic drop-out medium supplements without histidine (Sigma-Aldrich), and 10 g/L agar), and cultured 72 h at 28 ºC. Finally, for recombinant protein production, YEPS medium was used (20 g/L peptone, 10 g/L yeast extract, 10 g/L sorbitol, and 100 mM potassium phosphate buffer, pH 6), with daily addition of 6.5 mL/L of methanol as inducer. Cultures were incubated for 7 days, at 28 °C and 250 rpm, taking samples daily to measure protein production. All experiments were performed in triplicate.

**Purification and characterization of native and recombinant proteins**

All protein purifications were performed using an AKTA Purifier HPLC system (GE Healthcare Life Sciences). The native BGL-2 protein from *T. amestolkiae* was purified from protein extracts obtained after fungus growth in Mandels medium with 1% Avicel after 8 days of incubation, when maximal activity was detected. Culture supernatants were filtered, concentrated, and dialyzed though a 10 kDa cutoff membrane against 10 mM sodium phosphate pH 6, which is the buffer used across all the chromatographic steps required to purify the protein. The concentrate crude was first applied onto a Capto Adhere HiTrap cartridge (GE Healthcare Life Sciences) equilibrated with buffer using a 1 mL/min flow. Peak 1 eluted during the NaCl gradient (from 0 to 0.5 M), while peak 2 eluted once the gradient had finished, during reequilibration of the cartridge in the starting buffer. Fractions with high β-glucosidase activity were pooled, concentrated and tested for purity. The second purification step involved anion-exchange chromatography in a high-resolution Mono Q column (GE Healthcare Life Sciences) equilibrated with 10 mM sodium phosphate buffer, pH 6, at a flow rate of 0.8 mL/min. The retained proteins were eluted with a linear NaCl gradient (0 - 0.25 M over 50 min). Then the column was washed with 1 M NaCl for 7 min and equilibrated in the initial conditions. As above, the fractions with BGL activity were pooled, concentrated, and tested for purity. A third step of size exclusion chromatography on Superdex 75 HR 10/30 (GE Healthcare Life Sciences) was required to achieve BGL-2 purification. The column was equilibrated and the proteins eluted in the same buffer with 100 mM NaCl to avoid unspecific interactions, at a 0.3 mL/min flow.

In the case of the recombinant enzymes, the purification was achieved in a single chromatographic step. *P. pastoris* cultures producing the maximal BGL activity (8 days-old) were centrifuged and the supernatant was then concentrated, dialyzed and applied to a Capto Adhere HiTrap cartridge (GE Healthcare Life Sciences) in the same conditions as native protein. Once the NaCl gradient (from 0 to 0.5 M) was finished and the cartridge reequilibrated, a peak with pure BGL-2 isoforms was obtained.

In all chromatographic steps, protein concentration ($A_{280\ nm}$) and β-glucosidase activity, using *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG) as substrate, were measured. Protein homogeneity was checked after each purification step by 10% SDS-PAGE staining with Coomassie brilliant blue R-250. The approximate molecular mass of the proteins was calculated by this technique comparing the migration of the bands with those of the molecular weight markers provided by Bio-Rad. The accurate determination of the protein molecular mass was done by MALDI-TOF.

**Identification of BGL-2 by peptide mass fingerprinting**

Purified proteins were loaded on a 10% SDS-PAGE gel that, after the electrophoretic run, was stained with Sypro Ruby. Small protein pieces were digested to analyze tryptic peptide mixture (Shevchenko *et al*., 2007). MALDI-MS and MALDI-MS/MS data were obtained automatically in a mass spectrometer MALDI-TOF/TOF Autoflex III (Bruker Daltonics) equipped with a laser and a Smartbeam LIFT-MS/MS device. The data obtained were combined using the 3.0 BioTools (Bruker Daltonics) software, and mass values from trypsin, keratin, matrix or sodium adducts were removed. Data analysis was performed against the NCBInr database (National Center for Biotechnology Information non-redundant) with the 2.3 MASCOT search engine (Matrix Science). Relevant search parameters were set as follows: trypsin as enzyme, carbamidomethylation of cysteines as fixed modification, methionine oxidation as variable modification, 1 missed cleavage allowed, peptide tolerance of 50 ppm, and MS/MS tolerance of 0.5 Da. Protein scores greater than 75 were considered significant.

**Nucleic acid isolation, PCR and RT-PCR methods**

Mycelium from cultures grown in Mandels medium was used for genomic DNA and RNA extraction after filtration of 8 days old cultures with 0.8 μm nitrocellulose filters. DNA and RNA extraction were carried out using DNeasy Plant Mini Kit and RNeasy Plant Mini Kit (Qiagen) respectively, according to the manufacturer instructions. The extracted nucleic acids were quantified using a NanoDrop ND-100 (Thermo Scientific).

Isolated transcripts were converted to cDNA using the Superscript II Reverse Transcriptase RT-PCR kit (Invitrogen) using 50 µM random hexamers. PCR amplifications were performed in a thermocycler Mastercycler pro S (Eppendorf) using genomic DNA as template. Reaction mixtures were subjected to an initial denaturation at 95 °C for 5 min, followed by 36 cycles of amplification of 95 °C of denaturation for 45 s, 50 °C for 45 s of primer annealing step, and 72 °C for 3 min of elongation, followed by a final extension step at 72 °C for 10 min. The amplified sequences were separated in a 0.8% (w/v) agarose electrophoresis gel stained with GelRed, cut out, and purified using a QIAquick gel extraction kit (Qiagen). For amplifying *bgl2* gene from cDNA, two primers with restriction sites were used: *Fw-BGL2SnaBI* (5'TACGTACAGTCAGCTTCTTGGTCCGCAG3') and *RV-NotIBGL2* (5'GCGGCCGCCTATTGTAGGCATTGAGAATACCACTGATTC3'). PCR reaction mixtures contained: 1x PCR Buffer; 1.5 mM $MgCl_2$; 0.25 mM dNTPs; 0.25 mM forward and reverse primers; 100 ng of DNA template; and 0.05 U/mL of Taq polymerase.

**Plasmid construction, *E. coli* propagation and heterologous expression**

For *bgl2* heterologous expression, the amplification product was digested with *SnaBI* and *NotI* restriction enzymes, and was ligated to pPIC9 vector, previously digested with the same enzymes, using T4 phage ligase (Promega). Plasmids which contain *bgl2* wild type and *bgl2* retaining the third intron were built and sequenced using the BigDye Terminator v3.1 cycle sequencing kit and the automated ABI Prism 3730 DNAsequencer, in Secugen (Madrid, Spain).

The plasmids obtained from *E. coli* positive clones were linearized with *SalI*, and transformed into the KM71 *P. pastoris* strain by the lithium chloride transformation method, according to the *Pichia* expression kit (Invitrogen). The transformants were screened by histidine auxotrophy in YNB-His medium.

Selected clones on YNB-His plates were cultured on YPM medium (10 g/L yeast extract, 10 g/L peptone, and 10 g/L of agar; after autoclaving the medium, 15 mL/L sterile methanol was added), in order to identify positive transformants. The yeasts were cultured 24 h at 28 °C. The plates were covered with a solution of 50 mM 4-methylumbelliferyl-β-D-glucopyranoside (MUG) in 100 mM sodium acetate buffer and 0.8% agar. After solidifying, the plate was introduced in an oven at 50 °C for 20 min, and positive clones were revealed in a transilluminator Bio Rad Gel Doc XR. Clones with β-glucosidase activity hydrolyze MUG, releasing methylumbelliferyl, which is fluorescent under UV wavelength (Ramani *et al*., 2015).

**Physicochemical properties and homology modeling**

The accurate molecular mass and homogeneity of the pure enzyme were analyzed by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF) using an Autoflex III instrument (Bruker Daltonics).

In order to know if the β-glucosidases of *T. amestolkiae* were glycoproteins, the PAS (Periodic Acid-Schiff) staining method was used using the Schiff reagent (Sigma), following the manufacturer instructions. The glycoproteins exhibit visible red-purple bands after aproximately 20 min. *N*- and *O*-deglycosylation assays of purified proteins were carried out by using Endoglycosidase H or *O*-Glycosidase (Roche), according to the manufacturer instructions. Differences in molecular mass before and after deglycosylation were analyzed by SDS-PAGE electrophoresis, as described above. The possible glycosylation sites of BGL-2 were also analyzed with NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). Protein sequence was also submitted to the SignalP 4.1 server for identifying and locating the signal peptide, which was excluded from the mass prediction.

The isoelectric point of the native and recombinant proteins was determined by isoelectrofocusing on 5% polyacrylamide gels using pH 3 to 10 ampholytes (GE Healthcare), with 1 M $H_3PO_4$ and 1 M NaOH as anode and cathode buffers, respectively. The pH gradient was measured directly on the gel using a contact electrode (Crison). β-glucosidase activity was detected after incubation of the gels with 2 mM *p*-methylumbelliferyl-β-D-glucopyranoside (Sigma-Aldrich), with fluorescence visualized under UV light by use of the Gel Doc XR+ system (Bio-Rad).

The effect of pH on the BGLs was assessed using 100 mM Britton and Robinson buffer (Britton and Robinson, 1931), which can be adjusted to a broad range of pH (2-10) Optimal pH of the pure BGLs was tested using *p*NPG as substrate in the same buffer, and to study their stability to pH, the samples were incubated from pH 2 to pH 10 for 72 h at 4 °C. The effect of temperature on enzymes activity was analyzed subjecting solutions of the pure enzymes in 100 mM sodium acetate buffer, pH 4 to temperatures between 30 ºC and 80 °C for 10 min to determine their optimal temperature, and at 30 ºC and 70 ºC for 72 h to evaluate their thermostability.

Three dimensional models were generated by homology using the SWISS-MODEL server (Schwede *et al*., 2003), based on sequence similarity. The QMEAN index is used to select the best models, which are ideal when the number is close to zero. The PyMol v0.99 (Delano, 2004) program was used to visualize and analyze the structures.

**Circular dichroism spectroscopy**

Circular dichroism analysis of the native and recombinant BGL-2 forms was performed in order to compare their secondary structure. Measurements were carried out using a JASCO J-720 spectropolarimeter. Far-UV spectra (190-260 nm) were recorded in a 0.1 cm path length quartz cell at a protein concentration of 0.1 mg/mL in 10 mM phosphate buffer. The spectra from five scans were averaged and corrected for the baseline contribution of the buffer. The observed ellipticities were converted into mean residue ellipticities (θ) based on a mean molecular mass (per amino acid residue) of 110 Da.

**Protein quantification, enzyme assays and substrate specificity**

In this study all enzyme assays were performed in the presence of BSA (0.1%), a protein that does not affect the catalytic activity of the BGL but prevents missing activity when working at low enzyme concentrations (Nieto-Domínguez *et al*., 2015).

Total protein was estimated by the bicinchoninic acid assay (BCA) method, using bovine serum albumin as standard. BGL activity was usually determined at 60 °C versus *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG) in 50 mM acetate buffer pH 4, and 0.2% of BSA (standard assay). The reaction was stopped with 500 µL of sodium carbonate (2% w/v), and the release of *p*-nitrophenol (*p*NP) was measured in a spectrophotometer at 410 nm. One unit of BGL activity was defined as the amount of enzyme capable of hydrolyzing 1 µmol of *p*NPG to glucose and *p*NP per minute. The same conditions were used to measure hydrolysis of *o*-nitrophenyl-β-D-glucopyranoside (*o*NPG) and other nitrophenol derivatives.

BGLs were also incubated with different cellooligosacharides. Its activity was determined against cellobiose, cellotriose, cellotetraose, cellopentaose, and cellohexaose, in 10 min-reactions carried out in sodium acetate 100 mM, pH 4, mixing at 1200 rpm. The released glucose was measured using the Glucose-TR (Spinreact) commercial kit, according to the manufacturer's instructions.

The kinetic constants of the purified BGLs were determined by incubating the enzymes at their optimal pH and temperatures. The following substrates were analyzed over the range of concentrations stated in each particular case: *p*NPG (from 10 µM to 5 mM), *o*NPG (40 µM to 20 mM), cellobiose (80 µM to 40 mM), cellotriose (80 µM to 40 mM), cellotetraose (80 µM 40 mM), cellopentaose (40 µM to 20 mM), and cellohexaose (20 µM to 10 mM). The values of $K_m$ and $V_{max}$ were determined from Lineweaver-Burk linear equation of the Michaelis-Menten model, using the program SigmaPlot.

The activity of BGL-2, BGL-2* and BGL-2T* was also determined against different polysaccharides, prepared in 50 mM sodium acetate buffer

pH 4: 1.25% (w/v) Avicel (microcrystalline cellulose), 3% (w/v) CMC, and 3% (w/v) beechwood xylan. The reaction mixture was incubated in a heating block at 1200 rpm for 10 min. After the incubation time, the reducing substances released was determined by the Somogyi-Nelson method (Somogyi, 1945), measuring $A_{540\ nm}$.

A control without enzyme was included in all assays, to substrate the absorbance due to reactants. Also, a negative control with inactivated (boiled 20 min) enzyme was included to discard unspecific interactions with substrates. Celluclast 1.5L (Novozymes) was included as positive control in CMCase, Avicelase and xylanase assays.

**Cellulose binding assay**

In order to evaluate the ability of native BGL-2 and its recombinant forms to bind microcrystalline cellulose and the participation of the CBD in this interaction, adsorption tests on Avicel were performed (Chundawat *et al*., 2011). 20-30 ng of the purified enzymes were mixed with 500 µL of 1% Avicel (w/v) in 55 mM sodium acetate buffer pH 4. The reaction was carried out at 1200 rpm and 4 °C for 24 h. Aliquots were taken at different times (10 min, 1, 2, 3 and 24 h) after centrifuging the samples for 1 min at 14,000 rpm, before measuring residual BGL activity in the supernatants.

**Brewers spent grain saccharification**

The efficiency of the different BGL-2 forms as β-glucosidase supplements for the saccharification of lignocellulosic residues was tested following the release of glucose from brewers spent grain. Mixtures of Celluclast 1.5L and either purified BGL-2 enzymes or NS-50010, a commercial cocktail with high β-glucosidase activity, were used as catatalysts. Brewers spent grain (100 mg) was treated with 2 mL of an enzymatic solution in 100 mM sodium acetate buffer, pH 4, containing a total of 2 BGL U/mL: 1 U/mL from Celluclast 1.5L and 1 U/mL from the BGL sources compared (NS-50010 or purified BGL-2 enzymes). Reactions were performed in a heat block at 50 ºC and 1,200 rpm for 72 h. Free glucose was measured at different reaction times using the Glucose-TR commercial kit (Spinreact).

**Declarations**

**Authors' contributions**

JM, JG, MN and LdE contributed to design the research and participated in the experiments. JM, LdE and JG worked with native BGL-2. JM developed the work with recombinant enzymes. JB performed the homology modeling of the enzyme. JM drafted the manuscript. MJM and LdE co-coordinated the research and reviewed and edited the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**
The authors thank the Proteomics and Genomics facility at CIB, Novozymes for providing the commercial cocktails, and Heineken for supplying brewers spent grain.

**Competing interests**
The authors declare that they have no competing interests.

**Availability of supporting data**
*T. amestolkiae* whole genome shotgun project is deposited at DDBJ/ENA/GenBank under the accession number MIKG00000000. BGL-2 sequence is deposited in GenBank under the accession number KM393203.

**Funding**
This work has been funded by projects RTC-2014-1777-3 and BIO2015-68387-R from MINECO and S2013/MAE-2907 from Comunidad de Madrid. J.A. Méndez-Líter thanks its fellowship to Tatiana Pérez de Guzmán el Bueno Foundation. M. Nieto-Domínguez thanks its FPU fellowship to MINECO.

**Authors' information**
Not applicable

## REFERENCES

Alvira, P., Tomás-Pejó, E., Ballesteros, M., Negro, M.J., 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. Bioresour. Technol. 101, 4851-4861.

Amorim, H.V., Lopes, M.L., de Castro Oliveira, J.V., Buckeridge, M.S., Goldman, G.H., 2011. Scientific challenges of bioethanol production in Brazil. Appl. Microbiol. Biotechnol. 91, 1267.

Beckham, G.T., Bomble, Y.J., Matthews, J.F., Taylor, C., Resch, M.G., Yarbrough, J.M. *et al*., 2010. The *O*-Glycosylated linker from the *Trichoderma reesei* family 7 cellulase is a flexible, disordered protein. Biophys. J. 99, 3773-3781.

Bhatia, Y., Mishra, S., Bisaria, V.S., 2002. Microbial beta−glucosidases: Cloning, properties, and applications. Crit. Rev. Biotechnol. 22, 375-407.

Birch, P.R.J., Sims, P.F.G., Broda, P., 1995. Substrate-Dependent Differential Splicing of Introns in the Regions Encoding the Cellulose-Binding Domains of 2 Exocellobiohydrolase-I-Like Genes in *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. 61, 3741-3744.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J., 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382, 769.

Chundawat, S.P.S., Bellesia, G., Uppugundla, N., Costa Sousa, L., Gao, D., Cheh, A.M. *et al*., 2011. Restructuring the Crystalline Cellulose Hydrogen Bond Network Enhances Its Depolymerization Rate. J. Am. Chem. Soc. 133, 11163-11174.

de Eugenio, L.I., Méndez-Líter, J.A., Nieto-Domínguez, M., Alonso, L., Gil-Muñoz, J., Barriuso, J. *et al*., 2017. Differential β−glucosidase expression as a function of carbon source availability in *Talaromyces amestolkiae*: a genomic and proteomic approach. Biotechnol. Biofuels. 10, 161.

Delano, W.L., 2004. Use of PYMOL as a communications tool for molecular science. Abstr. Papers. Am. Chem. Soc. 228, U313-U314.

Gilkes, N.R., Claeyssens, M., Aebersold, R., Henrissat, B., Meinke, A., Morrison, H.D. *et al*., 1991. Structural and Functional-Relationships in 2 Families of Beta−1,4-Glycanases. Eur. J. Biochem. 202, 367-377.

Guan, X., Chaffey, P.K., Zeng, C., Greene, E.R., Chen, L., Drake, M.R. *et al*., 2015. Molecular-scale features that govern the effects of *O*-glycosylation on a carbohydrate-binding module. Chem. Sci. 6, 7185-7189.

Guo, B., Sato, N., Biely, P., Amano, Y., Nozaki, K., 2016. Comparison of catalytic properties of multiple β-glucosidases of *Trichoderma reesei*. Appl. Microbiol. Biotechnol. 100:4959–68.

Gusakov, A.V., 2011. Alternatives to *Trichoderma reesei* in biofuel production. Trends Biotechnol. 29, 419-425.

Hong, J., Tamaki, H., Kumagai, H., 2007. Cloning and functional expression of thermostable β−glucosidase gene from *Thermoascus aurantiacus*. Appl. Microbiol. Biotechnol. 73, 1331-1339.

Igarashi, K., Tani, T., Kawai, R., Samejima, M., 2003. Family 3 beta−glucosidase from cellulose-degrading culture of the white-rot fungus *Phanerochaete chrysosporium* is a glucan 1,3-beta−glucosidase. J. Biosci. Bioeng. 95, 572-576.

Jeya, M., Joo, A.R., Lee, K.M., Tiwari, M.K., Lee, K.M., Kim, S.H. *et al*., 2010. Characterization of β−glucosidase from a strain of *Penicillium purpurogenum* KJS506. Appl. Microbiol. Biotechnol. 86, 1473-1484.

Karkehabadi, S., Helmich, K.E., Kaper, T., Hansson, H., Mikkelsen, N.E., Gudmundsson, M. *et al*., 2014. Biochemical Characterization and Crystal Structures of a Fungal Family 3 beta−Glucosidase, Cel3A from *Hypocrea jecorina*. J. Biol. Chem. 289, 31624-31637.

Karnaouri, A., Topakas, E., Paschos, T., Taouki, I., Christakopoulos, P., 2013received. Cloning, expression and characterization of an ethanol tolerant GH3 β−glucosidase from *Myceliophthora thermophila*. Peer J. 1, e46.

Lin, Y., Tanaka, S., 2006. Ethanol fermentation from biomass resources: current state and prospects. Appl. Microbiol. Biotechnol. 69, 627-642.

Liu, D., Zhang, R., Yang, X., Zhang, Z., Song, S., Miao, Y. *et al*., 2012. Characterization of a thermostable β−glucosidase from *Aspergillus fumigatus* Z5, and its functional expression in *Pichia pastoris* X33. Microb. Cell. Fact. 11, 25-25.

Lymar, E.S., Li, B., Renganathan, V., 1995. Purification and Characterization of A Cellulose-Binding Beta−Glucosidase from Cellulose-Degrading Cultures of *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. 61, 2976-2980.

Maeda, R.N., Serpa, V.I., Rocha, V.A.L., Mesquita, R.A.A., Anna, L.M.M.S., de Castro, A.M. *et al*., 2011. Enzymatic hydrolysis of pretreated sugar cane bagasse using *Penicillium funiculosum* and *Trichoderma harzianum* cellulases. Process Biochem. 46, 1196-1201.

Mandels, M., Weber, J., 1969. Production of Cellulases. Adv. Chem. 95. pp. 391–414.

Marsh, J.W., Denis, J., Wriston, J.C., 1977. Glycosylation of *Escherichia−Coli* L-Asparaginase. J. Biol. Chem. 252, 7678-7684.

Martínez, A.T., Ruiz-Dueñas, F.J., Martínez, M.J., del Río, J.C., Gutierrez, A., 2009. Enzymatic delignification of plant cell wall: from nature to mill. Curr. Opin. Biotechnol. 20, 348-357.

Maruyama, N., Katsube, T., Wada, Y., Oh, M.H., De la Rosa, A.P.B., Okuda, E. *et al*., 1998. The roles of the *N*-linked glycans and extension regions of soybean beta–conglycinin in folding, assembly and structural features. Eur. J. Biochem. 258, 854-862.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A comprehensive review. Renew. Sustain. Energy Rev. 14, 578-597.

Nieto-Domínguez, M., de Eugenio, L.I., Barriuso, J., Prieto, A., Fernández de Toro, B., Canales-Mayordomo, Á. *et al*., 2015. Novel pH-Stable Glycoside Hydrolase Family 3 β–Xylosidase from *Talaromyces amestolkiae*: an Enzyme Displaying Regioselective Transxylosylation. Appl. Environ. Microbiol. 81, 6380-6392.

Notenboom, V., Boraston, A.B., Chiu, P., Freelove, A.C.J., Kilburn, D.G., Rose, D.R., 2001. Recognition of cello-oligosaccharides by a family 17 carbohydrate-binding module: an X-ray crystallographic, thermodynamic and mutagenic study1. J. Mol. Biol. 314, 797-806.

Payne, C.M., Knott, B.C., Mayes, H.B., Hansson, H., Himmel, M.E., Sandgren, M. *et al*., 2015. Fungal Cellulases. Chem. Rev. 115, 1308-1448.

Pei, X., Zhao, J., Cai, P., Sun, W., Ren, J., Wu, Q. *et al*., 2016. Heterologous expression of a GH3 β–glucosidase from *Neurospora crassa* in *Pichia pastoris* with high purity and its application in the hydrolysis of soybean isoflavone glycosides. Protein Expr. Purif. 119, 75-84.

Pei, J., Pang, Q., Zhao, L., Fan, S., Shi, H., 2012. *Thermoanaerobacterium thermosaccharolyticum* β-glucosidase: a glucose-tolerant enzyme with high specific activity for cellobiose. Biotechnol. Biofuels.5:31.

Pozzo, T., Pasten, J.L., Karlsson, E.N., Logan, D.T., 2010. Structural and Functional Analyses of beta–Glucosidase 3B from *Thermotoga neapolitana*: A Thermostable Three-Domain Representative of Glycoside Hydrolase 3. J. Mol. Biol. 397, 724-739.

Ramani, G., Meera, B., Vanitha, C., Rajendhran, J., Gunasekaran, P., 2015. Molecular cloning and expression of thermostable glucose-tolerant β–glucosidase of *Penicillium funiculosum* NCL1 in *Pichia pastoris* and its characterization. J. Ind. Microbiol. Biotechnol. 42, 553-565.

Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31, 3381-3385.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M., 2007. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protocols. 1, 2856-2860.

Singhania, R.R., Patel, A.K., Sukumaran, R.K., Larroche, C., Pandey, A., 2013. Role and significance of beta–glucosidases in the hydrolysis of cellulose for bioethanol production. Bioresour. Technol. 127, 500-507.

Singhania, R.R., Saini, J.K., Saini, R., Adsul, M., Mathur, A., Gupta, R. *et al*., 2014. Bioethanol production from wheat straw via enzymatic route employing *Penicillium janthinellum* cellulases. Bioresour. Technol. 169, 490-495.

Singhania, R.R., Sukumaran, R.K., Patel, A.K., Larroche, C., Pandey, A., 2010. Advancement and comparative profiles in the production technologies using solid-state and submerged fermentation for microbial cellulases. Enz. Microb. Technol. 46, 541-549.

Somogyi, M., 1945. A New Reagent for the Determination of Sugars. J. Biol. Chem. 160, 61-68.

Srisodsuk, M., Reinikainen, T., Penttila, M., Teeri, T.T., 1993. Role of the Interdomain Linker Peptide of *Trichoderma reesei* Cellobiohydrolase-I in Its Interaction with Crystalline Cellulose. J. Biol. Chem. 268, 20756-20761.

Sun, Y., Cheng, J.Y., 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour. Technol. 83, 1-11.

Suzuki, K., Sumitani, J.i., Nam, Y.W., Nishimaki, T., Tani, S., Wakagi, T. *et al*., 2013. Crystal structures of glycoside hydrolase family 3 beta−glucosidase 1 from *Aspergillus aculeatus*. Biochem. J. 452, 211-221.

Toonkool, P., Metheenukul, P., Sujiwattanarat, P., Paiboon, P., Tongtubtim, N., Ketudat-Cairns, M. *et al*., 2006. Expression and purification of dalcochinase, a β−glucosidase from *Dalbergia cochinchinensis* Pierre, in yeast and bacterial hosts. Protein Expr. Purif. 48, 195-204.

Uchiyama, T., Miyazaki, K., Yaoi, K., 2013. Characterization of a novel β-glucosidase from a compost microbial metagenome with strong transglycosylation activity. J. Biol. Chem. 288:18325–34.

Viegas, A., Bras, N.F., Cerqueira, N.M., Fernandes, P.A., Prates, J.A., Fontes, C.M. *et al*., 2008. Molecular determinants of ligand specificity in family 11 carbohydrate binding modules - an NMR, X-ray crystallography and computational chemistry approach. FEBS J. 275, 2524-2535.

Vogt, G., Woell, S., Argos, P., 1997. Protein thermal stability, hydrogen bonds, and ion pairs 1. J. Mol. Biol. 269, 631-643.

Wallecha, A., Mishra, S., 2003. Purification and characterization of two β−glucosidases from a thermo-tolerant yeast *Pichia etchellsii*. Biochim. Biophys. Acta. Proteins. Proteom. 1649, 74-84.

Wang, Y., Yuan, H., Wang, J., Yu, Z., 2009. Truncation of the cellulose binding domain improved thermal stability of endo-β−1,4-glucanase from *Bacillus subtilis* JA18. Bioresour. Technol. 100, 345-349.

Xia, W., Xu, X., Qian, L., Shi, P., Bai, Y., Luo, H. *et al*., 2016. Engineering a highly active thermophilic β−glucosidase to enhance its pH stability and saccharification performance. Biotechnol. Biofuels. 9, 147.

Yang, X., Ma, R., Shi, P., Huang, H., Bai, Y., Wang, Y. *et al*., 2014. Molecular characterization of a highly-active thermophilic β–Glucosidase from *Neosartorya fischeri* P1 and its application in the hydrolysis of soybean isoflavone glycosides. PLoS One. 9 e,106785.

# SUPPLEMENTARY INFORMATION

**Table S1.** Comparison of the kinetic parameters reported or calculated for different fungal β-glucosidases using *p*NPG as model substrate.

| Organism and reference | Enzyme | $K_m$ | $V_{max}$ | $k_{cat}$ | $k_{cat}/K_m$ |
|---|---|---|---|---|---|
| **Native BGLs** | | | | | |
| *Trichoderma reesei* (Guo *et al*., 2016) | Cel3A | N | 41.0 | | |
| | Cel3B | N | 36.0 | | |
| | Cel1A | N | 2.9 | | |
| Metagenome from compost (Uchiyama *et al*., 2013) | Td2f2 | 0.39 | | 12.0 | 30.6 |
| *Thermoanaerobacterium thermosaccharolyticum (Pei et al., 2012)* | | 0.63 | 64.0 | 55.4[a] | 88.0[a] |
| *Penicillium purpurogenum* (Jeya *et al*., 2010) | | 5.10 | 934.0 | 1395.1[a] | 273.5[a] |
| **BGLs expressed in *P. pastoris*** | | | | | |
| *Thermoascus aurantiacus* (Hong *et al*., 2007) | BGLI | 0.22 | 71.7 | | |
| *Aspergillus fumigatus* (Liu *et al*., 2012) | rBgl3 | 1.76 | 131.4 | 284.8 | 161.7 |
| *Penicillium funiculosum* (Ramani *et al*., 2015) | rBgl4 | 2.50 | 3332.0 | 7219.0[a] | 2887.6[a] |
| *Neurospora crassa* (Pei *et al*., 2016) | BGL2 | 0.21 | 143.2 | 181.2[a] | 94.8[a] |
| *Myceliophthora thermophila* (Karnaouri *et al*., 2013) | MtBgl3a | 0.39 | 97.7 | 146.5[a] | 375.7[a] |
| *Saccharomycopsis fibuligera* and *T. reesei* (Guo *et al*., 2016) | SfBGL1/TrBGL1 | 0.20/0.21 | | 7.2/10.6 | 34.8/50.2 |
| *Neosartorya fischeri* (Yang *et al*., 2014) | NfBGL1 | 0.51 | 2172.0 | 2853.0 | 5594.0 |
| **BGLs from this work** | | | | | |
| ***T. amestolkiae*** | **BGL-2** | **0.41** | **299.7** | **485.8** | **1167.8** |
| | **BGL-2*** | **0.19** | **261.1** | **444.1** | **2243.1** |
| | **BGL-2T** | **0.34** | **649.1** | **874.0** | **2563.2** |

[a]Calculated from data provided in the original article. N: data not provided in original article.

**ATG**TATGCTGCATTCATCTTACTCCTCGCATCGGCCATTCCCGCGGTCAACGCCCAGTCAGCTTCTTGGTCCGCAGCC
TACAGTAAAGCCACAGCTGCTTTGGGCAAACTATCTCAAAATGACAAGATTGGTATGGTGACAGGCGTCGGATGGGGA
AAAGGTCCATGTGTTGGTAACACCGCTGCACCTTCGGGCATCTCGTATCCATCACTCTGTGTTCAGGATAGTCCCCTA
GGTGTTCGCTATGCGAACCCCGTCACAGCGTTCCCGGCGGGAACAAATGCTGGATTGACTTGGGATCGGACCTTGATG
AACCAGAGAGGCGCTGCTATTGGAGCAGAATCCAAGGGATTAGGTGTCAATGTTCAGTTGGGCCCTGTGGCAGGACCT
TTAGGCAAGATCGCGCAGGGCGGTCGTGGTTGGGAAGGTTTTGGGACAGATCCATACCTCAGTGGTGTTGCAATGATC
GAGACTATTACTGGTATGCAAAGTTCTGGTACCCAGGCATGCGCAAAGGTAAGCACATCGCTGCAATCGAAAGCCTTG
TATGAGAAAAGTCTAATCTTCGACAGCACTTTATTGGTAACGAGCAAGAGTTAAACCGAGAGTCGATGAGTTCTAACA
TTGACGATCGTACTTTACATGAGCTATATCTGTGGCCCTTCGCTGATGCCGTCCGTGCCAATGTTGCCAGTGTGATGT
GCTCCTACAACCAAATCAATGGATCATTTTCTTGTGAGAATGAAGCATCAATGACAGGCATTTTAAAAAAGGAGCTCG
GCTTTCCCGGGTATATTATGTCTGATTGGGATGCGCAGCACACAACAGTCACTAGTGCTAACTCTGGGCTCGATATGA
CCATGCCAGGCAGTGATTATAGTGATACGCCGAGCAGTGTTCTTTGGGGCCAAAATCTGGCTAATGCTATCTCAAGTG
GTCAAGTTGCCCAGTCACGACTGGATGATATGGTAACTCGAATTTTGGCTGCTTGGTATTTGGTTGGCCAGGATCAAG
GTTACCCTGCCGTTGCCTTTAGCTCTTGGAACGGTGGCCAAGCAAGTGTCAATGTCACATCAAACCACAACGTAGTTG
CCCATGCAGTCGCTCGTGATTCTATCGTTCTTCTTAAGAATACTAACGGCACGCTTCCGTTGAAAAAACCATCAAGCA
TCGCTATTATTGGCACTGACGCCCAGACAAACCCTTCCGGACCAAACGCTTGTACTGATCGTGGTTGTGATACTGGAA
CCTTAGCTATGGGTTGGGGCAGCGGAACTTGCCAATTCCCTTACTTGACAGATCCTCTGACAGCCATAAAATCTCAAG
CGGCCAGCGATGGCACTACAATCACGACTAGCATCAGTGACGATGGCAGTGCTGGGGCGTCAGTTGCTCAAAACGCCG
AGTATGCAATCGTTTTTATCAACTCCGACTCTGGGGAAGGGTAAGTTTATATCTTTGGCAGACTCCAATAGCTTTTAA
TAATTGAATATTAGTTACATCACCGTCGAAGGCGTCGCTGGCGACCGCAACAACCTCGATCCATGGCACAGCGGTAAT
GCACTGGTACAATCCGTTGCCGCAGTCAATAAAAATACCATTGTCGTCATCCACAGTGTTGGGCCCGTCATTCTTGAA
ACCATACTGGCGCAACCCAACGTAGTAGCGGTGGTATGGGCCGGCTTACCAGGACAAGAGAGTGGTTCAGCCCTTACA
GATATCCTCTATGGAAGTACAGCTCCCAGTGGCAAATTAACCTACACGATTGCTAAACAAGCCTCCGACTACGGCACT
GCAGTAGTCAGTGGCAACGACAACTATCCTGAAGGGCTTTTCATTGATTACCGACACTTCGATCAAAGTAATATTCAA
CCTCGATATGAATTCGGCTATGGACTGTCATATACAACCTTTGGCTACACAAACTTGGCAATTGGTATTACGGTTTCT
ACTGGCCCGACTACAGGTCAAACTGTCCCCGGTGGACCTTCTGATCTTTTTGCGTCTGTTGGAACTGTTACGGTGCAG
GTCACAAACACAGGCAGCGTCGCAGGATCAGAAGTTGCCCAACTCTATATTGGATTACCATCGTCAGCACCGTCATCG
CCGCCAAAACAGTTGCGTGGGTTTGATAAGCTTTCCCTCGCTGCTGGAGCTAGTGGTACAGCAACTTTTGATTTGACA
CGAAGAGATTTGAGTTACTGGGACGTATCACAGCAAAAGTGGGTGGTTCCAAGCGGAGCATTTACCATATATGTTGGA
GCGTCTAGTAGGGATATCAGATTGCAGGGGACATTTACGGTAGGT**TAA**ACAATATGGTCAATCATGAGACTAGATTCA
GGTGAGCGATGCTGACTGTGATTAAGCCCGGAGGTAGCTCGACCACTTCTACTATAAGTTCATCTAAGACATCGACTA
CTACTAGCACTTCGGCTACTACCAGCACATCGGTTACTACCAGCAGCAGTACTACGATAAAGACTACCACAACTAGCT
CGACCACATCATCTGCCGGGCCAACACAGACCCCGTATGGACAGTGTGGTGGACAGGGCTGGAGCGGTCCTACAGTTT
GTTCATCCGGCTGGACTTGTAAGGTGACGAATCAGTGGTATTCTCAATGCCTACAA**TAG**

**Figure S1.** Sequence of *bgl2* from *T. amestolkiae*. Introns are indicated in grey. The CBD coding region is represented after the third intron. Start and stop transcription codons are marked in bold. The predicted region for the signal peptide is underlined. The TAA stop codon present in third intron that prematurely ends transcription, producing BGL-2T*, is underlined and in bold grey.

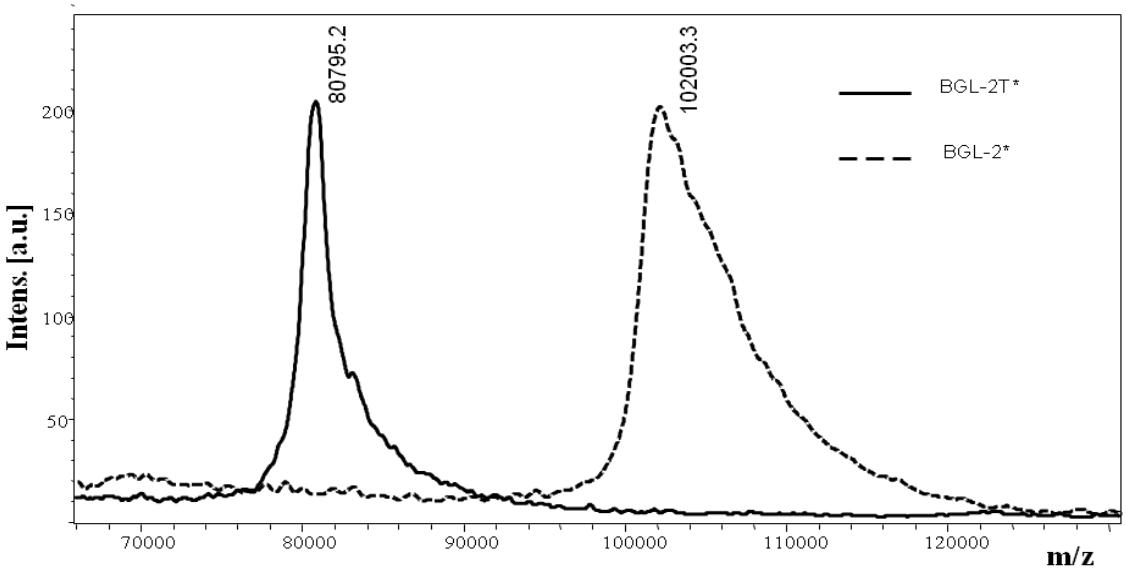



**Figure S2.** MALDI-TOF mass spectra of BGL-2T* and BGL-2*.

MYAAFILLLASAIPAVNAQSASWSAAYSKATAALGKLSQNDKIGMVTGVGWGKGPCVGNTAAPSGISYPS
LCVQDSPLGVRYANPVTAFPAGTNAGLTWDRTLMNQRGAAIGAESKGLGVNVQLGPVAGPLGKIAQGGRG
WEGFGTDPYLSGVAMIETITGMQSSGTQACAKHFIGNEQELNRESMSSNIDDRTLHELYLWPFADAVRAN
VASVMCSYNQINGSFSCENEASMTGILKKELGFPGYIMSDWDAQHTTVTSANSGLDMTMPGSDYSDTPSS
VLWGQNLANAISSGQVAQSRLDDMVTRILAAWYLVGQDQGYPAVAFSSWNGGQASVNVTSNHNVVAHAVA
RDSIVLLKNTNGTLPLKKPSSIAIIGTDAQTNPSGPNACTDRGCDTGTLAMGWGSGTCQFPYLTDPLTAI
KSQAASDGTTITTSISDDGSAGASVAQNAEYAIVFINSDSGEGYITVEGVAGDRNNLDPWHSGNALVQSV
AAVNKNTIVVIHSVGPVILETILAQPNVVAVVWAGLPGQESGSALTDILYGSTAPSGKLTYTIAKQASDY
GTAVVSGNDNYPEGLFIDYRHFDQSNIQPRYEFGYGLSYTTFGYTNLAIGITVSTGPTTGQTVPGGPSDL
FASVGTVTVQVTNTGSVAGSEVAQLYIGLPSSAPSSPPKQLRGFDKLSLAAGASGTATFDLTRRDLSYWD
VSQQKWVVPSGAFTIYVGASSRDIRLQGTFTPG*GSSTTSTISSSKTSTTTSTSATTSTSVTTSSSTTIKT*
*TTTSSTTSSAGPTQTPYGQCGGQGWSGPTVCSSGWTCKVTNQWYSQCLQ*

**Figure S3.** Protein sequence of BGL-2. The region underlined at the beginning corresponds to the predicted signal peptide, the catalytic residues are marked in black, the possible glycosylation sites are shown in grey. The missing residues in BGL-2T* are indicated in italics. CBD of the complete form is highlighted inside a rectangle.

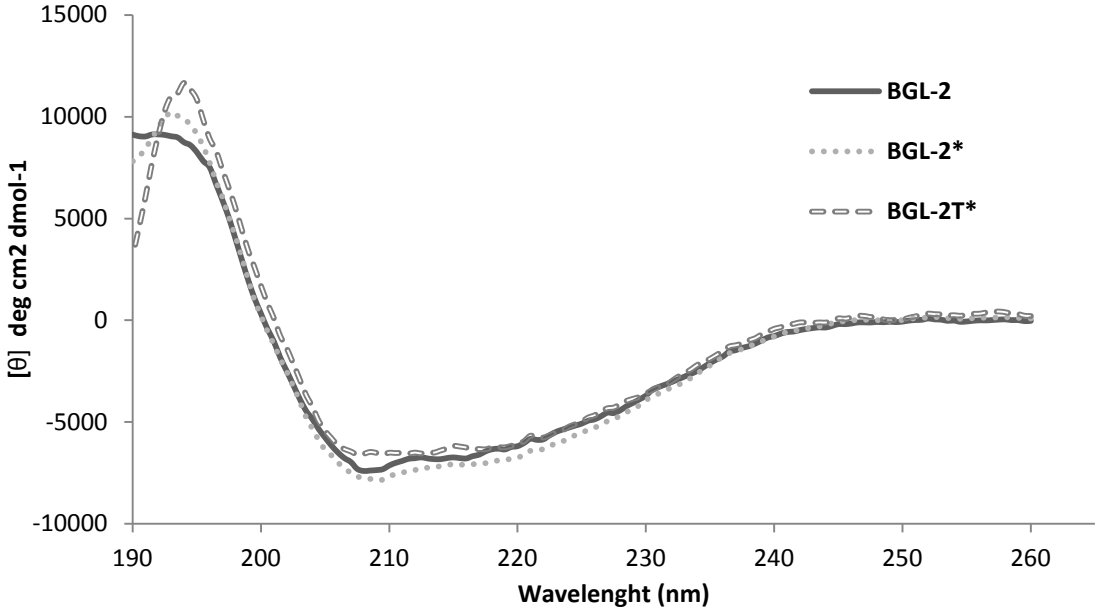



**Figure S4.** Circular dichroism spectra (far UV spectrum) of purified BGL-2 isoforms.

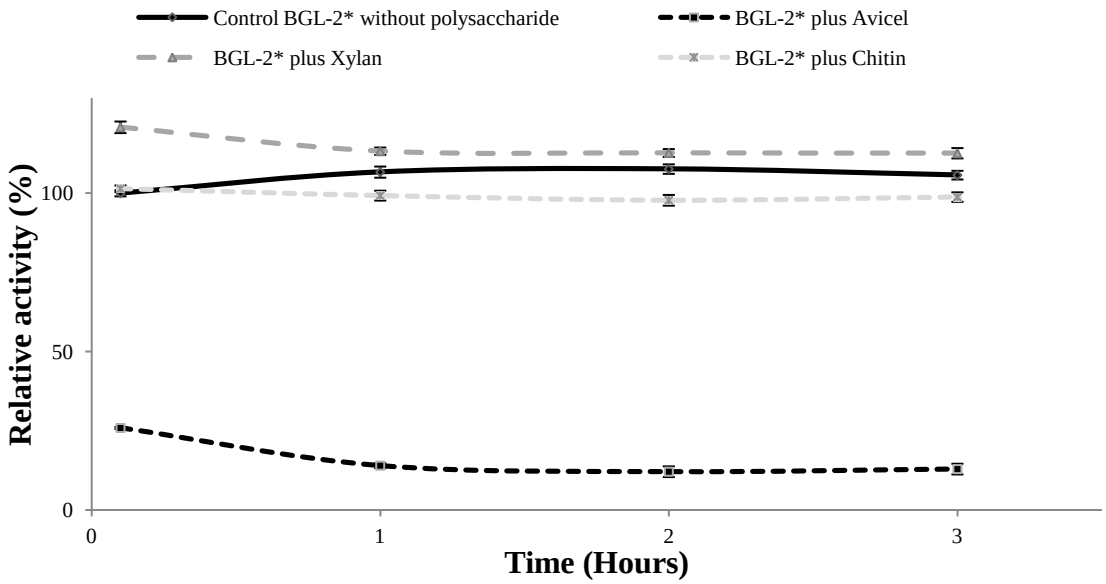



**Figure S5.** Cellulose-binding assays of BGL-2* against xylan and chitin, compared with Avicel.

# CHAPTER 3

**The β-glucosidase secreted by *Talaromyces amestolkiae* under carbon starvation: A versatile catalyst for biofuel production from plant and algal biomass.**

**Méndez-Líter, J.A.**, de Eugenio, L.I., Prieto, A., Martínez, M.J., 2018. The β-glucosidase secreted by *Talaromyces amestolkiae* under carbon starvation: a versatile catalyst for biofuel production from plant and algal biomass. Biotechnol. Biofuels. 11, 123-123. doi: 10.1186/s13068-018-1125-9

Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

## ABSTRACT

**Background**: In the last years, the most outstanding trend for obtaining high added-value components and second-generation (2G) biofuels consisted on exploitation of plant biomass. But recently, 3G biofuels, based in algae biomass, have emerged as a great alternative for production of energy.

**Results:** In this work, a versatile β-glucosidase from the ascomycete fungus *Talaromyces amestolkiae* has been purified, characterized, and heterologously expressed. The synthesis of this β-glucosidase (BGL-3) was not induced by cellulose, and the presence of a specific carbon source is not required for its production, which is uncommon for β-glucosidases. BGL-3, which was obtained from a basal medium with glucose as carbon source, was profusely secreted under carbon starvation conditions, which was corroborated by qRT-PCR assays. BGL-3 was purified from *T. amestolkiae* cultures in one step, and biochemically characterized. The enzyme showed high thermal stability, and very high efficiency on *p*NPG ($K_m$ of 0.14 mM and $V_{max}$ of 381.1 U/mg), cellobiose ($K_m$ of 0.48 mM and $V_{max}$ of 447.1 U/mg) and other cellooligosaccharides. Surprisingly, it also showed remarkable ability to hydrolyze laminarin, a β-1,3-glucan present in algae. The recombinant enzyme, obtained in the yeast *Pichia pastoris*, exhibited kinetic and physicochemical properties similar to those found for the native protein. Enzyme efficiency was examined in wheat straw saccharification processes, in which BGL-3 worked better supplementing Celluclast 1.5L than the commercial cellulase cocktail N-50010. Besides, BGL-3 hydrolyzed laminarin more efficiently than a commercial laminarinase.

**Conclusions:** A very efficient 1,4-β-glucosidase, which also showed activity over 1,3-β-glucose bonds, has been produced, purified and characterized. This is the first report of such versatility in a 1,4-β-glucosidase. The application of this enzyme for saccharification of wheat straw and laminarin and its comparison with commercial enzymes, suggest that it could be an interesting tool for the production of 2G and 3G biofuels.

**Keywords**: Fungi; Glycosyl hydrolases; Saccharification; Laminarinase; Biofuels; Carbon Starvation.

# BACKGROUND

The increase in world population, the expanding economy, and the non-renewable nature of fossil fuels are among the main reasons threatening the supply of energy to cover our needs. Industrial production of biofuels is a promising alternative to reduce our current dependence on petroleum (Naik *et al*., 2010).

In this context, cellulose is the most abundant potential source of renewable energy on earth. Its transformation into glucose is considered the key step in the production of biofuels from lignocellulosic biomass, and it determines the performance of the rest of the process. This transformation requires the synergistic action of three enzyme classes, endoglucanases, cellobiohydrolases, and β-glucosidases (BGLs), to hydrolyze the β-1,4 bonds of cellulose. Biological conversion of cellulosic biomass is a green way to produce second-generation (2G) ethanol and other chemicals, and it does not compete with human food resources used for first-generation bioethanol production, thus ending the controversial issues of food-versus-fuel and use of arable lands (Sun and Cheng, 2002).

Recently, 3G biofuels, from algae biomass, have received considerable attention (John *et al*., 2011) for their advantages over traditional production pathways for biofuels: i) algae do not need agricultural lands, since they can grow in swamp areas; ii) they do not require fresh water, an important factor taking into account that fresh water is limited. Moreover, algae are not a seasonal crop like corn, since they can be cultured all year, making it a more sustainable energy source than first and second generation biofuels. Algae could be applied to obtain a wide spectrum of products such as biodiesel, bioethanol, biohydrogen and biomethane (Hannon *et al*., 2010). Bioethanol production from algae has a notable potential due to the absence of lignin and hemicellulose, as compared to lignocellulosic plants, which could favor the degradation of polysaccharides to get glucose for fermentation processes. Numerous species have been studied for this objective. For example, green algae including *Spirogyra* sp. and *Chlorococcum* sp. accumulate high levels of polysaccharides like starch, which could be fermented to bioethanol (Jones and Mayfield, 2012). One of the most promising species is represented by brown algae, of the genus *Laminaria*, for containing high quantities of laminarin. This polysaccharide was first discovered in *Laminaria digitata* and it later showed to be the main food reserve of this type of algae, being very abundant in their fronds. Structurally, laminarin is a predominantly linear β-(1,3)-glucan that can contain few branches of mannitol or glucose attached at *O*-6 positions of the main chain, and then it is a potentially nice candidate to produce bioethanol 3G (Hou *et al*., 2015). With this perspective, enzymes used in 2G bioethanol processes, like cellulase

cocktails, including β-glucosidases, could represent a good alternative for efficient liberation of glucose from algae polysaccharides.

Traditionally, relatively pure forms of commercial cellulose and pretreated plant biomass have been used as carbon sources for the induction of cellulolytic enzymes (Singhania *et al*., 2013). However, although these have been reported as the best ways to produce this type of enzymes, the substrates are usually expensive. Thus, the discovery of organisms that secrete robust and efficient cellulases without these requirements could be an interesting progress for cheapening the enzyme cost.

Fungal species from *Penicillium* and *Talaromyces*, its perfect state, are well known for producing a wide variety of β-glucosidases, with good characteristics, as thermostability, glucose tolerance, and high efficiency towards diverse substrates ( Jeya *et al*., 2010; Gao *et al*., 2013; Ramani *et al*., 2015; Xia *et al*., 2016). Some recent reports deal with *T. amestolkiae*, that secretes high levels of cellulases growing in media with different carbon sources (de Eugenio *et al*., 2017), and a β-glucosidase (BGL2) with a cellulose binding domain from this fungus, has proved to efficiently degrade cellulose from brewers' spent grain (Méndez-Líter *et al*., 2017). This work presents the production, characterization, and catalytic versatility of a novel β-glucosidase from this fungus and its role for bioethanol 2G and 3G production.

## RESULTS AND DISCUSSION

### BGL-3 production and purification

The potential of *T. amestolkiae* as a β-glucosidase producer has been recently reported, revealing that it secreted at least two different β-glucosidases (de Eugenio *et al*., 2017). One of them was induced exclusively by cellulosic substrates (g3821), while the other one (g377) was produced in the four carbon sources tested (Avicel, glucose, xylan or pretreated wheat straw). This phenomenon is peculiar since, usually, cellulolytic enzymes require cellulose or its derivatives to be induced, and they are repressed by glucose or other easily metabolizable carbon sources (Amore *et al*., 2013).

In the current work, the β-glucosidase activity increased drastically when the carbon source (glucose) was consumed, reaching its maximal level at 7 days of incubation (1.8 U/mL). The crude enzyme, obtained from the culture medium after ultrafiltration and dialysis, was loaded in a HiTrap Capto Adhere cartridge and three peaks with β-glucosidase activity were detected. During the chromatographic run, a small peak that eluted during the first part of the NaCl gradient (13% of the total BGL activity) was separated from a big one that eluted with 100% NaCl and contained the

bulk of the retained proteins, but only around 4% of the BGL activity. However, most of the activity (around 83%) eluted during column re-equilibration with the buffer.

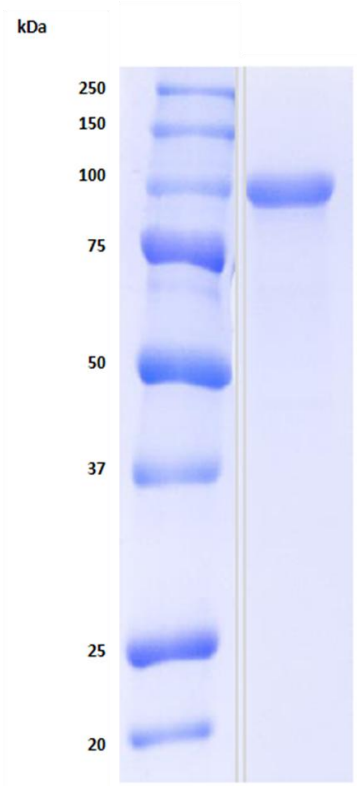



**Figure 1.** SDS-PAGE of BGL-3 from *T. amestolkiae* purified from Mandels medium using glucose as carbon source.

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the peak with the highest β-glucosidase activity showed a single protein band of around 100 kDa (figure 1), demonstrating that a protein, named as BGL-3, was purified to homogeneity in a single purification step. This fact could be explained because the cartridge used for separation contains a multimodal anion exchange separation bed, combining the properties of traditional anion exchangers with the different types of intermolecular forces related with the hydrophobicity of BGL-3. The final purification yield after concentration, dialysis, and purification, was 41% (Table 1). Peptide mass fingerprint analysis of BGL-3 disclosed that this enzyme was β-glucosidase g377, the major protein produced in cultures with glucose as carbon source.

**Table 1.** Purification of the BGL-3 secreted in Mandels medium with glucose as carbon source by *T. amestolkiae* cultures.

| BGL-3 Purification | | | | |
|---|---|---|---|---|
| Step | Total protein (mg) | Total activity (U) | Specific activity (U/mg) | Yield (%) |
| Crude extracts | 1.2 | 25.9 | 21.5 | 100.0 |
| HiTrap Capto Adhere | 0.7 | 10.8 | 42.4 | 41.7 |

## Physico-chemical properties of BGL-3

The molecular mass of the BGL-3 monomer, determined by MALDI-TOF, was 107.0 kDa, which coincided fairly well with the SDS-PAGE data (approximately 100 kDa). As these techniques cause dissociation of non-covalent aggregates, the native protein was analyzed by size exclusion chromatography. In these conditions, a molecular mass around 222 kDa, was calculated, suggesting that BGL-3 is a non-covalent, functional dimer. The molecular mass expected from the amino acid sequence of BGL-3 (89.5 kDa) was lower than the empirical value observed for the monomer of this protein, which can be due to protein glycosylation. This modification has already been reported for other extracellular glycosyl hydrolases from this fungus (Nieto-Domínguez *et al*., 2015). Subtracting the theoretical mass value from that determined by MALDI-TOF, an approximate glycosyl content of 16% can be deduced.
The pI (isoelectric point) value determined by isoelectric focusing (7.4) and that theoretically predicted (4.8) are different, and glycosylation can be partially responsible of this fact (Marsh *et al*., 1977), as the glycosyl chains can produce a shielding of surface charges. However, this difference in pI could also be due to other reasons like the uneven distribution of charged amino acids in the protein, with more negative charges being buried in the protein core.

The effect of temperature and pH on BGL-3 activity were also evaluated (figure 2). Optimum activity of BGL-3 was obtained at pH 4 and 70 °C. Higher temperatures and basic pHs produced a fast inactivation of the enzyme. Protein stability was measured for a 72-h period, observing that BGL-3 was active in a pH range from 2 to 7, and between 30-50 °C, losing activity quickly at higher temperatures.

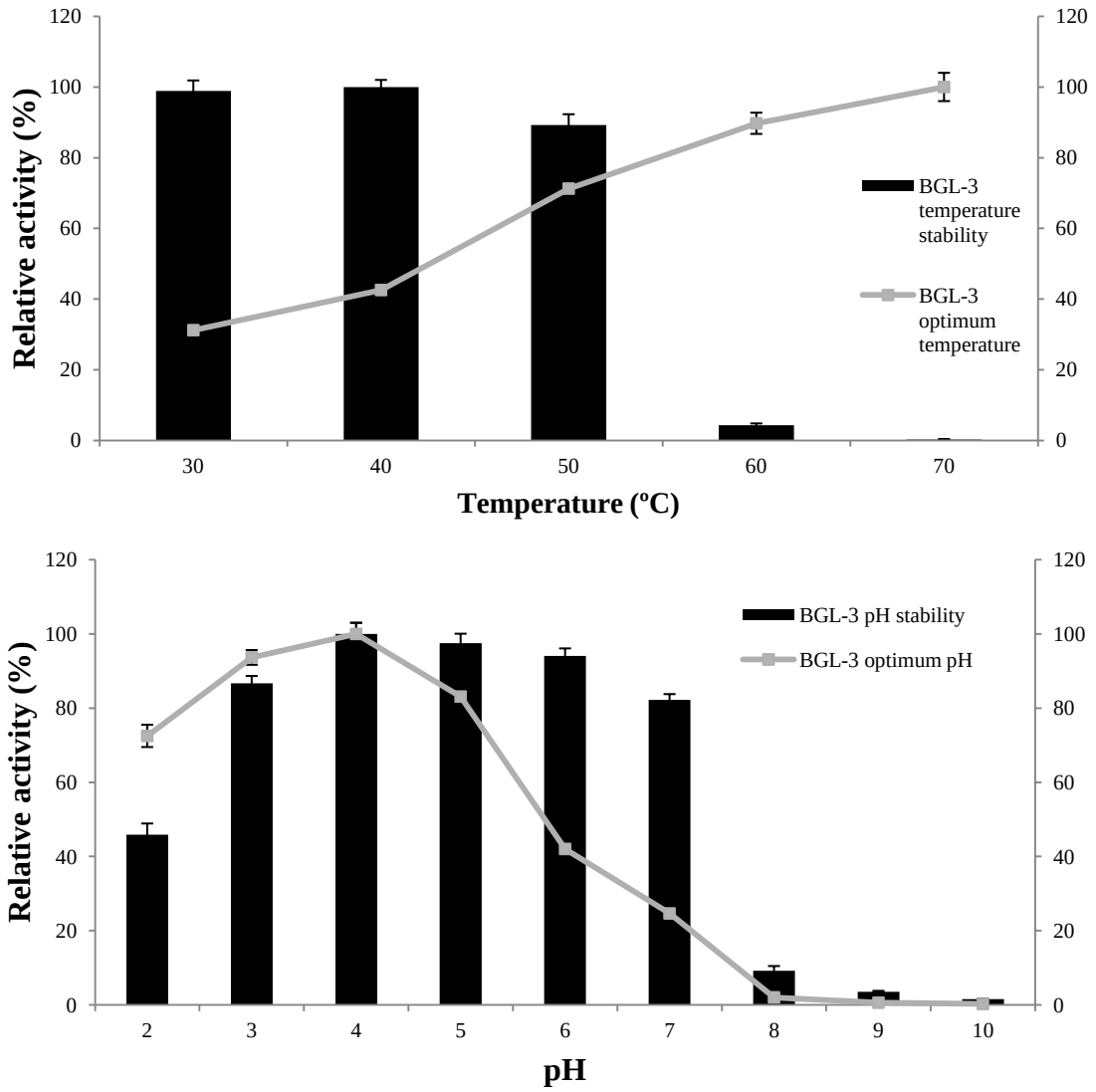



**Figure 2.** Stability and optimal pH and temperature of BGL-3, using 0.1% pNPG as substrate. Lines indicate the optimum temperature and pH values for enzyme activity; bars show its stability after 72 h in a range of temperatures from 30 °C to 70 °C or in a range of pH from 2 to 10. For temperature stability assays, the enzyme was incubated with sodium acetate buffer pH 4.0, 100 mM. For pH stability tests, the enzyme was incubated at the appropriate pH in Britton-Robinson buffer, 100 mM at 4° C. All reactions were performed in triplicate.

Temperature and pH are essential factors for enzymes and establishing their optimum conditions is very important to evaluate a potential industrial application. Many studies have been conducted on the search of β-glucosidases with enhanced properties, such as thermostability (Hong *et al*., 2007; Karnaouri *et al*., 2013; Yang *et al*., 2014). Among the biochemical characteristics evaluated for BGL-3, its main advantage is its good thermal tolerance, since this enzyme was fully active at 50 °C for 72 h.

## Substrate specificity

In hydrolysis reactions, BGL-3 showed very high efficiency and good affinity on *p*-nitrophenyl-α-D-glucopyranoside (*p*NPG), *o*-nitrophenyl-α-D-glucopyranoside (*o*NPG) and cellooligosaccharides from cellobiose to cellohexaose (Table 2). It is also remarkable that it was very active towards substrates without β-1,4 linkages, like laminaribiose and gentiobiose. However, only a residual activity was observed towards 4-nitrophenyl-β-D-xylopyranoside, maltose, and 4-nitrophenyl-α-D-glucopyranoside and no activity was detected on 4-nitrophenyl-β-D-galactopyranoside, 4-nitrophenyl-α-D-galactopyranoside, 4-nitrophenyl-α-L-rhamnopyranoside, 4-nitrophenyl-β-D-fucopyranoside, lactose or sucrose. The BGL-3 activity was also assayed on polysaccharides: Avicel, carboxymethyl cellulose (CMC), beechwood xylan and laminarin (Table S1). It showed a low but remarkable activity on Avicel, CMC and xylan and, unexpectedly, the enzyme had high activity against laminarin (Table 2).

**Table 2.** Kinetic parameters of BGL-3 from *T. amestolkiae* against different substrates. All reactions were performed in triplicate.

| Substrate | $K_m$ (mM) | $V_{max}$ (U/mg) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}.s^{-1}$) |
|---|---|---|---|---|
| ***p*NPG** | 0.14±0.1 | 381.1±4.1 | 1,359.4 | 9,710.2 |
| ***o*NPG** | 0.14±0.1 | 142.1±3.6 | 506.9 | 3,425.2 |
| **Cellobiose** | 0.48±0.1 | 447.1±7.0 | 1,594.6 | 3,308.4 |
| **Cellotriose** | 0.80±0.1 | 275.7±4.2 | 983.3 | 1,216.9 |
| **Cellotetraose** | 0.35±0.1 | 373.0±5.0 | 1,330.4 | 3,779.5 |
| **Cellopentaose** | 0.32±0.1 | 408.6±6.9 | 1,457.2 | 4,442.7 |
| **Cellohexaose** | 0.57±0.1 | 374.1±3.0 | 1,334.2 | 2,304.4 |
| **Laminaribiose** | 6.34±0.2 | 290.8±5.3 | 1,037.1 | 163.4 |
| **Gentiobiose** | 9.77±0.3 | 299.2±3.8 | 1,067.2 | 109.1 |
| **Laminarin from *L. digitata*** | 1.21±0.1 | 142.4±1.0 | 2,53.9 | 211.5 |
| **Laminarin from *L. hyperborea*** | 1.1±0.2 | 139.0±2.0 | 247.9 | 225.3 |

This β-(1,3)-glucan is specifically hydrolyzed by laminarinases (EC 3.2.1.6). To further study this interesting feature, the conditions for BGL-3 crystallization are currently being set. However, a preliminary 3D model of BGL-3 based on a β-glucosidase from *Aspergillus aculeatus* was constructed and compared to that of a barley β-D-glucan glucohydrolase isoenzyme in complex with 4'-nitrophenyl 3I-thiolaminaritrioside (SMTL ID 1j8v.1) (figure S2). Alignment of both models disclosed that the active

site from BGL-3 matched the amino acids involved in substrate binding in the laminarinase. This preliminary analysis suggests that one of the possible explanations to the laminarinase activity of BGL3 is the structural similarity between its active site and that of some strict laminarinases, but further experiments are required and will be developed in order to determine it.

β-glucosidases have traditionally been divided into three groups: cellobiases, which have high substrate specificity towards cellobiose, aryl-β-glucosidases, with very high specificity towards synthetic substrates as *p*NPG, and β-glucosidases with broad substrate specificity, acting on these two types of substrates and other oligosaccharides. Most β-glucosidases are placed in this last category, but it is remarkable that most of these enzymes are quite more active on *p*NPG than on cellobiose and other cellooligosaccharides, which are their natural substrates (Singhania *et al*., 2013). For production of 2G bioethanol, where their main substrates are cellulosic oligosaccharides, the search for efficient cellobiases is still a challenge. To our knowledge, the BGL-3 from *T. amestolkiae* is the most efficient BGL against cellooligosaccharides of three or more glucose units, and only rBgl4 from *Penicillium funiculosum* has similar efficiency for degradation of cellobiose (Table 3).

**Table 3.** Comparison of the kinetic parameters of *T. amestolkiae* BGL-3 with those reported for other fungal β-glucosidases, using cellobiose as substrate. All reactions were performed in triplicate.

| **Organism** | **Enzyme** | $K_m$ | $k_{cat}$ | $k_{cat}/K_m$ | **Ref** |
|---|---|---|---|---|---|
| *Penicillium purpurogenum* | | 5.1 | 1,395[a] | 273[a] | (Jeya *et al*., 2010) |
| *Talaromyces leycettanus* | BGL3A | 10.4 | 786 | 75 | (Xia *et al*., 2016) |
| *Aspergillus fumigatus* | rBgl3 | 2.2 | 114 | 52 | (Liu *et al*., 2012) |
| *Penicillium funiculosum* | rBgl4 | 1.2 | 4,513[a] | 3,610[a] | (Ramani *et al*., 2015) |
| *Myceliophthora thermophila* | MtBgl3a | 2.6 | 46[a] | 17[a] | (Karnaouri *et al*., 2013) |
| *Aspergillus oryzae* | HGT-BG | 7.0 | 252 | 36 | (Riou *et al*., 1998) |
| ***T. amestolkiae*** | **BGL-3** | **0.5** | **1,594** | **3,308** | **This work** |

[a]Calculated from data provided in the original article

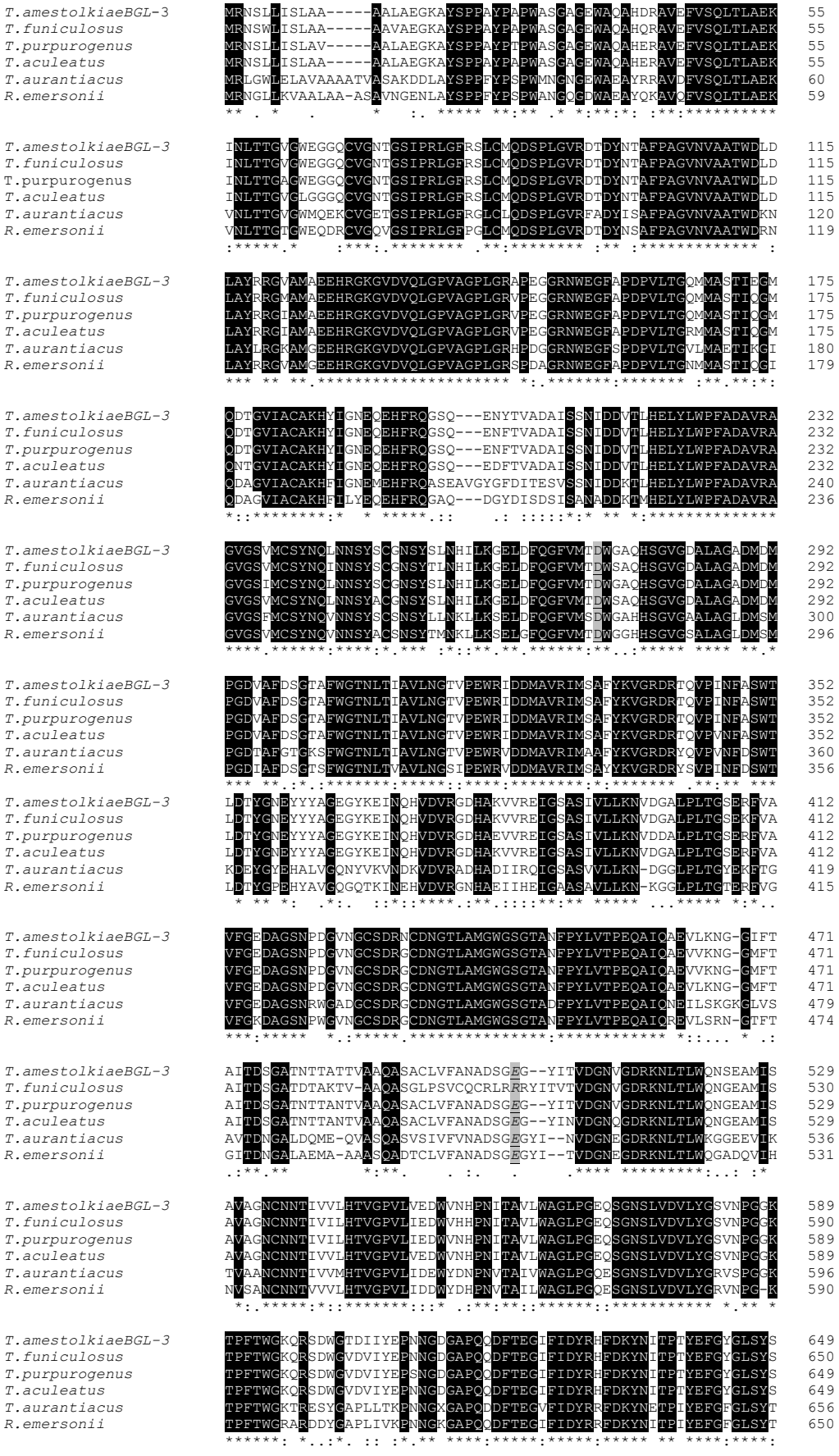



**Figure 3.** Protein sequence alignment of BGL-3 and other GH3 β-glucosidases, using Clustal Omega. The black-highlighted residues indicate fully-conserved regions. The catalytic residues are indicated in grey.

**DNA and amino acid sequence of BGL-3**

*T. amestolkiae* genome (de Eugenio *et al*., 2017) was searched to deduce the sequence of BGL-3, that was classified as a β-glucosidase from the GH3 family. Introns and exons identification was performed by comparison with transcriptomic data from similar sequences using BlastN. According to these data, the *bgl3* gene contains three introns, with a sequence of 2571 bp, coding for a protein of 857 amino acids (figure S1). A homology search with BLASTP revealed 94%, 95%, 95%, 69%, and 74% amino acid identity of BGL-3 with the BGLs of *Talaromyces funiculosus* (AFU91382.1), *Talaromyces purpurogenus* (ACV87737.1), *Talaromyces aculeatus* (AGA96121.1), *Thermoascus aurantiacus* (ABX79553.1), and *Rasamsonia emersonii* (XP_013330184.1), respectively. These data suggest that BGL-3-like proteins are well conserved among ascomycetes (figure 3).

**Expression of *bgl-3* gene under carbon starvation**

In a previous work, we demonstrated that the fungus secreted detectable amounts of BGL when glucose was used as carbon source, although this activity appeared upon glucose depletion (after 24 h of incubation). Besides, we detected repression of BGL activity when glucose pulses were added (de Eugenio *et al*., 2017). These results suggested that *T. amestolkiae* also released some β-glucosidase activity in the absence of cellulose. To confirm this, BGL activity and biomass were monitored at very short times in cultures with glucose (figure S3). The fungus reached its maximum growth after 12-24 h. BGL activity was detected in very low amounts during the first hours of incubation, and increased drastically once the culture reached the stationary phase.

To further study how BGL-3 is synthesized during a prolonged carbon starvation period, a transcriptomic analysis was performed. The biomass of *T. amestolkiae* was monitored in the culture over 7 days of incubation. The residual biomass determined in 7-day old cultures was around 30% of the initial amount. A qRT-PCR analysis of *bgl3* indicated that its expression increases over time (figure 4), reaching a maximum transcription level on day 7, which coincided in time with the highest level of extracellular BGL activity. This confirms the accumulation of BGL-3 on carbon- deficient cultures. It is interesting to note that *bgl3* expression grows 2.1-fold from day 1 to day 7, showing that the fungus increased the synthesis of this protein when the scarcity of carbon source persists over time. Few studies have been performed to investigate the effects of carbon starvation on fungal cultures. White *et al.* (2002) remarked the presence of extracellular hydrolase activity in these conditions. The induction of hydrolases, including glycosidases, has been proposed as a key event in aging of fungal cultures during carbon shortage. The most detailed works

developed with fungi have been conducted with *Aspergillus niger*. The studies of Nitsche *et al*. (2012), and Van Munster *et al*. (2014), revealed the ability of filamentous fungi to produce enzymes when the carbon source was consumed. The transcriptomic analysis of carbon-starved cultures of *A. niger* confirmed the expression of genes encoding CAZymes, including those responsible for the presence of BGL activity, which agrees with the production of BGL-3 by *T. amestolkiae* under carbon deprivation.

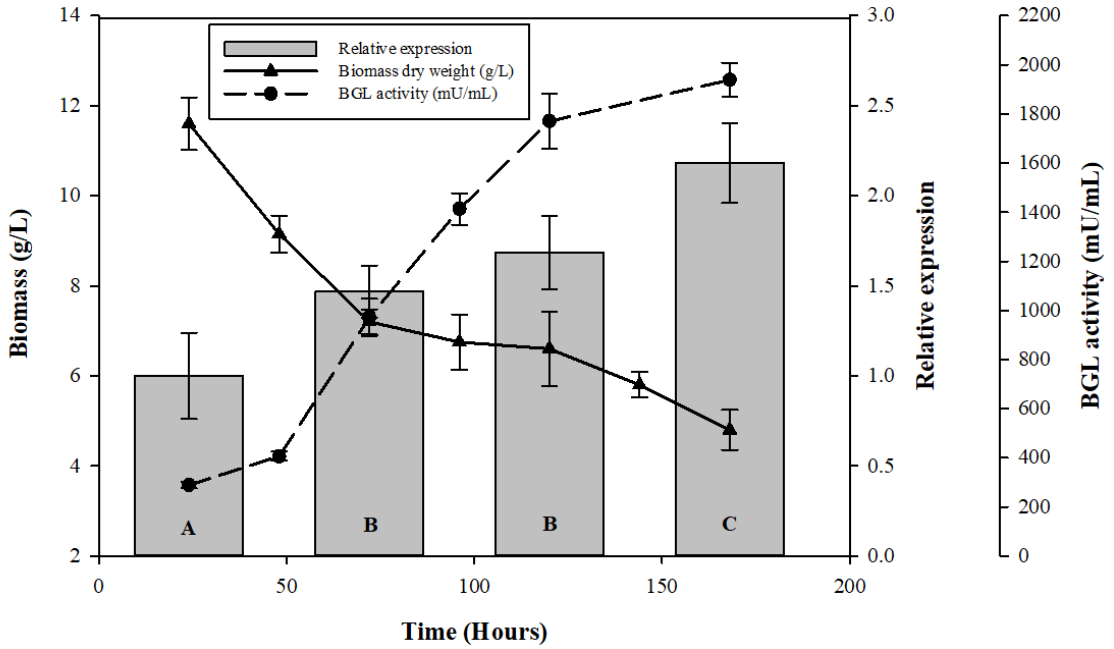



**Figure 4.** Comparison of qRT-PCR analysis of *bgl-3* expression and biomass over time in cultures with Mandels medium supplemented with glucose. Samples were taken periodically to determine β-glucosidase activity against 0.1% *p*NPG, relative expression by qRT-PCR and biomass dry weight. Data represent the mean of six replicates. Statistically significant differences between means were determined by the Tukey HSD test. Different letters in qRT -PCR results (A, B, C and D) indicate significant differences at a P value of 0.05.

## Heterologous expression of BGL-3 in *P. pastoris*

Due to the interesting properties of BGL-3 in terms of kinetic constants and physicochemical properties, the protein was expressed in the yeast *P. pastoris* with the aim of improving its production levels. After *P. pastoris* transformation, several positive clones were screened for β-glucosidase activity, selecting clones 34 and 20 as the highest producers in liquid cultures. Maximal β-glucosidase activity (8.1 and 7.9 U/mL, respectively) was found in 7-day-old YEPS cultures. These values are 4.5-fold higher than the total β-glucosidase activity detected in cultures of *T. amestolkiae* in Mandels medium with glucose (figure 5). The concentration of BGL-3 in *P. pastoris* supernatants was 21.2 mg/L for the clone with the highest production. The first trial to purify the recombinant BGL-3 (BGL-3*) was done using the same protocol applied for purification of the native

enzyme. However, BGL-3* was abnormally distributed throughout all fractions recovered, and it was necessary to change the procedure. The chromatographic separation on this cartridge relies on the combination of anionic exchange and hydrophobic interactions, and thus the different glycosylation patterns of the native and the recombinant enzyme can be responsible for their different retention in this bimodal column. BGL-3* was completely purified in two steps: anion-exchange chromatography on a HiTrap QFF cartridge, followed by size exclusion chromatography. In spite of this, the overall yield for production and purification of the recombinant enzyme was 7.5-fold higher than for the native protein (54% of initial activity, Table S2). As revealed by SDS-PAGE, BGL-3* had a molecular mass slightly superior as that of the native form, probably due to increased glycosylation in the yeast.

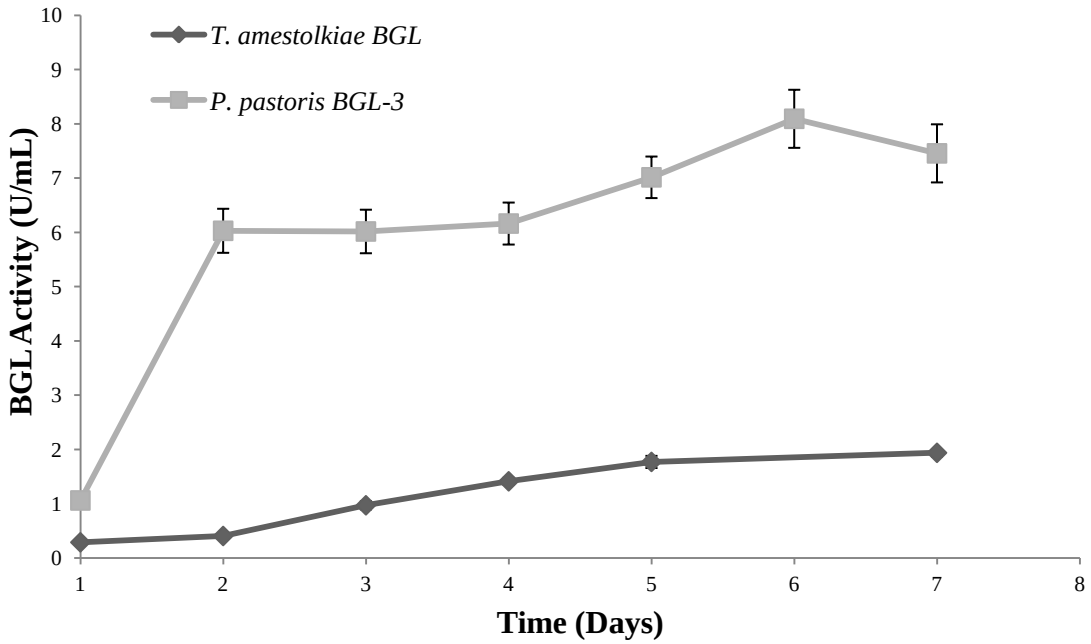



**Figure 5.** β-glucosidase production by *T. amestolkiae* or *P. pastoris*. Samples were taken daily and activity was measured against 0.1% (w/v) *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG, Sigma) in sodium acetate buffer 100 mM, pH 4.0, at 70 °C. All reactions were performed by triplicates.

The biochemical and kinetic properties of the recombinant protein were virtually the same as those of the native BGL-3 in terms of pH, temperature, and substrate specificity. We checked with special interest its activity on laminarin, concluding that the kinetic constants of BGL-3* on this β-1,3 glucan were very similar to those of the native enzyme, which confirms their catalytic versatility.

**BGL-3 for wheat straw saccharification**

Efficient conversion of polysaccharides into glucose prior to yeast fermentation is an essential step for ethanol production, and BGLs are the key enzymes to achieve full depolymerization. The effectiveness of BGL-3* in the 2G process was evaluated following the saccharification of wheat straw slurry, one of the main lignocellulosic feedstocks used for production of bioethanol. The glucose released from the substrate using only Celluclast 1.5L (model source of cellulases with low β-glucosidase activity), was compared with the values determined for the same cocktail supplemented either with BGL-3* or NS-50010, a commercial cocktail rich in β-glucosidase. The results reflected the synergistic action of both BGL preparations with the cellulose cocktail (figure 6), but supplementation of Celluclast 1.5L with BGL-3 increased the degradation yield around 37% while with NS-50010 the increment was of only 17%.

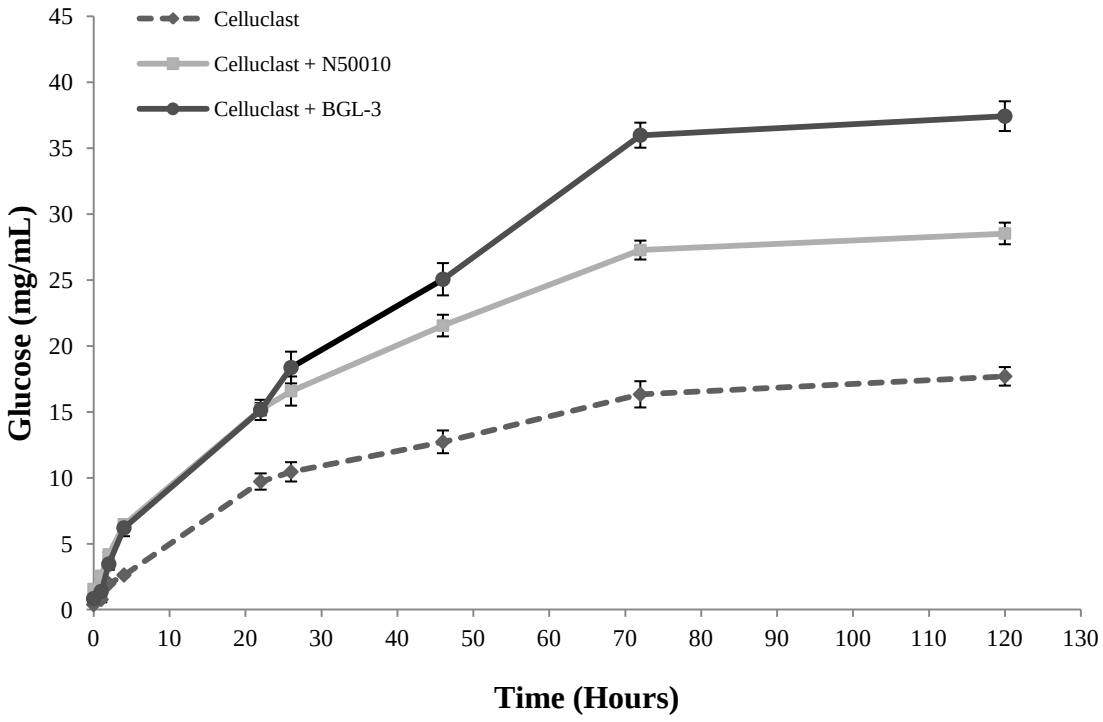



**Figure 6.** Glucose released from saccharification of wheat straw slurry. For the saccharification of 100 mg of wheat straw slurry, 1 U of BGL activity from Celluclast (the basic enzymatic cocktail) was supplemented with 1 U of either BGL activity from NS50010 (commercial BGL), or BGL-3 from *T. amestolkiae.* Glucose was determined in the samples after saccharification for 5 days (50 ºC, 1,200 rpm). All reactions were performed in triplicate.

In view of these data, and remembering that NS-50010 is commercially sold as a β-glucosidase-rich preparation, the potential relevance of BGL-3* from *T. amestolkiae* for saccharification of lignocellulosic biomass must be emphasized.

**BGL-3 for laminarin saccharification**

The versatility of the native and recombinant forms of the BGL-3 from *T. amestolkiae* prompted us to test the activity of the recombinant enzyme on hydrolysis of laminarin from two *Laminaria* species: *L. digitata* and *L. hyperborea.* Its efficiency on these substrates was compared with that of a commercial laminarinase (β-1,3 glucanase from *Helix pomatia*). Both the purified BGL-3* and the commercial β-1,3 glucanase hydrolyzed efficiently both laminarin samples in less than 24 h (figure 7), but BGL-3* worked better than the commercial enzyme at the same dose (3 U/mL).

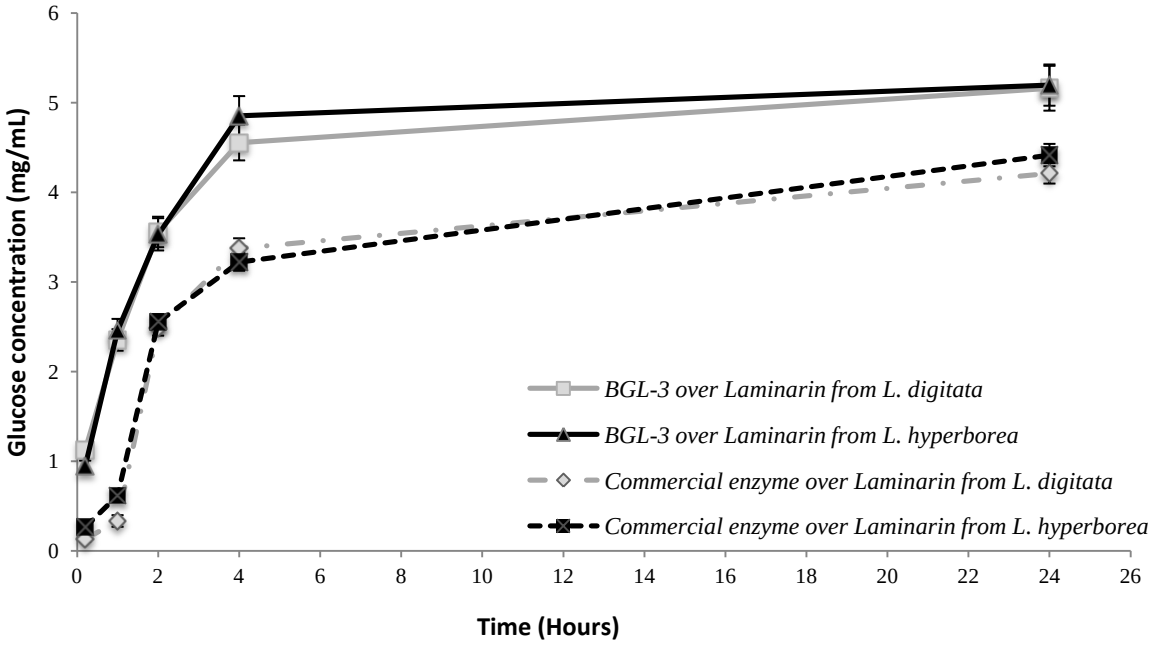



**Figure 7.** Glucose released from saccharification of laminarin from *L. digitata* and *L. hyperborea* using BGL-3 from *T. amestolkiae* and a commercial laminarinase. 100 mg of laminarin were incubated for 24 h with 3 U of laminarinase activity, at 50 ºC and 1,200 rpm. All reactions were performed in triplicate.

It should be noted that although the total enzymatic activity added to the reactions was identical, the dry weight of laminarinase necessary to get this activity was 10-fold higher (0.01 mg/mL of BGL-3* *vs.* 0.1 mg/mL of β-glucanase from *Helix pomatia*). These data showed unequivocally that BGL-3* is more efficient for laminarin hydrolysis than an enzyme whose activity is specific for this substrate.

The versatility of some β-glucosidases to hydrolyze short oligosaccharides with β-1,2, β-1,3, β-1,4 or β-1,6 linkages has been widely reported ( Jeya *et al*., 2010; Gao *et al*., 2013; Ramani *et al*., 2015; Xia *et al*., 2016). However, the ability of these enzymes to degrade a polysaccharide like laminarin is much more exceptional and in spite of the

higher affinity of the BGL-3 from *T. amestolkiae* against β-1,4 oligosaccharides, this catalyst releases a considerable amount of glucose from laminarin. Its efficiency has been compared with those of endo- and exo-β-1,3-glucanases. The results gathered in Table 4 show that the catalytic efficacy of BGL-3 on the laminarin samples tested was in the range of those reported for many enzymes that use laminarin as their natural substrate.

The unexpectedly high activity of BGL-3 on a β-1,3 polysaccharide, together with its production in the absence of a carbon source, suggest its possible physiological role in *T. amestolkiae* cell wall metabolism. β-1,3-glucans are the major components of fungal cell walls of ascomycetes and basidiomycetes (Ruel and Joseleau, 1991; Bowman and Free, 2006). Thus, a logical explanation can be that the fungus secretes BGL-3 for its autolysis, trying to find an alternative carbon source upon starvation from its own cell walls. This hypothesis is similar to that developed by Igarashi *et al*. (Igarashi *et al*., 2003), who reported an extracellular β-1,3 glucosidase from *Phanerochaete chrysosporium* and proposed its relationship with fungal cell wall metabolism. This process of autophagy was presented as a mechanism necessary to obtain glucose from the β-glucans present in its cell wall, sustaining fungal metabolism in the absence of an external carbon source.

**Table 4.** Comparison of the kinetic constants of BGL-3 with other laminarin-degradative enzymes, using laminarin as substrate.

| **Organism** | **Enzyme** | $K_m$ | $k_{cat}$ | $k_{cat}/K_m$ | **Reference** |
|---|---|---|---|---|---|
| *Podospora anserina* | PaGluc131A | 4.1[a] | 51.9[a] | 12.6 | (Lafond *et al*., 2012) |
| *Pseudoalteromonas sp. Strain BB1* | ExoP | 0.7 | 204.9 | 290.9 | (Nakatani *et al*., 2010) |
| *Talaromyces emersonii* | Exo-1,3-β-glucanase | 1.6 | 29.5 | 17.8 | (O'Connell *et al*., 2011) |
| Paddy soil microbial metagenome | Umcel9y-1 | 47.3 | 127.5 | 2.6 | (Zhou *et al*., 2016) |
| *Aspergillus fumigatus* | 1,3-β-glucanase | 0.2 | 56.9[a] | 219.1[a] | (Fontaine *et al*., 1997) |
| *Vibrio campbelli* | LamN | 4.0 | 0.8 | 0.1 | (Wang *et al*., 2015) |
| *Barley β-D-glucan* | ExoII | 0.1 | 28.0 | 230.0 | (Hrmova and Fincher 2017) |
| | ExoI | 0.1 | 73.0 | 740.0 | (Hrmova and Fincher 2017) |
| ***T. amestolkiae* BGL-3** | **BGL-3[b]** | **1.2** | **253.9** | **211.5** | **This work** |
| | **BGL-3[c]** | **1.1** | **247.9** | **225.3** | **This work** |

[a]Calculated from data provided in the original article; [b]Substrate: laminarin from *L. digitata.* [c]Substrate: laminarin from *L. hyperborea*

## CONCLUSIONS

In this study, a very efficient β-glucosidase has been purified from *T. amestolkiae* cultures in a basal medium with glucose as carbon source, and characterized. This enzyme did not require cellulosic substrates to be produced, and was secreted in high amounts under carbon starvation. BGL-3 is a versatile enzyme able to hydrolyze β-1,4 oligosaccharides and β-1,3-glucans. Due to its outstanding properties, the enzyme was produced in *P. pastoris*. The purification yield of the recombinant protein was 7.5-fold higher than that of the native BGL-3 from *T. amestolkiae*, and both enzymes showed similar kinetic and physicochemical properties. This enzyme hydrolyzes efficiently lignocellulosic substrates and laminarin, and this catalytic versatility could be of great interest for depolymerization of different glucans for 2G and 3G bioethanol production. In summary, its easy production and purification, high efficiency, and versatility make BGL-3 an interesting candidate for biorefinery processes.

## MATERIAL AND METHODS

### Microorganism and culture media

The ascomycete *T. amestolkiae* is deposited in the Collection of the Institute Jaime Ferrán of Microbiology (IJFM) at the Centro de Investigaciones Biológicas, with the access number A795. Fungal strains were maintained in tubes with PDA (Potato dextrose agar) medium, stored at 4 °C and periodically reseeded in PDA plates, incubated at 28 °C.

To obtain spore suspensions from this culture, agar pieces (1 $cm^2$) were cut and added to a tube containing 5 mL of a solution of 1% NaCl and 0.1% Tween 80. The mixture was used to inoculate 250 mL flasks with 50 mL of CSS medium (40 g/L glucose, 0.4 g/L $FeSO_4 \cdot 7H_2O$, 9 g/L $(NH_4)_2SO_4$, 4 g/L $K_2HPO_4$, 26.3 g/L corn steep solid, 7 g/L $CaCO_3$, and 2.8 ml/L soybean oil), incubating at 28 °C and 250 rpm for 5 days. These cultures were used as pre-inoculum.

### Production and purification of BGL-3

For BGL-3 production, 2 mL from the CSS cultures of *T. amestolkiae* were inoculated in 250-mL Erlenmeyer flasks containing 50 mL of Mandels medium (Mandels and Weber, 1969) with 1% of glucose as carbon source, and incubated in an Innova 4330 orbital shaker (New Brunswick Scientifics) at 28 °C and 250 rpm. All culture media were prepared with autoclaved distilled water.

When maximal β-glucosidase activity was detected in the supernatants (7 days), the cultures were cropped and centrifuged at 10,000 × *g* for 30 min to separate mycelium and supernatant, which was vacuum-filtered through filter paper and nitrocellulose membrane discs (Millipore) of 0.8, 0.45 and 0.22 μm to complete clarification. This treated supernatant was further concentrated using a tangential flow filtration system (7518-02 Masterflex, from Millipore) equipped

with a 10 kDa polysulfone membrane (Membrane Cassette, Filtron) and an ultrafiltration cell (Amicon, Millipore) with a 10 kDa cutoff polysulfone membrane (Millipore). Protein purifications were performed using an ÄKTA Purifier HPLC system (GE Healthcare Life Sciences). The crude extract was first dialyzed in 10 mM sodium phosphate, pH 6.0, and applied onto a Capto Adhere HiTrap cartridge (GE Healthcare Life Sciences) equilibrated with the same buffer at 2 mL/min. Retained proteins were eluted by using a linear gradient of NaCl in the same buffer (0 to 0.25 M of NaCl in 30 min) and 100% NaCl (15 min). Finally, the cartridge was equilibrated with the initial buffer.

The purified BGL-3 was dialyzed against acetate buffer, pH 4.0, and its homogeneity confirmed by SDS-PAGE using 10% gels stained with Coomassie brilliant blue R-250.

### *bgl-3* gene sequencing and real-time quantitative qRT-PCR analysis

In order to identify *bgl-3* sequence, a BLASTP against predicted proteins of *T. amestolkiae* was carried out. The gene sequences of the best hits were used as queries to run a local BLASTN against the assembled genome. An alignment between the gene and the best hits of the BLAST search was done to identify possible introns in the sequences of other BGLs. With the predicted coding sequence, the presence of a possible signal peptide was analyzed with the signalP server, and the putative mature gene of BGL-3, without introns and signal peptide, was translated to protein by using the ExPASy Bioinformatics resource portal (ProtParam tool), in order to obtain the theoretical molecular mass and isoelectric point of BGL-3.

Primers for qRT-PCR were designed based on *bgl-3* sequence (BGL-3FWQPCR (TTCGTATCATGTCTGCATTC) and BGL-3RVQPCR (ATTCTTGAGGAGAACAATGC)). 18S rRNA was chosen for normalization of expression across all treatments (Dankai *et al*., 2015) (primers 18sFW (ATTGGAGGGCAAGTCTGGTG) and 18sRV (CCAGTGAAGGCCATGGGATT)).

RNA was extracted from *T. amestolkiae* cultures growing in 1% of glucose using Trizol reagent (Chomczynski and Sacchi, 1987). One-step qRT-PCR was performed using total RNA preparations treated with a Turbo DNA–free kit (Ambion). Brilliant III Ultra-Fast SYBR® Green qRT-PCR Master Mix, from Agilent, was used for qRT-PCR reactions. Each reaction was performed according to the manufacturer's instructions, adding 5 ng of the respective RNA.

Reactions were done in a LightCycler® 96 detection system, and analyzed with LightCycler® 96 SW. The running method consisted of several steps: 50 °C for 10 min, 95 °C for 3 min, 40 cycles of 95 °C for 10 s, and 60 °C for 20 s. All reactions were performed six times. The amplification efficiency for each primer pair was determined with serial dilutions from an RNA sample (100 ng RNA/µL) with at least five dilution points. The relative quantification of PCR products was calculated by the comparative $2^{-\Delta\Delta CT}$ (cycle threshold) method.

**Cloning and expression of *bgl-3* in *P. pastoris***

RNA was isolated from fungal cultures by using trizol reagent, as explained before. The isolated transcripts were converted to cDNA using the Superscript II Reverse Transcriptase RT-PCR kit (Invitrogen) using 50 µM random hexamers. PCR amplifications were performed in a thermocycler Mastercycler pro S (Eppendorf) using genomic DNA as template. Primers were designed based on the nucleotide sequence of the *bgl3* gene from *T. amestolkiae* genome (GenBank accession no. MIKG00000000), but excluding the region corresponding to the signal peptide. Restriction sites for *XhoI* and *NotI* were respectively added to the forward and reverse primers (BG3FWXHOI: 5′-ATCTCGAGAAAAGATACTCTCCTCCAGCTTACCCT-3′, and BG3 RV NOTI: 5′-ATGCGGCCGCATGCCCAATCTTCAAAGCCAA -3′). Reaction mixtures were initially subjected to denaturation at 95 °C for 5 min, followed by 36 cycles of amplification consisting of: denaturation at 95 °C for 45 s, primer annealing at 55 °C for 45 s, and elongation at 72 °C for 3 min, followed by a final extension step at 72 °C for 10 min. The PCR product was ligated to the yeast expression vector pPICzα (Invitrogen), and it was used for transforming *P. pastoris* X-33 after linearization with *SacI* (New England Biolabs). Transformed colonies were grown on YPD medium plates (10 g/L Yeast extract, 20 g/L peptone, 20 g/L glucose and 10 g/L of agar) with 100 µg/mL of zeocin as selection marker. Since it is considered that the better the zeocin resistance, the better the protein production, the scored transformants were re-screened for resistance to a zeocin concentration of 1 mg/mL, selecting the clones with the highest tolerance for protein production.

**Production and purification of recombinant BGL-3**

To prepare a fresh inoculum, the selected clones were grown overnight in 250 mL flasks with 50 mL of YEPS medium at 28 °C and 250 rpm. Then, recombinant protein production was carried out in 2-L flasks with 400 mL of YEPS medium (20 g/L peptone, 10 g/L yeast extract, 10 g/L sorbitol). Cultures were incubated at 28 °C and 250 rpm for 7 days with daily addition of 5 g/L methanol. Samples were periodically taken to measure β-glucosidase activity.

For BGL-3* purification, 7 day-old cultures were harvested and centrifuged at 10,000 × *g* and 4 °C for 20 min. The supernatant was first concentrated by tangential filtration and finally concentrated and dialyzed against 10 mM phosphate buffer (pH 6.0) using a 50-kDa cutoff membrane (Merck-Millipore). BGL-3* was purified after two chromatographic steps. First, a QFF Hi Trap cartridge (GE Healthcare) equilibrated with phosphate buffer pH 6.0 was used. Elution of the bound proteins was carried out by applying a linear gradient from 0 to 0.25 M of NaCl in 25 min, at 2 mL/min. The column was then washed with 10 mL of 1 M NaCl and re-equilibrated using 10 mL of the starting buffer. Fractions with β-glucosidase activity were collected, dialyzed and concentrated. To complete the purification of BGL-3*, the sample from the

previous stage was analyzed by size exclusion chromatography on Superose 12 column (GE Healthcare Life Sciences). To avoid unspecific interactions, the same buffer (plus 100 mM NaCl) was used for column equilibration and proteins elution, at a flow of 0.5 mL/min.

**Protein quantification, enzyme assays and substrate specificity**

Total protein was estimated by the BCA method using bovine serum albumin as standard, measuring the absorbance of the sample at 280 nm in a Nanodrop (Thermo Fisher Scientific). The β-glucosidase standard reaction was performed using 0.1% (w/v) *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG, Sigma), at 70 °C, in sodium acetate buffer 100 mM, pH 4.0. Other nitrophenyl derivatives, as *p*NPX (*p*-nitrophenyl-β-D-xilopyranoside), β-*p*NPgal (*p*-nitrophenyl-β-D-galactopyranoside), α-*p*NPG (p-nitrophenyl-α-D-glucopyranoside), α-*p*NPgal (*p*-nitrophenyl-α-D-galactopyranoside), *p*-nitrophenyl-α-L-rhamnopyranoside, *p*-nitrophenyl-β-D-fucopyranoside were assayed analyzing *p*NP release. The reactions were stopped after 10 min by adding 2% (w/v) $Na_2CO_3$, and the *p*NP released was spectrophotometrically measured at 410 nm. One BGL activity unit was defined as the amount of enzyme capable of releasing 1 micromole of *p*NP per min (the molar extinction coefficient of *p*NP is 15,200 $M^{-1} \cdot cm^{-1}$).

β-glucosidase activity on cellobiose, gentiobiose, laminaribiose, cellotriose, cellotetraose, cellopentaose, and cellohexaose, maltose, sucrose and lactose, was quantified by measuring the glucose released from these compounds after enzyme hydrolysis, using the Glucose-TR commercial kit (Spinreact), according to the manufacturer's instructions. Reactions were performed in sodium acetate 100 mM, pH 4.0, incubating in a heating block for 10 min at 1,200 rpm. Then, the reactions were stopped by heating at 100 ºC for 5 min.

The activity of BGL-3 was also determined against different polysaccharides, all of them prepared in 50 mM sodium acetate buffer, pH 4.0: 1.25% Avicel (microcrystalline cellulose), 3% carboxymethyl cellulose (CMC), 1% laminarin from *L. digitata* and *L. hyperborea* and 3% beechwood xylan. The substrates were incubated with BGL-3 in a heating block at 60 ºC and 1,200 rpm for 10 min. The released reducing sugars were determined by the Somogyi-Nelson method (Somogyi, 1945). A degree of polymerization of 25 units has been previously described for laminarin (Fontaine *et al*., 1997).

The kinetic constants of the purified BGL-3 were determined against *p*NPG (over a range of concentrations from 10 μM to 5 mM), *o*NPG (40 μM to 20 mM), cellobiose (80 μM to 40 mM), gentiobiose (80 μM to 40 mM), laminaribiose (80 μM to 40 mM), cellotriose (80 μM to 40 mM), cellotetraose (80 μM 40 mM), cellopentaose (40 μM to 20 mM), and cellohexaose (20 μM to 10 mM). The values of $K_m$ and $V_{max}$ were determined using the program SigmaPlot, based in the Michaelis-Menten model.

$K_i$ for BGL-3 was calculated using *p*NPG as substrate, in the presence of different concentrations of glucose (0, 2.5, 5 and 10 mM).
All enzymatic assays were performed including 0.1% BSA, a protein which does not affect the catalytic activity of the BGL-3, to prevent the activity loss when working with low enzyme concentrations (Nieto-Domínguez *et al*., 2015).

**Physicochemical properties**

To obtain the peptide mass fingerprint of the protein, the sample was run in a SDS-PAGE gel as explained before, excising the BGL-3 band. After tryptic digestion (Shevchenko *et al*., 2007), the peptides' mixture was analyzed in a MALDI-TOF/TOF Autoflex III (Bruker Daltonics) equipped with a laser and a Smartbeam LIFT-MS/MS device. The data from MS and MS/MS experiments were combined using the 3.0 BioTools (Bruker Daltonics) software and searched against the NCBInr database using 2.3 MASCOT as the search engine (Matrix Science). Relevant search parameters were: trypsin as enzyme, carbamidomethylation of cysteines as fixed modification, methionine oxidation as variable modification, 1 missed cleavage allowed, peptide tolerance of 50 ppm, and MS/MS tolerance of 0.5 Da. Protein scores greater than 75 were considered significant.

The molecular mass of the native BGL-3 was determined both by size exclusion chromatography using a Superose 12 column (GE Healthcare), and by MALDI-TOF in the instrument described above. Isoelectric point (pI) was determined by isoelectrofocusing (IEF) in 5% (w/v) polyacrylamide gels, prepared with Pharmalyte (pH 3.0–10.0) as carrier ampholytes (GE Healthcare), using a Mini Protean III Cell system (Bio-Rad). 1 M $H_3PO_4$ and 1 M NaOH were the anode and cathode buffers, respectively. The pH gradient was directly measured on the gel using a contact electrode (Crison). The activity of BGL-3 was tested in zymograms after IEF, incubating the gel with 2 mM *p*-methylumbelliferyl-β-D-glucopyranoside (Sigma-Aldrich) for 10 min, and observing the gel under UV light with a Gel Doc XR+ system (Bio-Rad) to detect free 4-methylumbelliferone.

The optimal values of pH and temperature and the stability of BGL-3 were evaluated with *p*NPG as model substrate, measuring the residual activity after the treatments in standard conditions. The buffer Britton-Robinson (100 mM) was used to study the effect of pH on BGL-3 activity, adjusting different aliquots to pH values from 2.0 to 10.0. BGL-3 was incubated at 4 °C and different pH values for 3 days. After this time, a standard BGL reaction was performed to determine its optimal pH. Temperature assays were done between 30-80 °C using solutions of BGL-3 in acetate pH 4.0. Its thermostability was analyzed in the same temperature range for 72 h, taking aliquots at different incubation times to measure the residual activity.

**Saccharification of wheat straw slurry and laminarin**

Enzymatic saccharification was tested in samples of wheat straw slurry from steam explosion (kindly provided by Abengoa). For saccharification, 100 mg of wheat straw slurry were treated with 2 U/mL of BGL activity in 100 mM sodium acetate buffer, pH 4 (final volume of 2 mL), incubating in a heat block at 50 ºC and 1200 rpm, for 120 h. The sources of BGL activity tested were Celluclast 1.5L (Novozymes), a basal cocktail for biomass degradation with low BGL activity, NS-50010 (Novozymes), which is a β-glucosidase-rich cocktail, and the purified BGL-3. The control sample contained Celluclast 1.5L as the unique source of BGL activity. To compare the efficiencies of NS-50010 and BGL-3, 1 U/mL of BGL activity from Celluclast 1.5L was supplemented with 1 U/mL of either NS-50010 or the purified BGL-3. The glucose released was measured in sample supernatants at different time intervals, using the Glucose-TR commercial kit (Spinreact), according to the manufacturer's instructions.

Similarly, the release of glucose from laminarin, from *Laminaria digitata* (Sigma-Aldrich) and *Laminaria hyperborea* (Koch-light laboratories), was evaluated. The reaction mixtures contained 100 mg of laminarin in 10 mL of 100 mM sodium acetate buffer, pH 4.0 and 3U/mL of laminarinase activity. This was provided by either the purified BGL-3 or a commercial β-1,3-glucanase from *Helix pomatia* (Sigma-Aldrich). Reactions were performed in a heat block at 50 ºC and 1,200 rpm for 24 h, measuring the glucose release at different times as detailed above.

**Declarations**

**Authors' contributions**

JM, LdE contributed to design the research and participated in the experiments. JM drafted the manuscript. MJM, AP and LdE co-coordinated the research and reviewed and edited the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

The authors thank the Proteomics and Genomics facility at CIB, Novozymes for providing the commercial cocktails, and Abengoa for supplying pretreated wheat straw. J.A. Méndez thanks its fellowship to Tatiana Pérez de Guzmán el Bueno Foundation. We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI). We are indebted to A. Serrano for her help using protein structure software.

**Competing interests**

The authors declare that they have no competing interests.

**Availability of supporting data**

*T. amestolkiae* whole genome shotgun project is deposited at DDBJ/ENA/GenBank under the accession number MIKG00000000. BGL-3 sequence is deposited in GenBank under the accession number KM393202.1.

**Funding**
This work has been funded by projects S2013/MAE-2907 from Comunidad de Madrid and RTC-2014-1777-3 from MINECO and BIO2015-68387-R from MINECO/FEDER.

**Authors' information**
Not applicable

## REFERENCES

Amore, A., Giacobbe, S., Faraco, V., 2013. Regulation of cellulase and hemicellulase gene expression in fungi. Curr. Genomics. 14, 230-249.

Bowman, S.M., Free, S.J., 2006. The structure and synthesis of the fungal cell wall. BioEssays. 28, 799-808.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156-159.

Dankai, W., Pongpom, M., Vanittanakom, N., 2015. Validation of reference genes for real-time quantitative RT-PCR studies in *Talaromyces marneffei*. J. Microbiol. Methods 118, 42-50.

de Eugenio, L.I., Méndez-Líter, J.A., Nieto-Domínguez, M., Alonso, L., Gil-Muñoz, J., Barriuso, J. *et al*., 2017. Differential β−glucosidase expression as a function of carbon source availability in *Talaromyces amestolkiae*: a genomic and proteomic approach. Biotechnol. Biofuels. 10, 161.

Fontaine, T., Hartland, R.P., Beauvais, A., Diaquin, M., Latge, J.P., 1997. Purification and characterization of an endo-1,3-β-glucanase from *Aspergillus fumigatus*. Eur. J. Biochem. 243, 315-321.

Gao, L., Gao, F., Zhang, D., Zhang, C., Wu, G., Chen, S., 2013. Purification and characterization of a new β-glucosidase from *Penicillium piceum* and its application in enzymatic degradation of delignified corn stover. Bioresour. Technol. 147, 658-661.

Hannon, M., Gimpel, J., Tran, M., Rasala, B., Mayfield, S., 2010. Biofuels from algae: challenges and potential. Biofuels. 1, 763-784.

Hong, J., Tamaki, H., Kumagai, H., 2007. Cloning and functional expression of thermostable β−glucosidase gene from *Thermoascus aurantiacus*. Appl. Microbiol. Biotechnol. 73, 1331-1339.

Hou, X., Hansen, J., Bjerre, A.B., 2015. Integrated bioethanol and protein production from brown seaweed Laminaria digitata. Bioresour. Technol. 197, 310-317.

Hrmova, M., Fincher, G.B.2017. Barley beta−D-glucan exohydrolases. Carbohydr. Res. 305:209-21.

Igarashi, K., Tani, T., Rie, K., Masahiro, S., 2003. Family 3 β-glucosidase from cellulose-degrading culture of the white-rot fungus *Phanerochaete chrysosporium* is a glucan 1,3-β-glucosidase. J. Biosci. Bioeng. 95, 572-576.

Jeya, M., Joo, A.R., Lee, K.M., Tiwari, M.K., Lee, K.M., Kim, S.H. *et al*., 2010. Characterization of β−glucosidase from a strain of *Penicillium purpurogenum* KJS506. Appl. Microbiol. Biotechnol. 86, 1473-1484.

John, P., Gs, A., Nampoothiri, K.M., Pandey, A. 2011. Micro and macroalgal biomass: A renewable source for bioethanol. Bioresour. Technol. 102(1):186–93.

Jones, C.S., Mayfield, S.P., 2012. Algae biofuels: versatility for the future of bioenergy. Curr. Opin. Biotechnol. 23, 346-351.

Karnaouri, A., Topakas, E., Paschos, T., Taouki, I., Christakopoulos, P. 2013. Cloning, expression and characterization of an ethanol tolerant GH3 β–glucosidase from *Myceliophthora thermophila*. Peer J. 1, e46.

Lafond, M., Navarro, D., Haon, M., Couturier, M., Berrin, J.G., 2012. Characterization of a broad-specificity β-glucanase acting on β-(1,3)-, β-(1,4)-, and β-(1,6)-glucans That defines a new glycoside hydrolase family. Appl. Environ. Microbiol. 78, 8540-8546.

Liu, D., Zhang, R., Yang, X., Zhang, Z., Song, S., Miao, Y. *et al*., 2012. Characterization of a thermostable β–glucosidase from *Aspergillus fumigatus* Z5, and its functional expression in *Pichia pastoris* X33. Microb. Cell. Fact. 11, 25-25.

Mandels, M., Weber, J. 1969. Production of cellulases. Adv. Chem. Ser. 95:391.

Marsh, J.W., Denis, J., Wriston, J.C., 1977. Glycosylation of *Escherichia–coli* L-asparaginase. J. Biol. Chem. 252, 7678-7684.

Méndez-Líter, J.A., Gil-Muñoz, J., Nieto-Domínguez, M., Barriuso, J., de Eugenio, L.I., Martínez, M.J., 2017. A novel, highly efficient β–glucosidase with a cellulose-binding domain: characterization and properties of native and recombinant proteins. Biotechnol. Biofuels. 10, 256.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A comprehensive review. Renew. Sustain. Energy Rev. 14, 578-597.

Nakatani, Y., Lamont, I.L., Cutfield, J.F., 2010. Discovery and characterization of a distinctive exo-1,3/1,4-β-glucanase from the marine bacterium *Pseudoalteromonas sp*. strain BB1. Appl. Environ. Microbiol. 76, 6760-6768.

Nieto-Domínguez, M., de Eugenio, L.I., Barriuso, J., Prieto, A., Fernández de Toro, B., Canales-Mayordomo, A. *et al*., 2015. Novel pH-stable glycoside hydrolase family 3 β-xylosidase from *Talaromyces amestolkiae*: an enzyme displaying regioselective transxylosylation. Appl. Environ. Microbiol. 81, 6380-6392.

Nitsche, B.M., Jörgensen, T.R., Akeroyd, M., Meyer, V., Ram, A.F. 2012. The carbon starvation response of *Aspergillus niger* during submerged cultivation: Insights from the transcriptome and secretome. BMC Genomics. 13, 380-380.

O'Connell, E., Piggott, C., Tuohy, M., 2011. Purification of exo-1,3-beta–glucanase, a new extracellular glucanolytic enzyme from *Talaromyces emersonii*. Appl. Microbiol. Biotechnol. 89, 685-696.

Ramani, G., Meera, B., Vanitha, C., Rajendhran, J., Gunasekaran, P., 2015. Molecular cloning and expression of thermostable glucose-tolerant β–glucosidase of *Penicillium funiculosum* NCL1 in *Pichia pastoris* and its characterization. J. Ind. Microbiol. Biotechnol. 42, 553-565.

Riou, C., Salmon, J.M., Vallier, M.J., Gunata, Z., Barre, P. 1998. Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from *Aspergillus oryzae*. Appl. Environ. Microbiol. 64, 3607-3614.

Ruel, K., Joseleau, J.P., 1991. Involvement of an extracellular glucan sheath during degradation of populus wood by *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. 57, 374-384.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M., 2007. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protocols. 1, 2856-2860.

Singhania, R.R., Patel, A.K., Sukumaran, R.K., Larroche, C., Pandey, A., 2013. Role and significance of beta−glucosidases in the hydrolysis of cellulose for bioethanol production. Bioresour. Technol. 127, 500-507.

Somogyi, M., 1945. A new reagent for the determination of sugars. J. Biol. Chem. 160, 61-68.

Sun, Y., Cheng, J.Y., 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour. Technol. 83, 1-11.

van Munster, J.M., Daly, P., Delmas, S.p., Pullan, S.T., Blythe, M.J., Malla, S. *et al*., 2014. The role of carbon starvation in the induction of enzymes that degrade plant-derived carbohydrates in *Aspergillus niger*. Fungal. Genet. Biol. 72, 34-47.

Wang, Z., Robertson, L., Liu, C., Liu, J., Johnson, B., Leary, D., Compton, J., Vuddhakul, V., Legler, P., Vora, G. 2015. A novel vibrio beta−glucosidase (LamN) that hydrolyzes the algal storage polysaccharide laminarin, FEMS Microbiol. Ecol.

White, S., Workman, M., Berry, R., Mcneil, B. 2002. The autolysis of industrial filamentous fungi. Crit. Rev. Biotechnol. 22(1):1–4.

Xia, W., Xu, X., Qian, L., Shi, P., Bai, Y., Luo, H. *et al*., 2016. Engineering a highly active thermophilic β−glucosidase to enhance its pH stability and saccharification performance. Biotechnol. Biofuels. 9, 147.

Yang, X., Ma, R., Shi, P., Huang, H., Bai, Y., Wang, Y. *et al*., 2014. Molecular characterization of a highly-active thermophilic β−glucosidase from *Neosartorya fischeri* P1 and its application in the hydrolysis of soybean isoflavone glycosides. PLoS One 9, e106785.

Zhou, Y., Wang, X., Wei, W., Xu, J., Wang, W., Xie, Z. *et al*., 2016. A novel efficient β-glucanase from a paddy soil microbial metagenome with versatile activities. Biotechnol. Biofuels. 9, 36.

# SUPPLEMENTARY INFORMATION

**Table S1.** Specific activity of BGL-3 against other glycosides, disaccharides and polysaccharides.

| Substrate | Specific activity (U/mg) |
|---|---|
| ***p*NPX (*p*-nitrophenyl-β-D-xylopyranoside)** | 7.58±0.24 |
| **β-*p*NPgal (*p*-nitrophenyl-β-D-galactopyranoside)** | 0.06±0.01 |
| **α-*p*NPG (*p*-nitrophenyl-α-D-glucopyranoside)** | 6.26±0.12 |
| **α-*p*NPgal (*p*-nitrophenyl-α-D-galactopyranoside)** | 0 |
| ***p*-nitrophenyl-α-L-rhamnopyranoside** | 0 |
| ***p*-nitrophenyl-β-D-fucopyranoside** | 0 |
| **Maltose** | 3.42±0.14 |
| **Sacarose** | 0.26±0.01 |
| **Lactose** | 0.26±0.12 |
| **1.25% Avicel** | 7.01±0.27 |
| **3% Xylan** | 12.61±0.39 |
| **3% Carboxymethyl cellulose (CMC)** | 10.56±0.21 |

**Table S2.** Purification of recombinant BGL-3

| **BGL-3* Purification** | | | | |
|---|---|---|---|---|
| **Step** | **Total protein (mg)** | **Total activity (U)** | **Specific activity (U/mg)** | **Yield (%)** |
| Crude extracts | 22.17 | 287.36 | 19.96 | 100.00 |
| HiTrap QFF | 2.83 | 156.43 | 55.27 | 54.43 |
| Superose 12 | 0.21 | 43.63 | 207.78 | 54.27 |

ATGCGGAACAGTTTATTGATTTCGCTTGCTGCGGCAGCACTTGCCGAGGGCAAGGCC**TAC**TCTCCTCCAGCTTACCCT
GCTCCCTGGGCCAGTGGCGCCGGGGAATGGGCTCAAGCTCATGACAGAGCAGTCGAGTTCGTTTCGCAATTGACCTTG
GCCGAGAAGATAAACCTGACGACTGGTGTTGG**GTACGTTGTGATACCGTTGAGCAAAATAAAGACATGGGACTGACAA**
**TTGGTAG**ATGGGAGGGTGGACAATGTGTCGGTAACACTGGAAGCATTCCCCGCCTGGGATTCCGCAGCCTCTGTATGC
AGGATTCACCGCTCGGCGTGAGAGACA**GTATGTCGTGCCCTGTAGTCTCTTGCCTCTTACTTCTTACCGTCCAAATAC**
**AAAAAGAGAAAAAAAAAAAGGTTGAATTATATAGCTAACAGTTTTATGTTTCTAG**CTGACTACAATACTGCCTTCCCT
GCTGGCGTCAATGTCGCCGCTACCTGGGATCTCGATCTTGCATACCGGCGCGGTGTAGCCATGGCTGAGGAACACCGT
GGCAAAGGTGTGGATGTTCAGCTTGGACCCGTTGCTGGTCCGCTAGGAAGAGCACCAGAGGGTGGCCGTAATTGGGAA
GGCTTTGCACCCGACCCCGTTTTGACTGGTCAGATGATGGCAAGCACTATTGAAGGAATGCAGGATACCGGTGTGATT
GCTTGTGCAAAGCACTATATCGGTAATGAGCAAGAGCACTTTCGTCAGGGCTCCCAGGAAAATTATACAGTCGCTGAT
GCTATCAGCTCGAACATCGACGATGTTACTTTGCACGAGTTGTACCTGTGGCCGTTTGCCGATGCGGTTAGGGCAGGT
GTTGGTTCCGTCATGTGTTCTTACAATCAATTGAACAACAGTTATTCCTGCGGCAACAGCTACAGTTTGAACCACATT
CTTAAGGGAGAACTCGACTTTCAAGGATTCGTCATGACCGACTGGGGTGCTCAGCATTCTGGTGTGGGCGATGCTTTG
GCCGGTGCGGATATGGATATGCCTGGTGATGTGGCTTTCGACAGTGGAACTGCTTTCTGGGGTACTAACTTGACAATT
GCCGTGCTCAATGGAACTGTTCCTGAATGGCGTATTGACGACATGGCCGTTCGTATCATGTCTGCATTCTATAAGGTT
GGTCGTGATCGTACCCAGGTCCCCATCAACTTTGCTAGCTGGACTCTGGATACCTATGGCAATGAATACTACTACGCC
GGCGAGGGCTACAAGGAAATCAACCAGCACGTTGATGTACGTGGTGACCACGCCAAAGTTGTCCGTGAAATCGGCAGT
GCCAGCATTGTTCTCCTCAAGAATGTTGACGGCGCTCTTCCGTTGACTGGCTCCGAGAGGTTTGTCGCAGTTTTCGGA
GAGGATGCTGGCTCCAATCCTGATGGTGTCAATGGTTGCTCTGACCGTAACTGTGATAACGGTACCTTGGCTATGGGA
TGGGGTAGTGGTACTGCCAACTTCCCTTACCTAGTTACTCCTGAACAAGCTATCCAAGCCGAGGTCTTGAAGAATGGC
GGAATATTTACTGCTATTACCGACAGCGGCGCCACCAATACTACAGCCACGACCGTGGCTGCTCAAGCCTC**GTAAGTA**
**CTGGTCGTAAGTAGAACATAGCAATGGCTCAGATTACTGACTTTGTTTTAATTAG**GGCTTGCCTAGTGTTTGCCAATG
CAGACTCCGGCGAGGGATACATCACCGTTGACGGAAACGTGGGAGATCGTAAGAATTTGACATTATGGCAGAACAGTG
AAGCTATGATCTCGGCCGTTGCAGGTAACTGTAACAACACCATAGTAGTTCTTCATACTGTTGGACCTGTTCTCGTTG
AGGACTGGGTCAACCATCCCAACATCACTGCTGTTTTGTGGGCAGGTTTGCCTGGAGAGCAGAGCGGAAACTCTTTGG
TTGATGTTCTTTACGGCAGCGTCAACCCCGGAGGCAAGACTCCTTTCACTTGGGGCAAGCAACGTTCTGACTGGGGAA
CCGATATCATCTACGAACCCAACAACGGAGATGGTGCTCCTCAGCAGGACTTCACCGAGGGTATCTTCATTGACTACC
GACACTTTGATAAATACAATATTACTCCCACTTACGAGTTTGGTTATGGTCTCAGTTACAGCACCTTCTCTTTCTCAA
ATCTCCAGGTGACTCCTCTCGCTGCTTCGCCTTACAAACCAGCCACAGGTCACAGCGGTCCCGCACCTGTCCTGGGCA
AGGTTTTGAACGCCACGGCTTATCTTTTCCCTAACTACATCAAACGCATTGAAGCTTTCATTTACCCATGGCTTAACT
CCACTGATCTGAGGACTTCCTCTGGTGATCCAAATTACGGTTGGTCTACTTCCAAATACGTGCCTGACGGTGCTCAAG
ACGGATCTCCGCAACCTGTCAACCCCGCTGGTGGTGCCCCTGGTGGCAACCCTGCGCTGTATGACCCTGTTGCCGAAA
TCAGAGTGACTGTCAAGAACACCGGAAAGGTCGCTGGTGTTGAAGTGCCTCAGCTCTATGTCTCGCTCGGTGGCCCCT
CCGATGCGCCTAAGGTTCTTCGTGGCTTTGGCCGCCTTTCTCTCGGTGCTGGCGAGGAGGCTCAGTGGACTGCCACTT
TGACCCGACGTGACGTTTCTAATTGGGACACTGTCAGCCAGAACTGGGTTGTCTCAAACTACACCAAGACTGTCTATG
TCGGCAACTCTTCTCGCAACTTGCCGCTCCAGCAGACTTTGGCTTTGAAGATTGGGCAT**TAA**

**Figure S1.** DNA sequence of *bgl3*. The predicted signal peptide is underlined. Predicted introns are indicated in red.

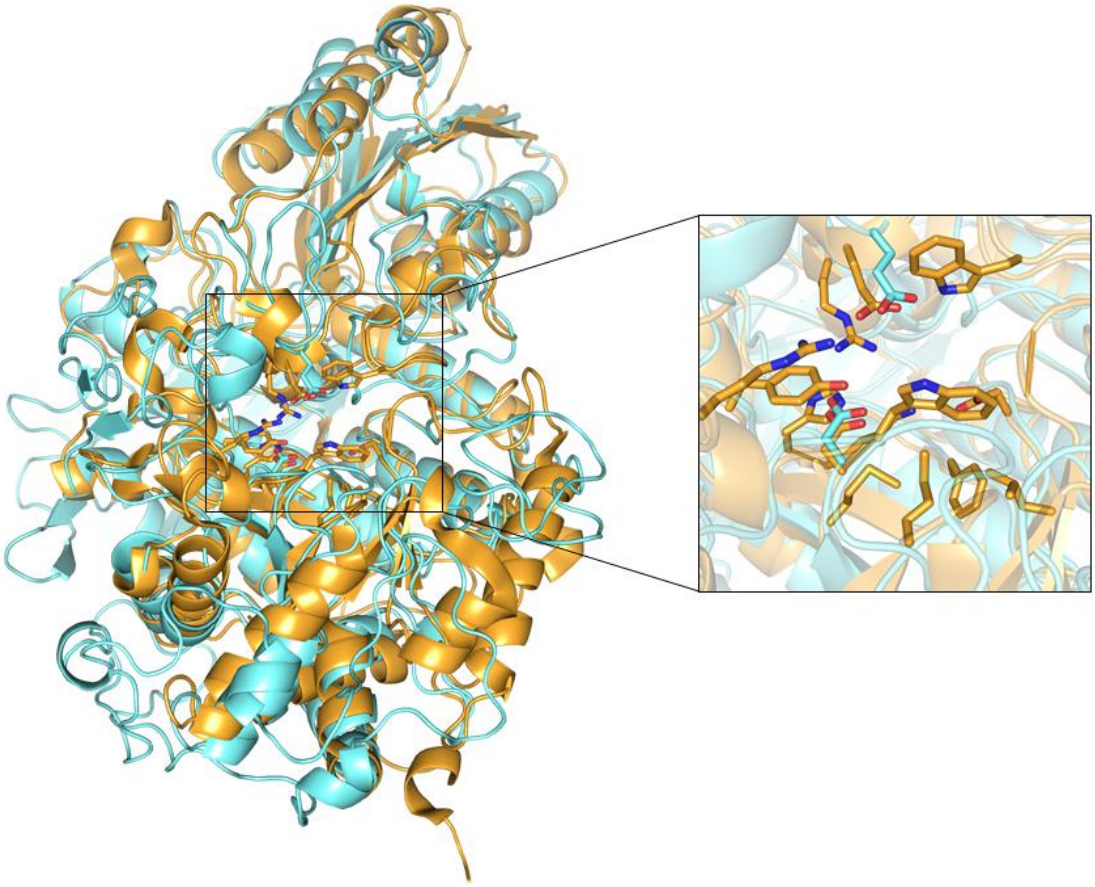


**Figure S2:** Structural comparison between BGL-3 (cyan) and a barley beta-D-glucan glucohydrolase isoenzyme in complex with 4'-nitrophenyl 3I-thiolaminaritrioside (SMTL ID 1j8v.1, orange). The BGL-3 model was generated using the SWISS-MODEL server, based on sequence similarity, using a β-glucosidase from *Aspergillus aculeatus* as template (SMTL ID 4iib.1). Q Mean, coverage and sequence identity for this model were -0.47, 0.96 and 65.75% respectively. The PyMol v0.99 program was used to visualize, analyze and align the structures. Active site residues D254 and E484 on BGL-3 are marked as cyan sticks, residues implied in substrate binding on glucohydrolase appear highlighted as orange sticks.

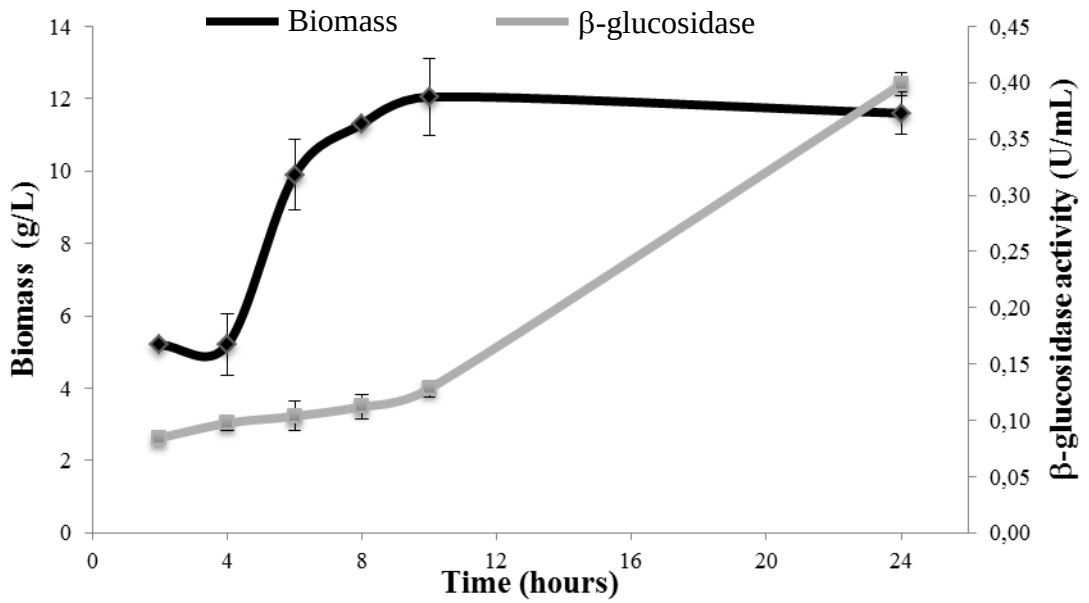



**Figure S3.** Relation between fungal growth (in Mandels medium with glucose as carbon source) and β-glucosidase activity, during the first 24 h of culture.

# CHAPTER 4

**Transglycosylation products generated by *Talaromyces amestolkia*e GH3 β-glucosidases: Effect of hydroxytyrosol, vanillin and its glucosides on breast cancer cells.**

**Méndez-Líter, J.A[1].**, Tundidor, I[2]., Nieto-Domínguez, M[1]., de Toro, B.F[3]., González Santana, A[4]., de Eugenio, L.I[1]., Prieto, A[1]., Asensio, J.L[4]., Cañada, F.J[3]., Sánchez, C[2]., Martínez, M.J[1]., 2019. Transglycosylation products generated by *Talaromyces amestolkiae* GH3 β-glucosidases: effect of hydroxytyrosol, vanillin and its glucosides on breast cancer cells. Microb. Cell Fact. 18, 97. doi: 10.1186/s12934-019-1147-4

[1]Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain.
[2] Department of Biochemistry and Molecular Biology, Complutense University, Madrid, Spain; Instituto de Investigación Hospital 12 de Octubre, Madrid, Spain.
[3] Department of Chemical and Physical Biology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain.
[4]Glycochemistry and Molecular recognition group. Instituto de Química Orgánica General (IQOG-CSIC), Calle Juan de la Cierva, 3, 28006 Madrid, Spain.

## ABSTRACT

**Background:** Transglycosylation represents one of the most promising approaches for obtaining novel glycosides, and plant phenols and polyphenols are emerging as one of the best targets for creating new molecules with enhanced capacities. These compounds can be found in diet and exhibit a wide range of bioactivities, such as antioxidant,
antihypertensive, antitumor, neuroprotective and anti-inflammatory, and the eco-friendly synthesis of glycosides from these molecules can be a suitable alternative for increasing their health benefits.
**Results:** Transglycosylation experiments were carried out using different GH3 β-glucosidases from the fungus *Talaromyces amestolkiae*. After a first screening with a wide variety of potential transglycosylation acceptors, mono-glucosylated derivatives of hydroxytyrosol, vanillin alcohol, 4-hydroxybenzyl alcohol, and hydroquinone were detected. The reaction products were analyzed by thin-layer chromatography, high-pressure liquid chromatography, and mass spectrometry. Hydroxytyrosol and vanillyl alcohol were selected as the best options for transglycosylation optimization, with a final conversion yield of 13.8 and 19% of hydroxytyrosol and vanillin glucosides, respectively. NMR analysis confirmed the structures of these compounds. The evaluation of the biological effect of these glucosides using models of breast cancer cells, showed an enhancement in the anti-proliferative capacity of the vanillin derivative, and an improved safety profile of both glucosides.
**Conclusions:** GH3 β-glucosidases from *T. amestolkiae* expressed in *P. pastoris* were able to transglycosylate a wide variety of acceptors. Between them, phenolic molecules like hydroxytyrosol, vanillin alcohol, 4-hydroxybenzyl alcohol, and hydroquinone were the most suitable for its interesting biological properties. The glycosides of hydroxytyrosol and vanillin were tested, and they improved the biological activities of the original aglycons on breast cancer cells.


**Keywords**: Transglycosylation; β-glucosidases; Glycosyl hydrolases; Hydroxytyrosol; Vanillin; Glucosides; Breast cancer cells.

## BACKGROUND

Glycosyl hydrolases (GHs) are enzymes that hydrolyze glycosidic linkages, and are essential in nature for the multiplicity of roles that they play. The huge diversity of natural carbohydrates is directly correlated with the wide variety of GH-type activities reported so far (Cantarel *et al*., 2009). The traditional concept of glycosidases mostly refers to their hydrolytic capacities, and their current applications are mainly associated to the degradation of lignocellulosic biomass and food industry (Linares-Pasten *et al*., 2014). Retaining glycosyl hydrolases act through a double-displacement mechanism involving the formation of a covalent glycosyl-enzyme intermediate, which is subsequently cleaved upon nucleophilic attack by water (hydrolysis reactions). However, they can also produce new glycosidic bonds when alternative nucleophiles to water participate as acceptors, through a mechanism named transglycosylation (Edelman, 1956). As a result of the latter reaction, a new glycoside is synthesized by the transfer of a sugar unit (eg: glucose, galactose, xylose, fructose…) to the nucleophilic acceptor.

The ability of GHs to catalyze transglycosylation reactions has made these enzymes a suitable alternative to chemical approaches for obtaining different glycosides (Xu *et al*., 2016). These enzymatic conversions allow a more environmentally friendly synthesis of these compounds, compared to the current chemical approaches, which usually imply the formation of toxic byproducts, need more steps, and display lower regio- and stereoselectivity (Danby and Withers, 2016). On the other hand, glycosidases outstand for having a complete stereoselectivity and a remarkably greater regioselectivity (Danby and Withers, 2016).

Many beneficial effects of glycosides have been reported, for instance: increasing solubility of the original compound (Kometani *et al*., 1994; Torres *et al*., 2011; Woo *et al*., 2012), making the new compound safer (Prodanovic *et al*., 2005) or improving its stability (Yamamoto *et al*., 1990; Woo *et al*., 2012; Dinesh *et al*., 2016). In general, enzymatic transglycosylation by glycosidases is considered as a good and eco-friendly alternative over traditional chemical synthesis for obtaining novel molecules with added value.

Phenolic compounds are molecules that possess one or more aromatic rings with one or more hydroxyl groups. They are broadly distributed in the Plant Kingdom, and are the most abundant secondary metabolites in plants, ranging from simple molecules such as aromatic acids, to highly polymerized substances like tannins (Bravo, 2009). Recently, the health effects of many phenolic compounds have come to the attention of nutritionists, mainly due to their antioxidant capacity and their potential effects in the prevention of various diseases associated to oxidative stress. The demonstrated effects of these plant metabolites, in terms of cardiovascular, neurodegenerative diseases and cancer (Dai and Mumper, 2010), have postulated them as an interesting option for preventing or even treating diseases. Nevertheless, one of the main

disadvantages of these molecules is their usually poor bioavailability when administered to humans.

*Talaromyces amestolkiae* has been postulated as a great GH producer (Nieto-Domínguez *et al*., 2015; de Eugenio *et al*., 2017) whose highly efficient β-glucosidases (BGLs) from the family GH3 deserve special attention. In this context, BGL-2 was the first fully characterized BGL that possessed a cellulose-binding domain (CBD), and it displayed a very high efficiency against cellobiose and other oligosaccharides. A truncated version, without the CBD region (BGL-2T) keeps the efficiency in hydrolytic reactions (Méndez-Líter *et al*., 2017). The other GH3 β-glucosidase characterized from this fungus, named BGL-3, is a versatile enzyme produced under carbon starvation, that hydrolyzes efficiently typical 1,4-β-glucosidase substrates, but also shows high activity over 1,3-β-glucose bonds (Méndez-Líter *et al*., 2018).

The aim of this work is to study the transglycosylation profile of these β-glucosidases (BGL-2, BGL-2T and BGL-3) using different acceptors, and to test the properties of the most interesting glycosides as potential antitumor agents in preclinical models of cancer.

## RESULTS AND DISCUSSION

### The BGLs of *T. amestolkiae* are versatile tools for transglycosylation

The cellulolytic system of *T. amestolkiae* has shown an outstanding hydrolytic efficiency over cellulose, but the potential of some of these enzymes in the synthesis of novel compounds by transglycosylation has not been evaluated so far. Many GHs can catalyze this kind of reaction, and numerous strategies have been developed in order to obtain added-value compounds by transglycosylation (van Rantwijk *et al*., 1999; Urrutia *et al*., 2013; Piedrabuena *et al*., 2016; Nieto-Domínguez *et al*., 2016; Nieto Dominguez *et al*., 2017; Míguez *et al*., 2018). In order to test the transglycosylation activity of the GH3 β-glucosidases BGL-2, BGL-2T and BGL-3, we first performed a screening of potential acceptors, according to the methodology previously developed (Blanchard and Withers, 2001). The 70 compounds tested (listed in Table 1) encompass a wide variety of alcohols including sugars, sterols, phenolic compounds, or amino acids. It is very important to remark that this method allows analyzing a great number of compounds, but the most interesting positive hits identified for each enzyme in this screening should be subjected to a second screening by Thin Layer chromatography to corroborate the results, thus discarding false positives. Out of the 70 potential acceptors assayed, 32, 31 and 35 were determined as positive hits for BGL-2, BGL-2T and BGL-3, respectively (Table 1). This number of positive hits for each BGL was high compared to those previously reported for other GHs (Blanchard and Withers, 2001), but relatively low if compared to what has been recently published for a β-xylosidase from *T. amestolkiae* (Nieto-Domínguez *et al*., 2016): It is remarkable that the three assayed enzymes were able to produce glucosides of

several phenolic compounds, many of them with reported antioxidant properties, such as epigallocatechin gallate (EGCG), hydroxytyrosol, vanillyl alcohol, 4-hydroxybenzyl alcohol, or hydroquinone. However most of them have a poor bioavailability, which limits their use in many industrial or biomedical applications. In this regard, their glycosylation may provide a solution to this problem by converting them into more soluble conjugates. Other positive hits for the three enzymes were detected when 4-nitrophenol galacto-, gluco- and xylo-pyranosides, and also disaccharides like lactose or melibiose, were used as acceptors. These results open up the possibility for synthesizing different oligosaccharides in a regio- and stereoselective manner. Many of these molecules have shown a great prebiotic effect, which is related with health benefits (Macfarlane *et al*., 2006; Rastall and Gibson, 2015). In addition, the synthesis of disaccharide derivatives of 4-nitrophenol seems to be promising, since they could be used as novel substrates for other GHs.

In the case of BGL-3, the glycosylation of L-serine, L-threonine, and L-tyrosine is very remarkable. Protein glycosylation is considered of fundamental importance (Spiro, 2002), thus the formation of glycosidic bonds with amino acids may have many applications, from serving as an assembly for synthetic glycopeptides that can be used for triggering tumor-cell-specific immune response, to acting as ligands of carbohydrate-binding proteins or as enzyme substrates or inhibitors (Seitz, 2000).

It is also worth noting that, although the three assayed BGLs belong to the same glycosyl hydrolase family, some differences can be found among their positive acceptor profiles (Table 1). This diversity in the transglycosylation profiles, even among enzymes that are considerably similar in terms of sequence, generates expectations of wider but selective applications via different enzyme variants. These variations could seem surprising among BGL-2 forms, although they agree with the different hydrolytic efficiencies over *p*-nitrophenol sugars and cellooligosaccharides shown by BGL-2 and BGL-2T, two enzymes that only differ on the presence or absence of the CBD.

**Table 1**. Inhibition recovery of *T. amestolkiae* BGLs in the presence of different molecules. The hits with higher recovered activity than no-acceptor control were considered potential acceptors of transglycosylation. In the table, acceptor efficiency was defined by symbols, being +++ which represents the best efficiency, and – if there is no activity recovered.

| **Molecule** | **BGL-2** | **BGL-2T** | **BGL-3** |
|---|---|---|---|
| 1-Butanol | + | - | - |
| 1-Heptanol | - | - | + |
| 1-Propanol | + | - | - |
| 2,4-Dinitrophenol | - | - | - |
| 2,6-Dihydroxynaphthalene | - | - | + + |
| 2-Butanol | - | + | - |
| 2-Mercaptoethanol | - | + + + | + |
| *2*-Nitrophenyl β-D-glucopyranoside | - | - | - |
| 2-Propanol | + + + | + | + |
| 3,3-Diphenyl propanol | - | + | + + |
| 4-Cresol | + | + + | - |
| 4-Hydroxybenzyl alcohol | + | + + + | + |
| 4-Methylumbilliferyl β-D-xylopyranoside | + + + | + + | |
| 4-Nitrophenol | - | - | + + + |
| *4*-Nitrophenyl α-arabinopyranoside | - | + | - |
| *4*-Nitrophenyl α-D-glucopyranoside | + | + + | + |
| *4*-Nitrophenyl α-D-rhamnopyranoside | - | - | + + |
| 4-Nitrophenyl β-D-fucopyranoside | - | - | - |
| *4*-Nitrophenyl β-D-galactopyranoside | + + | + + + | + + + |
| *4*-Nitrophenyl β-D-glucopyranoside | + + + | + + + | + + + |
| *4*-Nitrophenyl β-D-xylopyranoside | + + + | + + + | + + + |
| L-Arabinose | - | - | - |
| Arabitol | - | + + + | - |
| Ascorbic acid | - | - | - |
| Catechol | - | + | + + |

| | | | |
|---|---|---|---|
| Cellobiose | - | - | - |
| Cinnamyl alcohol | + | - | + + + |
| Cyclohexanol | - | - | + |
| Dulcitol | + + + | + | + + + |
| EGCG | + + + | + + + | + + + |
| Ergosterol | + + + | + | + |
| Ethanol | - | + | - |
| Eugenol | - | - | + + + |
| Ferulic acid | - | - | - |
| D-Fructose | - | - | - |
| D-Galactose | + | + | - |
| Gallic acid | + + | - | - |
| Gentiobiose | - | - | - |
| D-Glucose | - | - | - |
| Glycerol | + | - | + + + |
| Guaiacol | - | - | - |
| Hydroquinone | + | + | + |
| Hydroxytyrosol | + + + | + + + | + + + |
| *myo*-Inositol | + + + | + + + | + + + |
| Lactose | + + + | + + + | + + + |
| Maltose | - | + + | + + + |
| Mannitol | + + | + + + | - |
| D-Mannose | - | - | - |
| Melibiose | + | + | + + |
| Menthol | - | - | + |
| Methanol | + | - | - |
| Naphthol | + | + | + |
| Phenol | + + + | + + + | + |
| Propargyl alcohol | - | - | - |
| Quercetin | - | - | - |

| | | | |
|---|---|---|---|
| Raffinose | + | + | + + + |
| Resveratrol | + | - | - |
| D-Ribose | - | - | - |
| L-Serine | - | + + | + + + |
| Sorbitol | + + + | - | + + |
| Sorbose | - | - | - |
| Sucrose | - | + | - |
| L-Threonine | - | - | + |
| L-Trehalose | + | - | - |
| L-Tyrosine | - | - | + + + |
| Vanillyl alcohol | + | + + | + |
| Xylitol | + + + | - | - |
| D-Xylose | - | - | - |
| α-Tocopherol | - | - | - |
| β-Sitosterol | - | - | - |

## Screening of transglycosylation products by TLC

After the first screening, several phenolic antioxidants with potential biotechnological applications and solubility limitations that could be avoided by glycosylation were submitted to a second screening to eliminate false positive acceptors. These compounds were EGCG, hydroxytyrosol, hydroquinone, 4-hydroxybenzyl alcohol, and vanillyl alcohol. Non identifiable glycosylated derivatives were detected for EGCG, which discarded it as an effective transglycosylation acceptor in this system, in spite of the positive results of the preliminary screening, which ratifies the necessity of performing a second assay to confirm BGL acceptors. However, the other phenolic compounds evaluated did show their corresponding glucoside bands on TLC (figure 1).

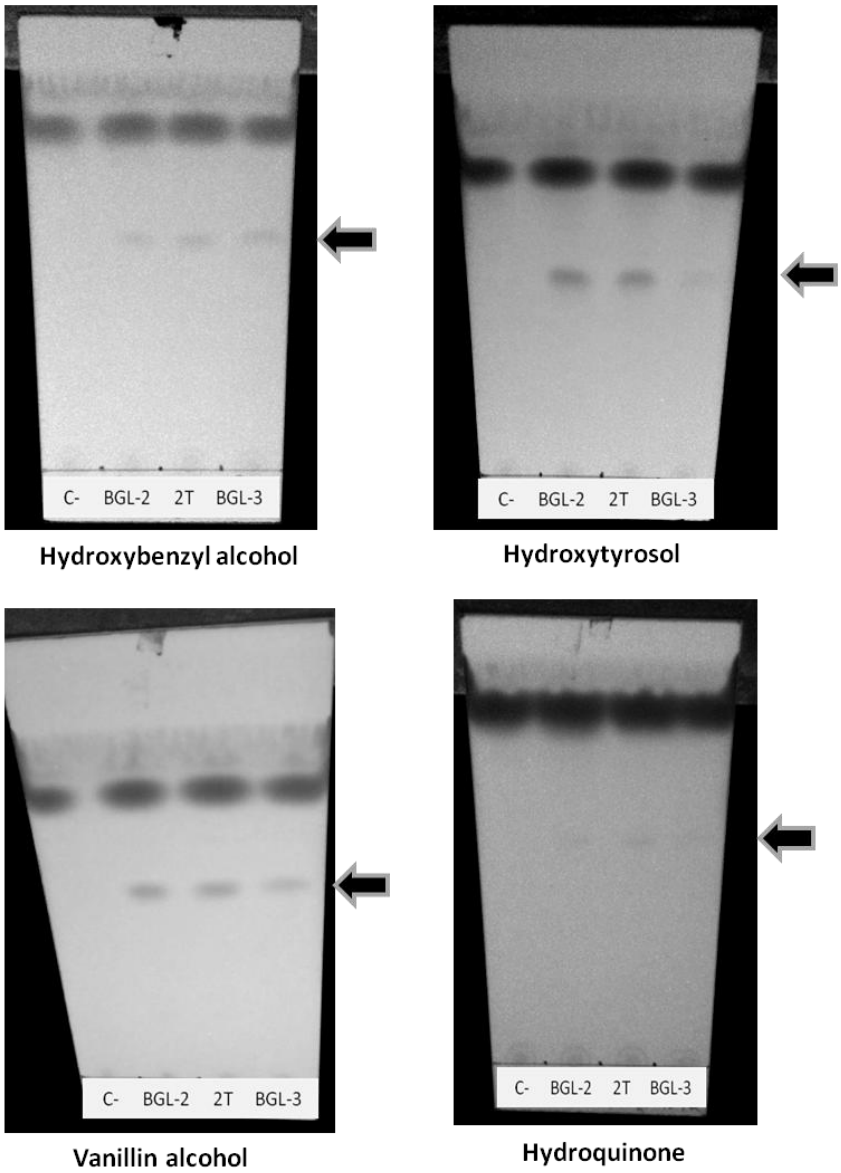



**Figure 1.** Thin layer chromatography analysis of transglycosylation reactions of hydroxytyrosol, vanillyl alcohol, hydroxybenzyl alcohol and hydroquinone. Arrows point at the reaction products.

**HPLC and MS analysis of the transglycosylation products and selection of the most efficient BGL**

In order to determine which of the evaluated BGLs displayed a higher yield in glucoside synthesis, the transglycosylation products of hydroxytyrosol, hydroquinone, 4-hydroxybenzyl alcohol and vanillyl alcohol (figure 2) were analyzed and quantified by HPLC. The peaks from these glucosides were detected by their absorbance at 270 nm, since all of the selected compounds contain aromatic rings and, at this wavelength, the remaining cellobiose or other side products, such as glucose, did not interfere with the analysis.

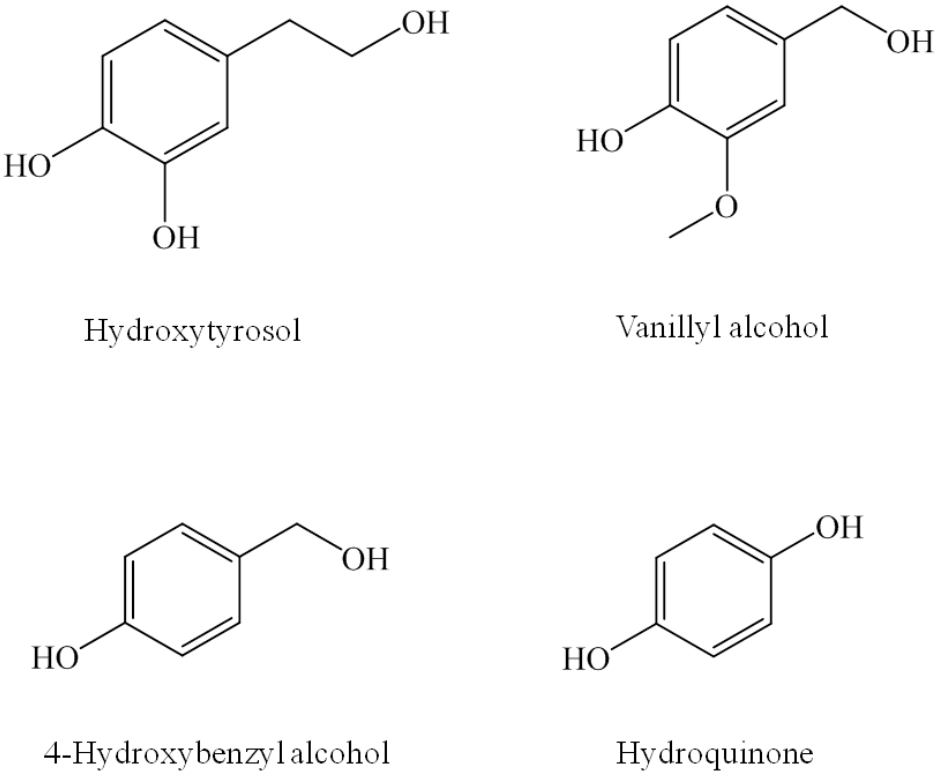



**Figure 2.** Molecular structures of the positive acceptor hits of transglycosylation.

The results (Table 2) showed that yields obtained with BGL-2 and BGL-2T were very similar in every case, whereas transglycosylation yield catalyzed with BGL-3 was lower. In addition, these analyses corroborated the higher efficiency as acceptors of hydroxytyrosol, 4-hydroxybenzyl alcohol and vanillyl alcohol over hydroquinone, since the corresponding product yield was 5-10-fold higher. Since BGL-2 was produced with a better yield and was easier to purify than BGL-2T (Méndez-Líter *et al*., 2017), this enzyme was selected for scaling up the transglycosylation reaction for the four selected glucosides. The presence of the expected mono-glucosylated derivative for each acceptor was corroborated by electrospray ionization-mass spectrometry (ESI-MS) (Table 3).

All the synthesized products are potentially interesting from a biotechnological point of view, since their phenolic precursors already outstand for their applications. Several studies have remarked that hydroxytyrosol has

antioxidant, anti-proliferative and anti-inflammatory activities, and beneficial effects on the cardiovascular system by preventing oxidative stress (Hu *et al*., 2014; Nieto Dominguez *et al*., 2017). Hydroquinone is an interesting molecule as a water-soluble reducing agent, but its main applications are related to skin depigmentation, although its side effects on human health are an increasing concern (Diana Draelos, 2007). With this idea in mind, an approximation from a prodrug point of view could be really interesting, with the hydroquinone being potentially inactivated in its glycosylated form, and then activated after the action of endogenous glycosidases. Although vanillyl alcohol has been scarcely used, vanillin has a wide variety of applications as a flavoring agent in food or beverages (Gallage and Møller, 2015), but also for having potential anti-proliferative and neuroprotective effects (Lirdprapamongkol *et al*., 2005; Kim *et al*., 2011). The anti-angiogenic, anti-inflammatory and anti-nociceptive applications of 4-hydroxybenzyl alcohol make it one of the best-known phenolic compounds isolated from plants (Lim *et al*., 2010; Laschke *et al*., 2011;). Out of the four products synthesized by transglycosylation with BGL-2, and on the basis of the interesting properties of their corresponding precursors and the good transglycosylation yields obtained in previous experiments, the glucosides of hydroxytyrosol and vanillyl alcohol were selected to optimize their production and to test their biological activity.

**Table 2.** Comparison of the transglycosylation activity of BGLs from T. amestolkiae. Final product yields are given in mM concentration.

| **Glycosylated product** | **BGL-2** | **BGL-2T** | **BGL-3** |
|---|---|---|---|
| Hydroxytyrosol | 2.41 | 2.56 | 0.89 |
| Vanillyl alcohol | 4.09 | 3.65 | 1.72 |
| Hydroquinone | 0.35 | 0.42 | 0.28 |
| 4-hydroxybenzyl alcohol | 1.63 | 1.54 | 0.83 |

**Table 3.** ESI-MS data for the products obtained by transglycosylation catalyzed by BGL-2. All the glycosides were detected as $Na^+$ adducts.

| Glycoside | Intensity | m/z |
|---|---|---|
| Hydroxytyrosol | 40,201 | 339.0 |
| Vanillyl alcohol | 240,349 | 339.0 |
| 4-Hydroxybenzyl alcohol | 110,093 | 309.0 |
| Hydroquinone | 26,153 | 295.0 |

## Optimization of glucosides synthesis by response surface methodology

The reaction conditions for hydroxytyrosol and vanillyl alcohol glucosides were optimized using a response surface method, the Box-Behnken design, in order to improve their production. The matrix of experiments generated comprised 50 reactions for each compound.

The results were analyzed using the Design-Expert® software. The optimum conditions for producing each transglycosylation product were determined by quadratic equations and can be seen in the supplementary material. Both models predicted the glucoside production as a function of the concentrations of hydroxytyrosol and vanillyl alcohol, the concentration of cellobiose, the enzyme units, the temperature and the time of reaction. The analysis of variance test performed by the software validated the experimental data. The Design Expert software allows selecting "maximum production yield" or "maximum conversion rate" as the parameters for optimization of the experimental conditions. As a general rule, the highest productions were obtained when donor and acceptor are added at their maximum concentrations, as it has been reported previously for other β-glucosidases (Bohlin *et al.*, 2013). However, maximum conversion rates relative to initial acceptor concentrations were only reached when the donor was added at the maximum concentration tested and the acceptor was in low amounts.

The conditions for maximum production generated by the software required the same reaction mixtures for each compound: 4 U of β-glucosidase activity of BGL-2, 350 mM of cellobiose as donor, 195 mM of hydroxytyrosol or vanillyl alcohol as acceptors, 50 mM acetate pH 4 and 0.1% BSA. They were incubated at 50° C for 5 h. However, for maximum relative conversion of the acceptors, a concentration of hydroxytyrosol and vanillyl alcohol of 32 mM was used, maintaining the above settings for the remaining variables. We obtained a "maximum production yield" value of 2.55 g/L (8 mM) for hydroxytyrosol glucoside, and 3.8 g/L (12 mM) for vanillyl glucoside, and "maximum conversion rate" of 13.8% and 19% for hydroxytyrosol and the vanillyl alcohol

glucosides, respectively. These glucosides were purified by HPLC, as described in materials and methods prior to their characterization by NMR.

**Solubility of new glycosides and characterization by NMR**

The aqueous solubility of each glycoside at room temperature was compared with that of the respective aglycon. Solubility detected for hydroxytyrosol was of 88.84 mg/mL, and the glycoside had its solubility increased up to 254.75 mg/mL. On the other hand, vanillyl alcohol showed a solubility of 16.12 mg/mL, which was improved by transglycosylation to a value of 165.32 mg/mL. This data confirmed the improvement of the solubility of the novel molecules, making them more bio-available. A similar effect were detected recently in a work that developed the α-glycosylation of pterostilbene (González-Alfonso *et al*., 2018).

$^{1}H$ and $^{13}C$-NMR experiments were carried out with the purified hydroxytyrosol and vanillyl glucosides produced in the reactions to assign their regiochemistry. The HMBC spectrum (figure 3) showed a correlation between the anomeric carbon (C1'') and the carbon C2' of the hydroxytyrosyl residue (see table 4 with the 1H and 13C NMR chemical shifts).

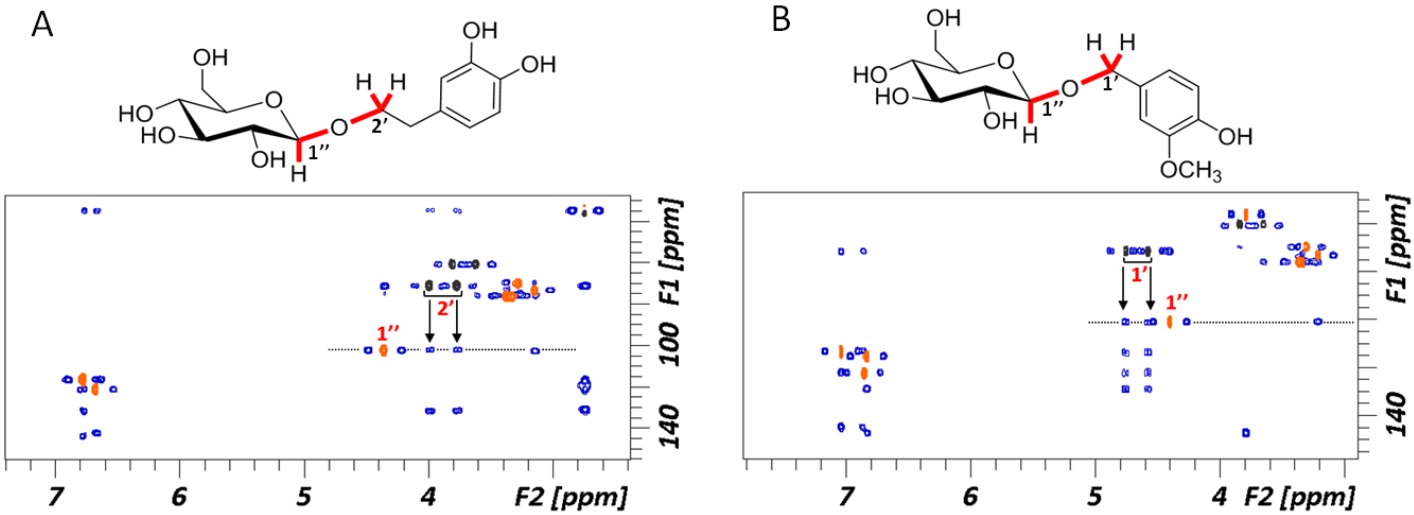



**Figure 3.** Superimposition of 1H-13C HSQC (orange/black) and HMBC (blue). Arrows represent the key cross peaks correlations corresponding to the connectivity between the anomeric position of glucose and the hydroxytyrosyl (A) and vanillyl (B) residues. Atoms are numbered in agreement with Table 4.

**Table 4.** Chemical shifts for hydroxytyrosol and vanillyl alcohol glucosides.

| | Hydroxytyrosol glucoside | | | Vanillyl glucoside | |
|---|---|---|---|---|---|
| | $^{1}$H (ppm) | $^{13}$C (ppm) | | $^{1}$H (ppm) | $^{13}$C (ppm) |
| **1'' Glu** | 4.36 | 102.20 | **1'' Glu** | 4.39 | 100.94 |
| **2'' Glu** | 3.15 | 73.25 | **2'' Glu** | 3.20 | 73.02 |
| **3'' Glu** | 3.35 | 76.38 | **3'' Glu** | 3.35 | 75.94 |
| **4'' Glu** | 3.27 | 69.65 | **4'' Glu** | 3.30 | 69.61 |
| **5'' Glu** | 3.35 | 76.38 | **5'' Glu** | 3.35 | 75.94 |
| **6'' Glu** | 3.62<br>3.81 | 60.60 | **6'' Glu** | 3.65<br>3.82 | 60.31 |
| **1 HT** | --- | 128. 97 | **1 Van** | --- | 128. 97 |
| **2 HT** | 6.77 | 113.29 | **2 Van** | 7.03 | 113.29 |
| **3 HT** | --- | 147.46 | **3 Van** | --- | 147.46 |
| **4 HT** | --- | 144.86 | **4 Van** | --- | 144.86 |
| **5 HT** | 6.77 | 115.37 | **5 Van** | 6.83 | 115.37 |
| **6 HT** | 6.67 | 121.36 | **6 Van** | 6.85 | 121.36 |
| **1' HT** | 2.74 | 55.98 | **$OCH_3$ Van** | 3.78 | 55.98 |
| **2' HT** | 3.77<br>3.99 | 71.39 | **1' Van** | 4.56<br>4.76 | 71.39 |

Moreover, the anomeric proton presents a coupling constant value of 8 Hz, pointing out the formation of the derivative through a β-linkage in accordance with the retaining β-glucosidase activity of the enzyme. Therefore, we can assure that the anomeric position is attached to the hydroxytyrosyl residue through the aliphatic chain. The same procedure was used to characterize the vanillyl glucoside, confirming that the anomeric position is connected to the vanillyl residue through the aliphatic chain (figure 3). This preference for the aliphatic chain versus the phenolic OH was also reported in a β-xylosidase of the same fungus, for the synthesis of hydroxytyrosol xyloside (Nieto-Domínguez *et al*., 2016).

The fact that BGL-2 incorporates the glucosyl residue preferentially at the aliphatic hydroxyl could represent an advantage for the biological activity of these novel glucosides. The phenolic hydroxyl groups are the responsible for the antioxidative and free radical-scavenging activity of phenolic compounds, whose scavenging mechanism consists in donating the hydrogen atom of the phenolic -OH to free radicals, thus blocking their propagation during the oxidation process. The presence of a second and a third -OH in the phenolic ring could increase the antioxidant potential (Torres de Pinedo *et al*., 2007). In this

sense, it has been reported that the alkyl chain of caffeic acid and its derivatives, which are antioxidants with similar structure to the phenols evaluated in the current work, could have a role in stabilizing the radical formed during oxidation. However, its exact contribution remains uncertain (Silva *et al*., 2000). Hence, the effect of glycosylation on bioactivity of the phenolic compounds should be further evaluated.

It is necessary to remark that both compounds can be found in some natural sources, as the vanillin glucoside in *Dendrotrophe frutescens* (Hou *et al*., 2013) or the hydroxytyrosol glucoside in *Prunus grayana* (Shimomura *et al*., 1986). But the low yields obtained (0.032% and 0.067% of initial biomass according to both articles respectively), and the complex methodology, using a considerable amounts of organic solvents, that is required for obtaining both glycosides directly from plant extracts, makes transglycosylation an easier and efficient alternative in comparison.

**Antitumor potential of hydroxytyrosol and vanillin glucosides**

It has been previously reported that hydroxytyrosol and vanillin and their derivatives can have an impact on different hallmarks of cancer (Casaburi *et al*., 2012; Kapoor, 2013).

In this work, we analyzed if their glycosylation affects their antitumor activity, using the human breast cancer cell line MCF-7 and the human non-tumoral mammary epithelial cell line MCF-10A. We should remark that the action of vanillyl glucoside was compared with that of vanillin (instead of vanillyl alcohol) for its superior biological effects.

Results from crystal violet assays revealed that both, hydroxytyrosol and its glucoside remarkably reduced the viability of MCF-7 cells in a concentration-dependent manner (figure 4 (2) B). The concentrations inducing 50 % decrease in cell viability ($IC_{50}$) were very similar in both cases (figure 4 (2) B). This reduction was observed as soon as 24 h post compound addition, and it was accompanied by significant morphological changes in the cells (i.e. cell shrinkage, partial detachment and formation of apoptotic bodies) (figure 4 (1) A), suggesting apoptotic cell death. To investigate the safety profile of hydroxytyrosol glucoside and to compare it with the original phenol, we analyzed their effect on the viability of the non-transformed MCF10-A cell line. These cells were less sensitive to both, hydroxytyrosol and its glucoside, than their tumoral counterparts at every tested concentration (figure 4 (2) C).

The impact of this glucoside on the viability of non-tumoral cells was negligible, and therefore lower than that observed with the original compound (figure 4 (2) C). Overall, these results suggest that the glucoside has antitumor activity, with similar efficacy than free hydroxytyrosol but with a safer profile (i.e. less toxic on non-transformed cells).

It has been previously reported that hydroxytyrosol and hydroxytyrosol-rich olive leaves extracts decrease the viability of MCF-7 cells (Goulas *et al*., 2009; Han *et al*., 2009; Bouallagui *et al*., 2010), as well as of other breast and

non-breast cancer cell lines (Casaburi *et al*., 2012). In most of these studies, this effect was attributed to the phenol's capacity to inhibit cell proliferation and promote their apoptosis. Although some molecular mechanisms to explain this phenomenon have been proposed, the initial stress signals remain unknown. A number of recent reports concur that, although hydroxytyrosol is a compound well known for its antioxidant properties, under certain conditions, it can promote pro-oxidant effects and induce anti-proliferative and pro-apoptotic reactions in cancer cell lines through $H_2O_2$ generation (Fabiani *et al*., 2009; Fabiani *et al*., 2012). In addition, the same studies suggest that the *orto*-dihydroxy phenolic group present in this molecule, which is the main structural feature responsible for its free radical-scavenging activity, was also fundamental for the reported pro-oxidant effect.

**A**

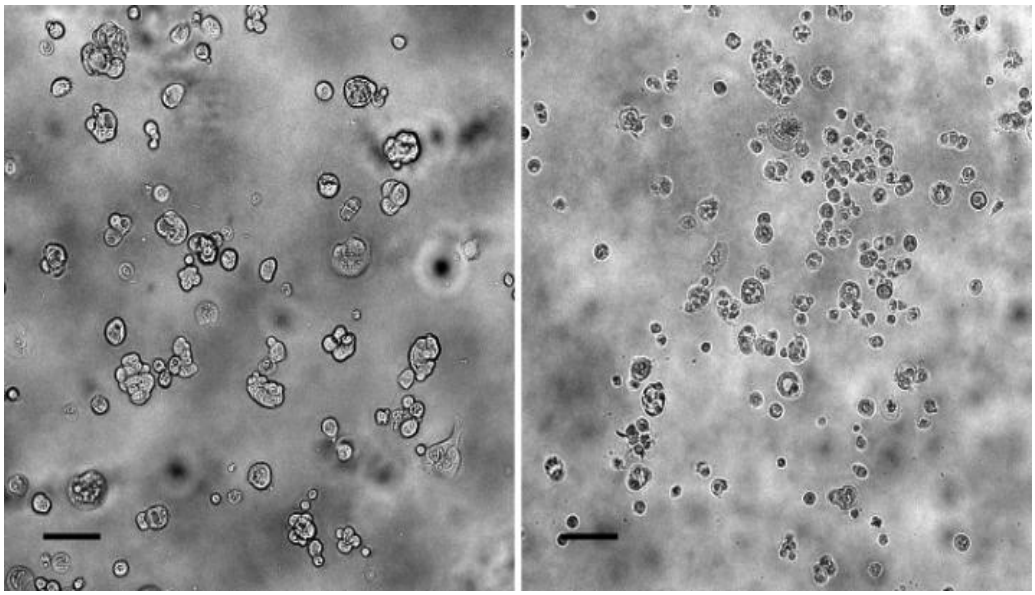


**Figure 4 (1).** Effect of hydroxytyrosol and its glucoside on the viability of (B) MCF-7 and (C) MCF-10A cells after 24 h of compound addition. Cell viability was determined by crystal violet assay. Data represent mean ± SD of three independent experiments. * $p < 0.05$ vs. vehicle-treated cells. (C) Representative image of MCF-7 cell morphology after 24 h treatment with 160 µM hydroxytyrosol glucoside (right panel) or the corresponding vehicle (left panel). Scale bar, 100 µm.

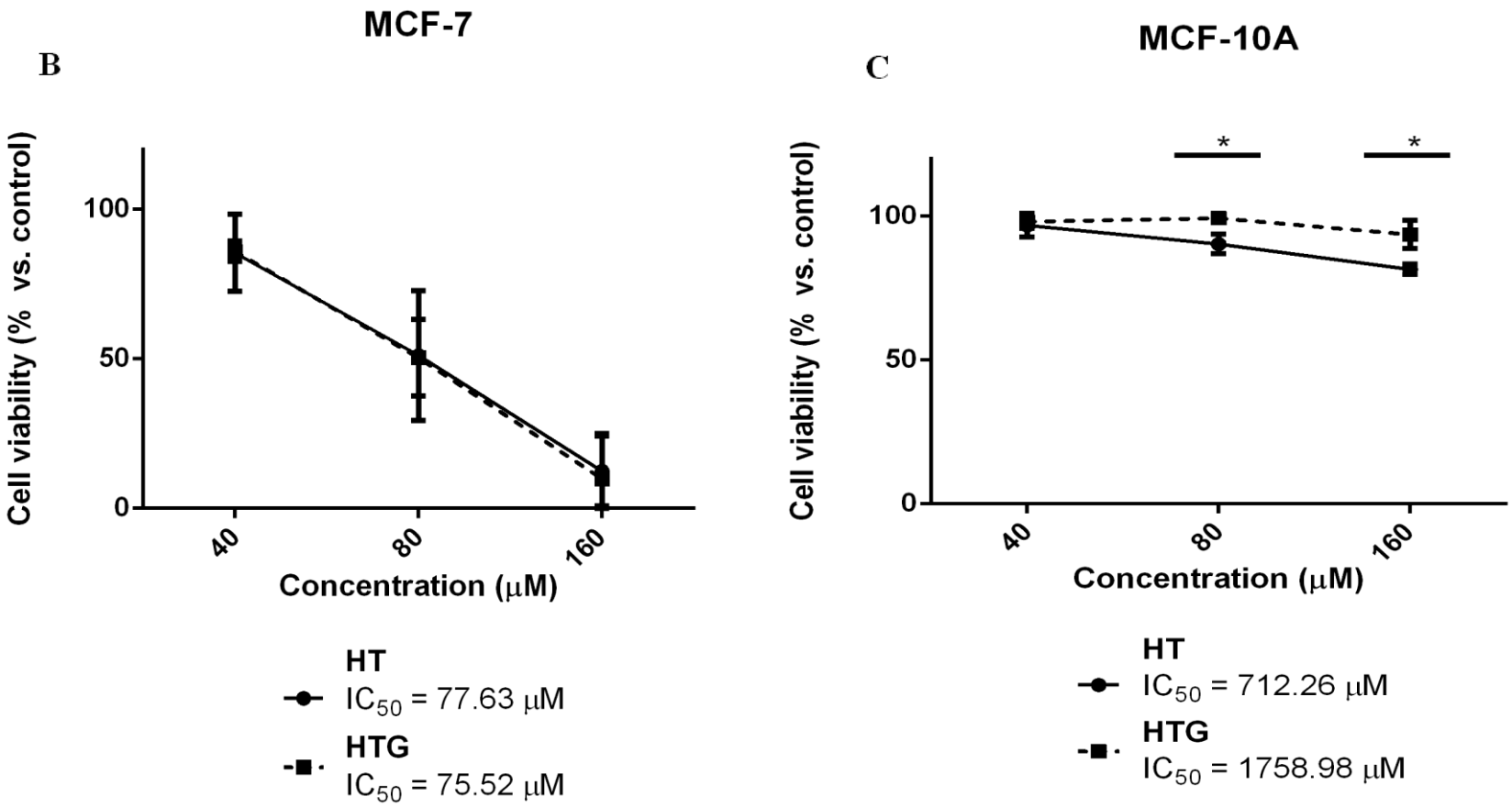



**Figure 4 (2).** Effect of hydroxytyrosol and its glucoside on the viability of (B) MCF-7 and (C) MCF-10A cells after 24 h of compound addition. Cell viability was determined by crystal violet assay. Data represent mean ± SD of three independent experiments. * $p < 0.05$ vs. vehicle-treated cells. (A) Representative image of MCF-7 cell morphology after 24 h treatment with 160 μM hydroxytyrosol glucoside (right panel) or the corresponding vehicle (left panel). Scale bar, 100 μm.

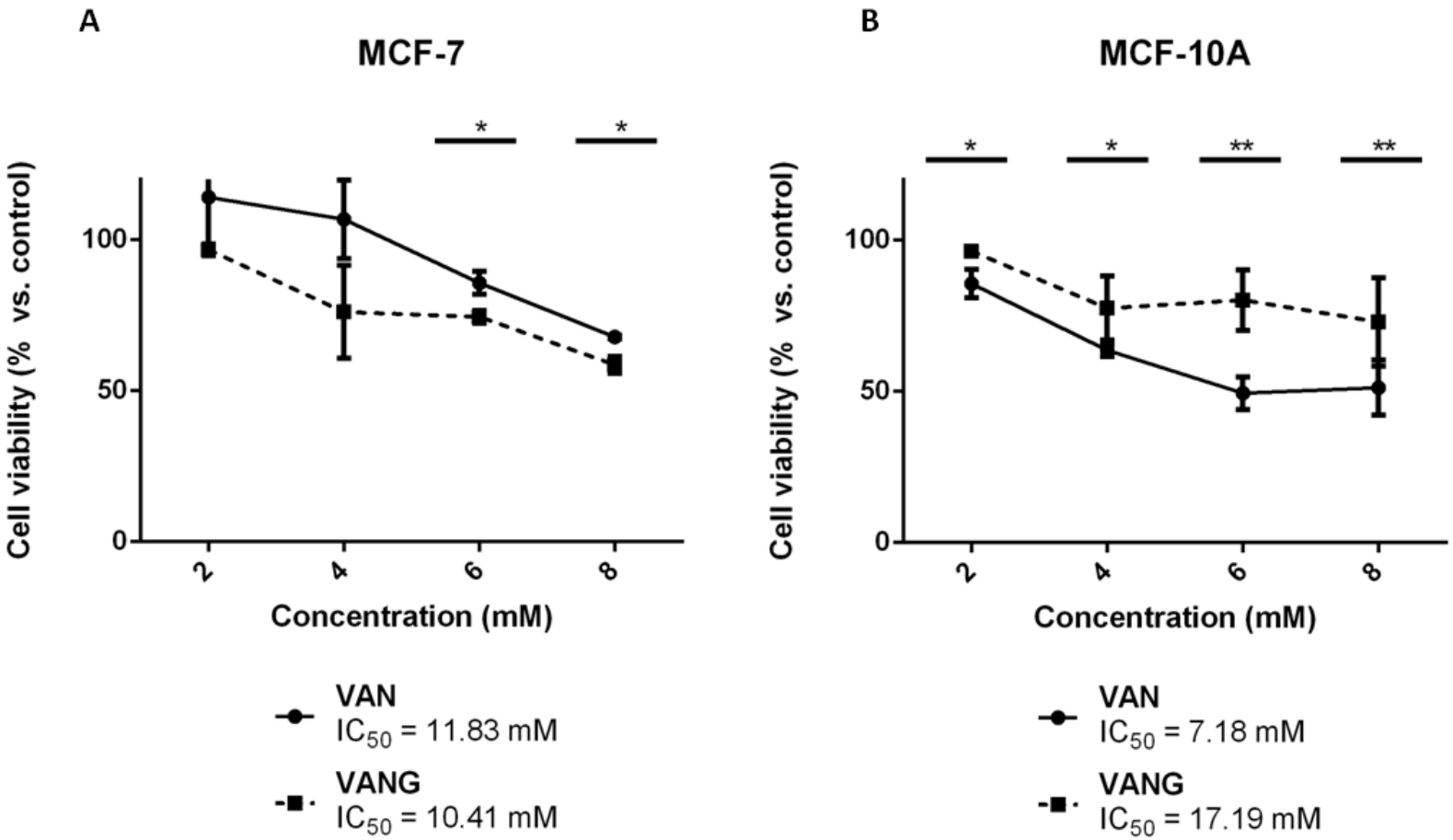



**Figure 5.** Effect of vanillin and vanillin glucoside on the viability of (A) MCF-7 and (B) MCF-10A cells after 72 hours of compound addition. Cell viability was determined by crystal violet assay. Data represent mean ± SD of two (A) and three (B) independent experiments. * $p < 0.05$; ** $p < 0.01$ vs. vehicle-treated cells.

As in the hydroxytyrosol glucoside, the phenolic OH is not involved in the transglycosylation reactions, there is no reason to think that the pro-oxidant properties may significantly differ from those of hydroxytyrosol. This would explain the almost identical decrease in the viability of MCF-7 cells observed upon addition of both compounds. Yet, further studies are needed to deepen our understanding on the molecular events underlying hydroxytyrosol or hydroxytyrosol glucoside-triggered cytotoxicity in the MCF-7 cancer cell line.

MCF-10A non-tumor breast epithelial cells, on the other hand, showed to be considerably more resistant to the deleterious effect of hydroxytyrosol and hydroxytyrosol glucoside than MCF-7 cells. It is interesting to point out that many human cancer cells present a highly oxidative state due to decreased antioxidant protective enzyme levels compared to their normal tissue counterparts. Therefore, cancer cells may be more sensitive to any generated reactive oxygen species (ROS) within the cells (Han *et al*., 2009; Manda *et al*., 2015). The ability to trigger cell death specifically in cancer cells while not affecting non-cancerous cells is the basis for any potential antitumor compound and thus, hydroxytyrosol and its glucoside seems to be exceptional candidates to be so. In line with this notion, it is worth highlighting that while MCF10-A cells treated with hydroxytyrosol start displaying apoptotic features when challenged with concentrations up to 160 µM, hydroxytyrosol glucoside-treated ones barely show signs of cytotoxicity at all.

With respect to the potential antitumor action of vanillin and its glucoside, we monitored the effects of compound addition in MCF-7 and MCF-10A cells for several days. Figure 5 (A) shows a concentration-dependent reduction in the viability of MCF-7 cancer cells in response to both compounds. The glycosylated derivative showed a slightly (but statistically significant) higher efficacy than the parental compound (figure 5 A). It should be noted that these differences in cell viability between treated and untreated cells became visible only 72 h after compound exposure. In addition, no morphological differences could be detected by light microscopy as a consequence of compound treatment throughout the experiment (data not shown), suggesting that these compounds may be acting in MCF-7 cells as inhibitors of cell proliferation rather than as apoptotic stimuli.

In this case, and unlike the hydroxytyrosol pair, both vanillin and vanillyl glucoside significantly decreased the viability of non-tumoral MCF-10A cells (figure 5 B). However, the glycosylated derivative showed a less toxic profile than vanillin. Many lines of evidence support the potential beneficial effect of vanillin against human cancers (Kapoor, 2013). Some authors have attributed its anti-invasive, anti-metastatic, anti-angiogenic and selectively cytotoxic actions to its ability to behave as a pro-oxidant instead of as an antioxidant (Bezerra *et al*., 2016), as we previously mentioned in the case of hydroxytyrosol. However, the mechanism through which vanillin modulates their selective anticancer effects has not been clearly delineated.

Our results showed no preferential cytotoxicity towards cancer cells after vanillin exposure. We hypothesize that the high concentrations used in this study (in the mM range) exerted toxic effects in both cell lines. As a matter of fact, aldehydes rarely accumulate in high concentrations in biological systems because of their high chemical reactivity. The natural vanillin biosynthesis pathway in the vanilla orchid, *Vanilla planifolia*, has an elegant solution to cope with the toxicity issue by glycosylation of vanillin to vanillin-β-D-glucoside (Gallage and Møller, 2015). Therefore, it is not surprising that vanillin glucoside displayed a weaker cytotoxic response on non-tumoral MCF-10A cells than vanillin.

In conclusion, glycosylation of vanillin generates a more effective anti-tumor compound that is less toxic to healthy cells than its non-glucosylated counterpart.

## CONCLUSIONS

In this work, the transglycosylation profile of three GH3 β-glucosidases (BGL-2, BGL-2T and BGL-3) was studied. The analysis showed that these enzymes may display transglucosylation in a variety of acceptors. BGL-2 was selected as model enzyme to corroborate the biotechnological potential of these enzymes because it displayed the highest transglucosylation yield. The results have shown that the glucosylated derivatives of hydroxytyrosol and vanillyl alcohol have more effective and/or safer profiles than hydroxytyrosol and vanillin when added to breast cancer cell cultures. The potential of this enzyme may lay the foundations for the design of new therapeutic tools for the management of cancer.

## METHODS

### β-glucosidase production and purification

BGL-2, BGL-2T, and BGL-3, from *T. amestolkiae* were heterologously expressed in *Pichia pastoris* and purified as previously reported (Méndez-Líter *et al*., 2017; Méndez-Líter *et al*., 2018). Enzymatic activity was determined spectrophotometrically at 410 nm by the release of *p*-nitrophenol (*p*NP) using *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG) as substrate. One unit of β-glucosidase activity was defined as the release of 1 μmol of *p*NP per minute. Bovine serum albumin (BSA) was always added to every reaction (hydrolysis or transglycosylation), at final concentration of 0.1%, in order to get reproducible results and prevent loss of enzyme activity.

### Screening for potential transglycosylation acceptors

A library of 70 compounds, which can be seen in Table 1, was tested in a preliminary screening as potential transglycosylation acceptors by *T.*

*amestolkiae* β-glucosidases. Stock solutions of these compounds were prepared in distilled water at 100 mM.

The assay was performed as described by Blanchard and Withers (Blanchard and Withers, 2001). The first step consists on inhibiting the enzymes using 5 mM 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-D-glucopyranoside, dissolved in 50 mM sodium phosphate (pH 6), at room temperature for 2 h. After inhibition, the sample was dialyzed against sodium phosphate pH 6, and an aliquot was added to each well of a 96-well plate. The final reaction mix was composed of each inactivated enzyme (in a dosage that would correspond to 2 U of β-glucosidase), 50 mM sodium phosphate buffer (pH 6), 0.1 % BSA and 10 mM of each tested compound. Controls of non-inhibited enzyme and inhibited enzyme without acceptor were included in triplicate.

The plate was incubated for 1 h to allow the potential transglycosylation acceptors to reverse the enzymatic inhibition. After this time, *p*NPG was added to each well, and the change in absorbance at 410 nm was monitored for 30 min at 40 °C. The compounds that showed higher absorbance than the controls without acceptor were considered potential hits of transglycosylation.

**Analysis of the transglycosylation products by TLC and HPLC**

After the aforementioned screening, the most interesting acceptors were assayed at higher scale. Transglycosylation standard reactions were carried out using 2.5 U/mL of each BGL in acetate buffer (pH 4.0), which is the optimum for BGL-2 variants and BGL-3 (Méndez-Líter *et al*., 2017; Méndez-Líter *et al*., 2018).

100 mM cellobiose was used as donor and the different acceptors were tested in a concentration of 20 mM. The reactions were incubated for 1 h at 40 °C at 1200 rpm. Controls for each acceptor without enzyme were also carried out. The synthesis of glucosides was followed by thin layer chromatography (TLC) in order to confirm the positive hits. TLCs were carried out by using silica gel G/UV254 polyester sheets, (0.2 mm thickness and 40 × 80 mm plate size) provided by Macherey–Nagel, using ethyl acetate/methanol/water 10:2:1 (v/v) as running solution. Detection of compounds and glucosides was performed under 254 nm UV light.

After TLC screening, the transglycosylation products were also analyzed by High Performance Liquid chromatography (HPLC) on an Agilent 1200 series LC instrument equipped with a ZORBAX Eclipse XDB-C18 column (Agilent). The system was equilibrated in acetonitrile/$H_2O$ (9:91 v/v), both containing 0.1 % acetic acid, with a flow of 2 mL/min, and the reaction products were separated isocratically for 8 min in the same buffer. Then, the mobile phase was changed to 95:5 acetonitrile/$H_2O$, for washing the column for 3 min and the system was finally re-equilibrated to initial conditions for 4 min. The products' peaks were detected by monitoring the absorbance at 270 nm. The peaks were quantified by referencing to a calibration curve of their phenolic precursors.

The purification of the selected glucosides was performed in the same HPLC equipped with a semi-preparative column (Mediterranea Sea 18 TR-010006, Teknokroma). The purification conditions were the same reported for HPLC analysis. After collection, the products were lyophilized and stored at -20° C.

**Analysis of the reaction products by mass spectrometry**

The reaction mixtures were analyzed by conventional mass spectrometry, performed on a HCT Ultra ion trap. The samples were analyzed by ESI-MS with methanol as ionizing phase in the positive reflector mode, and data were processed with the Masshunter Data Acquisition B.05.01 and Masshunter Qualitative Analysis B.07.00 software (Agilent Technologies).

**Optimization of transglycosylation catalyzed by BGL-2 by response surface methodology**

The reaction conditions for the production of hydroxytyrosol and vanillin glucosides were optimized by a response surface methodology approach. Design-Expert® software version 10.0.1.0 (Stat-Ease Inc. MN, USA) was used for generating a Box-Behnken design matrix and for data analysis. The parameters selected for building the model for glucoside production were the concentrations of the donor (cellobiose) and the selected acceptors (hydroxytyrosol and vanillyl alcohol), enzyme dosage, reaction time and temperature. The reactions were carried out at pH 4, the optimum for the enzyme (Méndez-Líter *et al*., 2018). After performing the required reactions, the software generates a polynomial quadratic equation from the obtained data, which reflects the effect of the independent variables on the response, and the highest production and yield expected.

**Solubility of novel glycosides**

A saturated solution of hydroxytyrosol, vanillyl alcohol, and its glucosylated counterparts, were prepared in water and incubated at room temperature, at 500 rpm, during 2 h. Then, the solution was centrifuged, filtered, and analyzed with the HPLC, using the same protocol as previously described in the paragraph “Analysis of the transglycosylation products by TLC and HPLC”.

**Nuclear Magnetic Resonance (NMR)**

NMR was used to confirm the structure and regiochemistry of the glucosides of hydroxytyrosol and vanillyl alcohol synthesized by transglycosylation. Samples were prepared in 500 μL of deuterated water ($D_2O$) for analysis. NMR spectra were acquired at 298 K, using a Bruker AVANCE 600 MHz spectrometer equipped with a cryogenic probe. For spectral assignment one dimensional 1D $^{1}$H-NMR spectra, $^{1}$H-$^{13}$C HSQC and HMBC experiments were acquired using corresponding pulse sequences included in TOPSIN acquisition software (Bruker).

**Cell lines and cultures**

The human breast adenocarcinoma cell line MCF-7 was obtained from the American Type Culture Collection (ATCC) and maintained in Eagle's Minimum Essential Medium (MEM) (Gibco by Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) (Gibco), 10 µg/mL insulin (SAFC Biosciences by Sigma-Aldrich), 2 mM L-glutamin and 50 U/mL of a penicillin-streptomycin solution (Lonza). The human non-cancerous mammary epithelial cell line MCF-10A was obtained from the ATCC and cultured in Dulbecco's Modified Eagle's Medium (DMEM)/Ham's F-12 (1 : 1) (Gibco) supplemented with 5% horse serum (HS) (Gibco), 10 µg/mL insulin (SAFC Biosciences), 0.5 µg/mL hydrocortisone (Sigma-Aldrich), 20 ng/mL epidermal growth factor (EGF) (Gibco), 10 µg/mL cholera toxin (Sigma-Aldrich) and 50 U/mL of a penicillin-streptomycin solution (Lonza). All cell lines were validated in the Genomics Core Facility at Alberto Sols Biomedical Research Institute (Madrid, Spain).

**Evaluation of cell viability by crystal violet staining.**

MCF-7 and MCF-10A cells were seeded in 96-well plates at a density of $4\times10^3$ cells/well and allowed to attach to the plastic surface for 24 h. The medium was then replaced with fresh medium supplemented with hydroxytyrosol or hydroxytyrosol glucoside (40, 80 and 160 µM) and incubated for 24 h, or with 2, 4, 6 and 8 mM vanillin or vanillin glucoside for 72 h. Distilled water was used in both cases as vehicle. Following treatment, cells were incubated with 0.1 % crystal violet (Panreac, Barcelona, Spain) for 20 min in agitation. The plate was then gently washed with tap water and the crystals were resuspended in methanol. Cell viability was determined by reading absorbance at 570 nm with a microtiter plate reader (Rayto Life and Analytical Sciences Co., Ltd., Shenzhen, China) and expressed as percentage versus vehicle-treated cells, set at 100 %.

**Declarations**

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

All authors read and approved the final manuscript.

**Availability of data and material**

*T. amestolkiae* whole genome shotgun project is deposited at DDBJ/ENA/GenBank under the accession number MIKG00000000. BGL-2 sequence is deposited in GenBank under the accession number KM393203. BGL-3 sequence is deposited in GenBank under the accession number KM393202.1.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**
This work has been funded by Projects RTC-2014-1777-3 and BIO2015-68387-R from MINECO and S2013/MAE-2907 from Comunidad de Madrid. J.A. Méndez-Líter thanks its fellowship to Tatiana Pérez de Guzmán el Bueno Foundation.

**Author contributions**
JAML drafted the manuscript and performed the experiments excluding NMR and mass anticancer biological assays. MND helped with the design of response surface methodology and helped analyzing the data. IT and CS designed and developed the antiproliferative assays and analyzed the data. BFT and FJC designed and performed the NMR assays and analyzed the data. AGS, and JLA, synthesized the 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-D-glucopyranoside used for the screening of potential transglycosylation acceptors. AP, LIE and MJM coordinated the study, helped to draft, and critically review the manuscript.

**Additional information**
Supplementary information accompanies

**Competing financial interests**
The authors declare no competing financial interests.

## REFERENCES

Bezerra, D.P., Soares, A.K.N., de Sousa, D.o.P., 2016. Overview of the role of vanillin on redox status and cancer development. Oxid. Med. Cell. Longev. 2016, 9734816.

Blanchard, E., Withers, S., 2001. Rapid screening of the aglycone specificity of glycosidases: Applications to enzymatic synthesis of oligosaccharides. Chem. Biol. 8, 627-633.

Bohlin, C., Praestgaard, E., Baumann, M.J., Borch, K., Praestgaard, J., Monrad, R.N. *et al*., 2013. A comparative study of hydrolysis and transglycosylation activities of fungal β–glucosidases. Appl. Microbiol. and Biotechnol. 97, 159-169.

Bouallagui, Z., Han, J., Isoda, H., Sayadi, S., 2010. Hydroxytyrosol rich extract from olive leaves modulates cell cycle progression in MCF-7 human breast cancer cells. Food Chem. Toxicol. 49, 179-184.

Bravo, L., 2009. Polyphenols: Chemistry, Dietary Sources, Metabolism, and Nutritional Significance. Nutr. Rev. 56, 317-333.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2009. The carbohydrate active enZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 37, D233-D238.

Casaburi, I., Puoci, F., Chimento, A., Sirianni, R., Ruggiero, C., Avena, P. *et al*., 2012. Potential of olive oil phenols as chemopreventive and therapeutic agents against cancer: A review of in vitro studies. Mol. Nutr. Food Res. 57, 71-83.

Dai, J., Mumper, J., 2010. Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. Molecules. 15, 7313-7352.

Danby, P.M., Withers, S.G., 2016. Advances in enzymatic glycoside synthesis. ACS Chem. Biol. 11, 1784-1794.

de Eugenio, L.I., Méndez-Líter, J.A., Nieto-Domínguez, M., Alonso, L., Gil-Muñoz, J., Barriuso, J. *et al*., 2017. Differential β−glucosidase expression as a function of carbon source availability in *Talaromyces amestolkiae*: a genomic and proteomic approach. Biotechnol. Biofuels. 10, 161.

Diana Draelos, Z., 2007. Skin lightening preparations and the hydroquinone controversy. Dermatol. Ther. 20, 308-313.

Dinesh, K., V, Priya Ranjan Prasad Verma and Sandeep Kumar Singh, 2016. Emerging technologies for improving bioavailability of polyphenols. Curr. Nutr. Food Sci. 12, 12-22.

Edelman, J., 1956. The Formation of Oligosaccharides by Enzymic Transglycosylation 189-232. Adv. Enzymol. Relat. Subj. Biochem. 17:189-232.

Fabiani, R., Fuccelli, R., Pieravanti, F., De Bartolomeo, A., Morozzi, G., 2009. Production of hydrogen peroxide is responsible for the induction of apoptosis by hydroxytyrosol on HL60 cells. Mol. Nutr. Food Res. 53, 887-896.

Fabiani, R., Sepporta, M.V., Rosignoli, P., De Bartolomeo, A., Crescimanno, M., Morozzi, G., 2012. Anti-proliferative and pro-apoptotic activities of hydroxytyrosol on different tumour cells: the role of extracellular production of hydrogen peroxide. Eur. J. Nutr. 51, 455-464.

Gallage, J., Møller, B., 2015. Vanillin-bioconversion and bioengineering of the most popular plant flavor and its de novo biosynthesis in the *Vanilla orchid*. Mol. Plant 8, 40-57.

González-Alfonso, L.J., Rodrigo-Frutos, D., Belmonte-Reche, E., peñalver, P., Poveda, A., Jiménez-Barbero, J. *et al*., 2018. Enzymatic synthesis of a novel pterostilbene α–glucoside by the combination of cyclodextrin glucanotransferase and amyloglucosidase. Molecules. 23(6): 1271.

Goulas, V., Exarchou, V., Troganis, A.N., Psomiadou, E., Fotsis, T., Briasoulis, E. *et al*., 2009. Phytochemicals in olive-leaf extracts and their antiproliferative activity against cancer and endothelial cells. Mol. Nutr. Food Res. 53, 600-608.

Han, J., Talorete, T.P.N., Yamada, P., Isoda, H., 2009. Anti-proliferative and apoptotic effects of oleuropein and hydroxytyrosol on human breast cancer MCF-7 cells. Cytotechnol. 59, 45-53.

Hou, X., Sun, M., Gao, H., Cui, Y., Xiao, K., 2013. Chemical constituents from the stems of Dendrotrophe frutescens. Biochem. Sys. Ecol. 51, 156-159.

Hu, T., He, X.W., Jiang, J.G., Xu, X.L., 2014. Hydroxytyrosol and its potential therapeutic effects. J. Agric. Food Chem. 62, 1449-1455.

Kapoor, S., 2013. Multiorgan anticarcinogenic effects of vanillin. J. Med. Food. 16.

Kim, I.S., Choi, D.K., Jung, H.J., 2011. Neuroprotective effects of vanillyl alcohol in gastrodia elata blume through suppression of oxidative stress and anti-apoptotic activity in toxin-induced dopaminergic MN9D cells. Molecules. 16 (7): 5349-5361.

Kometani, T., Terada, Y., Nishimura, T., Takii, H., Okada, S., 1994. Transglycosylation to hesperidin by cyclodextrin glucanotransferase from an alkalophilic *bacillus* species in alkaline pH and properties of hesperidin glycosides. Biosci. Biotechnol. Biochem. 58, 1990-1994.

Laschke, W., van Oijen, E.V., Scheuer, C., Menger, M., 2011. In vitro and in vivo evaluation of the anti-angiogenic actions of 4-hydroxybenzyl alcohol. Br. J. Pharmacol. 163, 835-844.

Lim, E., Jung, H., Park, E., Kang, H., 2010. Anti-angiogenic, anti-inflammatory and anti-nociceptive activity of 4-hydroxybenzyl alcohol. J. Pharm. Pharmacol. 59, 1235-1240.

Linares-Pasten, J.A., Andersson, M., Karlsson, E.N., 2014. Thermostable Glycoside Hydrolases in Biorefinery Technologies. Curr. Biotechnol. 3, 26-44.

Lirdprapamongkol, K., Sakurai, H., Kawasaki, N., Choo, M.K., Saitoh, Y., Aozuka, Y. *et al*., 2005. Vanillin suppresses in vitro invasion and in vivo metastasis of mouse breast cancer cells. Eur. J. Pharmacol. Sci. 25, 57-65.

Macfarlane, S., Macfarlane, G.T., Cummings, J.H., 2006. Review article: prebiotics in the gastrointestinal tract. Alim. Pharmacol. Ther. 24, 701-714.

Manda, G., Isvoranu, G., Comanescu, M.V., Manea, A., Debelec Butuner, B., Korkmaz, K.S., 2015. The redox biology network in cancer pathophysiology and therapeutics. Redox. Biol. 5, 347-357.

Méndez-Líter, J.A., Gil-Muñoz, J., Nieto-Domínguez, M., Barriuso, J., de Eugenio, L.I., Martínez, M.J., 2017. A novel, highly efficient β−glucosidase with a cellulose-binding domain: characterization and properties of native and recombinant proteins. Biotechnol. Biofuels. 10, 256.

Méndez-Líter, J.A., de Eugenio, L.I., Prieto, A., Martínez, M.J., 2018. The β−glucosidase secreted by *Talaromyces amestolkiae* under carbon starvation: a versatile catalyst for biofuel production from plant and algal biomass. Biotechnol. Biofuels. 11, 123.

Míguez, N., Ramírez-Escudero, M., Gimeno-Pérez, M., Poveda, A., Jiménez-Barbero, J., Ballesteros, A.O. *et al*., 2018. Fructosylation of hydroxytyrosol by the α−fructofuranosidase from *Xanthophyllomyces dendrorhous*: Insights into the molecular basis of the enzyme specificity. Chem. Cat. Chem. 10, 4878-4887.

Nieto-Dominguez, M., Eugenio, I.d., peñalver, P., Belmonte-Reche, E., Morales, J., Poveda, A. *et al*., 2017. Enzymatic synthesis of a novel neuroprotective hydroxytyrosyl glycoside. J. Agric. Food Chem. 65, 48, 10526-10533.

Nieto-Domínguez, M., Eugenio, I.d., Barriuso, J., Prieto, A., Fernández de Toro, B., Canales-Mayordomo, Á. *et al*., 2015. Novel pH-Stable Glycoside Hydrolase Family 3 β−Xylosidase from *Talaromyces amestolkiae*: an Enzyme Displaying Regioselective Transxylosylation. Appl. Environ. Microb. 81(18): 6380–6392.

Nieto-Domínguez, M., Prieto, A., Fernández de Toro, B., Cañada, F.J., Barriuso, J., Armstrong, Z. *et al*., 2016. Enzymatic fine-tuning for 2-(6-hydroxynaphthyl) β−d-xylopyranoside synthesis catalyzed by the recombinant β−xylosidase BxTW1 from *Talaromyces amestolkiae*. Microb. Cell Fact. 15, 171.

Piedrabuena, D., Míguez, N., Poveda, A., Plou, F.J., Fernández-Lobato, M., 2016. Exploring the transferase activity of Ffase from *Schwanniomyces occidentalis*, a α−fructofuranosidase showing high fructosyl-acceptor promiscuity. Appl. Microbiol.Biotechnol. 100, 8769-8778.

Prodanovic, R., Milosavic, N., Sladic, D., Zlatovic, M., Božic, B., Cirkovic Velickovic, T. *et al*., 2005. Transglucosylation of hydroquinone catalysed by α−glucosidase from baker's yeast. J. Mol. Catal. B Enz. 35, 142-146.

Rastall, A., Gibson, R., 2015. Recent developments in prebiotics to selectively impact beneficial microbes and promote intestinal health. Curr. Opin. Biotechnol. 32, 42-46.

Seitz, O., 2000. Glycopeptide synthesis and the effects of glycosylation on protein structure and activity. Chem. Bio. Chem. 1, 214-246.

Shimomura, H., Sashida, Y., Adachi, T., 1986. Phenolic glucosides from *Prunus grayana*. Phytochemistry. 26, 249-251.

Silva, F., Borges, F., Guimarães, C., Lima, F.C., Matos, C., Reis, S., 2000. Phenolic acids and derivatives: studies on the relationship among structure, radical scavenging activity, and physicochemical parameters. J. Agric. Food Chem. 48, 2122-2126.

Spiro, R., 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology. 4, 43-56.

Torres de Pinedo, A., peñalver, P., Morales, J., 2007. Synthesis and evaluation of new phenolic-based antioxidants: Structure-activity relationship. Food Chem. 103, 55-61.

Torres, P., Poveda, A., Jimenez-Barbero, J., Parra, J.L., Comelles, F., Ballesteros, A.O. *et al*., 2011. Enzymatic synthesis of α−glucosides of resveratrol with surfactant activity. Adv. Synth. Catal. 353, 1077-1086.

Urrutia, P., Rodríguez Colinas, B., Fernandez-Arrojo, L., Ballesteros, A., Wilson, L., Illanes, A. *et al*., 2013. Detailed analysis of galactooligosaccharides synthesis with β−galactosidase from *Aspergillus oryzae*. J. Agric. Food Chem. 61,

van Rantwijk, F., Woudenberg-van Oosterom, M., Sheldon, R., 1999. Glycosidase-catalysed synthesis of alkyl glycosides. J. Mol. Catal. B: Enz. 6, 511-532.

Woo, H.J., Kang, H.K., Nguyen, T.T.H., Kim, G.E., Kim, Y.M., Park, J.S. *et al*., 2012. Synthesis and characterization of ampelopsin glucosides using dextransucrase from *Leuconostoc mesenteroides* B-1299CB4: Glucosylation enhancing physicochemical properties. Enz. Microb. Technol. 51, 311-318.

Xu, L., Qi, T., Xu, L., Lu, L., Xiao, M., 2016. Recent progress in the enzymatic glycosylation of phenolic compounds. J. Carbohydr. Chem. 35, 1-23.

Yamamoto, I., Muto, N., Murakame, K., Suga, S., Yamaguchi, H., 1990. L-ascorbic acid alpha-glucoside formed by regioselective transglucosylation with rat intestinal and rice seed alpha−glucosidases: Its improved stability and structure determination. Chem. Pharmacol. Bull. 38, 3020-3023.

# SUPPLEMENTARY INFORMATION

**Supplementary data 1.** Equations for maximum production and maximum conversion for hydroxytyrosol and vanillyl glucosides.

Hydroxytyrosol glucoside highest production was adjusted to the following quadratic model equation: Production (g/L) = - 0.039701 + 0.028607 * [BGL-2] - 3.25588E-003 * [Cellobiose] -2.35913E-003 * [HT] + 6.25378E-003 * Temperature + 0.10520 * Time + 3.20163E-004 * [BGL-2] * [Cellobiose] + 9.05336E-004 * [BGL-2] * [HT] - 3.86338E-004 * [BGL-2] * Temperature - 5.00206E-003 * [BGL-2] * Time + 1.59294E-004 * [Cellobiose] * [HT] + 9.09881E-005 * [Cellobiose] * Temperature + 1.31978E-003 * [Cellobiose] * Time + 3.55835E-004 * [HT] * Temperature + 3.60724E-003 * [HT] * Time - 3.52302E-004 * Temperature * Time - 2.08988E-003 * $[BGL-2]^2$ - 1.95100E-005 * $[Cellobiose]^2$ - 6.83017E-004 * $[HT]^2$ - 1.41925E-004 * $Temperature^2$ - 0.050764 * $Time^2$.

Hydroxytyrosol glucoside highest conversion was adjusted to the following 2FI model equation: % conversion = + 2.90297 + 0.13209 * [BGL-2] - 0.045489 * [Cellobiose] - 0.017253 * [HT] -0.014176 * Temperature + 0.83868 * Time + 3.02077E-003 * [BGL-2] * [Cellobiose] -2.97914E-003 * [BGL-2] * [HT] -4.06229E-003 * [BGL-2] * Temperature - 0.062198 * [BGL-2] * Time - 3.79991E-004 * [Cellobiose] * [HT] + 8.56505E-004 * [Cellobiose] * Temperature + 0.012421 * [Cellobiose] * Time + 5.57314E-004 * [HT] * Temperature - 0.032373 * [HT] * Time - 6.90422E-003 * Temperature * Time

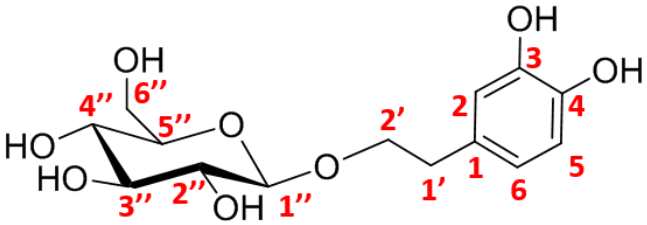
Vanillyl glucoside highest production also was adjusted to a quadratic model equation: Production (g/L) = -0.57667 + 0.061395 * [BGL-2] - 5.42711E-003 * [Cellobiose] + 2.40629E-004 * [Van] + 0.035012 * Temperature +0.21383 * Time + 5.14079E-004 * [BGL-2] * [Cellobiose] + 1.26264E-003 * [BGL-2] * [Van] - 7.16460E-004 * [BGL-2] * Temperature - 9.82830E-003 * [BGL-2] * Time + 2.99428E-004 * [Cellobiose] * [Van] + 1.28273E-004 * [Cellobiose] * Temperature + 1.96047E-003 * [Cellobiose] * Time + 3.10440E-004 * [Van] * Temperature + 6.00413E-003 * [Van] * Time - 3.65169E-003 * Temperature * Time - 3.26899E-003 * $[BGL-2]^2$ - 3.14314E-005 * $[Cellobiose]^2$ - 9.34766E-004 * $[Van]^2$ - 4.64102E-004* $Temperature^2$ -0.037594 * $Time^2$.

The maximum conversion rate for Vanillyl glucoside was determined by the following 2FI equation: % conversion = + 3.20209 + 0.22437 * [BGL-2] - 0.071574 * [Cellobiose] - 0.015999 * [Van] + 0.034461 * Temperature + 2.52053 * Time + 4.99532E-003 * [BGL-2] * [Cellobiose] - 5.90886E-003 * [BGL-2] * [Van] - 7.25258E-003 * [BGL-2] * Temperature - 0.085817 * [BGL-2] * Time - 5.53919E-004 * [Cellobiose] * [Van] + 1.13552E-003 * [Cellobiose] * Temperature + 0.019881 * [Cellobiose] * Time + 4.27289E-004 * [Van] * Temperature - 0.033424 * [Van] * Time - 0.048144 * Temperature * Time

[HT]: Hydroxytyrosol concentration. [Van]: Vanillyl alcohol concentration.

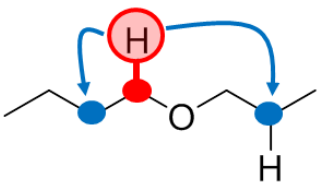
## **Supplementary data 2.** NMR study of hydroxytyrosol glucoside

We herein describe the assignment of the hydroxytyrosyl-glucose derivative studied by NMR.

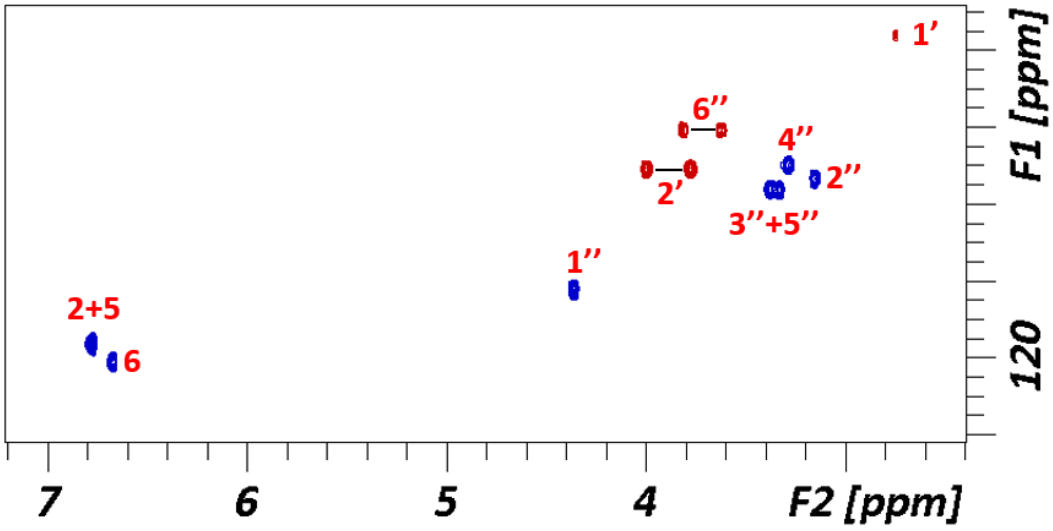



Structure of the hydroxytyrosyl-glucose derivative.

NMR spectroscopy: NMR experiments were acquired at 298 K, using a Bruker AVANCE 600 MHz spectrometer equipped with a cryogenic probe. 1D $^{1}H$ NMR spectra, $^{1}H$-$^{13}C$ HSQC and HMBC experiments were acquired to assign all the NMR signals. For 1D $^{1}H$, $^{1}H$-$^{13}C$ HSQC and HMBC experiments, the zg, hsqcedetgp, and hmbcgpndqf sequences were employed. HSQC correlates the chemical shifts of two types of nuclei (in this case $^{1}H$ and $^{13}C$) that are directly bonded, while HMBC points out the correlation through two or more chemical bonds.

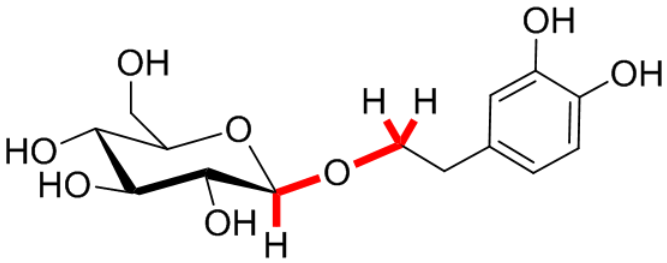


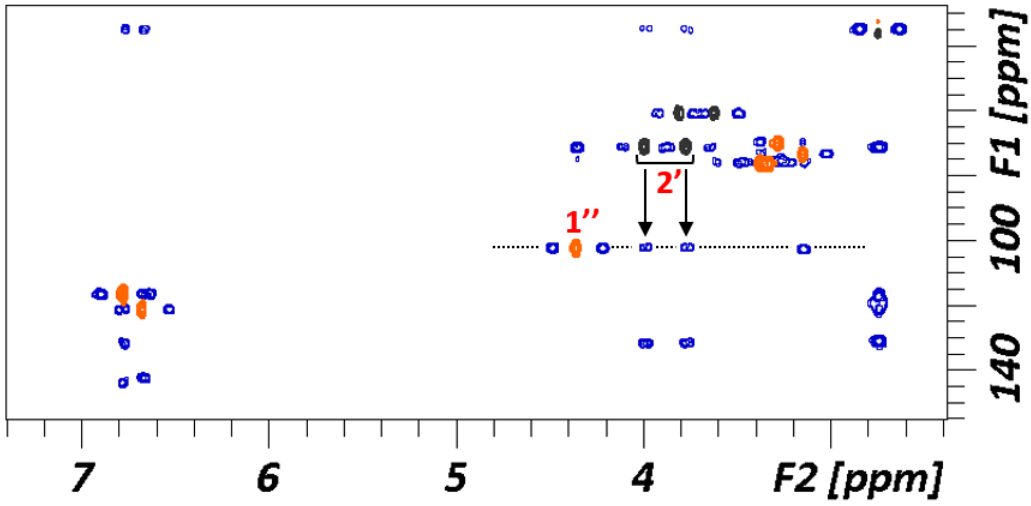

HSQC correlation (red) and HMBC correlation (blue).




Labelled $^{1}H$-$^{13}C$ HSQC of the hydroxytyrosyl-glucose derivative.

We employ HMBC to observe the correlation between the anomeric carbon (**C1''**) and the carbon **C2'** of the hydroxytyrosyl residue. Moreover, the anomeric proton presents a coupling constant value of 8 Hz, pointing out the formation of the derivative through a β-linkage.Therefore, we can assure that the anomeric position is bonded to the hydroxytyrosyl residue through the aliphatic chain.

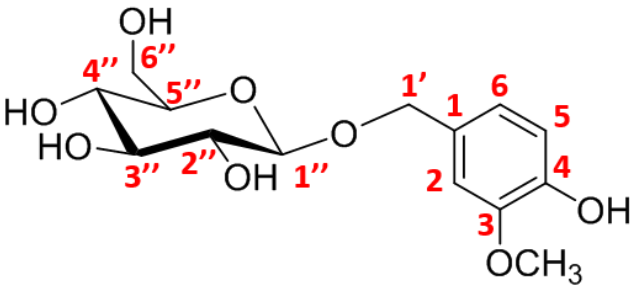



Superimposition of $^{1}H$-$^{13}C$ HSQC (orange/black) and HMBC (blue). Arrows represent the key cross peaks for the characterization of the molecule, the correlation of the anomeric position and the hydroxytyrosyl residue.

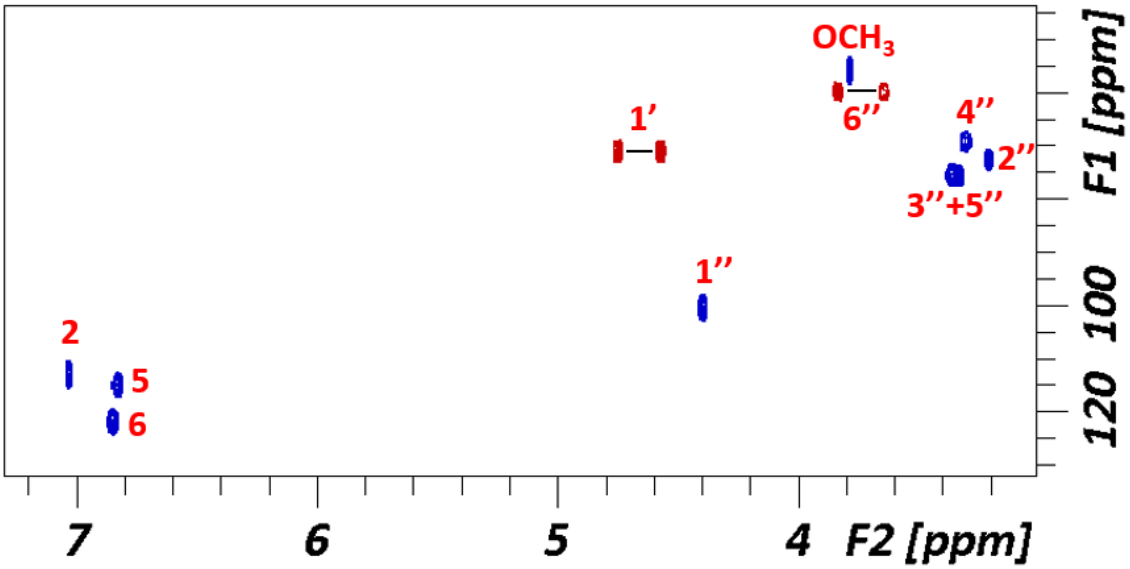
**Supplementary data 3.** NMR study of vanillin glucoside

We herein describe the assignment of the vanillyl-glucose derivative studied by NMR.



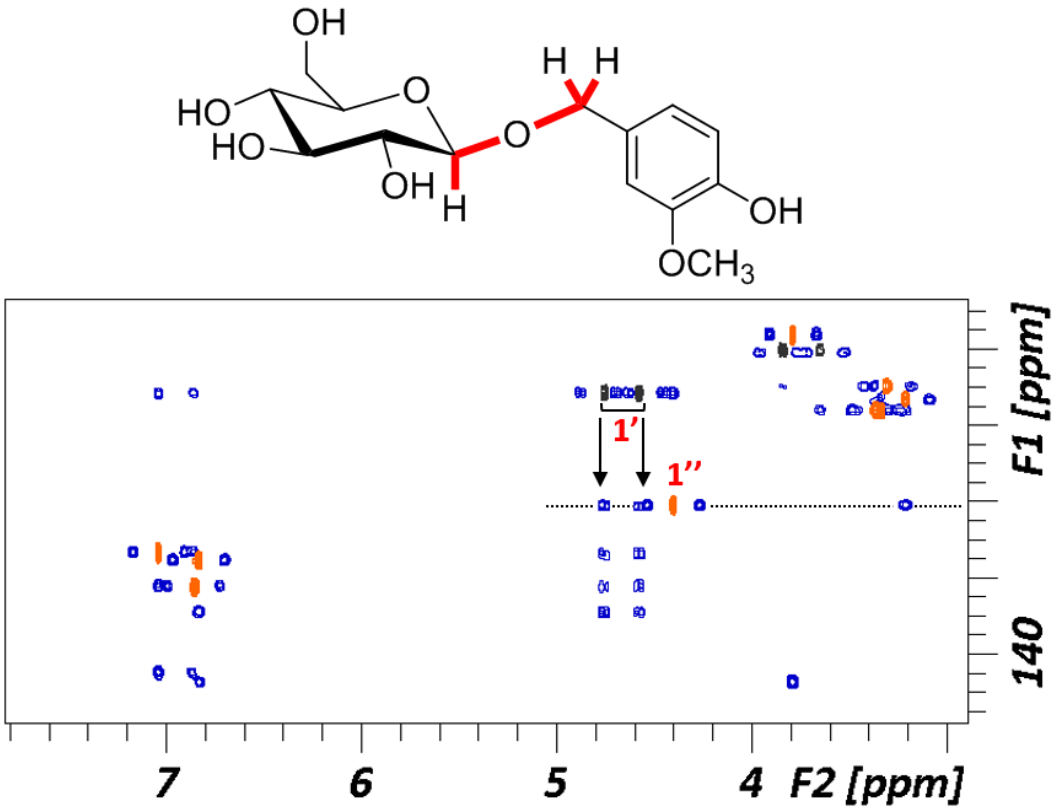

Structure of the vanillyl-glucose derivative.

NMR spectroscopy: NMR experiments were acquired at 298 K, using a Bruker AVANCE 600 MHz spectrometer equipped with a cryogenic probe. 1D $^{1}H$ NMR spectra, $^{1}H$-$^{13}C$ HSQC and HMBC experiments were acquired to assign all the NMR signals. For 1D $^{1}H$, $^{1}H$-$^{13}C$ HSQC and HMBC experiments, the zg, hsqcedetgp, and hmbcgpndqf sequences were employed.




Labelled $^{1}H$-$^{13}C$ HSQC of the vanillyl-glucose derivative.

We employ HMBC to observe the correlation between the anomeric carbon (**C1''**) and the carbon **C1'** of the vanillyl residue. Moreover, the anomeric proton presents a coupling constant value of 7.7 Hz, pointing out the formation of the derivative through a β-linkage. Therefore, we can assure that the anomeric position is bonded to the vanillyl residue through the aliphatic chain.




Superimposition of $^{1}H$-$^{13}C$ HSQC (orange/black) and HMBC (blue). Arrows represent the key cross peaks for the characterization of the molecule, the correlation of the anomeric position and the vanillyl residue.

# CHAPTER 5

**A glucotolerant β-glucosidase from the fungus *Talaromyces amestolkiae* and its conversion into a glucosynthase capable to glycosylate epigallocatechin gallate.**

**Méndez-Líter, J.A[1].**, Nieto-Domínguez, M[1]., de Toro, B.F[2]., González Santana, A[3]., Prieto, A[1]., Asensio, J.L[3]., Cañada, F.J[2]., de Eugenio, L.I[1]., Martínez, M.J[1].,
2019. A glucotolerant β-glucosidase from the fungus *Talaromyces amestolkiae* and its conversion into a glucosynthase capable to glycosylate epigallocatechin gallate.

(In preparation)

[1]Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain.
[2]Department of Chemical and Physical Biology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain.
[3]Glycochemistry and Molecular recognition group. Instituto de Química Orgánica General (IQOG-CSIC), Calle Juan de la Cierva, 3, 28006 Madrid, Spain.

## ABSTRACT

**Background:** The interest for finding novel β-glucosidases that can improve the yields for the production of second-generation (2G) biofuels is still very high. One of the most desired features for these enzymes is the glucose tolerance, which allows them to act optimally over elevated glucose concentrations. Besides, there is an additional attention in finding novel enzymatic alternatives for glycosides synthesis, and a mutated version of glycosidases, named glycosynthases, have gained much interest in recent years.

**Results:** In this work, a glucotolerant β-glucosidase (BGL-1) from the ascomycete fungus *Talaromyces amestolkiae* has been heterologously expressed in *Pichia pastoris*, purified, and characterized. The production of the enzyme in the yeast was very high, reaching 75 U/mL, and it was purified in just one step with a yield of 80%.

Although the enzyme showed good efficiency on *p*NPG ($K_m$=3.36 ± 0.7 mM, $k_{cat}$=898.31 $s^{-1}$), the activity detected over cellooligosaccharides, the natural substrates of the enzyme, was much lower, which could limit its exploitation in lignocellulose degradation processes. Interestingly, when examining the substrate specificity of BGL-1, it showed to be more active over sophorose, the β-1,2 disaccharide of glucose, than over cellobiose.

Besides, the transglycosylation profile of BGL-1 was examined, and, for expanding its synthesis capacities, it was converted into a glycosynthase. The mutant enzyme, named BGL-1-E521G, was able to use α-D-glucosyl-fluoride as donor in glycosylation reactions, and synthesized glycosylated derivatives of different *p*NP-sugars in a regioselective manner, but also of some phenolic compounds of industrial interest as epigallocatechin gallate.

**Conclusions:** In this work, we report the characterization of a novel glucotolerant 1,2-β-glucosidase, which also has a considerable activity over 1,4-β-glucose bonds, has been cloned in *P. pastoris*, produced, purified and characterized. In addition, the enzyme was converted in an efficient glycosynthase, which can transfer molecules of glucose to different interesting compounds. The interesting capacities of BGL-1 and its glycosynthase mutant, both in hydrolysis or in synthesis, suggest that it could be an enzyme with applications in industrial processes.

**Keywords**: Transglycosylation; Glycosynthases; Glycosyl hydrolases; Phenolic compounds.

## BACKGROUND

Lignocellulosic biomass is the most abundant material in the world to generate value added products for the benefits of mankind. It is composed by three distinct polymers, lignin, cellulose, and hemicelluloses. Cellulose is the major polysaccharide in plant cell wall and it is conformed as a linear polymer of glucose, linked by β-1,4 linkages. For a long time it was considered that this polysaccharide was exclusively degraded by the synergistic action of cellulolytic glycosyl hydrolases (GHs): endoglucanases (EGs), cellobiohydrolases (CBHs), and β-glucosidases (BGLs) (Martínez *et al*., 2005). However, this classic model underwent a big change with the discovery of lytic polysaccharide monooxigenases (LPMOs), auxiliary enzymes that break crystalline cellulose by oxidative depolymerization, creating new chains, and improving cellulose accessibility by glycosyl hydrolases (Martínez, 2016). The combination of endoglucanases, cellobiohydrolases and LPMOs generates cellobiose and other small cellooligosaccharides, which are degraded by β-glucosidases, converting them into glucose as the final step of cellulose degradation. BGLs are fundamental to degrade this polysaccharide since, if their activity is not high enough, oligosaccharides will accumulate in the medium, causing product inhibition of other cellulases and thus decreasing total yields (Sørensen *et al*., 2013). However, these enzymes are generally found in low proportion in commercial preparations, which are mostly produced by the fungus *Trichoderma reesei*, which usually represents the major bottleneck in cellulose degradation. Hence, many studies are focused on finding robust and efficient β-glucosidases, since enzyme cocktails must be supplemented with this activity to increase the efficiency of cellulose saccharification. BGLs are widely distributed in GHs families, but, according to the CAZY database, the two most predominant families are GH1 and GH3, being GH1 the family with the largest number of characterized BGLs (Lombard *et al*., 2014).

One of the most important characteristics of BGLs is that they are frequently inhibited by their own product, glucose. Therefore, glucotolerance is a feature always sought in novel BGLs, because the use of glucotolerant BGLs would probably decrease the enzyme amount needed for a complete lignocellulose hydrolysis, reducing process costs. Nevertheless, despite their fantastic hydrolytic capacity, glucotolerant GH3 BGLs are exceptional (Ramani *et al*., 2015), and this interesting property seems to be exclusive of BGLs from the GH1 family.

On the other side, the capacity of GHs for catalyzing transglycosylation reactions makes them a great biotechnological tool for the modification of several molecules of interest, through the addition of one or more sugar units to different compounds. Increasing the solubility of the original compound, making it safer, or improving its stability are among the many beneficial effects reported for glycosides (Dinesh *et al*., 2015; Xu *et al*., 2016). The application of biocatalysis to synthesize new glycosides has two main advantages: i) the

synthesis process is generally much less contaminant and eco-friendly, and ii) it is usually simpler, since the complex structure of oligosaccharides makes classical chemical synthesis difficult if the stereospecificity and regiospecificity of the glycosylation have to be controlled (Perugino *et al*., 2004).

Although GHs can be successfully used to synthesize glycoconjugates, the reaction yields are often poor. The newly-formed products are easily hydrolyzed by the same GH and, thus, the process is not economically viable at a large-scale. However, the hydrolytic capacity of these enzymes can be eliminated using rational design tools, leading to the production of a novel class of enzymes that possess only synthetic activity. In this context, glycosynthases are obtained by directed mutagenesis of the nucleophile amino acid placed in the active center of GHs (Mackenzie *et al*., 1998). These mutant enzymes catalyze the formation of glycosidic bonds using glycosyl fluoride donors. In the absence of the nucleophile residue the enzyme cannot degrade the produced glycosides and the reaction yields are much higher (Danby and Withers, 2016). The use of glycosynthases has been shown as an effective way to generate a wide variety of value added products, as oligosaccharides of nitrophenol-sugars, methylumbellyferyl-sugars (Perugino *et al*., 2004), oligosaccharides (Hayes and Pietruszka, 2017), and human milk oligosaccharides with prebiotic activity (Saumonneau *et al*., 2015).

The ascomycete *Talaromyces amestolkiae* was studied by its ability to degrade cellulose and hemicellulose. Sequencing and annotation of the genome of this fungus disclosed a high number of genes encoding GHs and specially BGLs, being its number of CAZymes significantly higher than those reported for other organisms used to produce commercial cellulolytic enzyme cocktails (de Eugenio *et al*., 2017). Two of the BGLs produced by *T. amestolkiae*, BGL-2 and BGL-3, belong to the GH3 family and have interesting peculiarities. BGL-2 is the major β-glucosidase secreted by this fungus in the presence of cellulosic inducers. Structurally, this enzyme has a cellulose binding domain, an unusual feature among BGLs (Méndez-Líter *et al*., 2017). On the other hand, *T. amestolkiae* produces BGL-3 in all carbon sources tested, which is uncommon for β-glucosidases, and it was isolated and characterized from a basal medium with glucose under carbon starvation conditions (Méndez-Líter *et al*., 2018). Both, BGL-2 and BGL-3, were efficiently used as BGL supplements of commercial cocktails for saccharification of lignocellulosic waste. However, the BGLs from the family GH1 produced by this fungus remain completely unexplored. In this work, we report the cloning and expression in *Pichia pastoris* of the β-glucosidase gene *bgl1* of this fungus, the purification and characterization of the recombinant BGL-1 protein, and its conversion into a glycosynthase, aimed to expand the biotechnological applications of this enzyme.

# RESULTS AND DISCUSSION

## Cloning, production, purification and biochemical characterization of BGL-1

*T. amestolkiae* has been recently postulated as a very interesting option for producing enzymatic cocktails rich in BGLs (de Eugenio *et al*., 2017). All the BGLs from this fungus characterized until date belong to GH3 family (Méndez-Líter *et al*., 2017; Méndez-Líter *et al*., 2018), which is usually considered as the one that groups the BGLs with better catalytic efficiency, although these enzymes have some limitations such as their low glucotolerance (Cao *et al*., 2015). In this sense, the GH1 family contains most of the glucose-tolerant BGLs characterized (Singhania *et al*., 2013). For this reason, investigating the presence of potential glucotolerant GH1 BGLs in *T. amestolkiae* genome and proteomes could increase the value of the cellulolytic system of this fungus.

In a previous work (de Eugenio *et al*., 2017), the secretome released by *T. amestolkiae* growing in different carbon sources was analyzed and, in every condition tested, one potential BGL from the family GH1 (protein g8384, renamed as BGL-1) was detected in very low amounts. Therefore, the *bgl1* gene was cloned and expressed in *P. pastoris* with the goal of increasing BGL-1 production in order to analyze its glucose tolerance, kinetic constants, and physicochemical properties. After identifying the DNA sequence, RNA was extracted from 7-day old cultures of *T. amestolkiae*, obtaining total cDNA by retrotranscription. Amplification of the sequence of the mature *bgl-1* gene concluded that the 1906 bp gene contains one intron, and encodes a 619 amino acids protein (figure S1).

The methylotrophic yeast *P. pastoris* has been widely used as one of the most efficient expression systems for heterologous expression of BGLs. Some of its most notable advantages include its ability to produce correctly folded protein at high levels, or to perform complex post-translational modifications (Cereghino and Cregg, 2000). A recombinant plasmid with the sequence of *bgl-1* without the intron was constructed using pPICZα as vector. Once transformed *P. pastoris* X-33, the transformants were screened to detect the best β-glucosidase producers. The maximal β-glucosidase activity was 75 U/mL, which is among the highest productions of BGLs reported in the literature. As can be seen in Table 1, this value is only surpassed by those found for the recombinant PtBglu3 from *Paecilomyces thermophila*, and bgl3A, from *Talaromyces leycettanus*. It is important to emphasize that the activity determined for this recombinant BGL-1 was 35-fold higher than the total β-glucosidase activity detected in cultures of *T. amestolkiae* (Méndez-Líter *et al*., 2017), which contains a mixture of BGL-1, BGL-2 and BGL-3. This value confirms the very high overexpression of BGL-1 in this system.

**Table 1.** Comparison of the heterologous production in *P. pastoris* of BGLs from different microorganisms.

| Enzyme Name | Microorganism | Production (U/mL) | References |
|---|---|---|---|
| rBgl3 | *Aspergillus fumigatus* | 4.9 | (Liu *et al*., 2012) |
| rBgl4 | *Penicillium funiculosum* | 52.8 | (Ramani *et al*., 2015) |
| bgl3A | *Talaromyces leycettanus* | 6,000.0 | (Xia *et al*., 2016) |
| PtBglu3 | *Paecilomyces thermophila* | 274.4 | (Yan *et al*., 2012) |
| Nfbgl1 | *Neosartorya fischeri* | 33.5 | (Yang *et al*., 2014) |
| MtBgl3a | *Myceliophtora termophila* | 41.0 | (Karnaouri *et al*., 2013) |
| Bgl3B | *Talaromyces leycettanus* | 1.5 | (Li *et al*., 2018) |
| BGL-2 | *T. amestolkiae* | 6.0 | (Méndez-Líter *et al*., 2017) |
| BGL-3 | *T. amestolkiae* | 8.1 | (Méndez-Líter *et al*., 2018) |
| **BGL-1** | ***T. amestolkiae*** | **75.0** | **This work** |

BGL-1 was purified in very high yield (around 80%) in just one step, by anion-exchange chromatography using a HiTrap QFF cartridge (Table 2), and 10 mM sodium phosphate buffer, pH 6.

**Table 2.** Purification of the β-glucosidase BGL-1 secreted from *P. pastoris* cultures.

| **BGL-1 Purification** | | | | |
|---|---|---|---|---|
| Step | Total Protein (mg) | Total Activity (U) | Specific Activity (U/mg) | Yield (%) |
| Crude extracts | 39.66 | 4,176.97 | 105.48 | 100.0 |
| HiTrap QFF Anion exchange | 11.82 | 3,354.15 | 282.83 | 80.1 |

The isoelectric point of the pure BGL-1 was determined to be 6.7 by isoelectrofocusing, and its molecular mass, measured by MALDI-TOF mass spectrometry, was 88.11 kDa. Considering that the theoretical mass was 23% lower (68.05 kDa), the differences found can be attributed to *P. pastoris* protein hyperglycosylation (Cereghino and Cregg, 2000).

Optimal activity of BGL-1 was detected at pH 4 and 60 °C (figure 1). These values are in the ranges reported for other native β-glucosidases (Singhania *et al*., 2013; Sørensen *et al*., 2013), although BGL-1 has the peculiarity of working unusually well at pH 3 and pH 4. This behavior was also observed in the two other known BGLs of this fungus (Méndez-Líter *et al*., 2017; Méndez-Líter *et al*., 2018), which may indicate that these enzymes from

*T. amestolkiae* are more tolerant to acidic pHs than most BGLs characterized to date.

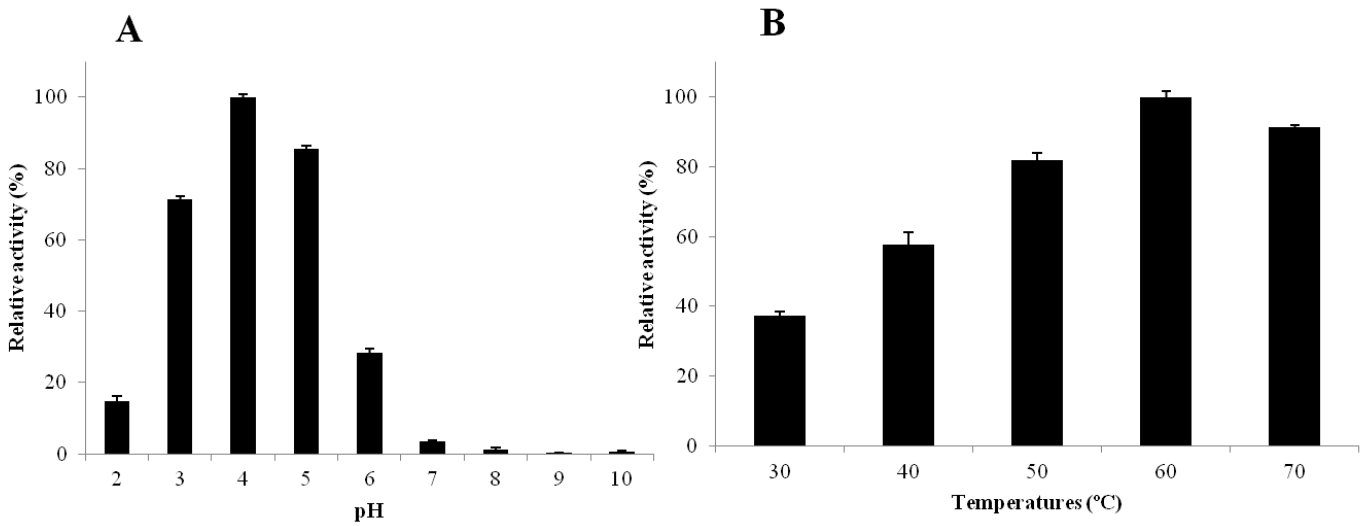



**Figure 1.** Effect on BGL-1 activity of: A) pH and B) temperature.

## Glucose tolerance and kinetic study of BGL-1

In general, most of the β-glucosidases used for cellulose degradation belong to the GH3 family, showing high catalytic efficiency values. However, they are often inhibited by glucose, with inhibition constant values lower than 0.1 M. In contrast, some GH1 β-glucosidases are much more glucotolerant than GH3 BGLs, but they usually have lower $k_{cat}/K_m$ values over cellooligosaccharides, which are their natural substrates (Sørensen *et al*., 2013; Cao *et al*., 2015). The inhibition constants of BGL-1 towards glucose were calculated, displaying two main features that should be highlighted. First, its $K_i$ value was very high (3.78 M), which to the best of our knowledge, corresponds to the second highest reported (Li *et al*., 2012). Second, the activity of BGL-1 was stimulated by low concentrations of glucose. This finding has been observed in other β-glucosidases, mostly belonging to the GH1 family (Cao *et al*., 2015), but also in few BGLs from the GH3 family (Ramani *et al*., 2015). The reason for this glucose-induced stimulation remains unknown, although it could be related either to an allosteric effect triggered by the binding of glucose to some part of the protein, or to an increased hydrolysis rate upon transglycosylation (Uchima *et al*., 2011). BGL-1 activity was improved by 1.18-fold in the presence of 0.25 M of glucose (figure 2). At this point, the activity begins to decrease, although BGL-1 retained 40% of its initial activity at 3 M glucose. Both characteristics may postulate BGL-1 as a candidate for industrial processes performed at high glucose concentrations.

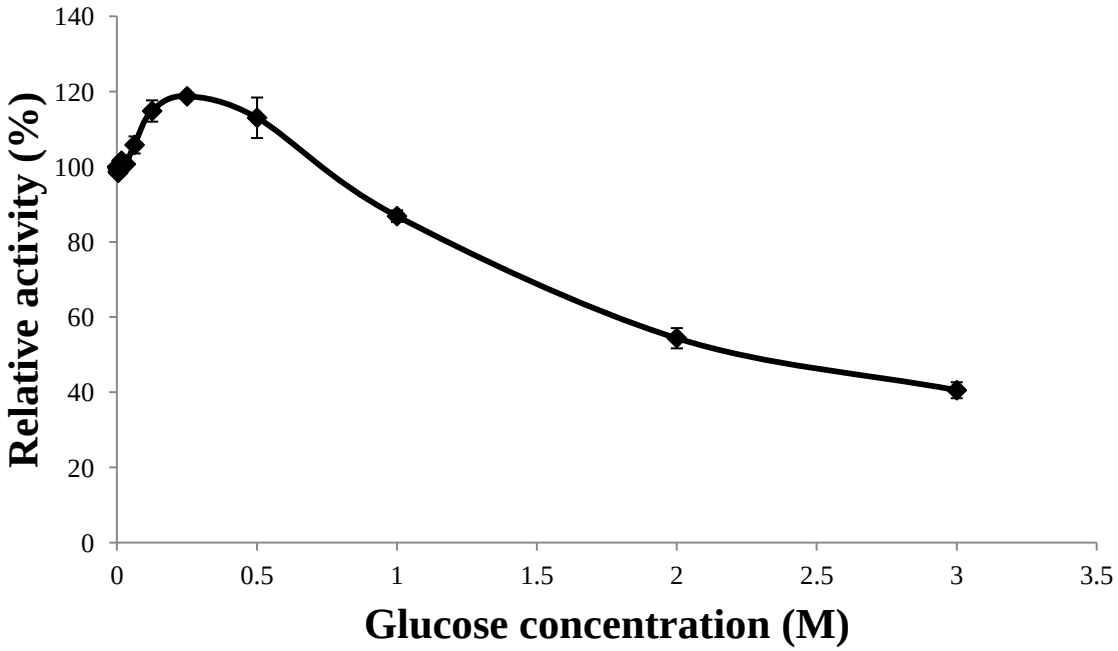



**Figure 2.** Inhibitory effect of glucose on the activity of BGL-1 in hydrolysis of *p*NPG.

β-glucosidases can be classified into three groups, according to their substrate preferences: cellobiases, which have high substrate specificity towards cellooligosaccharides, aryl-β-glucosidases, with very high specificity towards synthetic substrates as *p*NPG, and β-glucosidases with broad substrate specificity, that combine both activities (Singhania *et al*., 2013). In this sense, the kinetic constants showed that BGL-1 has very high efficiency and good affinity on *p*NPG and *o*NPG and, in addition, this enzyme displays remarkable activity against *p*-nitrophenyl-β-D-xylopyranoside (specific activity 5.3 U/mg), demonstrating to have some versatility. However, its catalytic efficiency against cellooligosaccharides, from cellobiose to cellohexaose is low, as occurs with other fungal BGLs from the GH1 family (Table 3). Thus, in contrast with the high affinity on oligosaccharides shown by BGL-2 and BGL-3, which could be classified as cellobiases from the family GH3, BGL-1 seems to be a member of the group of aryl-β-glucosidases, which are much more active over synthetic substrates.

**Table 3.** Kinetic constants of BGL-1 hydrolyzing different substrates.

| Substrate | $K_m$ (mM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1}.s^{-1}$) |
|---|---|---|---|
| ***p*NPG** | 3.36±0.7 | 898.31 | 267.35 |
| ***o*NPG** | 2.36±0.6 | 135.72 | 57.50 |
| **Cellobiose** | 20.36±3.4 | 137.77 | 6.76 |
| **Cellotriose** | 19.39±5.4 | 196.24 | 10.12 |
| **Cellotetraose** | 17.62±0.6 | 276.62 | 15.69 |
| **Cellopentaose** | 12.41±0.2 | 260.42 | 20.98 |
| **Cellohexaose** | 9.18±0.6 | 217.86 | 23.73 |

However, it is interesting to remark that, in spite of the relatively good $k_{cat}$ values of BGL-1 in the hydrolysis of cellooligosaccharides, the $K_m$ values are poor when compared with those observed for other BGLs from the GH3 family (Table 4). This result confirms that, as other BGLs from the GH1 family, BGL-1 has low affinity against these substrates, which could limit its applicability in hydrolysis processes.

**Table 4.** Comparison of the catalytic efficiency on cellobiose and glucotolerance reported for BGLs from the GH3 and GH1 families.

| Enzyme | $K_m$ (mM) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ | $K_i$ (mM) /Glucose tolerance |
|---|---|---|---|---|
| **GH3** | | | | |
| *Talaromyces leycettanus* Bgl3A (Xia *et al*., 2016) | 10.4 | 786 | 75.8 | 14 |
| *Talaromyces amestolkiae* BGL-2 (Méndez-Líter *et al*., 2017) | 1.1 | 630 | 567 | 1 |
| *Talaromyces amestolkiae* BGL-3 (Méndez-Líter *et al*., 2018) | 0.5 | 1594 | 3308 | 1.6 |
| *Penicillium funiculosum* rBgl4 (Ramani *et al*., 2015) | 1.2 | 4513 | 3610 | 60% of residual activity at 500 mM |
| **GH1** | | | | |
| Metagenomic library BGL mutant M3 (Cao *et al*., 2015) | 49.2 | 48.4 | 1.1 | 50% of residiual activity at 250 mM |
| Metagenomic library BGL mutant V174C (Cao *et al*., 2015) | 45.1 | 83.1 | 1.7 | 50% of residual activity at 3000 mM |
| Soil metagenomic library BGL (Li *et al*., 2012) | ND | ND | ND | 4280 |
| *Thermoanaerobacterium thermosaccharolyticum* BGL (Pei *et al*., 2012) | 7.9 | 120 | 13.3 | 600 |
| *Thermoanaerobacterium aotearoense* BGL (Yang *et al*., 2015) | 25.4 | 740.5 | 29.1 | 800 |
| *Actinomadura amylolytica* AaBGL1 (Yin *et al*., 2018) | 95.3 | 10.7 | 0.1 | 40% of residual activity at 2000 mM |
| *Actinomadura amylolytica* AaBGL2 (Yin *et al*., 2018) | 187.7 | 16.6 | 0.1 | 40% of residual activity at 500 mM |
| *Thermotoga thermarum* BGL (Zhao *et al*., 2013) | 35.5 | 19 | 0.5 | 1500 |
| **BGL-1*** | **20.4** | **137.8** | **6.8** | **3780** |

*ND: Not determined in the original article.

Finally, an interesting discovery was made when examining the regioselectivity in hydrolysis reactions catalyzed by BGL-1. Enzymatic activity was tested against cellobiose, sophorose, laminaribiose and gentiobiose. While the activity over laminaribiose and gentiobiose was very low, the activity over cellobiose and sophorose was considerably high, and BGL-1 was 5 fold more active on sophorose (535.82 U/mg) than over cellobiose (110.27 U/mg). These data confirmed that BGL-1 is a versatile β-1,2 BGL, since it is also able to hydrolyze β-1,4 bonds. It is interesting to remark that, until date, a very low number of β-1,2-glycosidases have been reported. For example, some glucanases and glucosidases induced by β-1,2-glucan have been discovered in *Acremonium sp*., a filamentous anamorphic fungus (Kitahata and Edagawa, 1987). However, their amino acid sequences have not been determined, which precludes their comparison with the BGL-1 from *T. amestolkiae*. Recently, a BGL with activity on β-1,2 bonds was reported in *Listeria innocua*, and related with the β-1,2-glucan metabolism in this bacteria (Nakajima *et al*., 2016), although the physiological role of these BGLs able to synthetize the β-1,2 disaccharide, remains poorly understood. However, it is well known that sophorose is the most powerful inducer of cellulases in *T. reesei*. A recent report describes the production of this disaccharide by transglycosylation catalyzed by intracellular BGLs of this fungus, and the regulatory role of another BGL that hydrolyzes this compound, triggering the synthesis of cellulases (Sternberg and Mandels, 1980; Zou *et al*., 2018). Therefore, taking into account that BGL-1 is produced by *T. amestolkiae* in all conditions assayed, and that this enzyme can hydrolyze and synthetize β-1,2 bonds, its physiological role could be related with the regulation of the induction of the cellulolytic system in this organism.

**Transglycosylation profile and regioselectivity of BGL-1**

In order to test the transglycosylation capacities of BGL-1 a screening with a variety of potential acceptors, like sugars, sterols, phenolic compounds, or amino acids (figure S2) was performed according to methodology previously developed (Blanchard and Withers, 2001) and also applied for studying the transglycosylation profile of other *T. amestolkiae* BGLs (Méndez-Líter *et al*., 2019). Unfortunately, in contrast with the good results obtained with the GH3 BGLs of this fungus, BGL-1 only showed potential for transglycosylating *p*-nitrophenol sugars derivatives, like *p*NPG, *p*NPGal, or *p*NPX, which discarded most of the potential acceptors of transglycosylation tested with the wild type enzyme. On the other hand, its regioselectivity was assessed analyzing by NMR the products of a transglycosylation reaction set up with *p*NPG as donor and $^{13}$C-labelled glucose as acceptor. The NMR spectra of the compounds detected in the mentioned reaction were compared with those from sophorose, cellobiose and laminaribiose, confirming their coincidence with the pattern from sophorose (figure 3). This result showed that BGL-1 transglycosylated only in the β-1,2 position.

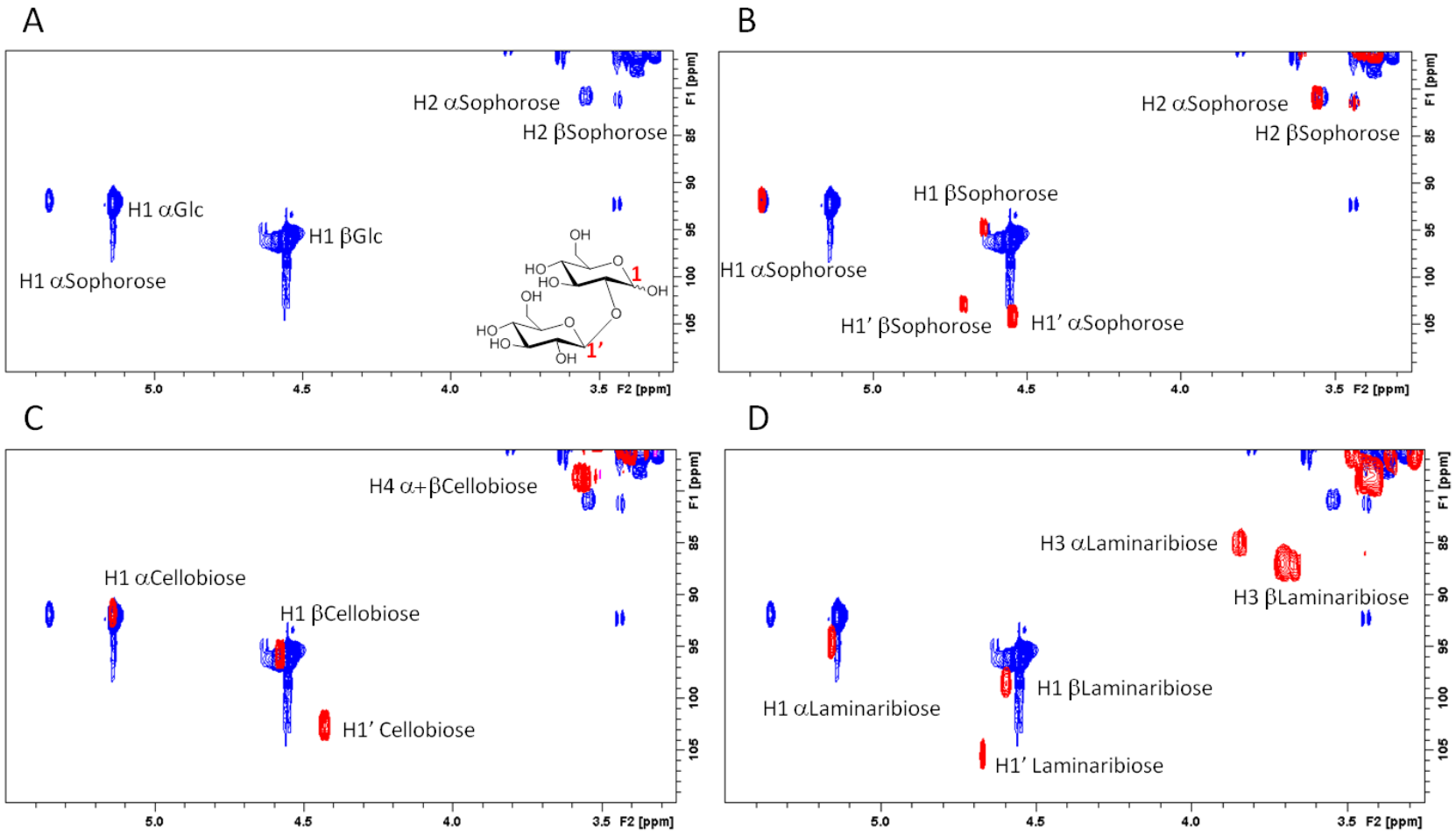



**Figure 3.** $^{1}$H-$^{13}$C HSQC NMR analysis to determine the regioselectivity of the β-glycosidase BGL-1 in transglycosylation. Detail of the anomeric region of the spectra of: A) the reaction mixture containing the products of transglycosylation of $^{13}$C-labelled glucose (acceptor) and *p*NPG (donor). The high intensity of the signals from unreacted $^{13}$C-glucose are due to the isotopic labelling. The structure of sophorose, identified as the main reaction product, is represented in the inset. B) Superimposition of the spectra of the reaction mixture (A, in blue) and commercial sophorose (in red). The signals from H1' of α- and β-sophorose are not observed in the spectrum of the reaction since the non-reducing residue comes from the donor. Only signals coming from the $^{13}$C-glucose (acceptor) can be observed (short acquisition NMR time). C) Superimposition of the spectra of the reaction mixture (A, in blue) and cellobiose (in red). D) Superimposition of the spectra of the reaction mixture (A, in blue) and laminaribiose (in red).

**Conversion of BGL-1 in glucosynthases by rational design**

With the aim of expanding the transglycosylation capacities of the recombinant BGL-1, it was converted into a glycosynthase. This kind of enzymes were first reported by Withers and coworkers (Mackenzie *et al*., 1998), who noticed that a mutated glycosyl hydrolase lacking its catalytic nucleophile can use activated glycosyl fluoride donors of the opposite anomeric configuration for synthesizing glycosides, without hydrolyzing the products. Historically, enzyme engineering has been successfully implemented to enhance the transglycosylation activity of glycosidases and, simultaneously, attenuate hydrolysis. This approach has been successfully applied for instance to convert glycosidases from GH1 family into glycosynthases, from GH36 in galactosynthases, or from GH29 in fucosynthases (Danby and Withers, 2016).

In this work, the replacement of the glutamic acid at position 521 of the protein for a glycine (BGL-1-E521G) or a serine (BGL-1-E521S) produced two novel versions of BGL-1. These mutations have been shown to be much more efficient than the alanine replacement in the synthesis of oligosaccharides and *p*-nitrophenol derivatives (Perugino *et al*., 2004). Both versions of the protein were produced in *P. pastoris* and purified with the same strategy used for BGL-1. The pure glucosynthases showed their ability to use α-GlcF for synthesizing glycosides, and the most efficient mutant was selected from a screening developed using 10 mM α-GlcF as donor and 10 mM *p*NPG as acceptor. Product formation was analyzed by HPLC, and the outcomes from this experiment confirmed that the glycine mutant was more efficient. The transglycosylation yield was 2-fold higher to that determined for the serine mutant. This is in concordance with data reported in the literature, which are explained considering that the rigid serine side-chain could be an obstacle in enhancing the departure of fluorine, which is instead stimulated in the glycine mutants (Moracci *et al*., 1998; Mayer *et al*., 2001; Jahn *et al*., 2003). In a similar way, other authors justify this different behavior between the mutants considering the lack of a side-chain in glycine, and thus of reduced steric hindrance compared with the alanine or serine side-chains (Perugino *et al*., 2003).

Once selected the glycine mutant, the kinetic parameters of this glycosynthase transglycosylating α-GlcF and *p*NPG, were determined. Since two molecules were used in the reaction, it was necessary to identify the kinetics for both substrates. The results revealed that the affinity of BGL-1-E590G for *p*NPG ($K_m$ 90.14 mM) was higher than for α-GlcF ($K_m$ 260.86 mM), although the catalytic constants were similar for both substrates ($k_{cat}$ 0.11 $s^{-1}$ and 0.08 $s^{-1}$, respectively). Few studies deal with kinetic studies with glycosynthases, but the results presented here agree with those reported (Kim *et al*., 2005), confirming that this new enzyme could be a valuable tool for transglycosylation reactions.

**Transglycosylation of selected acceptors and analysis of the products**

BGL-1-E521G was tested as the catalyst for transglycosylation of the acceptors *p*NPG, *p*NPX, *p*NPGal, vanillin, hydroxytyrosol, gallic acid, and epigallocatechin gallate (EGCG), using α-GlcF as the donor (figure 4).

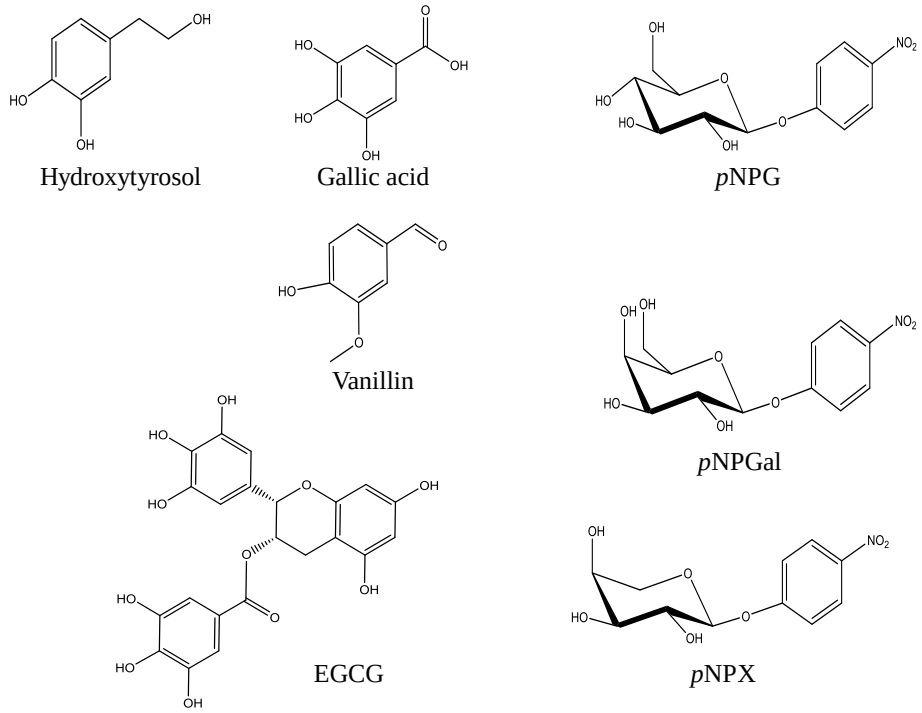



**Figure 4.** Molecular structure of the acceptors selected for transglycosylation catalysed by the glucosynthase BGL-1-E521G.

Most of the known glycosynthases use *p*NP-sugars as preferential acceptors of transglycosylation, generating a variety of products, from the expected *p*NP-disaccharides, to *p*NP-oligosaccharides of different length (Perugino *et al*., 2004). Besides these acceptors, phenolic compounds are very interesting targets for transglycosylation, because of the possibility of obtaining value-added glycosides from this type of compounds. These molecules have a lot of beneficial properties related with human health, and have been reported to confer cardiovascular protection, ant to exert a positive effect in neurodegenerative diseases and cancer (Dai and Mumper, 2010). One of the main disadvantages of these substances when used in treatments is their low bioavailability, and their glycosylation has been proposed as a potential solution, improving their solubility. In this context, various studies have demonstrated the interesting properties of hydroxytyrosol, vanillin and gallic acid (Fabiani *et al*., 2012; Kapoor, 2013; Badhani *et al*., 2015), and EGCG has recently attracted attention as a

potential therapeutic agent (Nagle *et al*., 2006; Singh *et al*., 2011), even in its glycosylated forms (Gonzalez-Alfonso *et al*., 2019). Glycosynthases can be considered as a versatile platform for saccharides synthesis and, even in the chemical processes where glycosynthases cannot improve synthetic strategies, they can be incorporated to complement these approaches (Danby and Withers, 2016).

The first approach to detect the synthesis of glycosides from the acceptors listed above consisted in a TLC analysis of the reaction mixtures, identifying positive bands for each potential glycoside (figure S3). Then, the expected molecular weight of the newly synthesized molecules was confirmed by mass spectrometry (MS). All the molecules were detected in its sodium adduct form (Table 5).

**Table 5.** ESI-MS data glycosides obtained by transglycosylation catalyzed by BGL-1-E521G. All the glycosides were detected as $Na^+$ adducts.

| **Glycoside** | **Intensity** | **m/z** |
|---|---|---|
| EGCG-glucose | 533,662 | 643.2 |
| Vanillin-glucose | 43,977 | 337.1 |
| Hydroxytyrosol- glucose | 22,386 | 339.1 |
| Gallic acid-glucose | 48,294 | 355.1 |
| *p*NPX-glucose | 307,637 | 456.1 |
| *p*NPGal-glucose | 569,165 | 486.1 |
| *p*NPG-glucose | 430,001 | 486.2 |
| G2 | 449,035 | 365.1 |
| G3 | 428,105 | 527.2 |
| G4 | 287,540 | 689.3 |
| G5 | 59,761 | 851.3 |

It is interesting to highlight that MS analysis revealed the presence of other products in all reactions. The peaks corresponding to saccharides with two (G2), three (G3), four (G4) and five (G5) glucose units were detected, showing that this glucosynthase can also generate oligosaccharides. The presence of non-fluorinated derivatives of these molecules could be due to self-hydrolysis of α-GlcF or the fluorinated derivatives during overnight reactions. Besides, considering the ability of the native BGL-1 hydrolizing *p*NPX, and the capacity of BGL-1-E521G to interact with transglycosylation acceptors with xylose, it opens up the possibility of synthesizing oligosaccharides with xylose, or using D-xylosyl fluoride as potential donors of the reactions, which may expand the applications of the enzyme. The products detected with more intensity in the TLC and MS analysis (using as acceptors *p*NPG, *p*NPX, *p*NPGal, and EGCG) were analyzed by HPLC to determine the transglycosylation yield and to purify

the main glycosides for further NMR analysis. The conversions were 73.5% for *p*NPG, 89.8% for *p*NPX, and 36.6% for *p*NPGal, and, more interestingly, the glucoside of EGCG was obtained with a high yield of 48.8%. Besides, a second EGCG product was also detected and purified. The yield of 48.8% in transglycosylation of EGCG is among the best reported in the literature, although yields of 58% (Gonzalez-Alfonso *et al*., 2018) and 91% (Kim *et al*., 2016), have also been reported for the enzymatic synthesis of the same compound using a cyclodextrin glucanotransferase and a dextransucrase, respectively. However, no optimization of the production using BGL-1-E521G has been done, and the yields can be improved in future works.

**Products characterization by NMR**

The pure EGCG, *p*NPG, *p*NPX and *p*NPGal glucosides synthetized by transglycosylation catalyzed by BGL-1-E521G were analyzed by $^{1}$H and $^{13}$C-NMR in order to assign their regiochemistry. The spectra of the two EGCG-glucose derivatives showed a correlation between the anomeric position of glucose (1) and the carbons 3''+5'' of the gallate aromatic ring, indicating the position of the linkage between the phenolic and the sugar moieties in the glycoconjugates. In addition, the second sugar unit in the EGCG-disaccharide is attached to the *O*-2 of the first one through a β-linkage, as deduced from the value of the coupling constant of the anomeric proton (7 Hz). The assignment of the signals detected for both derivatives is presented in Figure 5 and Table 6 (NMR spectra can be seen in supplementary material, figures S5 and S6).

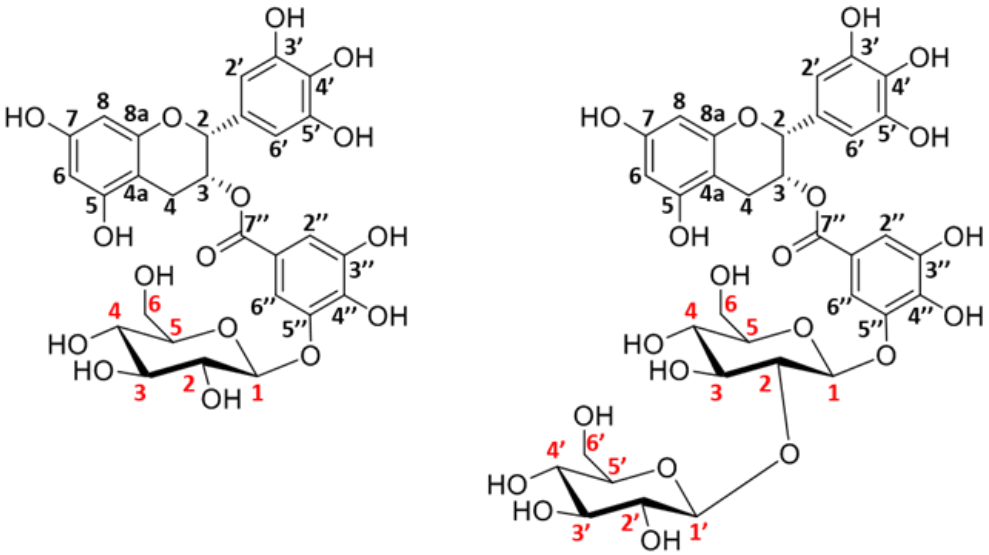


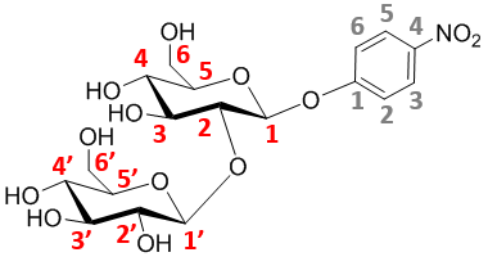
**Figure 5.** Structures, deduced from the NMR analysis, of the two glucosides produced by transglycosylation of EGCG with the synthase BGL-1-E521G. Every C atom in the molecules is numbered to clarify the identification of the signals.

**Table 6.** Chemical shifts for EGCG-glucose and EGCG-sophorose.

| EGCG-Glucose | | | EGCG-Sophorose | | |
|---|---|---|---|---|---|
| | $^{1}$H | $^{13}$C | | $^{1}$H | $^{13}$C |
| **2** | 5.08 | 77.06 | **2** | 5.03 | 77.05 |
| **3** | 5.60 | 68.89 | **3** | 5.57 | 68.71 |
| **4** | 2.86<br>3.01 | 24.91 | **4** | 2.84<br>2.97 | 24.86 |
| **4a** | --- | 99.04 | **4a** | --- | 99.00 |
| **5** | --- | 155.18 | **5** | --- | 155.34 |
| **6** | 6.08 | 95.72 | **6** | 6.06 | 96.01 |
| **7** | --- | 155.17 | **7** | --- | 155.34 |
| **8** | 6.08 | 95.72 | **8** | 6.06 | 96.01 |
| **8a** | --- | 155.18 | **8a** | --- | 155.21 |
| **1'** | --- | 106.33 | **1'** | --- | 106.39 |
| **2'** | 6.48 | 106.45 | **2'** | 6.47 | 106.40 |
| **3'** | --- | 145.42 | **3'** | --- | 145.51 |
| **4'** | --- | 132.10 | **4'** | --- | 132.62 |
| **5'** | --- | 145.44 | **5'** | --- | 145.65 |
| **6'** | 6.48 | 106.45 | **6'** | 6.47 | 106.40 |
| **1''** | --- | 120.59 | **1''** | --- | 120.58 |
| **2''/6''** | 7.09 | 108.61 | **2''/6''** | 7.06 | 108.61 |
| **3''** | --- | 144.30 | **3''** | --- | 144.59 |
| **4''** | --- | 139.73 | **4''** | --- | 139.90 |
| **5''** | --- | 144.27 | **5''** | --- | 144.59 |
| **6''/2''** | 7.06 | 112.27 | **6''/2''** | 7.05 | 112.18 |
| **7''** | --- | 166.50 | **7''** | --- | 166.43 |
| **1 Glc** | 4.97 | 100.33 | **1 Glc** | 5.09 | 99.22 |
| **2 Glc** | 3.53 | 72.60 | **2 Glc** | 3.78 | 80.98 |
| **3 Glc** | 3.54 | 75.32 | **3 Glc** | 3.69 | 75.39 |
| **4 Glc** | 3.49 | 68.65 | **4 Glc** | 3.53 | 68.29 |
| **5 Glc** | 3.22 | 75.97 | **5 Glc** | 3.21 | 75.63 |
| **6 Glc** | 3.41<br>3.66 | 59.73 | **6 Glc** | 3.41<br>3.66 | 59.83 |
| | | | **1' Glc** | 4.77 | 102.77 |
| | | | **2' Glc** | 3.25 | 73.85 |
| | | | **3' Glc** | 3.43 | 75.70 |
| | | | **4' Glc** | 3.36 | 69.36 |
| | | | **5' Glc** | 3.29 | 76.09 |
| | | | **6' Glc** | 3.25<br>3.53 | 60.33 |

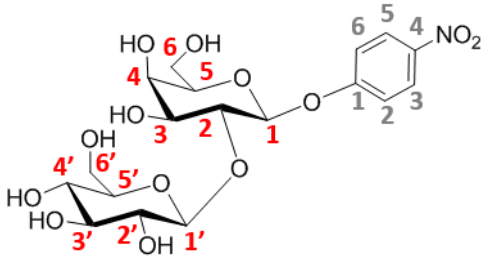
The results from the NMR analysis of the three *p*NP derivatives are shown in figures 6, 7 and 8 (NMR spectra can be seen in supplementary material figures S7, S8 and S9). All of them indicate that the synthase specifically forms *p*NP-disaccharides, incorporating the second sugar unit through a β-1,2 linkage. Finally, the different molecular species produced in a crude transglycosylation reaction mixture of α-GlcF (donor) and *p*NPG (acceptor) were also analysed by NMR. Interestingly, this sample showed significant heterogeneity as observed in the anomeric region of the $^{1}$H-$^{13}$C HSQC spectrum of the mixture (figure 9). The presence of some unreacted acceptor *p*NPG-glucose but not starting donor α-GlcF, was confirmed. Unexpectedly signals tentatively assigned to α-F-sophorose were observed

indicating that the α-GlcF itself with the α configuration could fit in the acceptor site. Besides, free glucose was also identified, confirming the auto-hydrolysis of the fluorinated substrate during the reaction. To note, this free glucose also worked as acceptor of α-GlcF, as deduced from the presence of sophorose among the reaction products. Again, the newly-synthetized disaccharides were linked by β-1,2 bonds, confirming the total regioselectivity of the synthase.



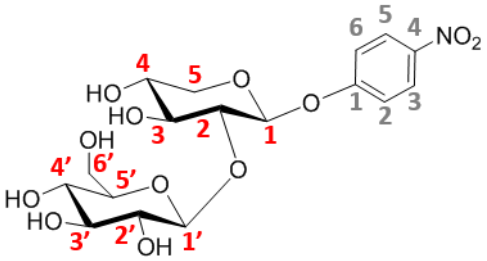

| | $^{1}$H | $^{13}$C | | $^{1}$H | $^{13}$C | | $^{1}$H | $^{13}$C |
|---|---|---|---|---|---|---|---|---|
| **1' Glu** | 4.77 | 102.83 | **1 Glu** | 5.41 | 97.90 | **1 pNP** | --- | 161.19 |
| **2' Glu** | 3.22 | 73.88 | **2 Glu** | 3.82 | 81.29 | **2 pNP** | 8.20 | 126.15 |
| **3' Glu** | 3.44 | 75.46 | **3 Glu** | 3.74 | 75.28 | **3 pNP** | 7.18 | 116.18 |
| **4' Glu** | 3.32 | 69.23 | **4 Glu** | 3.49 | 68.98 | **4 pNP** | --- | 142.52 |
| **5' Glu** | 3.27 | 76.03 | **5 Glu** | 3.63 | 76.13 | **5 pNP** | 7.18 | 116.18 |
| **6' Glu** | 3.26<br>3.48 | 60.20 | **6 Glu** | 3.68<br>3.86 | 60.71 | **6 pNP** | 8.20 | 126.15 |

**Figure 6.** Structure, deduced from NMR analysis, of the *p*NPG-glucoside produced by transglycosylation of *p*NPG with the synthase BGL-1-E521G. Every C atom in the molecules is numbered to clarify the identification of the signals. Chemical shifts (ppm) are indicated in the table.




| | $^{1}$H | $^{13}$C | | $^{1}$H | $^{13}$C | | $^{1}$H | $^{13}$C |
|---|---|---|---|---|---|---|---|---|
| **1' Glu** | 4.74 | 103.00 | **1 Gal** | 5.35 | 98.28 | **1 pNP** | --- | 161.43 |
| **2' Glu** | 3.20 | 73.68 | **2 Gal** | 3.98 | 79.60 | **2 pNP** | 8.20 | 126.05 |
| **3' Glu** | 3.45 | 75.42 | **3 Gal** | 3.90 | 72.22 | **3 pNP** | 7.19 | 116.18 |
| **4' Glu** | 3.30 | 69.00 | **4 Gal** | 3.97 | 68.02 | **4 pNP** | --- | 142.52 |
| **5' Glu** | 3.25 | 75.94 | **5 Gal** | 3.87 | 75.32 | **5 pNP** | 7.19 | 116.18 |
| **6' Glu** | 3.18<br>3.46 | 59.83 | **6 Gal** | 3.70<br>3.70 | 60.73 | **6 pNP** | 8.20 | 126.05 |

**Figure 7.** Structure, deduced from NMR analysis, of the *p*NPGal-glucose product obtained by transglycosylation of *p*NPGal with the synthase BGL-1-E521G. Every C atom in the molecule is numbered to clarify the identification of the signals. Chemical shifts (ppm) are indicated in the table.



| | $^{1}$H | $^{13}$C | | $^{1}$H | $^{13}$C | | $^{1}$H | $^{13}$C |
|---|---|---|---|---|---|---|---|---|
| **1' Glu** | 4.74 | 102.91 | **1 Xyl** | 5.40 | 98.39 | **1 pNP** | --- | 161.22 |
| **2' Glu** | 3.22 | 73.65 | **2 Xyl** | 3.82 | 80.95 | **2 pNP** | 8.20 | 126.06 |
| **3' Glu** | 3.44 | 75.36 | **3 Xyl** | 3.72 | 74.44 | **3 pNP** | 7.17 | 116.13 |
| **4' Glu** | 3.32 | 69.15 | **4 Xyl** | 3.71 | 68.71 | **4 pNP** | --- | 142.56 |
| **5' Glu** | 3.28 | 75.89 | **5 Xyl** | 3.49<br>3.98 | 64.94 | **5 pNP** | 7.17 | 116.13 |
| **6' Glu** | 3.25<br>3.48 | 60.01 | | | | **6 pNP** | 8.20 | 126.06 |

**Figure 8.** Structure, deduced from NMR analysis, of the *p*NPX-glucose product obtained by transglycosylation of *p*NPX with the synthase BGL-1-E521G. Every C atom in the molecule is numbered to clarify the identification of the signals. Chemical shifts (ppm) are indicated in the table.

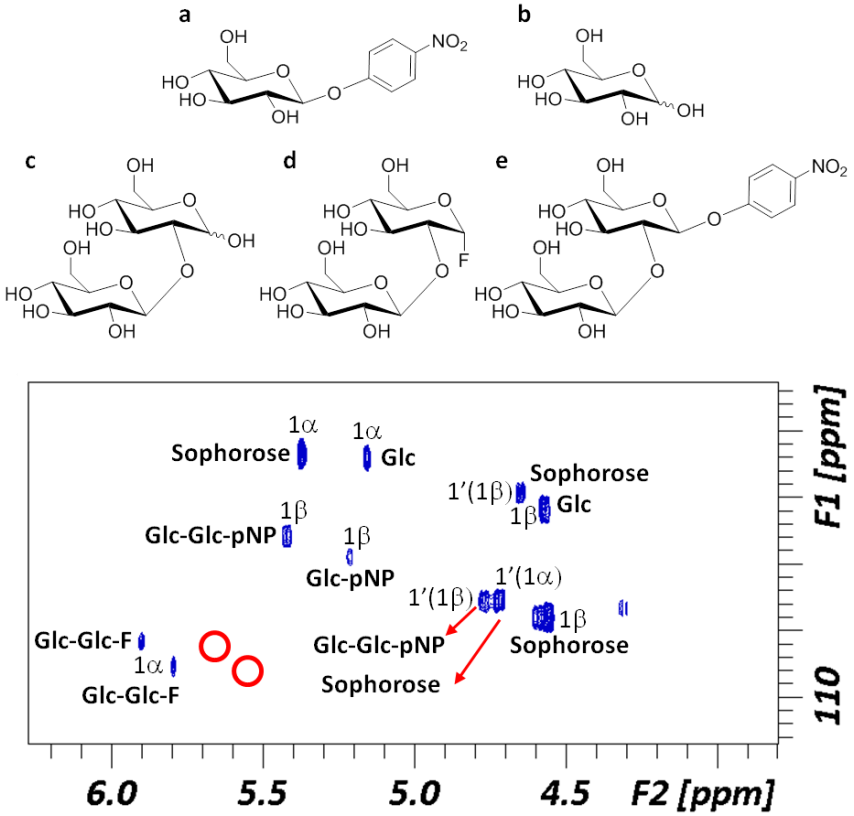



**Figure 9.** 1) Structures of *p*NPG (a), Glc (b), sophorose (c), sophorose-F (d) and sophorose-*p*NP (e). 2) Anomeric region of the $^{1}$H-$^{13}$C HSQC spectrum of a transglycosylation reaction mixture which pNPG as acceptor and α-Glc-F as donor catalyzed by BGL-1-E521G.The peaks are labeled with their corresponding assignment. Red circles indicate where the signals corresponding α-Glc-F, which has been fully consumed, should appear.

## CONCLUSIONS

The novel GH1 β-glucosidase BGL-1 from the ascomycete *T. amestolkiae* has been heterologously produced in *P. pastoris* in large amounts. The recombinant enzyme was easily purified in good yield (80%) and then characterized and evaluated in hydrolysis and transglycosylation reactions. The inhibition constant by glucose of BGL-1 was 3.78 M, confirming its outstanding glucotolerance. However, its low catalytic efficiency in hydrolysis of cellooligosaccharides may limit its applications in saccharification. BGL-1 was also tested as catalyst of the transglycosylation of a wide array of compounds showing poor efficiency in synthesis. To overcome this problem, the enzyme was converted into a glucosynthase by protein engineering, abolishing its hydrolytic activity. This novel biocatalyst showed to be regioselective, and catalyzed the transfer of glucose molecules not only to *p*-nitrophenol sugars, but also to interesting phenolic acceptors as EGCG, a compound with many potential medical applications. These findings postulate this mutant as a potential candidate to be used in industrial processes devoted to the synthesis of bioactive glycosides.

## MATERIALS AND METHODS

### Microorganism and culture conditions

The fungus *T. amestolkiae* A795, deposited in the IJFM collection, at the Centro de Investigaciones Biológicas (Madrid, Spain), was cultured in PDA (potato dextrose agar) plates at 28 °C and maintained at 4 ºC. *T. amestolkiae* was also cultured in Mandels medium, as reported before (Méndez-Líter *et al*., 2018), with 1% of Avicel as carbon source, for RNA extraction.

For plasmid propagation, *Escherichia coli* DH5α (Invitrogen) was grown at 37 ºC, overnight, in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, and 15 g/L agar) containing 100 mg/L ampicillin for selection of resistant colonies.

The heterologous expression of BGL-1 was performed using *P. pastoris* X33 strain (Invitrogen), previously grown in YPD medium plates (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose and 10 g/L of agar). The positive clones were screened in YPD containing 100 mg/L of zeocin as selection marker, and cultured 2-5 days at 28 ºC. Recombinant protein was produced in YEPS medium (20 g/L peptone, 10 g/L yeast extract, 10 g/L sorbitol, and 100 mM potassium phosphate buffer, pH 6), with daily addition of 10 mL/L of methanol as inducer. Cultures were incubated for 9 days, at 28 °C and 250 rpm, taking samples daily to measure BGL-1 production. All experiments were performed in triplicate.

**Nucleic acid isolation, enzyme mutagenesis and cloning in *Pichia pastoris***

In order to identify the bgl-1 DNA sequence, a TBLASTN was performed against the set of predicted proteins of T. amestolkiae, which were obtained in a previous work (de Eugenio *et al*., 2017). The gene sequences returned were used to run a local BLASTN against the assembled genome (de Eugenio *et al*., 2017). Once recognized the predicted sequence for *bgl-1*, the presence of a signal peptide in the BGL-1 protein sequence was examined using the SignalP server. RNA was extracted from 7-day old *T. amestolkiae* cultures growing in 1% of Avicel using Trizol reagent (Chomczynski and Sacchi, 1987). The isolated transcripts were converted to cDNA using the Superscript II Reverse Transcriptase RT-PCR kit, according to the manufacturer's instructions. PCR amplifications were performed in a thermocycler Mastercycler pro S (Eppendorf). The primers were designed based on the nucleotide sequence of the *bgl-1* gene identified in *T. amestolkiae* genome (GenBank accession no. MIKG00000000), excluding the region corresponding to the signal peptide. Restriction sites for *XhoI* and *NotI* were included in the forward and reverse primers respectively (BG1FWXHOI: 5′-ATCTCGAGAAAAGACAAGAGGTGTACATCACGACT-3′, and BG1RVNOTI: 5′- ATGCGGCCGCATATCCCAGCCCATTCCTCGC -3′). The PCR protocol was developed as follows: initial denaturation at 95 °C for 5 min, followed by 36 cycles of amplification: denaturation at 95 °C for 45 s, primer annealing at 55 °C for 45 s, and elongation at 72 °C for 2 min. A final extension step at 72 °C for 10 min were also carried out. The PCR product obtained in the last step was ligated to the yeast expression vector pPICzα (Invitrogen), and it was used to transform *P. pastoris* X-33. For improving transformation efficiency, the vector must be linearized with *SacI* (New England Biolabs). The lithium chloride method was used for transformation according to the manufacturer's instructions. Transformed colonies were grown on YPD medium plates with zeocin as selection marker. Positive clones were screened with 4-methylumbelliferyl β-D-glucopyranoside as described in Méndez-Líter *et al*., (2017).

**Conversion of BGL-1 into the two glycosynthase variants**

The plasmid pPICzα containing *bgl-1* gene was used to generate two new versions of the protein by directed mutagenesis, replacing the glutamic acid 521 by a glycine or a serine. The identification of the catalytic amino acids of BGL-1, was performed by alignment using clustal omega (see supplementary material, figure S4), with the BGLs sequences of bacterium *Clostridium cellulovorans*, fungus *Trichoderma reesei* and termite *Neotermes koshunensis* (Jeng *et al*., 2011). To do so, the Expand™ Long

Template PCR System (Roche) was used as described by the manufacturer. Primers BG1sfwSer (CCCTCGTCCTCAGCT CATTCGGTTTTCCCGTCTAC), BG1sRvSer (GTAGACGGGAAAACCGAATGAGCTGAGGACGAGGG), BG1sfwgly (CCCTCGTCCTCAGCGGATTCGGTTTTCCCGTCTAC) and BG1sRvgly (GTAGACGGGAAAACCGA ATCCGCTGAGGACGAGGG) were used for serine and glycine replacements, respectively. After PCR reaction, the product was digested by *DpnI* (New England Biolabs), in order to hydrolyze the parental methylated DNA that was used as template. Both new vectors were cloned into *P. pastoris* with the same method used before.

**Production and purification of BGL-1 and BGL-1 glycosynthase variants**

The selected positive *P. pastoris* clones were grown overnight in 250 mL flasks with 50 mL of YPD medium at 28 °C and 250 rpm to obtain the respective preinocula. Then, they were used for recombinant protein production in 2-L flasks with 400 mL of YEPS medium. Cultures were incubated at 28 °C and 250 rpm for 9 days with daily addition of 10 mL/L methanol.

For BGL-1 purification, 9 day-old cultures were harvested and centrifuged at 10,000 × *g* and 4 °C for 20 min. The supernatant was concentrated and dialyzed against 10 mM phosphate buffer (pH 6.0) using a 30-kDa cutoff membrane (Merck-Millipore). BGL-1 was purified in a single chromatographic step using an FPLC system (Äkta), using a 5 mL QFF HiTrap cartridge (GE Healthcare) equilibrated with phosphate buffer pH 6.0. Elution of the bound proteins was carried out by applying a 25 min-linear gradient from 0 to 0.3 M of NaCl, at 2 mL/min. The column was then washed with 10 mL of 1 M NaCl and re-equilibrated using 10 mL of the starting buffer. Fractions with β-glucosidase activity were dialyzed and concentrated. The glycosynthase variants of BGL-1 were purified using the same protocol.

**Protein quantification, enzyme assays and substrate specificity**

Total protein was calculated by 280 nm absorbance using a Nanodrop spectrophotometer (Thermo Fisher Scientific), and confirmed by the bicinchoninic acid assay (BCA) method.

The β-glucosidase standard reaction was performed using 3 mM (w/v) *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG, Sigma), at 60 °C, in sodium acetate buffer 100 mM, pH 4.0. The reaction was stopped after 10 min by adding 2% (w/v) $Na_2CO_3$, measuring the release of *p*-nitrophenol (*p*NP) in a spectrophotometer at 410 nm. One BGL activity unit was

defined as the amount of enzyme capable of releasing 1 micromole of *p*NP per min (the molar extinction coefficient of *p*NP is 15,200 $M^{-1} \cdot cm^{-1}$).

Glucose tolerance was determined in standard BGL reactions containing increasing glucose concentrations from 0.1 mM to 3 M. Glucotolerance was calculated comparing the activity measured in reactions without glucose, that was considered as the 100%, with those obtained for reactions with added glucose. For $K_i$ determinations, the concentrations of glucose used were 1, 1.25, and 1.5 M.

To prevent the activity loss when working with low enzyme concentrations, all enzymatic assays included 0.1% BSA, a protein that does not affect the catalytic activity of BGL-1 (Nieto Domínguez *et al*., 2015). The kinetic constants of the purified BGL-1 were determined against *p*NPG over a range of concentrations from 10 μM to 5 mM, *o*-nitrophenyl-β-D-glucopyranoside (*o*NPG, 40 μM to 20 mM), cellobiose (80 μM to 40 mM), cellotriose (80 μM to 40 mM), cellotetraose (80 μM 40 mM), cellopentaose (40 μM to 20 mM), and cellohexaose (20 μM to 10 mM). The parameters $K_m$ and $V_{max}$ were calculated using SigmaPlot. The reactions containing cellooligosaccharides were quantified by measuring the glucose released after hydrolysis, using the Glucose-TR commercial kit (Spinreact) according to the manufacturer's instructions. All reactions were carried out in sodium acetate 100 mM, pH 4.0, in a heating block for 10 min at 1200 rpm. Then, the reactions were stopped by heating at 100 ºC for 5 min.

Kinetic parameters were calculated for two transglycosylation experiments catalyzed by the BGL-1-E521G variant with D-glucosyl fluoride (α-GlcF) and *p*NPG as substrates. Each substrate was used in one experiment at a fixed concentration of 10 mM, and with varying concentrations in the other. When *p*NPG was examined, it concentrations ranged between 500 mM and 12.5 mM. When calculating kinetic constants α-GlcF, it was used in a range from 25 mM to 1 M.

BGL-1 activity towards cellobiose (β-D-glucopyranosyl-1,4-β-D-glucopyranose), sophorose (β-D-glucopyranosyl-1,2-β-D-glucopyranose), laminaribiose (β-D-glucopyranosyl-1,3-β-D-glucopyranose) and gentiobiose (β-D-glucopyranosyl-1,6-β-D-glucopyranose), was determined using 10 mM of the disaccharides in sodium acetate buffer 100 mM, pH 4.0, with the amount of enzyme conveniently diluted. Reactions were performed for 10 min at 60 °C and 1200 rpm, and quantified by measuring the glucose released using the Glucose-TR commercial kit (Spinreact).

**Physicochemical properties**

The molecular mass of the native BGL-1 was determined using MALDI-TOF according to the protocol explained in Méndez-Líter *et al*., (2018). Isoelectric point (pI) was determined by isoelectrofocusing (IEF),

and the gel was revealed with 4-methylumbelliferyl-β-D-glucopyranoside (Sigma Aldrich), following the same procedure explained in Méndez-Líter *et al*., (2017). Optimal pH was assayed for 10 min using *p*NPG as substrate and Britton-Robinson buffer (100 mM) in a range from 2 to 10. Optimal temperature was assayed using standard conditions but varying the temperature from 30-70 °C.

**Screening for transglycosylation acceptors of BGL-1**

In the search of potential transglycosylation acceptors for the wild type BGL-1, a library of 70 compounds (Supplementary material, Figure S2), was used to conduct a preliminary screening. The compounds were used to perform recovery inhibition assays as described in a previous work (Méndez-Líter *et al*., 2019), and those that produced higher absorbance reads than the controls without acceptor were considered potential hits for transglycosylation.

**Transglycosylation reactions catalyzed by the glycosynthases. Analysis of the products**

In these reactions, α-GlcF was prepared as described previously (Wadzinski *et al*., 2018), and used as donor in every reaction. The efficiency of both glycosynthase variants (BGL-1-E521G and BGL-1-E521S) was first compared in reactions containing 10 mM *p*NPG, 10 mM α-GlcF, 0.4 mg/mL of the corresponding mutant enzyme, and 50 mM acetate buffer pH 4. The reaction mixture was incubated at room temperature for 16 h at 500 rpm, and then analyzed by High Performance Liquid chromatography (HPLC) as explained below.

Other potential acceptors tested were: *p*NPG, *p*-nitrophenyl-β-D-xylopyranoside (*p*NPX), *p*-nitrophenyl-β-D-galactopyranoside (*p*NPGal), vanillin, hydroxytyrosol, gallic acid, and epigallocatechin gallate (EGCG). The standard transglycosylation reaction contained 20 mg/mL of α-GlcF, 5 mg/mL of each acceptor and 1 mg/ mL of BGL-1-E521G in acetate buffer 50 mM, pH 4, with 0.1% of BSA, and it was performed at room temperature for 16 h at 500 rpm. The synthesis of glycosides was first checked by thin layer chromatography (TLC) in silica gel G/UV254 polyester sheets, (0.2 mm thickness and 40 × 80 mm plate size) provided by Macherey–Nagel, using ethyl acetate/methanol/water 10:2:1 (v/v) as running solution. Substrates and glucosides were detected under 254 nm UV light, since all the potential acceptors possess an aromatic ring detectable at this wavelength.

The reaction mixtures were also analyzed by mass spectrometry in a HCT Ultra ion trap, in order to identify the expected products. The samples were analyzed by electrospray ionization-mass spectrometry (ESI-MS) with methanol as ionizing phase in the positive reflector mode, and data

were processed with the Masshunter Data Acquisition B.05.01 and Masshunter Qualitative Analysis B.07.00 software (Agilent Technologies).

Finally, the most interesting glycosides were purified by HPLC, in an Agilent 1200 series LC instrument equipped with a ZORBAX Eclipse plus C18 column (Agilent). The column was first equilibrated in a mix of acetonitrile and H2O with 0.1% acetic acid, with a flow of 2 mL/min, and the reaction products were separated isocratically in 8 min. For the glycosides of pNPG and pNPGal, they were purified with a proportion of 14:86 (v/v) acetonitrile:$H_2O$, while for the pNPX glycoside this proportion changed to 20:80 (v/v) and for the EGCG products it was 13:87 (v/v). After isocratic elution, the column was washed for 3 min with 95:5 acetonitrile:$H_2O$, and the system was finally re-equilibrated to the initial conditions for 4 min. Every product peak was detected by monitoring the absorbance at 270 nm, and quantification was done using a calibration curve of each non glycosylated parental. The fractions containing the glycosides were collected to be further analyzed by nuclear magnetic resonance (NMR) to determine their structure. After collection, the products were lyophilized and stored at -20 °C. The reactions conducted to determine the kinetic parameters of BGL-1-E521G were analyzed isocratically in the equilibration buffer, as described above for the glycosides of pNPG and pNPGal.

**Nuclear Magnetic Resonance**

The structure and regiochemistry of the purified glucosides of EGCG, *p*NPG, *p*NPX and *p*NPGal synthesized by BGL-1-E521G was elucidated by NMR. A transglycosylation reaction using BGL-1-E521G, with α-GlcF as donor and *p*NPG as acceptor, was also analyzed. Besides, a reaction using the native BGL-1, with pNPG as donor and 13C-labelled glucose as acceptor, was also studied to determine the regioselectivity of this β-glucosidase in transglycosylation. The samples were prepared in 500 μL of deuterated water (D2O). NMR spectra were acquired at 298 K, using a Bruker AVANCE 600 MHz spectrometer equipped with a cryogenic probe. 1D $^{1}H$ NMR spectra, $^{1}H$-$^{13}C$ HSQC and HMBC experiments were acquired to assign all NMR signals. For 1D $^{1}H$, $^{1}H$-$^{13}C$ HSQC, and HMBC experiments, the zg, zgpr, hsqcedetgp, and hmbcgpndqf sequences were employed. The analysis of the reaction mixture from transglycosylation with BGL-1 was performed just considering the anomeric region of the $^{1}H$-$^{13}C$ HSQC, as this sample was more heterogeneous because it contained substrates and products.

## Declarations

### Competing interests

The authors declare that they have no competing interests.

### Funding

This work has been funded by Projects RTI2018-093683-B-I00 from MCIU/AEI/FEDER, UE and S2018/EMT-4459 from Comunidad de Madrid. J.A. Méndez-Líter thanks the Tatiana Pérez de Guzmán el Bueno Foundation for his scholarship.

### Author contributions

JAML drafted the manuscript and performed the experiments excluding NMR assays. BFT and FJC designed and performed the NMR assays and analyzed the data. AGS and JLA, synthesized the α-D-glucosyl-fluoride used as donor in the glycosynthase reactions, and the 2,4-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucopyranoside used for the screening of potential transglycosylation acceptors. MND, AP, LIE and MJM coordinated the study, helped to draft, and critically reviewed the manuscript. All authors read and approved the final manuscript.

### Availability of supporting data

*T. amestolkiae* whole genome shotgun project is deposited at DDBJ/ENA/GenBank under the accession number MIKG00000000. BGL-1 sequence is deposited in GenBank under the accession number KM393204.1.

### Competing financial interests

The authors declare no competing financial interests.

## REFERENCES

Badhani, B., Sharma, N., Kakkar, R., 2015. Gallic acid: A versatile antioxidant with promising therapeutic and industrial applications. RSC Adv. 5, 27540-27557.

Blanchard, J.E., Withers, S.G., 2001. Rapid screening of the aglycone specificity of glycosidases: applications to enzymatic synthesis of oligosaccharides. Chem. Biol. 8, 627-633.

Cao, L.C., Wang, Z.J., Ren, G.H., Kong, W., Li, L., Xie, W. *et al*., 2015. Engineering a novel glucose-tolerant β−glucosidase as supplementation to enhance the hydrolysis of sugarcane bagasse at high glucose concentration. Biotechnol. Biofuels 8, 202-202.

Cereghino, J.L., Cregg, J.M., 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS microbiol. rev. 24, 45-66.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156-159.

Dai, J., Mumper, R.J., 2010. Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. Molecules 15, 7313-7352.

Danby, P.M., Withers, S.G., 2016. Advances in enzymatic glycoside synthesis. ACS Chem. Biol. 11, 1784-1794.

de Eugenio, L.I., Méndez-Líter, J.A., Nieto-Domínguez, M., Alonso, L., Gil-Muñoz, J., Barriuso, J. *et al*., 2017. Differential β−glucosidase expression as a function of carbon source availability in *Talaromyces amestolkiae*: a genomic and proteomic approach. Biotechnol. Biofuels 10, 161-161.

Fabiani, R., Sepporta, M.V., Rosignoli, P., De Bartolomeo, A., Crescimanno, M., Morozzi, G., 2012. Anti-proliferative and pro-apoptotic activities of hydroxytyrosol on different tumour cells: the role of extracellular production of hydrogen peroxide. Eur. J. Nutr. 51, 455-464.

Gonzalez-Alfonso, J.L., Leemans, L., Poveda, A., Jimenez-Barbero, J., Ballesteros, A.O., Plou, F.J., 2018. Efficient α−glucosylation of epigallocatechin gallate catalyzed by cyclodextrin glucanotransferase from *Thermoanaerobacter* species. J. Agric. Food Chem. 66, 7402-7408.

Gonzalez-Alfonso, J.L., Peñalver, P., Ballesteros, A.O., Morales, J.C., Plou, F.J., 2019. Effect of α−glucosylation on the stability, antioxidant properties, toxicity, and neuroprotective activity of (-)-epigallocatechin gallate. Front. Nutr. 6, 30-30.

Hayes, M.R., Pietruszka, J. 2017. Synthesis of glycosides by glycosynthases. Molecules 22, 1434.

Jahn, M., Stoll, D., Warren, R., Szabo, L., Singh, P., Gilbert, H. *et al*., 2003. Expansion of the glycosynthase repertoire to produce defined manno-oligosaccharides. Chem. Comm. 9, 1327-1329.

Jeng, W.Y., Wang, N.C., Lin, M.H., Lin, C.T., Liaw, Y.C., Chang, W.J. *et al*., 2011. Structural and functional analysis of three β−glucosidases from bacterium *Clostridium cellulovorans*, fungus *Trichoderma reesei* and termite *Neotermes koshunensis*. J. Struct. Biol. 173, 46-56.

Kapoor, S., 2013. Multiorgan anticarcinogenic effects of vanillin. J. Med. Food. 16, 777-777.

Karnaouri, A., Topakas, E., Paschos, T., Taouki, I., Christakopoulos, P., 2013. Cloning, expression and characterization of an ethanol tolerant GH3 β−glucosidase from *Myceliophthora thermophila*. PeerJ 1, e46-e46.

Kim, J., Nguyen, T.T.H., Kim, N.M., Moon, Y.H., Ha, J.M., Park, N. *et al*., 2016. Functional properties of novel epigallocatechin gallate glucosides synthesized by using dextransucrase from *Leuconostoc mesenteroides* B-1299CB4. J. Agric. Food Chem. 64, 9203-9213.

Kim, Y.W., Chen, H., Withers, S.G., 2005. Enzymatic transglycosylation of xylose using a glycosynthase. Carbohydr. Res. 340, 2735-2741.

Kitahata, S., Edagawa, S., 1987. Cyclic (1,2)-β−D-glucan-hydrolyzing enzymes from *Acremonium sp*.: Purification and some properties of endo-(1,2)-β−D-glucanase and β−D-glucosidase. Agric. Biol. Chem. 51, 2701-2708.

Li, G., Jiang, Y., Fan, X.J., Liu, Y.H., 2012. Molecular cloning and characterization of a novel β−glucosidase with high hydrolyzing ability for soybean isoflavone glycosides and glucose-tolerance from soil metagenomic library. Bioresour. Technol. 123, 15-22.

Li, X., Xia, W., Bai, Y., Ma, R., Yang, H., Luo, H. *et al*., 2018. A novel thermostable GH3 β−glucosidase from *Talaromyces leycettanus* with broad substrate specificity and significant soybean isoflavone glycosides-hydrolyzing capability. Biomed. Res. Int. 4794690-4794690.

Liu, D., Zhang, R., Yang, X., Zhang, Z., Song, S., Miao, Y. *et al*., 2012. Characterization of a thermostable β−glucosidase from *Aspergillus fumigatus* Z5, and its functional expression in *Pichia pastoris* X33. Microb. Cell Fact. 11, 25-25.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. D490-D495.

Mackenzie, L.F., Wang, Q., Warren, R.A., Withers, S.G., 1998. Glycosynthases: Mutant glycosidases for oligosaccharide synthesis. J. Am. Chem. Soc. 120, 5583-5584.

Martínez, A.T., 2016. How to break down crystalline cellulose. Science 352, 1050.

Martínez, Á.T., Speranza, M., Ruiz-Dueñas, F.J., Ferreira, P., Camarero, S., Guillén, F. *et al*., 2005. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int. Microb. 8, 194-205.

Mayer, C., Jakeman, D.L., Mah, M., Karjala, G., Gal, L., Warren, R.A.J. *et al*., 2001. Directed evolution of new glycosynthases from *Agrobacterium* β–glucosidase: a general screen to detect enzymes for oligosaccharide synthesis. Chem. Biol. 8, 437-443.

Méndez-Líter, J.A., Gil-Muñoz, J., Nieto-Domínguez, M., Barriuso, J., de Eugenio, L.I., Martínez, M.J., 2017. A novel, highly efficient β–glucosidase with a cellulose-binding domain: characterization and properties of native and recombinant proteins. Biotechnol. Biofuels. 10, 256.

Méndez-Líter, J.A., de Eugenio, L.I., Prieto, A., Martínez, M.J., 2018. The β–glucosidase secreted by *Talaromyces amestolkiae* under carbon starvation: a versatile catalyst for biofuel production from plant and algal biomass. Biotechnol. Biofuels. 11, 123-123.

Méndez-Líter, J.A., Tundidor, I., Nieto-Domínguez, M., de Toro, B.F., González Santana, A., de Eugenio, L.I. *et al*., 2019. Transglycosylation products generated by *Talaromyces amestolkiae* GH3 β–glucosidases: effect of hydroxytyrosol, vanillin and its glucosides on breast cancer cells. Microb. Cell Fact. 18, 97.

Moracci, M., Trincone, A., Perugino, G., Ciaramella, M., Rossi, M., 1998. Restoration of the activity of active-site mutants of the hyperthermophilic β–glycosidase from *Sulfolobus solfataricus*: Dependence of the mechanism on the action of external nucleophiles. Biochemistry. 37, 17262-17270.

Nagle, D.G., Ferreira, D., Zhou, Y.D., 2006. Epigallocatechin-3-gallate (EGCG): chemical and biomedical perspectives. Phytochemistry. 1849-1855.

Nakajima, M., Yoshida, R., Miyanaga, A., Abe, K., Takahashi, Y., Sugimoto, N. *et al*., 2016. Functional and structural analysis of a β–glucosidase involved in β–1,2-glucan metabolism in *Listeria innocua*. PLoS ONE 11, e0148870.

Nieto Domínguez, M., Eugenio, L., Barriuso, J., Prieto, A., Fernández de Toro Ronda, B., Canales-Mayordomo, Á. *et al*., 2015. Novel pH-stable glycoside hydrolase family 3 β–xylosidase from *Talaromyces amestolkiae*: an enzyme displaying regioselective transxylosylation. Appl. Environ. Microbiol. 81(18): 6380-6392.

Pei, J., Pang, Q., Zhao, L., Fan, S., Shi, H., 2012. *Thermoanaerobacterium thermosaccharolyticum* β–glucosidase: a glucose-tolerant enzyme with high specific activity for cellobiose. Biotechnol. Biofuels. 5, 31-31.

Perugino, G., Trincone, A., Giordano, A., van der Oost, J., Kaper, T., Rossi, M. *et al*., 2003. Activity of hyperthermophilic glycosynthases is significantly enhanced at acidic pH. Biochemistry 42, 8484-8493.

Perugino, G., Trincone, A., Rossi, M., Moracci, M., 2004. Oligosaccharide synthesis by glycosynthases. Trends Biotechnol. 22, 31-37.

Ramani, G., Meera, B., Vanitha, C., Rajendhran, J., Gunasekaran, P., 2015. Molecular cloning and expression of thermostable glucose-tolerant β–glucosidase of *Penicillium*

*funiculosum* NCL1 in *Pichia pastoris* and its characterization. J. Ind. Microbiol. Biotechnol. 42, 553-565.

Saumonneau, A., Champion, E., Peltier-Pain, P., Molnar-Gabor, D., Hendrickx, J., Tran, V. *et al*., 2015. Design of an α−L-transfucosidase for the synthesis of fucosylated HMOs. Glycobiology 26, 261-269.

Singh, B.N., Shankar, S., Srivastava, R.K., 2011. Green tea catechin, epigallocatechin-3-gallate (EGCG): mechanisms, perspectives and clinical applications. Biochem Pharmacol. 1807-1821.

Singhania, R.R., Patel, A.K., Sukumaran, R.K., Larroche, C., Pandey, A., 2013. Role and significance of beta−glucosidases in the hydrolysis of cellulose for bioethanol production. Bioresour. Technol. 127, 500-507.

Sørensen, A., Lübeck, M., Lübeck, P.S., Ahring, B.K., 2013. Fungal Beta−glucosidases: a bottleneck in industrial use of lignocellulosic materials. Biomolecules 3, 612-631.

Sternberg, D., Mandels, G.R., 1980. Regulation of the cellulolytic system in *Trichoderma reesei* by sophorose: induction of cellulase and repression of beta−glucosidase. J. Bacteriol. 144, 1197-1199.

Uchima, C.A., Tokuda, G., Watanabe, H., Kitamoto, K., Arioka, M., 2011. Heterologous expression and characterization of a glucose-stimulated β−glucosidase from the termite *Neotermes koshunensis* in *Aspergillus oryzae*. Appl. Microbiol. Biotechnol. 89, 1761-1771.

V., D., Verma, P., Singh, D., 2015. Emerging technologies for improving bioavailability of polyphenols. Curr. Nutr. Food Sci. 11, 1-1.

Wadzinski, T.J., Steinauer, A., Hie, L., Pelletier, G., Schepartz, A., Miller, S.J., 2018. Rapid phenolic O-glycosylation of small molecules and complex unprotected peptides in aqueous solvent. Nat. Chem. 10, 644-652.

Xia, W., Xu, X., Qian, L., Shi, P., Bai, Y., Luo, H. *et al*., 2016. Engineering a highly active thermophilic β−glucosidase to enhance its pH stability and saccharification performance. Biotechnol. Biofuels. 9, 147.

Xu, L., Qi, T., Xu, L., Lu, L., Xiao, M., 2016. Recent progress in the enzymatic glycosylation of phenolic compounds. J. Carbohydr. Chem. 35, 1-23.

Yan, Q., Hua, C., Yang, S., Li, Y., Jiang, Z., 2012. High level expression of extracellular secretion of a β−glucosidase gene (PtBglu3) from *Paecilomyces thermophila* in *Pichia pastoris*. Protein Expr. Purif. 84, 64-72.

Yang, F., Yang, X., Li, Z., Du, C., Wang, J., Li, S., 2015. Overexpression and characterization of a glucose-tolerant β−glucosidase from *T. aotearoense* with high specific activity for cellobiose. Appl. Microbiol. Biotechnol. 99, 8903-8915.

Yang, X., Ma, R., Shi, P., Huang, H., Bai, Y., Wang, Y. *et al*., 2014. Molecular characterization of a highly-active thermophilic β−glucosidase from *Neosartorya fischeri* P1 and its application in the hydrolysis of soybean isoflavone glycosides. PLoS ONE 9, e106785.

Yin, Y.R., Sang, P., Xian, W.D., Li, X., Jiao, J.Y., Liu, L. *et al*., 2018. Expression and characteristics of two glucose-tolerant GH1 β−glucosidases from *Actinomadura amylolytica* YIM 77502T for promoting cellulose degradation. Front. Microbiol. 9: 3149.

Zhao, L., Xie, J., Zhang, X., Cao, F., Pei, J., 2013. Overexpression and characterization of a glucose-tolerant β−glucosidase from *Thermotoga thermarum* DSM 5069T with high catalytic efficiency of ginsenoside Rb1 to Rd. J. Mol. Catal. B: Enz. 95, 62-69.

Zou, G., Jiang, Y., Liu, R., Zhu, Z., Zhou, Z., 2018. The putative β−glucosidase BGL3I regulates cellulase induction in *Trichoderma reesei*. Biotechnol. Biofuels. 11, 314.

# SUPPLEMENTARY INFORMATION

A

**ATG**GGATTTCTCTTTCTATTAGCGGCTGTCGCCGTTGTAGTACAGGCT**CAA**GAGGTGTACATCACGACTACTGGCTACACGGCGCGGCCACAATGCACTGAACCCCCAGCTACGCCAACTTTTC
GCTTCCAGTCGTTTTCATATGCATCGCTCAACGATACAATTCGGTGTGTTGTTCTCTTCTGTCGTACAATTTCCAATTAACACTCGCAGCTATGCAATCTCAGTCCCCTCTCCAACAACCACCC
ACACCTACGGACCAGCATACACAGATGCAGTAGCAAAACTGTCGACGAAGCTAACGACCACGACGTGGGGTAGTTGGGTACCAAGTCAGACCGTCATCTCTGCAACCGATACAGCGGACAAATA
CGGCCAAGCCGCGTGGTCATCGCAATGGCTACATGCAAGCTTGGCGAATTATACCAATATTGGTTTGTATACGACCACCGTCAATCCTACCCCATTACCTTCCAGTGAACTTGTTCTTCCACCA
CGAGACTATTTCGGTCCAACGGATTGCTACAACTTTCCGGAAGGATTCACGTTCGGTGTTGCCGGATCTGCTGCTCAGATCGAAGGTGCTATTGGCCTCGAAGGTCGTGCGCCGAGTATACTGG
AGAAATTGCTGCCAGATACAGAACCGCAGGATTACGTGACAAATGAGAACTACTACCTGTATAAGCAGGACATTCAGCGTCTTGCATCGATTGGCGTCAAGTATTACAGTTTCTCAATCTCGTG
GGGCCGCATTCTACCGTTCACCGTGCCCGGAAGCCCAATCAATGAACAAGGCCTCAAGCATTACAATGACTTGATCGATTATGTTTTGGAAGTTGGCATGGTACCTATTGTTACGATGCTCCAC
TTCGATACTCCACTGTATTTCATCAATGCCCTCTGCTGGTTACGTGGTGCCGGATATTGGATACCAGAATGGCGGGTATTGGAATAAAGAGTTTGTCGATTCGTTTGTGAATTACGGCAAGATCT
TGTTTACGCATTTTGCCGATCGTGTCCCGTTCTGGGTTACTATTAATGAGCCCCTGCTGTATGCGTTTAATTTCACAGGACTGGATAATGTTGTGCATGCCCACGCAGAACTGTATCACTTCTA
TCATGATACTTTGAATGGAACTGGAAAGGTGGGATTGAAATTGAACGACAATTTTGGCGTGCCCAAGCACCCTGAGAACCAGACCGAGATCGATGCCGCCAACCGTTTCAATGACATGCAGCTC
GGCGTCTTTGCGTATCCGATTTGTCTCGGCCAGCAATATCCTAAATCTATTCTCGACACTCTGCCAGGCGCCAAACCACTCAGTAAAAAGGAATTGGAGTACATCAGCCATACAACCGACTTCA
TCGGCATCGACGCATATACAGCCACCGTCATCTCCGTTCCCGCCGAAGGCATCGAATATTGCTCCAAACAGAACATGACCACCAACTCCTTGTATCCGTACTGCGTCACCCAAGAAACCGTAAA
TTCATATGGCTGGGATATTGGCTACCGCTCTCAATCCTACGTCTACATCACACCCACCTACCTCCGCGCATTCCTGTCTTATATCTGGAACACCTACAAAACGCCCCTCGTCCTCAGCGAATTC
GGTTTTCCCGTCTACGCTGAATCGACTCGCGATTTAGTCGACCAACTGTATGATTCACCCCGAAGTCAATATTATCTCTCGTTCATGTCCGAGGTACTCAAATCGATATGGGAGGATGGGGTGG
ATGTCATTGGGGCGATTGCGTGGAGCTTTATGGATAATTGGGAATTTGGGGATTATGCGCAGCAATTTGGGATGCAGGTTGTCAATCGGACGACGCAGGAGAGGTGGTTTAAGAAGAGCTTTTT
TGACATTGTGGATTTTGTGGGTGCGAGGAATGGGCTGGGATAT**TGA**

B

MGFLFLLAAVAVVVQAQEVYITTTGYTARPQCTEPPATPTFRFQSFSYASLNDTIRYAIS
VPSPTTTHTYGPAYTDAVAKLSTKLTTTTWGSWVPSQTVISATDTADKYGQAAWSSQWLH
ASLANYTNIGLYTTTVNPTPLPSSELVLPPRDYFGPTDCYNFPEGFTFGVAGSAAQIEGA
IGLEGRAPSILEKLLPDTEPQDYVTNENYYLYKQDIQRLASIGVKYYSFSISWGRILPFT
VPGSPINEQGLKHYNDLIDYVLEVGMVPIVTMLHFDTPLYFINASAGYVVPDIGYQNGGY
WNKEFVDSFVNYGKILFTHFADRVPFWVTINEPLLYAFNFTGLDNVVHAHAELYHFYHDT
LNGTGKVGLKLNDNFGVPKHPENQTEIDAANRFNDMQLGVFAYPICLGQQYPKSILDTLP
GAKPLSKKELEYISHTTDFIGIDAYTATVISVPAEGIEYCSKQNMTTNSLYPYCVTQETV
NSYGWDIGYRSQSYVYITPTYLRAFLSYIWNTYKTPLVLSEFGFPVYAESTRDLVDQLYD
SPRSQYYLSFMSEVLKSIWEDGVDVIGAIAWSFMDNWEFGDYAQQFGMQVVNRTTQERWF
KKSFFDIVDFVGARNGLGY-

**Figure S1.** DNA and protein sequences of BGL-1. A) DNA sequence of the *bgl-1* gene. The nucleotides that codify for the signal peptide are underlined. The first codon encoding the BGL-1 protein after the signal peptide, and the stop codon are magnified and marked in black. The sequence of the intron is highlighted in grey. B) Amino acid sequence of BGL-1, with the catalytic amino acids marked in light grey.

1-Butanol, 1-heptanol, 1-propanol, 2,4-dinitrophenol, 2,6-dihydroxynaphthalene, 2-butanol, 2-mercaptoethanol, 2-nitrophenyl β-D-glucopyranoside, 2-propanol, 3,3-diphenyl propanol, 4-cresol, 4-hydroxybenzyl alcohol, 4-methylumbilliferyl β-D-xylopyranoside, 4-nitrophenol, 4-nitrophenyl α-arabinopyranoside, 4-nitrophenyl α-D-glucopyranoside, 4-nitrophenyl α-D-rhamnopyranoside, 4-nitrophenyl β-D-fucopyranoside, 4-nitrophenyl β-D-galactopyranoside, 4-nitrophenyl β-D-glucopyranoside, 4-nitrophenyl β-D-xylopyranoside, l-arabinose, arabitol, ascorbic acid, catechol, cellobiose, cinnamyl alcohol, cyclohexanol, dulcitol, egcg, ergosterol, ethanol, eugenol, ferulic acid, D-fructose, D-galactose, gallic acid, gentiobiose, D-glucose, glycerol, guaiacol, hydroquinone, hydroxytyrosol, myo-inositol, lactose, maltose, mannitol, d-mannose, melibiose, menthol, methanol, naphthol, phenol, propargyl alcohol, quercetin, raffinose, resveratrol, D-ribose, L-serine, sorbitol, sorbose, sucrose, L-threonine, L-trehalose, L-tyrosine, vanillyl alcohol, xylitol, D-xylose, α-tocopherol, β-sitosterol

**Figure S2.** List of compounds tested in transglycosylation screening assay

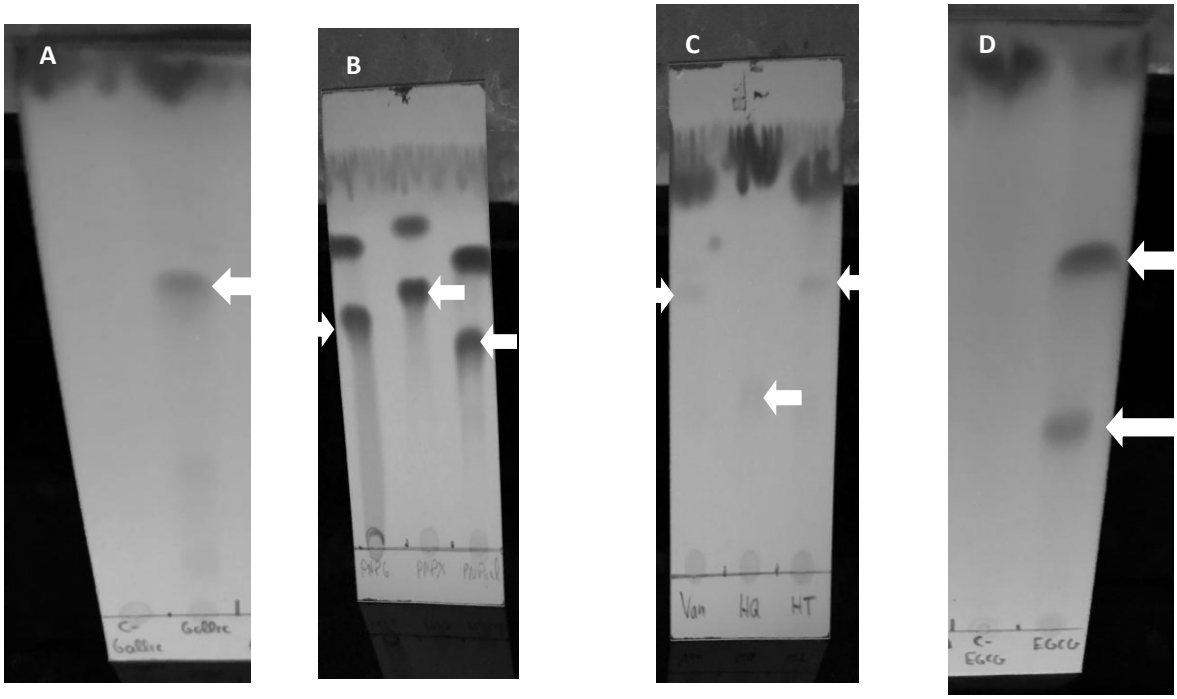



**Figure S3.** Thin layer chromatography of the different compounds in transglycosylation reactions using BGL-1-E521G. A) Gallic acid reaction. B) Nitrophenol sugars reaction. C) Vanillin, hydroquinone and hydroxytyrosol reactions. D) EGCG reaction. The new products detected are signaled with arrows.

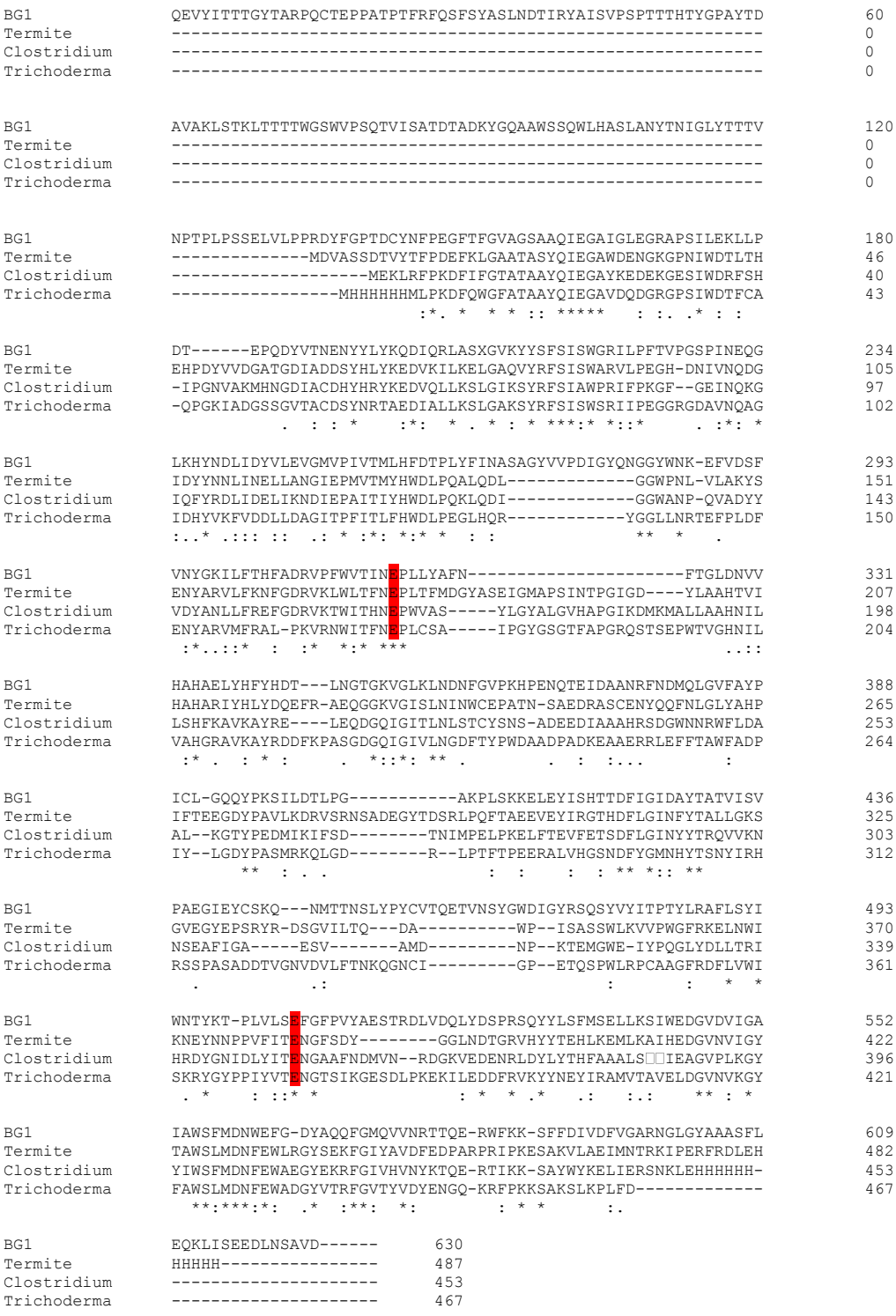



**Figure S4.** Alignment of BGL-1 amino acid sequence with the BGLs sequences of bacterium *Clostridium cellulovorans*, fungus *Trichoderma reesei* and termite *Neotermes koshunensis*. The catalytic amino acids are highlighted in red.

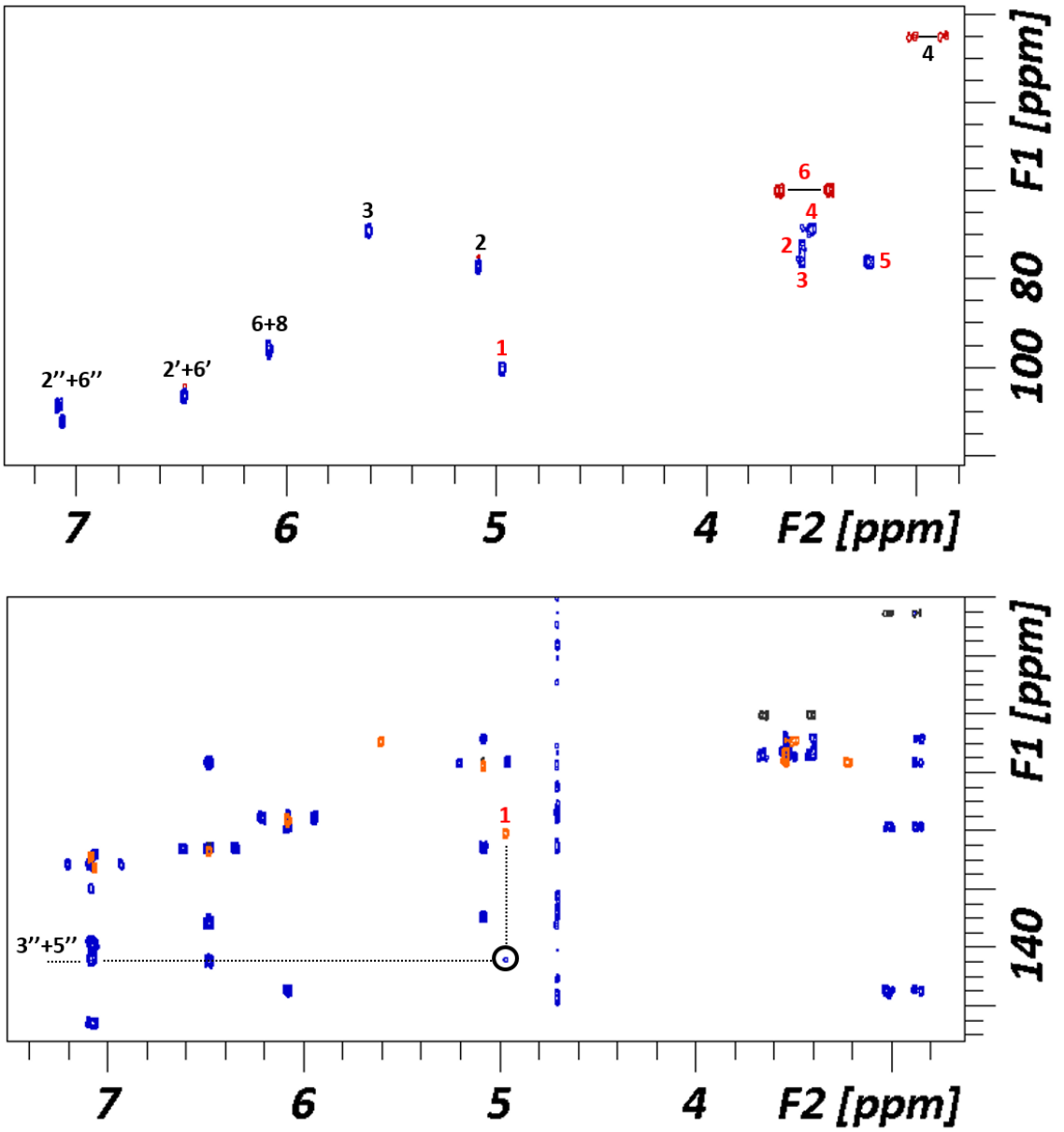



**Figure S5.** NMR Spectra of the EGCG glycoside. **Top:** Labelled $^{1}H$-$^{13}C$ HSQC of EGCG-glucose. **Bottom:** Superimposition of $^{1}H$-$^{13}C$ HSQC (black/orange) and HMBC (blue) spectra. The correlation of the anomeric position is highlighted.

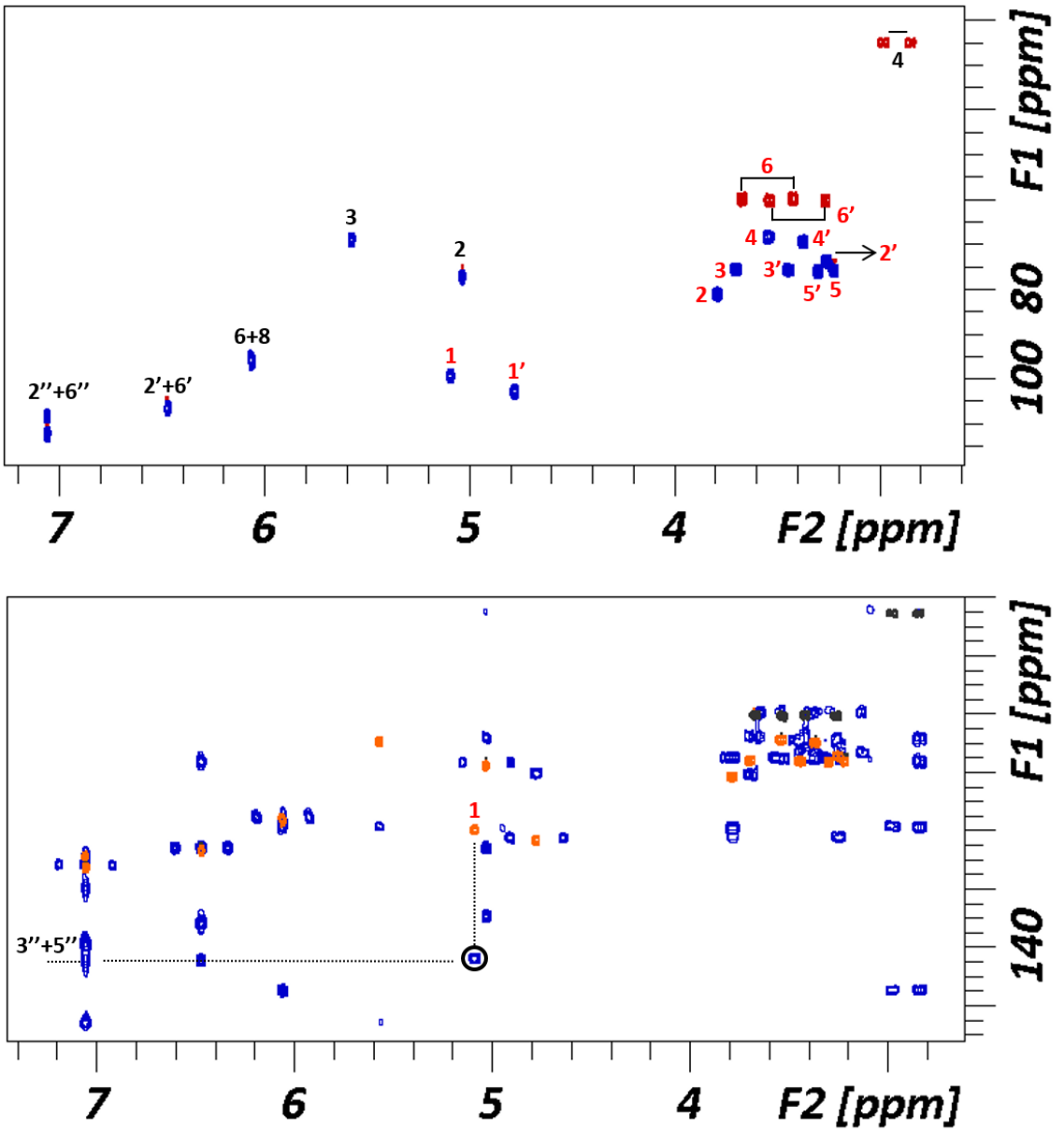



**Figure S6.** NMR Spectra of the EGCG plus sophorose. **Top:** Labelled $^{1}H$-$^{13}C$ HSQC of EGCG-sophorose. **Bottom:** Superimposition of $^{1}H$-$^{13}C$ HSQC (black/orange) and HMBC (blue) spectra. The correlation of the anomeric position is highlighted.

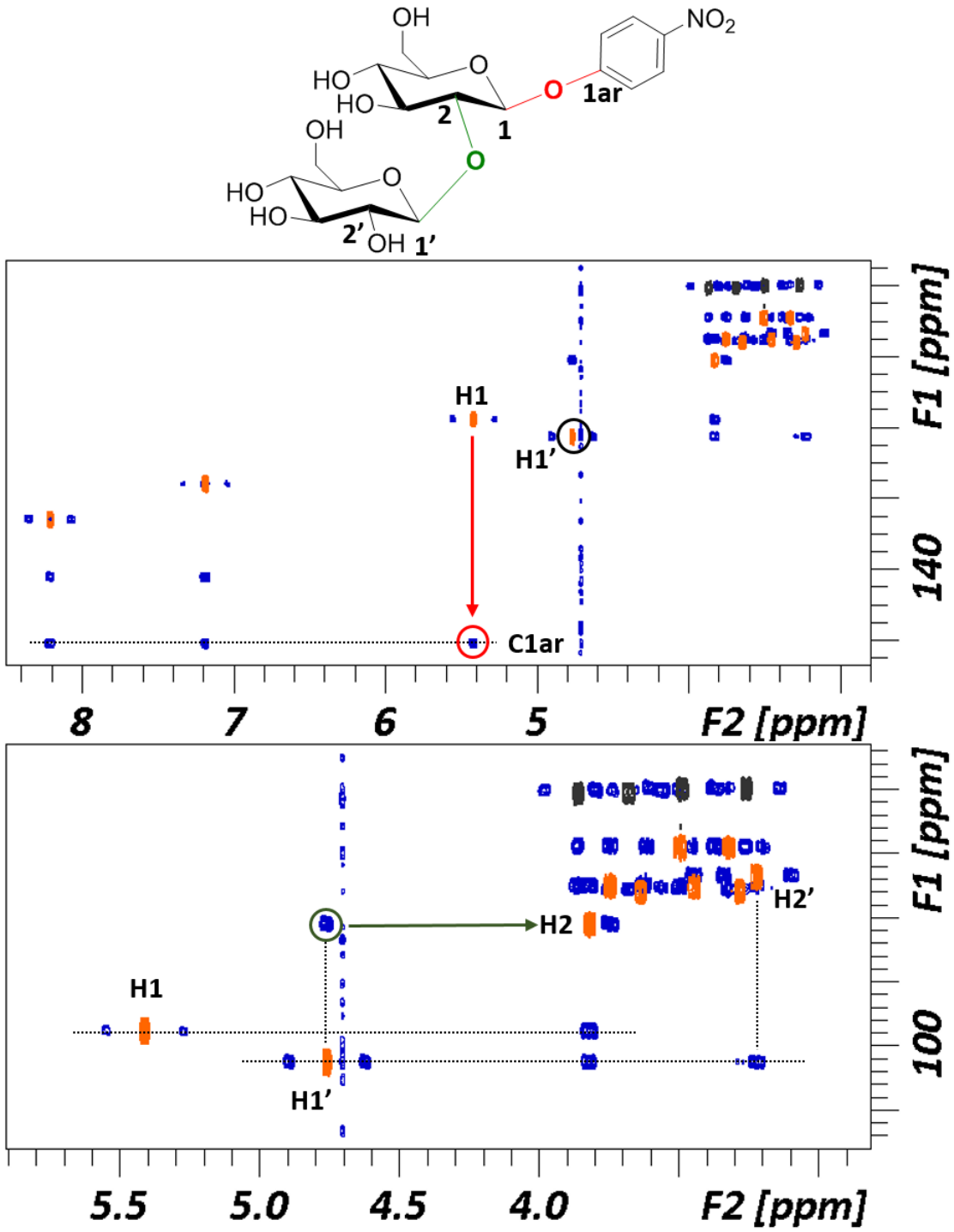



**Figure S7.** Top: structure of the Glucose-Glucose-*p*NP derivative. Superimposition of $^{1}H$-$^{13}C$ HSQC (orange/black) and HMBC (blue). Bottom: Expansion of the carbohydrate region. Arrows represent the key cross peaks for the characterization of the molecule. The lower panel shows the correlation of the anomeric H1' proton of the transferred glucose with the carbon at position 2 in the acceptor glucose for the disaccharide linkage (green), and the upper panel displays the correlation of the anomeric position of the acceptor glucose and the quaternary carbon of the *p*NP ring residue (red).

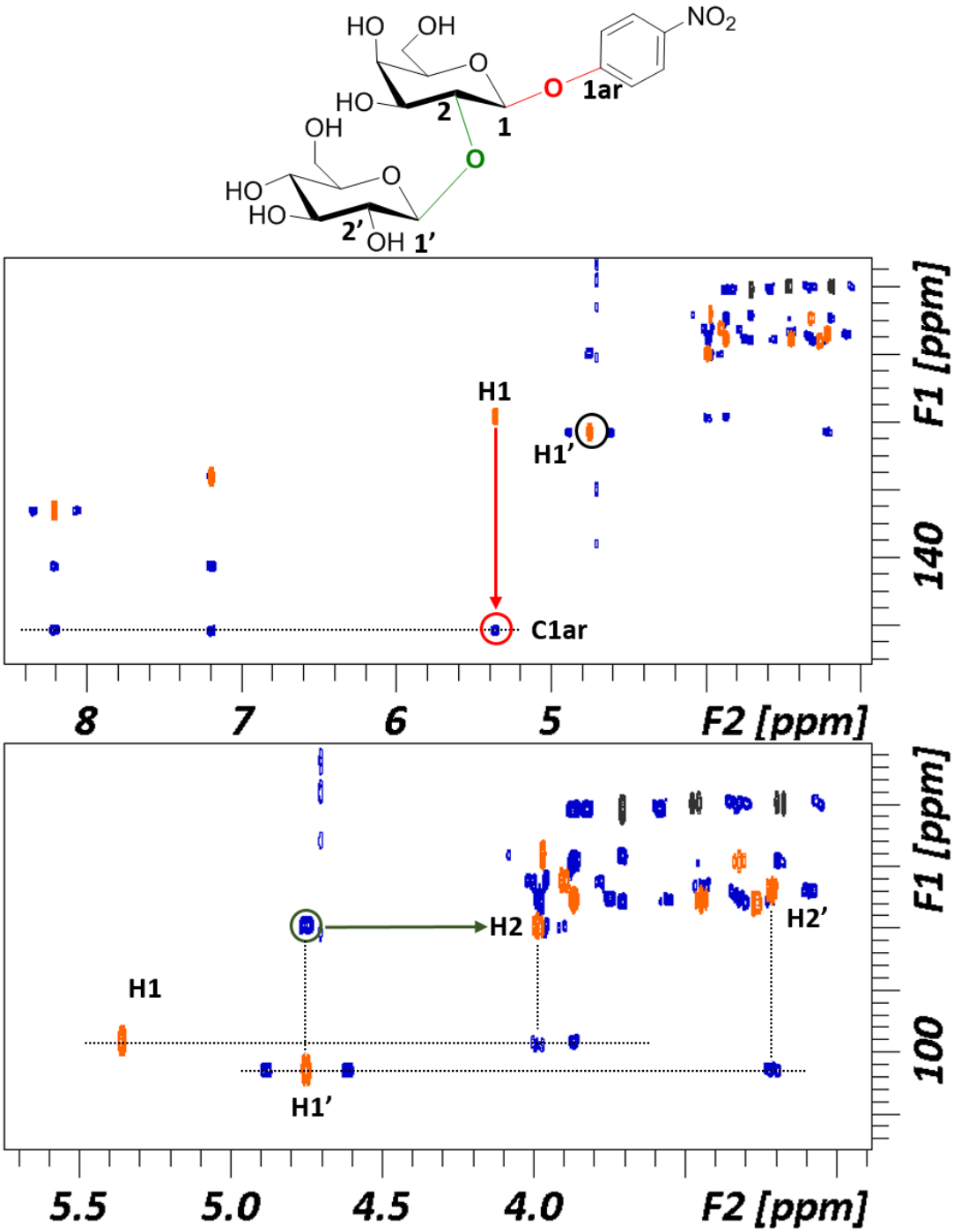



**Figure S8.** Top: structure of the Glucose-Galactose-*p*NP derivative. Superimposition of $^{1}H$-$^{13}C$ HSQC (orange/black) and HMBC (blue). Bottom: Expansion of the carbohydrate region. Arrows represent the key cross peaks for the characterization of the molecule. The lower panel shows the correlation of the anomeric H1' proton of the transferred glucose with the carbon at position 2 in the acceptor galactose for the disaccharide linkage (green), and the upper panel displays the correlation of the anomeric position of the acceptor galactose and the quaternary carbon of the *p*NP ring residue (red).

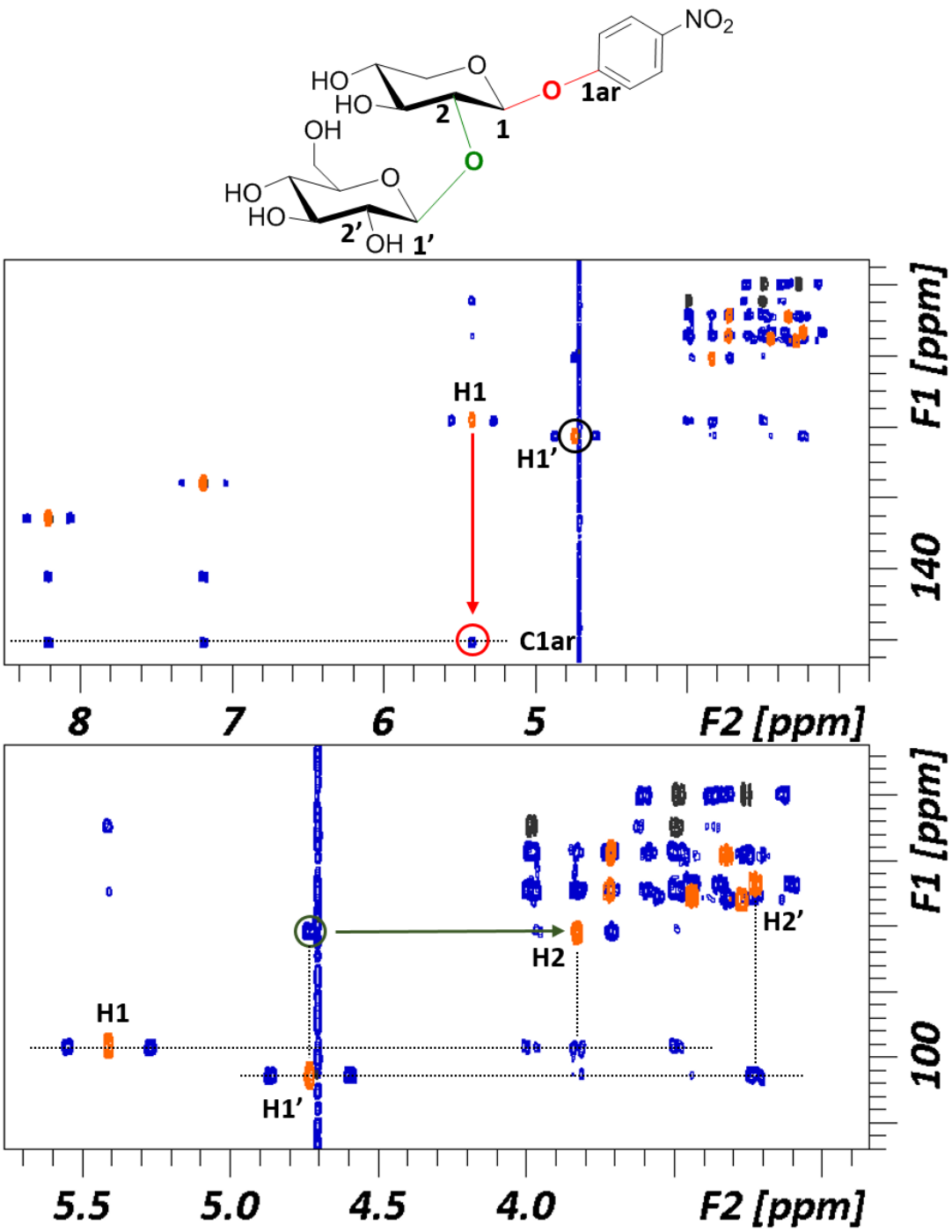



**Figure S9.** Top: structure of the Glucose-Xylose-*p*NP derivative. Superimposition of $^{1}H$-$^{13}C$ HSQC (orange/black) and HMBC (blue). Bottom: Expansion of the carbohydrate region. Arrows represent the key cross peaks for the characterization of the molecule. The lower panel shows the correlation of the anomeric H1' proton of the transferred glucose with the carbon at position 2 in the acceptor xylose for the disaccharide linkage (green), and the upper panel displays the correlation of the anomeric position of the acceptor xylose and the quaternary carbon of the *p*NP ring residue (red).

# DISCUSIÓN GENERAL

# DISCUSIÓN GENERAL

Como se ha mencionado a lo largo de la presente Tesis Doctoral, las GHs en general, y las BGLs en particular, son enzimas ampliamente distribuidas en los organismos vivos, que desempeñan papeles fundamentales en varios procesos biológicos, y han despertado una considerable atención en los últimos años debido a sus importantes funciones en diversos procesos biotecnológicos.

Este trabajo comenzó con el estudio del genoma y proteoma del hongo *T. amestolkiae* obtenido en diferentes fuentes de carbono, identificándose las principales CAZYmes presentes en cada una de las condiciones estudiadas. Concreamente, se analizó en detalle el perfil de producción de sus β-glucosidasas, enzimas clave para el aprovechamiento de los azúcares presentes en la biomasa vegetal. Estas enzimas fueron clonadas y expresadas heterólogamente en *P. pastoris*, y caracterizadas, analizándose su potencial para la sacarificación de material lignocelulósico. Tras esto, se ensayó su capacidad de transglicosilación y se diseñó una forma mutante de una de ellas para aumentar su rendimiento en este tipo de reacciones. A continuación, se discutirán de forma conjunta los principales resultados obtenidos en los distintos capítulos.

## 1. GHs y BGLs del hongo *T. amestolkiae*.

En la actualidad, el descubrimiento de nuevas enzimas activas sobre carbohidratos, con propiedades mejoradas que permitan desarrollar procesos biotecnológicos de una forma más eficiente, se encuentra fuertemente influenciado por la aplicación de las ciencias ómicas, como la genómica y la proteómica.

Como se describió en el capítulo 1, el genoma de *T. amestolkiae* contiene una gran cantidad de genes que codifican GHs. En el genoma de este hongo se identificaron 24 posibles genes codificantes de BGLs, cantidad muy superior a la encontrada en otros hongos actualmente utilizados para producir cócteles enzimáticos con actividad lignocelulósica, como *T. reesei*, *A. niger*, o *P. oxalicum*. De todos ellos, 9 codificarían posibles proteínas extracelulares, ya que en ellas se ha identificado la secuencia del péptido señal. Teniendo en cuenta que, en las condiciones ensayadas, tan sólo se han conseguido purificar y caracterizar 3 de estas 9 BGLs, el potencial de *T. amestolkiae* para producir nuevas BGLs podría ser aún mayor del descrito en este trabajo. Como se explica en el primer capítulo de esta Tesis Doctoral, cuando se cultiva el hongo en sustratos simples, como Avicel y glucosa, la cantidad de BGLs identificadas en el proteoma es menor que la que se detecta cuando el hongo crece en sustratos complejos, como xilano o *slurry* (denominación que se le da a la paja de

trigo tras un pretatramiento por *steam explosion*). Esto indica que materiales estructuralmente más complejos, heterogéneos y con ramificaciones, deben requerir enzimas complementarias a las que se necesitan para metabolizar sustratos más sencillos, y *T. amestolkiae* podría ser un organismo de gran interés para el estudio de estas enzimas.

En este sentido, el diseño de estrategias para buscar nuevas BGLs en este hongo, ya sea utilizando sustratos más complejos para su inducción o expresando heterólogamente las enzimas encontradas en su genoma, podría conducir a la identificación de nuevas BGLs.

Es interesante destacar que la producción de distintas isoenzimas en función de la fuente de carbono es un fenómeno que se ha observado en hongos del género *Aspergillus*, como *A. niger* (Singhania *et al*., 2011) o *A. terreus* (Nazir *et al*., 2010). Conseguir dilucidar las bases de la regulación de la producción de estas enzimas sería esencial para diseñar unas condiciones de cultivo que permitan producir la isoforma deseada en grandes cantidades. En el caso de *T. amestolkiae*, los resultados obtenidos en esta Tesis permiten establecer algunas condiciones específicas para su producción:

1. La BGL-1 se produce en todos los medios y condiciones estudiados, pero en niveles muy bajos. Esto podría indicar, o bien que es una enzima constitutiva, o que se produce bajo estímulos que se han dado en todos los secretomas estudiados, como es el caso de BGL-3. Para confirmar estas hipótesis, habría que realizar un estudio más profundo de su expresión bajo determinadas condiciones.
2. BGL-2 sólo se produce de forma significativa en medios con celulosa, como Avicel y *slurry* de paja de trigo, por lo que parece una enzima inducible en presencia de celulosa, bien esté en forma microcristalina o presente en materiales lignocelulósicos. Además, hay que destacar que cuando se analizó el ARN del hongo obtenido a partir de cultivos con Avicel se encontraron dos tránscritos distintos de esta enzima. Uno de ellos codificaría la enzima completa con un dominio de unión a celulosa (BGL-2), siendo esta la primera β-glucosidasa fúngica caracterizada con un CBD. El otro tránscrito codificaría una forma especial de la proteína sin motivo de unión a celulosa (BGL-2T). La versión truncada de BGL-2 podría haberse generado por un proceso de *splicing* alternativo, uno de los principales mecanismos de regulación genética en organismos eucariotas, que aumenta la diversidad de proteínas encontradas (Sammeth *et al*., 2008). BGL-2 tiene 3 intrones, estando uno de ellos próximo a un codón de terminación, justo delante del CBD. La forma truncada de BGL-2 podría explicarse por la retención de este intrón, al procesarse como se discute en el capítulo 2. La existencia de

procesos de retención de intrones, como reguladores de la traducción de un dominio de unión a carbohidratos, han sido descritos en celobiohidrolasas y endoglucanasas producidas por hongos (Birch *et al*., 1995; Baba *et al*., 2005; Costanzo *et al*., 2007). Para dilucidar la regulación de los 2 tránscritos que codifican BGL-2 en *T. amestolkiae*, sería necesario estudiar los mecanismos de regulación génica del hongo, en función de la fuente de carbono o del momento de crecimiento. Una posible aproximación para diferenciar el procesamiento, o retención de este intrón, sería el análisis transcriptoma del hongo, utilizando técnicas de PCR cuantitativa a distintos tiempo de incubación y en diferentes condiciones de cultivo. En la figura 1 se puede observar, de manera esquemática, la hipótesis planteada sobre el proceso de *splicing* que acabaría generando las dos versiones de BGL-2 encontradas.

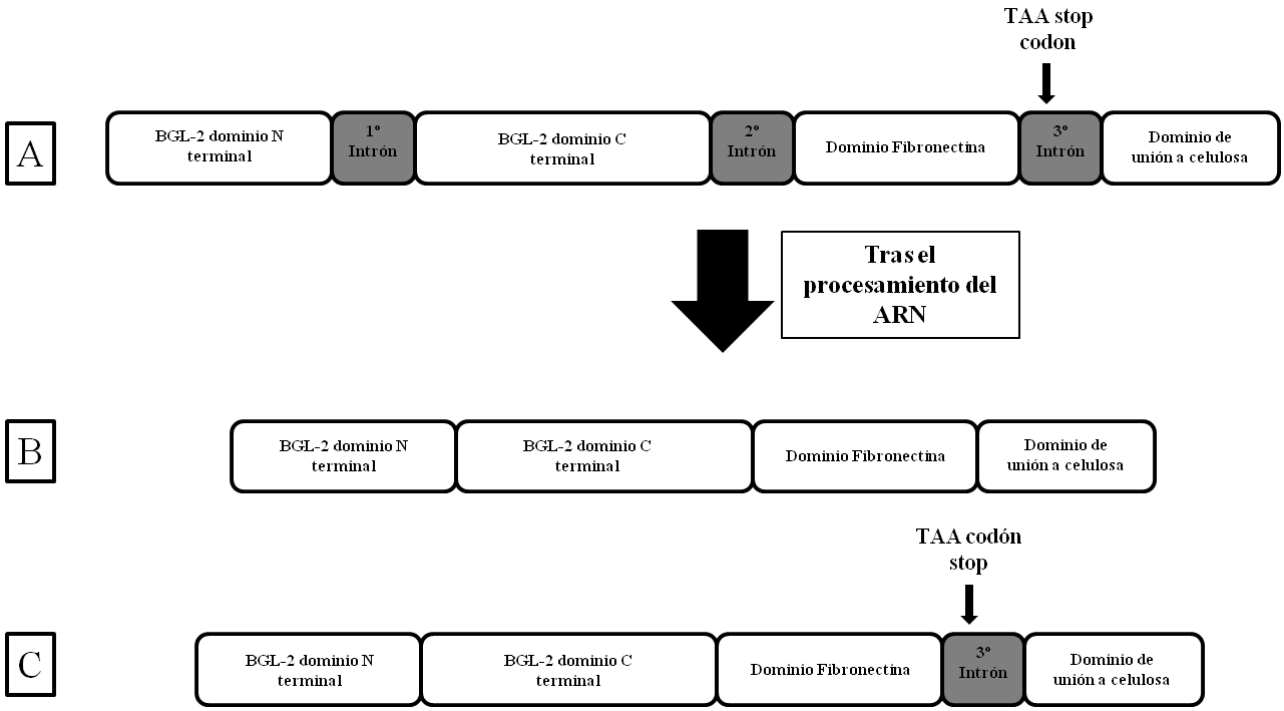



**Figura 1.** Representación esquemática de las diferentes partes del ARN mensajero procedente del gen *bgl2* antes y después del posible proceso de *splicing*. Los 4 exones están representados en blanco. Los 3 intrones están indicados en gris. A) Secuencia antes del procesamiento de ARN. B) Procesamiento que daría lugar a la secuencia que codificaría la proteína BGL-2 con el dominio de unión a celulosa. C) Procesamiento del ARN reteniendo el tercer intrón, que se traduciría en la proteína BGL-2T y carecería del dominio de unión a celulosa.

**3.** La producción de BGL-3 también tiene lugar en todos los medios y condiciones de cultivo estudiados, con niveles altos, pero solo cuando se agota la fuente de carbono, como se explica en el capítulo

3. Esto permitiría utilizar medios con fuentes de carbono muy asequibles para su producción, como es el caso de un gran número de residuos, o de monosacáridos simples.

## 2. BGLs en procesos de hidrólisis de biomasa lignocelulósica

La producción de bioetanol de primera generación (1G) está basada en el uso de productos agrícolas con alto contenido en carbohidratos, como maíz, remolacha o caña de azúcar, como fuente de azúcares fermentables (Naik *et al*., 2010). Aunque su implementación supuso un importante incremento en la producción mundial de bioetanol a principios del siglo XXI, la llegada de varias crisis alimentarias introdujo el dilema moral sobre el uso de alimentos para la obtención de combustibles. Así surgió la necesidad de desarrollar nuevos procesos en los que la obtención de biocombustibles no interfiriera con la cadena alimentaria (Escobar *et al*., 2009) y comenzaron los estudios para la producción de bioetanol de segunda generación (2G), en los que se utiliza la biomasa lignocelulósica como materia prima, en especial residuos agrícolas, para su producción. Desafortunadamente, la producción de este biocombustible tiene varios inconvenientes, ya que la hidrólisis de los polisacáridos de la pared celular de las plantas requiere pretratamientos que eliminen la lignina, el heteropolímero que protege los polisacáridos, lo que dificulta e incrementa los costes del proceso. En la actualidad, a pesar de los avances de los últimos años, los rendimientos del proceso de bioetanol 2G siguen siendo bajos y alcanzar una elevada viabilidad comercial requiere obtener otros productos de valor añadido a partir de la biomasa lignocelulósica. Este enfoque integrador, en el que se puedan utilizar y rentabilizar todos los componentes de la pared celular vegetal, recibe el nombre de biorrefinería (Ragauskas *et al*., 2006).

La producción de etanol de segunda generación (2G) es uno de los principales campos de aplicación de las celulasas, y más específicamente, de las BGLs. En relación con los procesos de conversión de biomasa, se considera que una buena BGL debe poseer las siguientes características: elevado rendimiento de hidrólisis, tolerancia a inhibición por producto, y estabilidad en las condiciones del proceso, para mantener tasas de conversión altas durante la hidrólisis (Sørensen *et al*., 2013). Todos estos parámetros varían enormemente entre las diferentes BGLs caracterizadas, como se describe a continuación. Poseer un elevado rendimiento de hidrólisis, o lo que es lo mismo, un alto valor en la eficiencia catalítica ($k_{cat}/K_m$), es un aspecto esencial para que una BGL pueda ser seleccionada en procesos industriales de degradación de biomasa. Hasta la fecha, la mayoría de las BGLs empleadas en la hidrólisis de la celulosa pertenecen a la familia GH3, porque estas enzimas poseen unos valores de eficiencia

catalítica más altos que las de otras familias. De las BGLs estudiadas en este trabajo, dos son de la familia GH3 y una de la GH1, y los resultados encontrados están de acuerdo con esto, ya que BGL-2 y BGL-3 son entre 20 y 100 veces más activas sobre celobiosa y otros celooligosacáridos que BGL-1. Sin embargo, aunque tanto BGL-2 como su versión truncada, BGL-2T, mostraron buenas capacidades hidrolíticas en procesos de sacarificación sobre bagazo de cerveza, BGL-3 mostró mayor eficiencia catalítica siendo la enzima más activa sobre celooligosacáridos descrita hasta la fecha, como se puede observar en las tablas comparativas de los capítulos 2, 3 y 5.

En cuanto a la glucotolerancia, es un factor esencial debido a que altas concentraciones de esta molécula provocan inhibición en estas enzimas. Como se ha mencionado en los capítulos 2, 3 y 5, las BGL de la familia GH3 de *T. amestolkiae* (BGL-2 y BGL-3) no se caracterizan por poseer una alta tolerancia a la glucosa. En cambio, BGL-1, ha demostrado ser altamente glucotolerante, e incluso se han detectado incrementos de actividad del 20% a concentraciones bajas de glucosa. Este comportamiento se ha descrito en otras enzimas de la familia GH1, siendo tolerantes o incluso estimuladas por altas concentraciones de glucosa (Cao *et al*., 2015). A pesar de su glucotolerancia, la BGL-1 de *T. amestolkiae* tiene preferencia por los enlaces β-1,2 del dimero de glucosa (soforosa) que sobre los β-1,4, lo que podría explicar su menor afinidad que BGL-2 y BGL-3 por los oligosacárdos de glucosa derivados de la celulosa. Una aproximación para solucionar el problema de la inhibición por glucosa que afecta, sobre todo, a las BGLs de la familia GH3, podría ser la realización de los procesos de sacarificación y fermentación simultáneos. En cualquier caso, la inhibición por producto de las BGLs para hidrolizar la lignocelulosa es un gran desafío a resolver.

Respecto a las condiciones del proceso, pH y temperatura son parámetros esenciales para el funcionamiento de las BGLs. Numerosas revisiones establecen que la mayoría de BGLs procedentes de hongos tienen un pH óptimo entre 4-5 y una temperatura óptima entre 60 - 75 ºC (Sørensen *et al*., 2013), aunque a temperaturas superiores y pHs más bajos, pierden su actividad rápidamente. Las BGLs de *T. amestolkiae*, tienen un perfil de pH óptimo en torno a 4, y la temperatura óptima oscila entre 60 ºC (BGL-1 y BGL-2) y 70 ºC (BGL-2T y BGL-3). En relación al perfil de temperaturas óptimas, las BGLs de *T. amestolkiae* se encuentran en el rango de las BGLs fúngicas. Sin embargo, si es reseñable la estabilidad a 50 ºC durante tres días que mostraron BGL-2T y BGL-3. En cuanto al pH hay que destacar que, tras los pretratamientos, el pH de la biomasa es bastante ácido (≤ 4) y debe ajustarse para el tratamiento enzimático. Las BGL de *T. amestolkiae* poseen características muy interesantes a este respecto, ya que mantienen una elevada actividad a pH 3, e incluso hasta

pH 2 (excepto BGL-1). Esto podría ser muy útil para su aplicación directa en procesos de sacarificaciones sobre residuos que hayan sido pretratados, especialmente los tratados por *steam explosion* en presencia de ácido diluido para conseguir solubilizar los xilanos y amentar los rendimientos de azúcares fermentables (Alvira *et al*., 2010).

## 3. *P. pastoris* como modelo de expresión de las BGL de *T. amestolkiae*

Con el fin de complementar los cócteles comerciales ya existentes, o formular mezclas idóneas de enzimas para la hidrólisis de la celulosa, estos catalizadores pueden expresarse heterólogamente en distintos hospedadores. Sin embargo, en muchas ocasiones pueden aparecer dificultades debido a las características de cada hospedador. Por ejemplo, un posible problema es la alteración del patrón de glicosilación de la enzima, lo que puede alterar seriamente su actividad y estabilidad (Jeoh *et al*., 2008). El uso de *P. pastoris*, como organismo eucariota hospedador, ha resultado ser un éxito para la producción de las proteínas BGL-1, BGL-2 y BGL-3 del hongo *T. amestolkiae*, secretando proteínas altamente eficientes y estables.

*P. pastoris* es una levadura que posee características deseables para producir enzimas recombinantes, tales como: i) la existencia de un promotor fuerte y altamente regulable por adición de metanol como fuente de carbono, lo que permite elevadas producciones de enzimas, ii) la presencia de modificaciones postraduccionales similares a las de los organismos eucariotas superiores, lo cual es una ventaja para expresar enzimas fúngicas glicosiladas y iii) la secreción de pocas proteínas al medio extracelular, lo que simplifica el proceso de purificación de las enzimas recombinantes.En este hospedador han sido expresadas con éxito otras BGLs, como por ejemplo las de *Aspergillus fumigatus* (Liu *et al*., 2012), *Thermoascus aurantiacus* (Hong *et al*., 2007) o *Penicillium funiculosum* (Ramani *et al*., 2015), presentándose en alguno de ellos una mejoría de las propiedades de las enzimas recombinantes con respecto a las formas nativas.

Por último, es interesante destacar que aunque los patrones de glicosilación de *P. pastoris* no son los mismos que los de algunas glicosil hidrolasas de *T. amestolkiae* (Nieto-Domínguez *et al*., 2016), esto no parece afectar los niveles de secreción de las BGLs de *T. amestolkiae*, que se producen en la levadura sin perder sus propiedades catalíticas.

## 4. Transglicosilación

La glicosilación de moléculas bioactivas presenta gran interés porque puede mejorar sus propiedades. Concretamente, puede aumentar su

solubilidad, lo que influye en su biodisponibilidad en entornos acuosos. Este hecho es especialmente relevante en moléculas con baja solubilidad como los antioxidantes. En el tercer capítulo de la Tesis se describe el incremento de la solubilidad por la glicosilación de hidroxitirosol y el alcohol vainillínico, demostrando que presentaron mayor actividad biológica que los compuestos sin glicosilar. Este fenómeno se ha descrito también con otros compuestos antioxidantes, moléculas que generalmente poseen una baja solubilidad (Salas *et al.*, 2011; Nieto-Domínguez *et al.*, 2017; González-Alfonso *et al*., 2018; Míguez *et al*., 2018).

De manera habitual, las BGLs catalizan la hidrólisis de los enlaces glicosídicos de oligosacáridos de cadena corta, aunque en determinadas condiciones pueden catalizar la reacción reversa, generando nuevos enlaces glucosídicos (Bohlin *et al*., 2010). Todas las BGLs producidas por el hongo *T. amestolkiae* fueron capaces de llevar a cabo reacciones de transglicosilación, aunque BGL-1 lo hace de forma muy limitada. En el capítulo 4 de la Tesis, se describe la actividad transglicosiladora de BGL-2, BGL-2T y BGL-3 sobre diferentes moléculas, entre las que destacan algunos antioxidantes fenólicos como el hidroxitirosol, el alcohol vainillínico, la hidroquinona y el 4-hidroxibenzil alcohol. En el capítulo 5, se estudia la capacidad sintética de la enzima BGL-1 y de su variante glicosintasa (BGL-1-E521G y E521S), encontrando que estas enzimas, como se había descrito previamente en otras glicosintasas (Perugino *et al*., 2004), tienen mayor actividad de transglicosilación. La variante BGL-1-E521G resultó ser la más eficaz y estudios preliminares indican que esta glicosintasa tiene un gran potencial en la síntesis de nuevos glicósidos, ya que no hidroliza los productos que genera. En este sentido, cabe destacar que esta enzima sintetiza, con gran eficacia, glicósidos de *p*-nitrofenoles y de EGCG, esta última una molécula bioactiva presente en el té verde.

Como se mencionó anteriormente, la capacidad de transferir carbohidratos a diferentes moléculas aceptoras puede asociarse con muchas aplicaciones de interés biotecnológico. Sin embargo, el papel de esta actividad en la fisiología fúngica todavía está lejos de ser comprendido en profundidad. Actualmente, existen dos teorías que intentan explicar la utilidad de estos fenómenos:

1. Una de las explicaciones más interesantes, está relacionada directamente con la inhibición por producto de estas enzimas. Generalmente la glucosa a altas concentraciones las inactiva (como se ha comentado anteriormente) y la transglicosilación podría servir para disminuir la concentración de glucosa, ya que ocurre a altas concentraciones tanto de moléculas aceptoras como donadoras, disminuyendo los niveles de glucosa libre para poder mantener las enzimas activas más tiempo.

2. Otra posible explicación es que la actividad de transglicosilación tenga un rol en la fisiología de estos organismos, como, por ejemplo, que alguna de las moléculas que se generan, de forma natural, en procesos de transglicosilación pudiera servir para inducir el sistema celulolítico de los microorganismos. Se ha probado que determinadas BGLs pueden sintetizar soforosa (dos moléculas de glucosa unidas por enlace β-1,2) a partir de celobiosa y glucosa. La soforosa induce el sistema celulolítico del hongo *T. reesei* (Fowler y Brown Jr, 1992), aunque debido a su alto precio su uso como inductor es limitado y no es atractivo ni rentable para procesos industriales.

En el capítulo 5 se ha demostrado que la BGL-1 de *T. amestolkiae* produce soforosa. Inicialmente se consideró que esta enzima era una β-1,4-BGL debido a su actividad sobre *p*NPG y otros celooligosacáridos. Sin embargo, estudios realizados en esta Tesis Doctoral, han demostrado que es una β-1,2-BGL, ya que posee considerablemente más actividad sobre soforosa que sobre celobiosa. Además, las reacciones de transglicosilacion que cataliza esta enzima dan como resultado enlaces glicosídicos β-1,2. Así, una posible explicación a que BGL-1 esté presente en todos los secretomas estudiados, podría ser su secreción constitutiva como inductor del sistema celulolítico, fenómeno que, como ya se ha comentado, se ha descrito en BGLs de otros hongos filamentosos como *T. reesei*.

Una de las principales ventajas que presentan las enzimas sobre la síntesis química es su regioselectividad. Sin embargo, a pesar del hecho de que las glicosidasas muestran mayor especificidad hacia un determinado grupo hidroxilo, en algunos trabajos se han detectado glucósidos minoritarios (Kato *et al*., 2002), lo que complica el estudio y la purificación del compuesto deseado y contribuye a disminuir el rendimiento global del proceso. En el presente trabajo, los glicósidos analizados (glucósidos de hidroxitirosol, vainillina y EGCG y soforosa) se han obtenido en reacciones regioselectivas, no generándose subproductos. La BGL-2 catalizó la síntesis de glicósidos con enlaces β-1,4, mientras que BGL-1, y su versión glicosintasa, formaron glicósidos con enlaces β-1,2. Esta gran regioselectividad podría facilitar su aplicación en procesos biotecnológicos.

## 5. Mejora biotecnológica de BGLs: Glicosintasas

El descubrimiento de nuevas BGLs con propiedades mejoradas es interesante, tanto para su aplicación en procesos de hidrólisis como de transglicosilación. Esto puede conseguirse mediante ingeniería genética cuando se dispone de la secuencia génica y de un buen sistema de

expresión heteróloga. De este modo, pueden modificarse sus propiedades fisico-químicas y catalíticas, así como su resistencia a agentes físicos y químicos (temperatura, pH, inhibidores), para mejorar su eficiencia en los procesos que se pretende desarrollar. Existen dos métodos que se han aplicado a este propósito: evolución dirigida, y diseño racional.

La evolución dirigida es una técnica que imita la evolución natural, y exige el análisis de un gran número de variantes de enzimas con propiedades mejoradas (Chirumamilla *et al.*, 2001; Cherry y Fidantsef, 2003). Esta técnica ha sido aplicada para mejorar la capacidad hidrolítica de algunas BGLs, procedentes principalmente de bacterias. Concretamente, se ha mejorado la termoestabilidad de una BGL de *Paenibacillus polymyxa* mediante la introducción de mutaciones generadas utilizando una PCR propensa a introducir errores en las secuencias (González-Blasco *et al*., 2000). Resultados similares se lograron con otra BGLs de *Pyrococcus furiosus*, generando una mejora de la hidrólisis de celobiosa a baja temperatura (Lebbink *et al*., 2000). Más recientemente, la combinación de genes de BGLs de *Thermobifida fusca* y *Paebibacillus polymyxa*, originó una enzima quimérica con mayor termoestabilidad que las enzimas parentales y un aumento del 94% en su $k_{cat}$ sobre celobiosa (Pei *et al*., 2011). También se han aplicado técnicas de mutagénesis aleatoria en bacterias para mejorar la glucotolerancia y el rendimiento en procesos de sacarificación con una BGL obtenida a partir de una librería de genes procedentes de metagenomas (Cao *et al*., 2015). Las estrategias de diseño racional también se han utilizado en BGLs fúngicas con la intención de mejorar procesos hidrolíticos, aunque estos estudios no son muchos debido al escaso número de estructuras de BGLs que se conocen. Aun así, recientemente se han cambiado determinados aminoácidos del sitio activo de una BGL de *T. reesei* mejorando significativamente su actividad y termoestabilidad (Lee *et al*., 2012).

Utilizando esta estrategia, se ha conseguido mejorar el potencial transglicosilador la enzima BGL-1. Las glicosil hidrolasas, en su forma nativa, pueden llevar a cabo reacciones de transglicosilación, pero su capacidad de degradar los productos formados hace que los rendimientos en la síntesis de nuevos compuestos estén limitados. Por ello, McKenzie y colaboradores (1998) desarrollaron mediante ingeniería genética variantes de estas enzimas, llamadas glicosintasas, que eran incapaces de degradar los productos formados. Dependiendo de los aminoácidos catalíticos que se cambien se obtienen distintas variantes que se denominan glicosintasas y tioglicoligasas (ver figura 3 de la introducción). La enzima BGL-1 ha sido exitosamente convertida en su variante glicosintasa, como recoge el capítulo 5 de esta tesis, donde se explica cómo es capaz de sintetizar glucósidos de varios nitrofenoles así como de varios antioxidantes fenólicos. El repertorio de glicosintasas disponibles ha aumentado

exponencialmente en los últimos años, generándose nuevas variantes que son capaces de generar enlaces α o β utilizando como moléculas donadoras galactosa, fucosa, arabinosa, manosa o incluso lactosa, unidas a una molécula de flúor (Hayes y Pietruszka, 2017).

La síntesis enzimática de carbohidratos utilizando glicosintasas es, hoy en día, una alternativa muy prometedora frente la síntesis química clásica. Aunque los rendimientos son comparables, el gran número de pasos de protección y desprotección de grupos funcionales, así como el elevado número de reacciones necesarias para conseguir los productos por síntesis química, hace que las glicosintasas se sitúen en posición ventajosa para la síntesis de oligosacáridos y glicósidos.

## 6. Perspectivas futuras en la búsqueda de nuevas aplicaciones para las BGLs

### 6.1. Valorización de nuevos residuos

La primera y segunda generación de biocombustibles ha tenido un enorme desarrollo en las últimas dos décadas. Sin embargo, recientemente se han empezado a desarrollar los de tercera generación (3G), en los que se utiliza principalmente biomasa de algas y cianobacterias (Hannon *et al*., 2010; Jones y Mayfield, 2012). Las algas tienen numerosas ventajas, por ejemplo, la rápida formación de biomasa, o su facilidad para ser cultivadas. Estos microorganismos han sido utilizados en numerosas aplicaciones, como la producción de biodiesel. Las algas del género *Laminaria* (también conocidas como algas marrones), menos explotadas biotecnológicamente hasta ahora, poseen un polisacárido de reserva, denominado laminarina, que podría ser hidrolizado enzimáticamente para producir glucosa, de manera similar a la actual despolimerización de la celulosa de las plantas, y ésta, mediante fermentación por levaduras, convertirse en bioetanol 3G. La laminarina es un polisacárido compuesto por monómeros de glucosa unidos por enlaces β-1,3 que posee ramificaciones β-1,6. En el capítulo 3 de esta tesis doctoral, se describe su completa sacarificación utilizando únicamente la enzima BGL-3 como catalizador. Aunque existen β-1,4-glucosidasas que son capaces de hidrolizar enlaces distintos a los β-1,4, la enorme eficiencia de BGL-3 para degradar por completo la laminarina, incluso con mayor rendimiento que utilizando una laminarinasa, no había sido descrita previamente en la literatura, y pone de manifiesto su gran versatilidad. Aparte de su aplicación directa para la degradación de polisacáridos, las investigaciones desarrolladas en la presente tesis doctoral podrían servir para descubrir nuevas vías para la revalorización de residuos. La generación actual de desechos a nivel mundial, es la consecuencia directa del uso ineficiente de los recursos naturales en las actividades humanas. La sociedad industrial ha centrado tradicionalmente su atención en la

producción de bienes y servicios, con el consumo de materias primas y la generación correspondiente de residuos (Giusti, 2009). En estos términos, es especialmente interesante el modo de producción de la enzima BGL-3 en el hongo *T. amestolkiae*. Como se explica en el capítulo 3, BGL-3 solo se produce cuando se agota la fuente de carbono en el cultivo. Esto también se puso de manifiesto en el capítulo 1, cuando se usó glicerol como fuente de carbono, como posible inductor de celulasas en el cultivo. Estos resultados abren un enorme abanico de posibilidades para revalorizar diferentes residuos que pueden ser utilizados como fuente de carbono por *T. amestolkiae* para producir crudos enzimáticos ricos en actividad BGL: en presencia de lignocelulosa la BGL-2 sería la enzima mayoritaria mientras en presencia de otras fuentes de carbono la mayoritaria sería la BGL-3.

## 6.2. Generación de variantes mejoradas de estas enzimas

La mayoría de las BGLs utilizadas en la hidrólisis de la celulosa pertenecen a la familia GH3, por su alta eficacia catalítica para hidrolizar celobiosa y oligosacáridos (Bohlin *et al*., 2010; Teugjas y Väljamäe, 2013). Sin embargo, como se ha mencionado en el apartado anterior, las enzimas de esta familia a menudo son inhibidas por la glucosa que se genera en el proceso de hidrólisis. Sin embargo, la mayoría de BGLs glucotolerantes pertenecen a la familia GH1, caracterizada por poseer una menor eficacia en la hidrólisis de celobiosa y oligosacáridos (Vallmitjana *et al*., 2001; Wierzbicka−Wos *et al*., 2013). Se puede considerar que una BGL "ideal" se podría obtener combinando las características que destacan en cada familia, utilizando técnicas de ingeniería genética que permitan mejorar la tolerancia a la glucosa de las GH3 o la capacidad de hidrólisis de las GH1. Para ello, una posible aproximación sería la llamada quimerogénesis, una técnica de evolución dirigida que consiste en poder mezclar dominios de dos o más proteínas distintas, para obtener al final una variante nueva, que es una quimera de las enzimas parentales. Esta técnica, como se ha comentado anteriormente, se ha aplicado para conseguir BGLs bacterianas con mayor estabilidad y eficacia que las enzimas parentales (Pei *et al*., 2011) y también se ha utilizado con éxito en generar nuevas lacasas fúngicas con propiedades mejoradas (Pardo Mendoza 2017). Desarrollar una variante con la glucotolerancia de las GH1 y la tasa de hidrólisis de las GH3, sería una interesantísima alternativa para su utilización en procesos industriales. Futuros trabajos con las BGLs de *T. amestolkiae* podrían centrarse en el desarrollo de nuevas variantes eficientes y glucotolerantes, realizando estudios estructura-función con BGL-1, BGL-2 y BGL-3, e intentando averiguar cuáles son los aminoácidos responsables de las propiedades de cada una, aunque no se descarta utilizar como base otras enzimas, incluso de la misma familia, que hayan sido mejoradas para tratar de conseguir mejores propiedades para las de *T. amestolkiae*.

### 6.3. Aprovechamiento de polisacáridos mediante cascadas enzimáticas

En cuanto al apartado de la actividad de transglicosilación, a pesar de las interesantes características de las 3 BGLs examinadas a lo largo de los capítulos de la presente tesis, su potencial biotecnológico puede estar limitado debido al elevado coste de las moléculas donadoras que se usan, ya sean celobiosa u otros celooligosacáridos, en el caso de BGL-2 y BGL-3, o glucosa activada con fluor en el caso de BGL-1-E521G.

Una alternativa interesante podría ser producir estos glucósidos directamente a partir de la celulosa que está presente en los residuos lignocelulósicos. Sin embargo, aunque tanto BGL-2 (y su forma truncada) como BGL-3 son activas frente a polisacáridos como el Avicel o la carboximetil celulosa, la velocidad de reacción es muy limitada. Una posibilidad para solventar esto podría ser generar una cascada enzimática, en la cual endoglucanasas y celobiohidrolasas hidrolizarían eficientemente la celulosa del residuo vegetal. Como consecuencia de la acción de esta cascada, quedarían en la reacción celobiosa y otros celooligosacáridos, que serían transferidos a una molécula aceptora de interés por acción de las BGLs. En la literatura hay diferentes trabajos que describen el uso de cascadas enzimáticas para intentar revalorizar la celulosa. Por ejemplo, Lee y colaboradores (2014), llevaron a cabo la conversión de la celulosa en hidroximetil furfural, utilizando celulasas e isomerasas, con un rendimiento de conversión del 46%. En estos términos, el hongo *T. amestolkiae* podría ser una interesante fuente de enzimas para poder realizar la cascada enzimática de transglicosilación. Actualmente ya se dispone de dos endoglucanasas caracterizadas, una de la familia GH5, y otra de la familia GH12 (de Eugenio *et al*., 2018), que podrían utilizarse para producir los oligosacáridos.

Además, como se puede observar en el capítulo 1 de la tesis, al analizar los proteomas de *T. amestolkiae*, entre las 10 proteínas extracelulares más abundantes detectadas se encuentra una celobiohidrolasa de la familia GH7, y otra endoglucanasa más, de la familia GH6. La aplicación combinada de las diferentes celulasas de este hongo, en tándem, debería conducir a una buena degradación del polisacárido, generándose moléculas que puedan ser donadoras en el proceso de transglicosilación (figura 2).

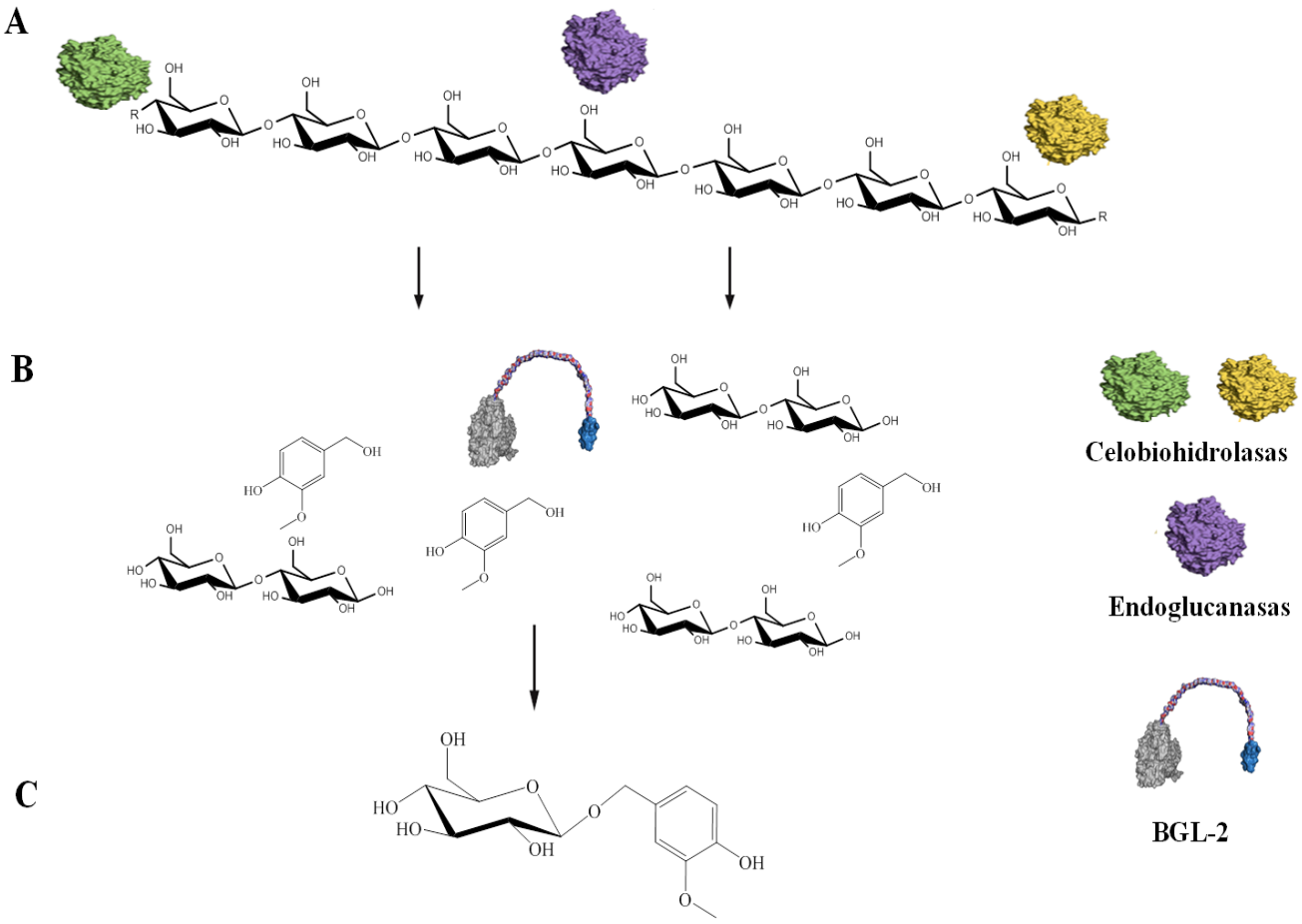



**Figura 2.** Esquema del funcionamiento de una hipotética cascada enzimática para la transglicosilación del alcohol vanillínico, combinando celobiohidrolasas y endoglucanasas con la enzima BGL-2. El apartado **A** muestra el paso inicial de la cascada enzimática, en la que se llevaría a cabo la degradación de la celulosa por endoglucanasas y celobiohidrolasas. En el apartado **B**, en un segundo paso, la BGL-2, utilizaría como molécula aceptora la celobiosa, generada en el primer paso, y catalizaría la transglicosilación del alcohol vanillínico, que se añadiría a la reacción (**C**).

En este sentido, cabe resaltar que en nuestro grupo de investigación se ha ensayado con éxito este protocolo para la producción de oligosacáridos a partir de paja de trigo, en este caso de xilooligosacáridos, utilizando una endoxilanasa y una β-xilosidasa que, en presencia del aceptor adecuado, produce un glicósido con propiedades antiproliferativas (Nieto-Domínguez *et al*., 2019).

Con el fin de obtener endoglucanasas y celobiohidrolasas en suficiente cantidad como para desarrollar estas aplicaciones, ya se están expresando estas enzimas en *P. pastoris*, esperando conseguir el mismo éxito en su producción que con las BGLs.

# CONCLUSIONES/CONCLUSIONS

# CONCLUSIONES

**1.** El estudio genómico y proteómico del hongo *T. amestolkiae* indica que tiene un gran número de glicosilhidrolasas implicadas en la degradación de la lignocelulosa, más que *T. reesei* y *A. niger*, hongos que se utilizan para producir cócteles enzimáticos para tratamiento de biomasa vegetal.

**2.** Los análisis proteómicos de los secretomas de *T. amestolkiae* revelaron la producción de dos BGLs mayoritarias. Una de ellas se produjo independientemente de la fuente de carbono utilizada en el cultivo (BGL-3), mientras la otra se induce por la presencia de celulosa en los cultivos (BGL-2). Los crudos de este hongo, en el que además de BGL existen otras enzimas implicadas en la degradación de la lignocelulosa, resultaron muy eficaces para sacarificar *slurry* de paja de trigo.

**3.** La levadura *P. pastoris* es un buen sistema de expresión heteróloga para las β-glucosidasas de *T. amestolkiae*. En todos los casos, se lograron cepas transformadas capaces de producir niveles de actividad superiores a los detectados en *T. amestolkiae*, y las proteínas purificadas igualaron o incluso mejoraron (en el caso de BGL-2) las propiedades catalíticas de las enzimas nativas.

**4.** La β-glucosidasa BGL-2 de *T. amestolkiae*, que pertenece a la familia GH3, presenta un dominio de unión a celulosa. Es la primera β-1,4-glucosidasa en la que se ha caracterizado este dominio, inusual en estas enzimas. Esta enzima, en sus distintas formas (nativa, recombinante y recombinante truncada), se utilizó para suplementar un cóctel celulolítico basal, Celluclast 1.5L (Novozymes), mostrando similar rendimiento en la sacarificación del bagazo de cerveza que cuando se utilizó en su suplementación una β-glucosidasa comercial (NS-50010).

**5.** La β-glucosidasa BGL-3, también de la familia GH3, es una enzima versátil, capaz de hidrolizar muy eficientemente todos los sustratos típicos de las β-1,4 glucosidasas y, además, glucanos con enlaces β-1,3, como la laminarina. La suplementación del cóctel basal, Celluclast 1.5L, con esta enzima mostró que es más eficiente que la β-glucosidasa comercializada por Novozymes (NS-50010), en la sacarificación de *slurry* de paja de trigo. Además, BGL-3 mejoró los rendimientos de una laminarinasa comercial hidrolizando laminarina.

**6.** Las distintas formas de BGL-2 y la BGL-3 mostraron la capacidad de transglicosilar una gran variedad de aceptores, entre los que se incluyen

compuestos fenólicos de gran interés para distintas aplicaciones. La BGL-2 se utilizó para obtener glucósidos de hidroxitirosol y alcohol vainillínico, compuestos potencialmente antiproliferativos y, utilizando modelos de células de cáncer de mama, se demostró una mejora en la capacidad antiproliferativa del derivado de vainillina y un perfil de seguridad mejorado de ambos glucósidos.

**7.** La β-glucosidasa BGL-1 de *T. amestolkiae*, enzima minoritaria en los cultivos, se expresó en *P. pastoris*. Esta enzima, que pertenece a la familia GH1, mostró una de las glucotolerancias más elevadas descubiertas hasta la fecha, aunque su afinidad por los celooligosacáridos fue menor que la encontrada en BGL-2 y BGL-3. Esto podría explicarse porque BGL-1 es más eficaz actuando sobre enlaces β-1,2 que sobre los β-1,4.

**8.** La BGL-1 recombinante mostró una capacidad transglicosiladora limitada. Sin embargo, sus variantes glucosintasas E521G y E521S fueron eficaces transglicosilando distintos compuestos. La glucosintasa E521G, con la que se obtuvieron mejores resultados, fue capaz de glicosilar eficazmente galato de epigalocatequina, un antioxidante fenólico que despierta un enorme interés en la actualidad por sus aplicaciones terapeúticas.

**9.** En este trabajo se ha puesto de manifiesto el potencial biotecnológico de las BGLs de *T. amestolkiae*, tanto para la hidrólisis de residuos lignocelulósicos, como para la síntesis de nuevos glicósidos de interés para distintos sectores industriales.

# CONCLUSIONS

**1.** The genomic and proteomic study of the fungus *T. amestolkiae* indicates that it can produce a large number of glycosyl hydrolases involved in the degradation of lignocellulose, more than other lignocellulolytic fungus like *T. reesei* and *A. niger*, that are used to produce enzymatic cocktails for vegetal biomass treatment.

**2.** Proteomic analysis of *T. amestolkiae* secretomes revealed the production of two major BGLs. One of them was produced independently of the carbon source used in the culture (BGL-3), while the other is induced by the presence of cellulose in the cultures (BGL-2). The enzymatic cocktails generated by this fungus were very effective in the saccharification of pretreated wheat straw.

**3.** The yeast *P. pastoris* is a good heterologous expression system for the β-glucosidases from *T. amestolkiae*. In all cases, the transformed strains were able to produce BGL activity levels higher than those detected in *T. amestolkiae*, and purified proteins matched or improved (in the case of BGL-2) the catalytic properties of native enzymes.

**4.** The β-glucosidase BGL-2 of *T. amestolkiae*, which belongs to the GH3 family, has a cellulose binding domain. This is the first 1,4-β-glucosidase in which this domain, unusual in these enzymes. The enzyme, in its different forms (native, recombinant and truncated recombinant), and a commercial β-glucosidase (NS-50010), were used to supplement a basal cellulolytic cocktail, Celluclast 1,5L (Novozymes), to saccharify brewers spent grain. The results showed similar saccharification yields using the basal cocktail with the enzymes of *T. amestolkiae* or the commercial BGLs provided by Novozymes.

**5.** The β-glucosidase BGL-3, also member of the GH3 family, was able to efficiently hydrolyze all typical substrates of β-1,4 glucosidases, but also glucans with β-1,3 bonds, such as laminarin. The supplementation of the basal cocktail, Celluclast 1,5L, with BGL-3 from *T. amestolkiae* demonstrated a higher efficiency than the commercial β-glucosidase from Novozymes (NS-50010), in the saccharification of pretreated wheat straw. In addition, this enzyme

improved the yields of a commercial laminarinase when hydrolyzing laminarin.

**6.** The different forms of BGL-2 and the BGL-3 showed the ability to transglycosylate a wide variety of acceptors, including phenolic compounds of great interest for different applications. BGL-2 was used to obtain hydroxytyrosol and vanillinic alcohol glycosides, which are potentially antiproliferative compounds. Using breast cancer cell models, an improvement in the antiproliferative capacity of the vanillin derivative was demonstrated, and the safety of both glycosides with respect to parental compounds was also improved.

**7.** The sequence of β-glucosidase BGL-1 from *T. amestolkiae*, a minoritary enzyme in secretomes, was obtained frrom the genome of the fungus. This enzyme, which belongs to the GH1 family, was produced in *P. pastoris*, and showed one of the highest glucotolerances discovered to date, although its affinity for cellooligosaccharides was much lower than that found in BGL-2 and BGL-3. This could be explained because BGL-1 is more effective acting on β-1,2 bonds than β-1,4.

**8.** BGL-1 showed a limited capacity for transglycosylation. However, its E521G and E521S glucosynthase variants were effective transglycosylating different compounds. In this sense, BGL-1-E521G, which showed better transglycosylation yields, was able to effectively glycosylate epigallocatechin gallate, a phenolic antioxidant that today arouses great interest today for its therapeutic applications.

**9.** Finally, it should be remarked that this work has revealed the biotechnological potential of *T. amestolkiae* BGLs, both for the hydrolysis of lignocellulosic residues, and for the synthesis of novel glycosides of interest for different industrial sectors.

# BIBLIOGRAFÍA

Alfani, F., Gallifuoco, A., Saporosi, A., Spera, A., Cantarella, M., 2000. Comparison of SHF and SSF processes for the bioconversion of steam-exploded wheat straw. J. Ind. Microbiol. Biot. 25, 184-192.

Alvira, P., Tomás-Pejó, E., Ballesteros, M., Negro, M.J., 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. Bioresource Technol. 101, 4851-4861.

Amaya, M.F., Watts, A.G., Damager, I., Wehenkel, A., Nguyen, T., Buschiazzo, A. *et al*., 2004. Structural insights into the catalytic mechanism of *Trypanosoma cruzi* trans-sialidase. Structure. 12, 775-784.

Amorim, H.V., Lopes, M.L., de Castro Oliveira, J.V., Buckeridge, M.S., Goldman, G.H., 2011. Scientific challenges of bioethanol production in Brazil. Appl. Microbiol. Biotechnol. 91, 1267.

Andlar, M., Rezic, T., Mardetko, N., Kracher, D., Ludwig, R., Santek, B., 2018. Lignocellulose degradation: An overview of fungi and fungal enzymes involved in lignocellulose degradation. Eng. Life. Sci. 18, 768-778.

Arantes, V., Saddler, J.N., 2010. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. Biotechnol. Biofuels. 3, 4.

Avellar, B.K., Glasser, W.G., 1998. Steam-assisted biomass fractionation: Process considerations and economic evaluation. Biomass. Bioener. 14, 205-218.

Azadi, P., Inderwildi, O.R., Farnood, R., King, D.A., 2013. Liquid fuels, hydrogen and chemicals from lignin: A critical review. Renew. Sust. Energ. Rev. 21, 506-523.

Baba, Y., Shimonaka, A., Koga, J., Kubota, H., Kono, T., 2005. Alternative splicing produces two endoglucanases with one or two carbohydrate-binding modules in *Mucor circinelloides*. J. Bacteriol. 187, 3045-3051.

Bader, G.D., Heilbut, A., Andrews, B., Tyers, M., Hughes, T., Boone, C., 2003. Functional genomics and proteomics: charting a multidimensional map of the yeast cell. Trends. Cell. Biol. 13, 344-356.

Bailey, J.E., Ollis, D.F.1986. Biochemical engineering fundamentals, McGraw-Hill.

Baruah, J., Nath, B., Sharma, R., Kumar, S., Deka, R., Baruah, D. *et al*., 2018. Recent trends in the pretreatment of lignocellulosic biomass for value-added products. Front. Energ. Res. 141.

Becker, E.G., Gusakov, A.V., Sinitsyn, A.P., 1991. Transglycosylation reactions catalysed by beta-glucosidase from *Aspergillus japonicus* in the presence of lignin model compounds. Prikl. Biokhim. Mikrobiol. 27, 482-485.

Beckham, G.T., Matthews, J.F., Bomble, Y.J., Bu, L., Adney, W.S., Himmel, M.E. *et al*., 2010. Identification of amino acids responsible for processivity in a family 1 carbohydrate-binding module from a fungal cellulase. J. Phy. Chem. B. 114, 1447-1453.

Berlin, A., Balakshin, M., Gilkes, N., Kadla, J., Maximenko, V., Kubo, S. *et al*., 2006. Inhibition of cellulase, xylanase and β-glucosidase activities by softwood lignin preparations. J. Biotechnol. 125, 198-209.

Bhat, M.K., Bhat, S., 1997. Cellulose degrading enzymes and their potential industrial applications. Biotechnol. Adv. 15, 583-620.

Bhatia, Y., Mishra, S., Bisaria, V.S., 2002. Microbial β-glucosidases: cloning, properties, and applications. Crit. Rev. Biotechnol. 22, 375-407.

Birch, P.R., Sims, P.F., Broda, P., 1995. Substrate-dependent differential splicing of introns in the regions encoding the cellulose binding domains of two exocellobiohydrolase I-like genes in *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. 61, 3741-3744.

Bissaro, B., Monsan, P., Fauré, R., O'Donohue, M.J., 2015. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. Biochem. J. 467, 17.

Blanchette, R.A., 1991. Delignification by wood-decay fungi. Annu. Rev. Phytopathol. 29, 381-403.

Bohlin, C., Olsen, S.N., Morant, M.D., Patkar, S., Borch, K., Westh, P., 2010. A comparative study of activity and apparent inhibition of fungal β-glucosidases. Biotechnol. Bioeng. 107, 943-952.

Bohlin, C., Praestgaard, E., Baumann, M.J., Borch, K., Praestgaard, J., Monrad, R.N. *et al*., 2013. A comparative study of hydrolysis and transglycosylation activities of fungal β-glucosidases. Appl. Microbiol. Biotechnol. 97, 159-169.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J., 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382, 769-781.

Burmeister, W.P., Cottaz, S., Rollin, P., Vasella, A., Henrissat, B., 2000. High resolution X-ray crystallography shows that ascorbate Is a cofactor for myrosinase and substitutes for the function of the catalytic base. J. Biol. Chem. 275, 39385-39393.

Butterworth, P.J., 2005. Lehninger principles of biochemistry (4th ed.): Nelson, D., and Cox, M.. Biochem. Mol. Biol. Educ., 33: 74-75.

Cairns, T.C., Nai, C., Meyer, V., 2018. How a fungus shapes biotechnology: 100 years of *Aspergillus niger* research. Fungal. Biol. Biotechnol. 5, 13-13.

Campbell, J.A., Davies, G.J., Bulone, V., Henrissat, B., 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem. J. 326 ( Pt 3), 929-939.

Cao, L.c., Wang, Z.j., Ren, G.h., Kong, W., Li, L., Xie, W. *et al*., 2015. Engineering a novel glucose-tolerant β-glucosidase as supplementation to enhance the hydrolysis of sugarcane bagasse at high glucose concentration. Biotechnol. Biofuels. 8, 202.

Castro, A., Pereira Jr, N., 2009. Production, properties and application of cellulases in the hydrolysis of agroindustrial residues. Quím. Nova. 33, 181-188.

Chávez, R., Bull, P., Eyzaguirre, J., 2006. The xylanolytic enzyme system from the genus *Penicillium*. J. Biotechnol. 123, 413-433.

Chen, W., Zhong, L.x., Peng, X.w., Wang, K., Chen, Z.f., Sun, R.c., 2014. Xylan-type hemicellulose supported palladium nanoparticles: a highly efficient and reusable catalyst for the carbon-carbon coupling reactions. Catal. Sci. Technol. 4, 1426-1435.

Cherry, J.R., Fidantsef, A.L., 2003. Directed evolution of industrial enzymes: an update. Curr. Opin. Biotechnol. 14, 438-443.

Cherubini, F., Strømman, A.H., 2011. Chemicals from lignocellulosic biomass: opportunities, perspectives, and potential of biorefinery systems. Biofuels. Bioprod. Bioref. 5, 548-561.

Chinnadurai, K., Muthukumarappan, K., Julson, J., 2008. Influence of high shear bioreactor parameters on carbohydrate release from different biomasses. Providence, Rhode Island.

Chirumamilla, R.R., Muralidhar, R., Marchant, R., Nigam, P., 2001. Improving the quality of industrially important enzymes by directed evolution. Mol. Cell. Biochem 224, 159-168.

Chundawat, S., Bellesia, G., Uppugundla, N., Sousa, L., Gao, D., Cheh, A. *et al*., 2011. Restructuring the crystalline cellulose hydrogen bond network enhances its depolymerization rate. J. Am. Chem. Soc. 133, 11163-11174.

Costanzo, S., Ospina-Giraldo, M.D., Deahl, K.L., Baker, C.J., Jones, R.W., 2007. Alternate intron processing of family 5 endoglucanase transcripts from the genus *Phytophthora*. Curr. Genet. 52, 115-123.

Culleton, H., McKie, V., de Vries, R.P., 2013. Physiological and molecular aspects of degradation of plant polysaccharides by fungi: what have we learned from *Aspergillus*? Biotechnol. J. 8, 884-894.

Danby, P.M., Withers, S.G., 2016. Advances in enzymatic glycoside synthesis. ACS Chem. Biol. 11, 1784-1794.

Davies, G.J., Wilson, K.S., Henrissat, B., 1997. Nomenclature for sugar-binding subsites in glycosyl hydrolases. Biochem. J. 321 ( Pt 2), 557-559.

Davies, G., Henrissat, B., 1995. Structures and mechanisms of glycosyl hydrolases. Structure 3, 853-859.

de Albuquerque de Carvalho, M.L., Carvalho, D.F., Barros Gomes, E., Nobuyuki Maeda, R., Melo Santa Anna, L.M., Castro, A.M.d. *et al*., 2014. Optimisation of cellulase production by *Penicillium funiculosum* in a stirred tank bioreactor using multivariate response surface analysis. Enzyme Res. 703291.

de Eugenio, L.I., Méndez-Líter, J.A., los Ríos, V., Prieto, A., Martínez, M.J., 2018. β-1,4-endoglucanases from *Talaromyces amestolkiae*: Production of glucooligosaccharides from different β-glucans. Biocatal. Biotransfor. 36, 68-77.

de Oliveira, J.M., de Graaff, L.H., 2011. Proteomics of industrial fungi: trends and insights for biotechnology. Appl. Microbiol. Biotechnol. 225-237.

de Vries, R.P., Visser, J., 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol. Mol. Biol. Rev. 65, 497-522.

de Vries, R.P., Riley, R., Wiebenga, A., Aguilar-Osorio, G., Amillis, S., Uchima, C.A. *et al*., 2017. Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. Genome Biol. 18, 28.

Doherty, W.O.S., Mousavioun, P., Fellows, C.M., 2011. Value-adding to cellulosic ethanol: Lignin polymers. Ind. Crop. Prod. 33, 259-276.

Dong, X., Dong, M., Lu, Y., Turley, A., Jin, T., Wu, C., 2011. Antimicrobial and antioxidant activities of lignin from residue of corn stover to ethanol production. Ind. Crop. Prod. 34, 1629-1634.

Doyle, S., 2011. Fungal proteomics: from identification to function. FEMS Microbiol. Lett. 321, 1-9.

Ebringerová, A., 2005. Structural diversity and application potential of hemicelluloses. Macromol. Symp. 232, 1-12.

Edelman, J., 1956. The Formation of Oligosaccharides by Enzymic Transglycosylation 189-232. Adv. Enzymol. Relat. Subj. Biochem. 17:189-232.

Eklöf, J.M., Brumer, H., 2010. The XTH gene family: An update on enzyme structure, function, and phylogeny in xyloglucan remodeling. Plant. Physiol. 153, 456.

Eyzaguirre, J., Hidalgo, M., Leschot, A., 2005. β-Glucosidases from filamentous fungi: Properties, structure, and applications. Handbook of Carbohydrate Engineering 645-685.

Farahi, R.H., Charrier, A.M., Tolbert, A., Lereu, A.L., Ragauskas, A., Davison, B.H. *et al*., 2017. Plasticity, elasticity, and adhesion energy of plant cell walls: nanometrology of lignin loss using atomic force microscopy. Sci. Rep. 7, 152.

Finkenstadt, V.L., Millane, R.P., 1998. Crystal structure of valonia cellulose Ib. Macromolecules. 31, 7776-7783.

Foreman, P.K., Brown, D., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S. *et al*., 2003. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J. Biol. Chem. 278, 31988-31997.

Fowler, T., Brown Jr, R.D., 1992. The *bgI1* gene encoding extracellular β-glucosidase from *Trichoderma reesei* is required for rapid induction of the cellulase complex. Mol. Microbiol. 6, 3225-3235.

Frutuoso , M.A., 2013. A single amino acid residue determines the ratio of hydrolysis to transglycosylation catalyzed by beta-glucosidases. Protein Pep. Lett. 20, 102-106.

Fujii, T., Hoshino, T., Inoue, H., Yano, S., 2014. Taxonomic revision of the cellulose-degrading fungus *Acremonium cellulolyticus* nomen nudum to *Talaromyces* based on phylogenetic analysis. FEMS Microbiol. Lett. 351, 32-41.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D. *et al*., 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. Nature. 422, 859-868.

Galbe, M., Zacchi, G., 2007. Pretreatment of lignocellulosic materials for efficient bioethanol production, in: Olsson, L. (Ed.), Biofuels. Springer Berlin Heidelberg, pp. 41-65.

Gao, D., Haarmeyer, C., Balan, V., Whitehead, T.A., Dale, B.E., Chundawat, S.P., 2014. Lignin triggers irreversible cellulase loss during pretreated lignocellulosic biomass saccharification. Biotechnol. Biofuels. 7, 175.

Gardner, K.H., Blackwell, J., 1974. The structure of native cellulose. Biopolymers. 13, 1975-2001.

Gibson, G.R., Probert, H.M., Loo, J.V., Rastall, R.A., Roberfroid, M.B., 2004. Dietary modulation of the human colonic microbiota: updating the concept of prebiotics. Nutr. Res. Rev. 17:259-275.

Gil Muñoz, J., 2015. Estudio de las β-glucosidasas del complejo celulolítico de *Talaromyces amestolkiae*: caracterización y aplicaciones biotecnológicas, Universidad Complutense de Madrid.

Gilbert, H.J., Knox, J.P., Boraston, A.B., 2013. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. Curr. Opin. Structl. Biol. 23, 669-677.

Giusti, L., 2009. A review of waste management practices and their impact on human health. Waste Manag. 29, 2227-2239.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H. *et al*., 1996. Life with 6000 Genes. Science 274, 546.

González-Alfonso, J.L., Rodrigo-Frutos, D., Belmonte-Reche, E., Peñalver, P., Poveda, A., Jiménez-Barbero, J. *et al*., 2018. Enzymatic synthesis of a novel pterostilbene a-Glucoside by the combination of cyclodextrin glucanotransferase and amyloglucosidase. Molecules. 23, 1271.

González-Blasco, G., Sanz-Aparicio, J., González, B., Hermoso, J.A., Polaina, J., 2000. Directed evolution of β-glucosidase A from *Paenibacillus polymyxa* to thermal resistance. J. Biol. Chem. 275, 13708-13712.

Gorke, J., Srienc, F., Kazlauskas, R., 2010. Toward advanced ionic liquids. Polar, enzyme-friendly solvents for biocatalysis. Biotechnol. Bioprocess Eng. 15, 40-53.

Goyari, S., Devi, S.H., Bengyella, L., Khan, M., Sharma, C.K., Kalita, M.C. *et al*., 2015. Unveiling the optimal parameters for cellulolytic characteristics of *Talaromyces verruculosus* SGMNP f3 and its secretory enzymes. J. Appl. Microbiol. 119, 88-98.

Gray, K.A., Zhao, L., Emptage, M., 2006. Bioethanol. Curr. Opin. Chem. Biol. 10, 141-146.

Guillén, F., Martinez, A.T., Martinez, M.J., 1990. Production of hydrogen peroxide by aryl-alcohol oxidase from the ligninolytic fungus *Pleurotus eryngii*. Appl. Microbiol. Biotechnol. 32, 465-469.

Guo, M., Jin, T., Nghiem, N.P., Fan, X., Qi, P.X., Jang, C.H. *et al*., 2018. Assessment of antioxidant and antimicrobial properties of lignin from corn stover residue pretreated with low-moisture anhydrous ammonia and enzymatic hydrolysis process. Appl. Biochem. Biotechnol. 184, 350-365.

Gusakov, A.V., 2011. Alternatives to *Trichoderma reesei* in biofuel production. Trends. Biotechnol. 29, 419-425.

Hannon, M., Gimpel, J., Tran, M., Rasala, B., Mayfield, S., 2010. Biofuels from algae: challenges and potential. Biofuels. 1, 763-784.

Harris, P.V., Welner, D., McFarland, K.C., Re, E., Navarro Poulsen, J.C., Brown, K. *et al*., 2010. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: Structure and function of a large, enigmatic family. Biochemistry. 49, 3305-3316.

Hawksworth, D.L., 1991. The fungal dimension of biodiversity: magnitude, significance, and conservation. Mycol. Res. 95, 641-655.

Hayashida, S., Mo, K., Hosoda, A., 1988. Production and characteristics of Avicel-digesting and non-Avicel-digesting cellobiohydrolases from *Aspergillus ficum*. Appl. Environ. Microbiol. 54, 1523-1529.

Hayes, M.R., Pietruszka, J., 2017. Synthesis of glycosides by glycosynthases. Molecules. 22, 1434.

Hendriks, A.T.W.M., Zeeman, G., 2009. Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour. Technol. 100, 10-18.

Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280 ( Pt 2), 309-316.

Henrissat, B., Davies, G., 1997. Structural and sequence-based classification of glycoside hydrolases. Curr. Opin.Struct. Biol. 7, 637-644.

Hildén, L., Johansson, G., 2004. Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. Biotechnol. Lett. 26, 1683-1693.

Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W. *et al.*, 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science. 315, 804.

Hong, J., Tamaki, H., Kumagai, H., 2007. Cloning and functional expression of thermostable β-glucosidase gene from *Thermoascus aurantiacus*. Appl. Microbiol. Biotechnol. 73, 1331-1339.

Hou, Y., Wang, T., Long, H., Zhu, H., Xu, M., 2007. Cloning, sequencing and expression analysis of the first cellulase gene encoding cellobiohydrolase 1 from a cold-adaptive *Penicillium chrysogenum* FS010. Acta. Biochim. Biophys. Sin. 39, 101-107.

Ivanova, G.S., Beletskaia, O.P., Okunev, O.N., Golovlev, E.L., Kulaev, I.S., 1983. Fractionation of the cellulase complex of the fungus *Aspergillus terreus*. Prikl. Biokhim. Mikrobiol. 19, 362-368.

Jahn, M., Marles, J., Warren, R.A., Withers, S.G., 2003. Thioglycoligases: mutant glycosidases for thioglycoside synthesis. Angew. Chem. Int. Ed. 42, 352-354.

Jahn, M., Withers, S.G., 2003. New approaches to enzymatic oligosaccharide synthesis: glycosynthases and thioglycoligases. Biocatal. Biotransfor. 21, 159-166.

Jenkins, J., Lo Leggio, L., Harris, G., Pickersgill, R., 1995. β-Glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β-strands four and seven. FEBS Lett. 362, 281-285.

Jeoh, T., Michener, W., Himmel, M.E., Decker, S.R., Adney, W.S., 2008. Implications of cellobiohydrolase glycosylation for use in biomass conversion. Biotechnol. Biofuels. 1, 10.

Jones, C.S., Mayfield, S.P., 2012. Algae biofuels: versatility for the future of bioenergy. Curr. Opin. Biotechnol. 23, 346-351.

Jørgensen, H., Eriksson, T., Börjesson, J., Tjerneld, F., Olsson, L., 2003. Purification and characterization of five cellulases and one xylanase from *Penicillium brasilianum* IBT 20888. Enzyme Microb. Technol. 32, 851-861.

Juers, D.H., Matthews, B.W., Huber, R.E., 2012. LacZ β-galactosidase: structure and function of an enzyme of historical and molecular biological importance. Protein Sci. 21(12): 1792-1807.

Juturu, V., Wu, J.C., 2014. Microbial cellulases: Engineering, production and applications. Renew. Sustain. Energy Rev. 33, 188-203.

Kamerling, J., 2007. Basic concepts and nomenclature recommendations in carbohydrate chemistry, in:Volume 1: Introduction to Glycoscience; Synthesis of Carbohydrates. Elsevier, pp. 1-37.

Karkehabadi, S., Helmich, K.E., Kaper, T., Hansson, H., Mikkelsen, N.E., Gudmundsson, M. *et al*., 2014. Biochemical characterization and crystal structures of a fungal family 3 β-glucosidase, Cel3A from *Hypocrea jecorina*. J. Biol. Chem. 31624-31637.

Kato, N., Suyama, S., Shirokane, M., Kato, M., Kobayashi, T., Tsukagoshi, N., 2002. Novel alpha-glucosidase from *Aspergillus nidulans* with strong transglycosylation activity. Appl. Environ. Microbiol. 68, 1250-1256.

Kirk, T.K., Farrell, R.L., 1987. Enzymatic "combustion": The microbial degradation of lignin. Annu. Rev. Microbiol. 41, 465-501.

Klemm, D., Heublein, B., Fink, H.P., Bohn, A., 2005. Cellulose: fascinating biopolymer and sustainable raw material. Angew. Chem. Int. Ed. 44, 3358-3393.

Kometani, T., Terada, Y., Nishimura, T., Takii, H., Okada, S., 1994. Transglycosylation to hesperidin by cyclodextrin glucanotransferase from an alkalophilic bacillus species in alkaline pH and properties of hesperidin glycosides. Biosci. Biotechnol. Biochem. 58, 1990-1994.

Koshland, J., 1953. Stereochemistry and the mechanism of enzymatic reactions. Biol. Rev. 28, 416-436.

Kuhad, R.C., Gupta, R., Singh, A., 2011. Microbial cellulases and their industrial applications. Enzyme. Res. 280696-280696.

Kumar, R., Wyman, C., 2009. Effects of cellulase and xylanase enzymes on the deconstruction of solids from pretreatment of poplar by leading technologies. Biotechnol. Prog. 25, 302-314.

Langan, P., Nishiyama, Y., Chanzy, H., 2001. X-ray structure of mercerized cellulose II at 1 Å resolution. Biomacromolecules. 2, 410-416.

Langsford, M.L., Gilkes, N.R., Singh, B., Moser, B., Miller, R.C., Warren, R.A.J. *et al*., 1987. Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. FEBS Lett. 225, 163-167.

Larsbrink, J., Izumi, A., Hemsworth, G.R., Davies, G.J., Brumer, H., 2012. Structural enzymology of *Cellvibrio japonicus* Agd31B protein reveals a-transglucosylase activity in glycoside hydrolase family 31. J. Biol. Chem. 43288-43299.

Laureano-Perez, L., Teymouri, F., Alizadeh, H., Dale, B.E., 2005. Understanding factors that limit enzymatic hydrolysis of biomass. Appl .Biochem. Biotechnol. 124, 1081-1099.

Lebbink, J.H., Kaper, T., Bron, P., van der Oost, J., de Vos, W.M., 2000. Improving low-temperature catalysis in the hyperthermostable Pyrococcus furiosus β-glucosidase CelB by directed evolution. Biochemistry. 39, 3656-3665.

Lee, H.L., Chang, C.K., Jeng, W.Y., Wang, A., Liang, p.h., 2012. Mutations in the substrate entrance region of β-glucosidase from Trichoderma reesei improve enzyme activity and thermostability. Protein. Eng. Des. Sel. 25.

Lee, Y.C., Dutta, S., Wu, K.C.W., 2014. Integrated, cascading enzyme-/chemocatalytic cellulose conversion using catalysts based on Mesoporous silica nanoparticles. Chem. Sus. Chem. 7, 3241-3246.

Leschine, S.B., 1995. Cellulose degradation in anarerobic environments. Annu. Rev. Microbiol. 49, 399-426.

Levasseur, A., Drula, E., Lombard, V., Coutinho, P.M., Henrissat, B., 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol. Biofuels. 6, 41.

Levy, J. 1966. The soft rot fungi: Their mode of action and significance in the degradation of wood, in: Preston, R.D. (Ed.), Adv. Bot. Res. Academic Press, pp. 323-357.

Li, Q., He, Y.C., Xian, M., Jun, G., Xu, X., Yang, J.M. *et al*., 2009. Improving enzymatic hydrolysis of wheat straw using ionic liquid 1-ethyl-3-methyl imidazolium diethyl phosphate pretreatment. Bioresour. Technol. 100, 3570-3575.

Liu, D., Zhang, R., Yang, X., Zhang, Z., Song, S., Miao, Y. *et al*., 2012. Characterization of a thermostable β-glucosidase from *Aspergillus fumigatus* Z5, and its functional expression in *Pichia pastoris* X33. Microb. Cell. Fact. 11, 25-25.

Lima, M.A., Oliveira-Neto, M., Kadowaki, M.A., Rosseto, F.R., Prates, E.T., Squina, F.M. *et al*., 2013. *Aspergillus niger* β-glucosidase has a cellulase-like tadpole molecular shape: insights into glycoside hydrolase family 3 (GH3) β-glucosidase structure and function. J. Biol. Chem. 32991-33005.

Linder, M., Teeri, T.T., 1997. The roles and function of cellulose-binding domains. J. Biotechnol. 57, 15-28.

Liu, G., Zhang, L., Wei, X., Zou, G., Qin, Y., Ma, L. *et al*., 2013. Genomic and secretomic analyses reveal unique features of the lignocellulolytic enzyme system of *Penicillium decumbens*. PloS ONE 8, e55185.

Liu, L., Qian, M., Song, P., Huang, G., Yu, Y., Fu, S., 2016. Fabrication of green lignin-based flame retardants for enhancing the thermal and fire retardancy properties of polypropylene/wood composites. ACS Sust. Chem. Eng. 4, 2422-2431.

Liu, Q.P., Sulzenbacher, G., Yuan, H., Bennett, E.P., Pietz, G., Saunders, K. *et al*., 2007. Bacterial glycosidases for the production of universal red blood cells. Nat. Biotechnol. 25, 454-464.

Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P., Henrissat, B., 2010. A hierarchical classification of polysaccharide lyases for glycogenomics. Biochem. J. 432, 437.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. D490-D495.

Lynd, L.R., Weimer, P.J., van Zyl, W.H., Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. 66, 506-577.

Mackenzie, L.F., Wang, Q., Warren, R.A., Withers, S.G., 1998. Glycosynthases: mutant glycosidases for oligosaccharide synthesis. J. Am. Chem. Soc. 120, 5583-5584.

Maeda, R.N., Serpa, V.I., Rocha, V.A.L., Mesquita, R.A.A., Anna, L.M.M.S., de Castro, A.M. *et al*., 2011. Enzymatic hydrolysis of pretreated sugar cane bagasse using *Penicillium funiculosum* and *Trichoderma harzianum* cellulases. Process Biochem. 46, 1196-1201.

Mao, X., Hong, Y., Shao, Z., Zhao, Y., Liu, Z., 2010. A novel cold-active and alkali-stable β-Glucosidase gene isolated from the marine bacterium *Martelella mediterranea*. Appl. Biochem.Biotechnol. 162, 2136-2148.

Mark, B.L., Vocadlo, D.J., Knapp, S., Triggs-Raine, B.L., Withers, S.G., James, M.N.G., 2001. Crystallographic evidence for substrate-assisted catalysis in a bacterial β-Hexosaminidase. J. Biol. Chem. 276, 10330-10337.

Martínez, Á.T., 2016. How to break down crystalline cellulose. Science 352, 1050.

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

Martínez, Á.T., Ruiz-Dueñas, F.J., Martínez, M.J., del Río, J.C., Gutiérrez, A., 2009. Enzymatic delignification of plant cell wall: from nature to mill. Curr. Opin. Biotechnol. 20, 348-357.

Martínez, Á.T., Speranza, M., Ruiz-Dueñas, F.J., Ferreira, P., Camarero, S., Guillén, F. *et al*., 2005. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int. Microbiol. 8, 195-204.

Martinez, D., Berka, R.M., Henrissat, B., Saloheimo, M., Arvas, M., Baker, S.E. *et al*., 2008. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat. Biotechnol. 26, 553-560.

Martinez, D., Larrondo, L.F., Putnam, N., Gelpke, M.D.S., Huang, K., Chapman, J. *et al*., 2004. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. Nat. Biotechnol. 22, 695-700.

Martins, L.F., Kolling, D., Camassola, M., Dillon, A.J.P., Ramos, L.P., 2008. Comparison of *Penicillium echinulatum* and *Trichoderma reesei* cellulases in relation to their activity against various cellulosic substrates. Bioresour. Technol. 99, 1417-1424.

McCarter, J.D., Stephen Withers, G., 1994. Mechanisms of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. 4, 885-892.

Merry, A.H., Merry, C.L.R., 2005. Glycoscience finally comes of age. EMBO Rep. 6, 900-903.

Míguez, N., Ramírez-Escudero, M., Gimeno-Pérez, M., Poveda, A., Jiménez-Barbero, J., Ballesteros, A.O. *et al*., 2018. Fructosylation of hydroxytyrosol by the α-fructofuranosidase from *Xanthophyllomyces dendrorhous*: Insights into the molecular basis of the enzyme specificity. Chem. Cat. Chem. 10, 4878-4887.

Miller, G.L., 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31, 426-428.

Mori, T., Fujita, S., Okahata, Y., 1997. Transglycosylation in a two-phase aqueous-organic system with catalysis by a lipid-coated β-D-galactosidase. Carbohydr. Res. 298, 65-73.

Mosier, N., Hendrickson, R., Ho, N., Sedlak, M., Ladisch, M.R., 2005. Optimization of pH controlled liquid hot water pretreatment of corn stover. Bioresour. Technol. 96, 1986-1993.

Mulat, D.G., Dibdiakova, J., Horn, S.J., 2018. Microbial biogas production from hydrolysis lignin: insight into lignin structural changes. Biotechnol.Biofuels. 11, 61.

Murphy, L., Bohlin, C., Baumann, M.J., Olsen, S.N., Sørensen, T.H., Anderson, L. *et al*., 2013. Product inhibition of five *Hypocrea jecorina* cellulases. Enz. Microb. Technol. 52, 163-169.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A comprehensive review. Renew. Sustain. Energy Rev. 14, 578-597.

Nazir, A., Soni, R., Saini, H.S., Kaur, A., Chadha, B.S., 2010. Profiling differential expression of cellulases and metabolite footprints in *Aspergillus terreus*. Appl. Biochem. Biotechnol. 162, 538-547.

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J. *et al*., 2017. Complex pectin metabolism by gut bacteria reveals novel catalytic functions. Nature 544, 65.

Nelson, N., 1944. A photometric adaptation of the Somogyi method for the determination of glucose. J. Biol. Chem. 153, 375-380.

Nicotra, F., 1997. Synthesis of C-glycosides of biological interest, In Driguez H and Thiem J (eds.), Glycoscience synthesis of substrate analogs and mimetics. Springer Berlin Heidelberg.

Nieto-Domínguez, M., de Eugenio, L.I., Barriuso, J., Prieto, A., Fernández de Toro, B., Canales-Mayordomo, Á. *et al*., 2015. Novel pH-stable glycoside hydrolase family 3 β-xylosidase from *Talaromyces amestolkiae*: an enzyme displaying regioselective transxylosylation. Appl. Environ. Microbiol. 6380-6392.

Nieto-Domínguez, M., de Eugenio, L.I., York-Durán, M.J., Rodríguez-Colinas, B., Plou, F.J., Chenoll, E. *et al*., 2017. Prebiotic effect of xylooligosaccharides produced from birchwood xylan by a novel fungal GH11 xylanase. Food Chem. 232, 105-113.

Nieto-Domínguez, M., 2017. Estudio de xilanasas fúngicas para el aprovechamiento de la biomasa lignocelulósica, Universidad Complutense de Madrid.

Nieto-Domínguez, M., Martínez-Fernández, J.A., de Toro, B.F., Méndez-Líter, J.A., Cañada, F.J., Prieto, A. *et al*., 2019. Exploiting xylan as sugar donor for the synthesis of an antiproliferative xyloside using an enzyme cascade. Microb. Cell Fact. 18, 174.

Nieto-Domínguez, M., Prieto, A., Fernández de Toro, B., Cañada, F.J., Barriuso, J., Armstrong, Z. *et al*., 2016. Enzymatic fine-tuning for 2-(6-hydroxynaphthyl) β-d-xylopyranoside synthesis catalyzed by the recombinant β-xylosidase BxTW1 from *Talaromyces amestolkiae*. Microb. Cell Fact. 15, 171.

Pandey, K.R., Naik, S.R., Vakil, B.V., 2015. Probiotics, prebiotics and synbiotics- a review. J. Food Sci. Technol. 7577-7587.

Papatheofanous, M.G., Billa, E., Koullas, D.P., Monties, B., Koukios, E.G., 1995. Two-stage acid-catalyzed fractionation of lignocellulosic biomass in aqueous ethanol systems at low temperatures. Bioresour. Technol. 54, 305-310.

Pardo Mendoza, I. 2017. Diseño e ingeniería de lacasas quiméricas por evolución dirigida para el tratamiento de biomasa vegetal, Universidad Complutense de Madrid.

Payne, C.M., Knott, B.C., Mayes, H.B., Hansson, H., Himmel, M.E., Sandgren, M. *et al*., 2015. Fungal Cellulases. Chem. Rev. 115, 1308-1448.

Pazarlioglu, N.K., Sariisik, M., Telefoncu, A., 2005. Treating denim fabrics with immobilized commercial cellulases. Process Biochem. 40, 767-771.

Pei, X.Q., Yi, Z.L., Tang, C.G., Wu, Z.L., 2011. Three amino acid changes contribute markedly to the thermostability of β-glucosidase BglC from *Thermobifida fusca*. Bioresour. Technol. 102, 3337-3342.

Pérez, J.A., Ballesteros, I., Ballesteros, M., Sáez, F., Negro, M.J., Manzanares, P., 2008. Optimizing liquid hot water pretreatment conditions to enhance sugar recovery from wheat straw for fuel-ethanol production. Fuel. 87, 3640-3647.

Perugino, G., Trincone, A., Rossi, M., Moracci, M., 2004. Oligosaccharide synthesis by glycosynthases. Trends Biotechnol. 22, 31-37.

Peterson, R., Nevalainen, H., 2012. *Trichoderma reesei* RUT-C30: Thirty years of strain improvement. Microbiology. 158, 58-68.

Prodanovic, R., Milosavic, N., Sladic, D., Zlatovic, M., Bozic, B., Velìkovic, T.C. *et al*., 2005. Transglucosylation of hydroquinone catalysed by a-glucosidase from baker's yeast. J. Mol. Catal. B: Enzym. 35, 142-146.

Quiroz Castañeda, R., Folch-Mallol, J., 2011. Proteínas que remodelan y degradan la pared celular vegetal: perspectivas actuales. Biotecnol. Apl. 28, 194.

Ragauskas, A.J., Williams, C.K., Davison, B.H., Britovsek, G., Cairney, J., Eckert, C.A. *et al*., 2006. The path forward for biofuels and biomaterials. Science 311, 484.

Ramani, G., Meera, B., Vanitha, C., Rajendhran, J., Gunasekaran, P., 2015. Molecular cloning and expression of thermostable glucose-tolerant β-glucosidase of *Penicillium funiculosum* NCL1 in *Pichia pastoris* and its characterization. J. Ind. Microbiol. Biotechnol. 42, 553-565.

Reese, E.T., Siu, R.G.H., Levinson, H.S., 1950. The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. J. Bacteriol. 59, 485-497.

Reeves, W.P., 1996. Organic Chemistry, Third Edition (Wade, Jr., L.G.). J. Chem. Educ. 73, A13.

Rempel, B.P., Withers, S.G., 2008. Covalent inhibitors of glycosidases and their applications in biochemistry and biology. Glycobiol. 18, 570-586.

de Vries, R., Tsang, A., Grigoriev, I.V. 2018. Fungal genomics: methods and protocols. Springer Science+Business Media, LLC, part of Springer Nature.

Roth, J., 2002. Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. Chem. Rev. 102, 285-304.

Rye, C.S., Withers, S.G., 2000. Glycosidase mechanisms. Curr. Opin.Chem. Biol. 4, 573-580.

Salas, P., Poveda, A., Jiménez-Barbero, J., Parra, J., Comelles, F., Ballesteros, A. *et al*., 2011. Enzymatic synthesis of a-glucosides of resveratrol with surfactant activity. Adv. Synth. Catal. 353, 1077-1086.

Salvachúa, D., Martinez, A.T., Tien, M., López-Lucendo, M., García, F., Ríos, V. *et al*., 2013. Differential proteomic analysis of the secretome of *Irpex lacteus* and other white-rot fungi during wheat straw pretreatment. Biotechnol. Biofuels 6, 115.

Salvachúa, D., Prieto, A., López-Abelairas, M., Lu-Chau, T., Martínez, A.T., Martínez, M.J., 2011. Fungal pretreatment: An alternative in second-generation ethanol from wheat straw. Bioresour. Technol. 102, 7500-7506.

Samson, R.A., Yilmaz, N., Houbraken, J., Spierenburg, H., Seifert, K.A., Peterson, S.W. *et al*., 2011. Phylogeny and nomenclature of the genus *Talaromyces* and taxa accommodated in *Penicillium* subgenus *Biverticillium*. Stud. Mycol. 70, 159-183.

Sánchez, C., 2009. Lignocellulosic residues: Biodegradation and bioconversion by fungi. Biotechnol. Adv. 27, 185-194.

Scheller, H.V., Ulvskov, P., 2010. Hemicelluloses. Annu. Rev. Plant Biol. 61, 263-289.

Schülein, M., 1997. Enzymatic properties of cellulases from *Humicola insolens*. J. Biotechnol. 57, 71-81.

Shewale, J.G., 1982. β-Glucosidase: Its role in cellulase synthesis and hydrolysis of cellulose. Int. J. Biochem. 14, 435-443.

Shi, J., Chinn, M.S., Sharma-Shivappa, R.R., 2008. Microbial pretreatment of cotton stalks by solid state cultivation of *Phanerochaete chrysosporium*. Bioresour. Technol. 99, 6556-6564.

Singh, A.P., Singh, T., 2014. Biotechnological applications of wood-rotting fungi: A review. Biomass Bioenerg. 62, 198-206.

Singhania, R.R., Patel, A.K., Sukumaran, R.K., Larroche, C., Pandey, A., 2013. Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. Bioresour. Technol. 127, 500-507.

Singhania, R.R., Sukumaran, R.K., Patel, A.K., Larroche, C., Pandey, A., 2010. Advancement and comparative profiles in the production technologies using solid-state and submerged fermentation for microbial cellulases. Enz. Microb. Technol. 46, 541-549.

Solov'eva, I.V., Okunev, O.N., Vel'kov, V.V., Koshelev, A.V., Bubnova, T.V., Kondrat'eva, E.G. *et al*., 2005. The selection and properties of *Penicillium verruculosum* mutants with enhanced production of cellulases and xylanases. Microbiology 74, 141-146.

Sørensen, A., Lübeck, M., Lübeck, P.S., Ahring, B.K., 2013. Fungal Beta-glucosidases: a bottleneck in industrial use of lignocellulosic materials. Biomolecules. 3, 612-631.

Srisodsuk, M., Reinikainen, T., Penttilä, M., Teeri, T.T., 1993. Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. J. Biol. Chem. 268, 20756-20761.

Sukumaran, R., Singhania, R., Pandey, A., 2005. Microbial cellulases: Production, applications and challenges. J.Sci. Ind. Res. 64, 832-844.

Sun, Y., Cheng, J., 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour. Technol. 83, 1-11.

Suto, M., Yachi, M., Kamagata, Y., Sasaki, H., Takao, S., Tomita, F., 1991. Cellulase induction by soluble acetyl cellobioses in *Penicillium purpurogenum*. J. Ferment. Bioeng. 72, 352-357.

Suzuki, K., Sumitani, J.i., Nam, Y.W., Nishimaki, T., Tani, S., Wakagi, T. *et al*., 2013. Crystal structures of glycoside hydrolase family 3 β-glucosidase 1 from *Aspergillus aculeatus*. Biochem. J. 452, 211-221.

Teeri, T.T., Koivula, A., Linder, M., Wohlfahrt, G., Divne, C., Jones, T.A., 1998. *Trichoderma reesei* cellobiohydrolases: why so efficient on crystalline cellulose? Biochem. Soc. Trans. 26, 173.

Terwisscha van Scheltinga, A.C., Armand, S., Kalk, K.H., Isogai, A., Henrissat, B., Dijkstra, B.W., 1995. Stereochemistry of chitin hydrolysis by a plant chitinase/lysozyme and x-ray structure of a complex with allosamidin evidence for substrate assisted catalysis. Biochemistry. 34, 15619-15623.

Teugjas, H., Väljamäe, P., 2013. Selecting β-glucosidases to support cellulases in cellulose saccharification. Biotechnol. Biofuels. 6, 105-105.

Torres, P., Poveda, A., Jimenez-Barbero, J., Parra, J.L., Comelles, F., Ballesteros, A.O. *et al*., 2011. Enzymatic synthesis of a-glucosides of resveratrol with surfactant activity. Adv. Synth. Catal. 353, 1077-1086.

Uitdehaag, J.C.M., Mosi, R., Kalk, K.H., van der Veen, B.A., Dijkhuizen, L., Withers, S.G. *et al*., 1999. X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the a-amylase family. Nat. Struct. Biol. 6, 432-436.

Vaaje-Kolstad, G., Westereng, B., Horn, S.J., Liu, Z., Zhai, H., Sørlie, M. *et al*., 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. Science. 330, 219.

Vafiadi, C., Topakas, E., Christakopoulos, P., Faulds, C.B., 2006. The feruloyl esterase system of *Talaromyces stipitatus*: Determining the hydrolytic and synthetic specificity of TsFaeC. J. Biotechnol. 125, 210-221.

Vaishnav, N., Singh, A., Adsul, M., Dixit, P., Sandhu, S.K., Mathur, A. *et al*., 2018. *Penicillium*: The next emerging champion for cellulase production. Bioresour. Technol. Rep. 2, 131-140.

Vallmitjana, M., Ferrer-Navarro, M., Planell, R., Abel, M., Ausín, C., Querol, E. *et al*., 2001. Mechanism of the family 1 β-glucosidase from *Streptomyces sp*: Catalytic residues and kinetic studies. Biochemistry. 40, 5975-5982.

Valmaseda, M., Martinez, M.J., Martinez, A.T., 1991. Kinetics of wheat straw solid-state fermentation with *Trametes versicolor* and *Pleurotus ostreatus*: Lignin and polysaccharide alteration and. Appl. Microbiol. Biotechnol. 35, 817-823.

Van Dyk, J.S., Pletschke, B.I., 2012. A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-factors affecting enzymes, conversion and synergy. Biotechnol. Adv. 30, 1458-1480.

van Hijum, S.A.F.T., Kralj, S., Ozimek, L.K., Dijkhuizen, L., Geel-Schutten, I.G.H., 2006. Structure-function relationships of glucansucrase and fructansucrase enzymes from lactic acid bacteria. Microbiol. Mol. Biol. Rev. 70, 157-176.

Wada, M., Heux, L., Isogai, A., Nishiyama, Y., Chanzy, H., Sugiyama, J., 2001. Improved structural data of cellulose IIII prepared in supercritical ammonia. Macromolecules. 34, 1237-1243.

Watkins, D., Nuruddin, M., Hosur, M., Tcherbi-Narteh, A., Jeelani, S., 2015. Extraction and characterization of lignin from different biomass resources. J. Mat. Res. Technol. 4, 26-32.

Wierzbicka-Wos, A., Bartasun, P., Cieslinski, H., Kur, J., 2013. Cloning and characterization of a novel cold-active glycoside hydrolase family 1 enzyme with β-glucosidase, β-fucosidase and β-galactosidase activities. BMC Biotechnol. 13, 22.

Witzler, M., Alzagameem, A., Bergs, M., Khaldi-Hansen, B.E., Klein, S.E., Hielscher, D. *et al*., 2018. Lignin-derived biomaterials for drug release and tissue engineering. Molecules. 23, 1885.

Wolfenden, R., Lu, X., Young, G., 1998. Spontaneous hydrolysis of glycosides. J. Am. Chem. Soc. 120, 6814-6815.

Woo, H.J., Kang, H.K., Nguyen, T.T.H., Kim, G.E., Kim, Y.M., Park, J.S. *et al*., 2012. Synthesis and characterization of ampelopsin glucosides using dextransucrase from *Leuconostoc mesenteroides* B-1299CB4: Glucosylation enhancing physicochemical properties. Enz. Microb. Technol. 51, 311-318.

.

Wyman, C., 1996. Handbook on bioethanol: production and utilization, CRC press, Applied Energy Technology Series.

Xiao, Z., Zhang, X., Gregg, D.J., Saddler, J.N., 2004. Effects of sugar inhibition on cellulases and β-glucosidase during enzymatic hydrolysis of softwood substrates. Appl. Biochem.Biotechnol. 115, 1115-1126.

Yamamoto, K., Li, S.C., Li, Y.T., 2000. Microbial glycosidases, p. 497-511. In Ernst B, Hart GW, and Sinaÿ P (eds.), Carbohydrates in chemistry and biology. Wiley-VCH Verlag GmbH.

Yan, T.R., Lin, Y.H., Lin, C.L., 1998. Purification and characterization of an extracellular β-glucosidase II with high hydrolysis and transglucosylation activities from *Aspergillus niger*. J. Agric. Food Chem. 46, 431-437.

Yeoman, C.J., Han, Y., Dodd, D., Schroeder, C.M., Mackie, R.I., Cann, I.K.O., 2010. Thermostable enzymes as biocatalysts in the biofuel industry. Adv. Appl. Microbiol. 1-55.

Yip, V.L.Y., Varrot, A., Davies, G.J., Rajan, S.S., Yang, X., Thompson, J. *et al*., 2004. An unusual mechanism of glycoside hydrolysis involving redox and elimination steps by a Family 4 β-Glycosidase from *Thermotoga maritima*. J. Am. Chem. Soc. 126, 8354-8355.

Zhao, Z., Liu, H., Wang, C., Xu, J.R., 2013. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. BMC Genomics. 14, 274.

# ANEXO

Además de los artículos recogidos en este manuscrito, durante el periodo dedesarrollo de esta tesis doctoral se han elaborado otras dos publicaciones en las que ha participado el doctorando:

1. de Eugenio, L. I., Méndez-Líter, J. A., de los Ríos, V., Prieto A., Martínez, M. J., 2018. β-1,4-endoglucanases from *Talaromyces amestolkiae*: Production of glucooligosaccharides from different β-glucans, Biocatal. and Biotransfor. 36:1, 68-77, DOI: 10.1080/10242422.2017.1306741

2. Nieto-Domínguez, M., Martínez-Fernández, J. A., de Toro, B. F., Méndez-Líter, J. A., Cañada, F. J., Prieto, A., de Eugenio, L.I., Martínez, M. J., 2019. Exploiting xylan as sugar donor for the synthesis of an antiproliferative xyloside using an enzyme cascade. Microb. Cell Fact. 18(1), 174. doi:10.1186/s12934-019-1223-9