



UNIVERSIDAD DE CHILE – FACULTAD DE CIENCIAS – ESCUELA DE PREGRADO

“Aproximación a la validación de método bioinformático de termoestabilización, basado en la enzima exo-celobiohidrolasa I Cel7C del hongo *Phanerochaete chrysosporium*, mediante reemplazo de aminoácidos por cisteínas en su dominio catalítico y su expresión en *Pichia pastoris*.”

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

NATALIA BEATRIZ ARANDA SOTO

Directora del Seminario de Título:
Dra. Oriana Salazar Aguirre

Co-directora del Seminario de Título:
Dra. María Elena Lienqueo Contreras

Profesora Patrocinante:
Dra. Claudia Stange Klein

Mayo 2019
Santiago - Chile



UNIVERSIDAD DE CHILE – FACULTAD DE CIENCIAS – ESCUELA DE PREGRADO

INFORME DE APROBACIÓN SEMINARIO DE TITULO

Se informa a la Escuela de Pregrado de la Facultad de Ciencias, de la Universidad de Chile que el Seminario de Título, presentado por la **Srta. Natalia Beatriz Aranda Soto**.

“Aproximación a la validación de método bioinformático de termoestabilización, basado en la enzima exo-celobiohidrolasa I Cel7C del hongo *Phanerochaete chrysosporium*, mediante reemplazo de aminoácidos por cisteínas en su dominio catalítico y su expresión en *Pichia pastoris*.”

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

Dra. Oriana Salazar Aguirre

Director Seminario de Título: _____

Dra. María E. Lienqueo Contreras

Co-Director Seminario de Título: _____

Dra. Claudia Stange Klein

Patrocinante Seminario de Título: _____

Comisión Revisora y Evaluadora

Presidente Comisión: _____

Evaluador: _____

Santiago de Chile, 23 de Mayo 2019.



BIOGRAFÍA

Nací el 14 de mayo de 1993. Desde chiquitita me gustaba jugar a ser científica y hacía experimentos inventados por mi.

Siempre supe que mi futuro estaría relacionado a la ciencia. En el colegio me fue muy bien, salí como la mejor alumna. Cuando postulé a la Universidad, no quedé a través de la PSU, quedé gracias a una beca que me gané por mi desempeño en el colegio. Haber quedado en la Universidad de Chile, y en la carrera que quería en ese entonces, fue un regalo de la vida. En la universidad no solía ser una de las mejores alumnas, pero me fue bien. Siento que mi paso por la universidad fue un desafío, donde me enfrenté a facetas de la vida que amé con pasión y otras que se me hicieron muy difíciles, pero pude conocerme mejor, hacerme más fuerte, valorar mis capacidades, y fue un periodo en el cual principalmente, pude definir más claramente qué camino es el que quiero seguir.

DEDICATORIA

Le dedico este trabajo a mi mamá y su gran esfuerzo por sacarme adelante, y a mis abuelos maternos que también fueron una base muy importante en mi crianza.

AGRADECIMIENTOS

Mis agradecimientos son dedicados principalmente a mi profesora guía, la Dra. Oriana Salazar y al centro CeBiB, por darme la oportunidad de realizar mi Seminario de Título en dicho lugar, con el apoyo incondicional de mi profesora, su paciencia y su buena disposición ante cualquier requerimiento que yo tuviera.

También agradezco a mis seres queridos por estar siempre apoyándome y orientándome en los momentos en que me sentía perdida. (Mamá, abuelos, pololo, papá, tíos y primos)

Agradezco enormemente a la Universidad también por los lindos amigos que me entregó, que también han sido un apoyo fundamental y son con los que mejor nos hemos podido entender, ponernos en el lugar del otro y dar fuerzas en este proceso. (Nopito, Conisi, Cony Blanco, Engelina, Joaco, Sebita, Rama, Nyni, Gabo, Maca, Alonso)

INDICE DE CONTENIDOS

INTRODUCCIÓN.....	1
Rol de la Celulosa en la Industria.....	1
Síntesis y Estructura de la Celulosa	3
Celulasas: Complejo Enzimático Degradador de Celulosa	6
Familia 7 de las Glicosil Hidrolasas	8
Cel7C de <i>Phanerochaete chrysosporium</i>	9
Celulasas en la Industria Biotecnológica.....	12
Celulasas Termoestables.....	14
Microorganismos Termófilos.....	15
Ingeniería de Proteínas	16
<i>Antecedentes Para la Termoestabilización de Cel7C de P. chrysosporium</i>	17
Propuesta.....	19
Objetivos	20
MATERIALES Y MÉTODOS	21
Material Biológico.....	21
Técnicas de Biología Molecular.....	22
PCR.....	22
Ligación del DNA a los vectores	24
PCR de Colonias y Secuenciación de Plasmidios	25
Transformación y Selección de Clones Recombinantes	27
<i>Escherichia coli</i> TOP10	27
<i>Pichia pastoris</i> KM71	28
Expresión de la Proteína Cel7C y sus Variedades Mutantes.....	29
Confirmación de la Presencia de la Secuencia de cel7C y cel7C_M1M2, y de su Expresión.....	29
SDS-PAGE.....	30
Zimogramas	30
Ensayos de Actividad.....	31
RESULTADOS	32
Diseño experimental de Mutagénesis de Sitio Dirigido	32
Implementación de mutaciones en la secuencia codificante para el dominio catalítico de Cel7C	34
Generación de vectores de expresión derivados de pPIC9k.....	40
Expresión de las versiones silvestre y mutadas de Cel7CDC en <i>P. pastoris</i>	46
DISCUSIÓN.....	55
CONCLUSIONES.....	65
BIBLIOGRAFÍA.....	66
ANEXO.....	a

INDICE DE TABLAS

Tabla 1. Cepas de microorganismos, y plasmidios de clonación y de expresión utilizados en este trabajo.	21
Tabla 2. Caracterización y descripción de los partidores utilizados para los diferentes PCR.	22
Tabla 3. Descripción de los fragmentos generados por los diferentes partidores utilizados en la de mutagénesis sitio dirigida mediante PCR.	23
Tabla 4. Programa de PCR mutación sitio dirigida.	24
Tabla 5. Colonias de <i>E. coli</i> TOP10 con los vectores derivados de pBluescript II KS (+), secuenciadas.	40
Tabla 6. Colonias de <i>E. coli</i> TOP10 con los vectores derivados de pPIC9k, secuenciadas.	45
Tabla 7. Colonias de <i>P. pastoris</i> transformadas con los vectores derivados de pPIC9k, secuenciadas.	51

INDICE DE FIGURAS

Figura 1. Consumo de biomasa por la industria forestal en Chile según producto, periodo 2000-2015.	2
Figura 2. Representación estructural de la Celulosa.	5
Figura 3. Sinergismo entre Endoglucanasas (EG), Celobiohidrolasas (CBH) y β-Glucosidasas (BGL) en un sistema fúngico de degradación de celulosa.	7
Figura 4. Metodología de mutagénesis sitio dirigida mediante PCR.	17
Figura 5. Representación en 3D de Cel7CDC y de las seis mutaciones propuestas para rigidizar la estructura de la proteína.	19
Figura 6. Diseño experimental para la obtención de vectores derivados de pBluescript II SK (+) para clonación de los insertos.	33
Figura 7. Amplificación de los fragmentos utilizados para realizar mutación sitio dirigida.	34
Figura 8. Fragmentos de cel7CDC mutados, purificados.	35
Figura 9. Resultados de la transformación de <i>E. coli</i> Top10 con vectores recombinantes pBluescript II SK (+) que contenían las versiones silvestre y mutadas de cel7CDC.	36
Figura 10. PCR de colonias de las transformaciones de <i>E. coli</i> TOP10 con vectores derivados de pBluescript II SK (+).	37
Figura 11. Análisis de restricción de vectores recombinantes derivados de pBluescript II SK (+) con la enzima SmaI.	39
Figura 12. Extracto del alineamiento de secuenciaciones de plasmidios derivados de pBluescript II KS (+).	40
Figura 13. Diseño estratégico de los partidores para unir el gen cel7CDC y sus variedades al vector pPIC9k.	41
Figura 14. Resultados de la transformación de <i>E. coli</i> Top10 con vectores pPIC9K recombinantes que contenían las versiones silvestre y mutadas de cel7CDC.	43
Figura 15. PCR de colonias de las transformaciones de <i>E. coli</i> TOP10 con vectores derivados de pPIC9k.	43
Figura 16. Análisis de restricción de vectores recombinantes derivados de pPIC9k con la enzima EcoRV.	45
Figura 17. Detalle del alineamiento de secuenciaciones de plasmidios derivados de pPIC9k.	46

Figura 18. Inserción del vector pPIC9k al genoma de <i>P. pastoris</i> por recombinación homóloga.	48
Figura 19. Resultados de la transformación de <i>P. pastoris</i> KM71 con vectores pPIC9K recombinantes que contenían las versiones silvestre y mutadas de cel7CDC.	50
Figura 20. PCR de colonias de las transformaciones de <i>P. pastoris</i> KM71 con vectores derivados de pPIC9k.	50
Figura 21. Detalle del alineamiento de la secuenciación de productos de PCR, proveniente de <i>P. pastoris</i> KM71 transformadas con vectores derivados de pPIC9k.	51
Figura 22. SDS-PAGE de sobrenadantes de cultivo de <i>P. pastoris</i> KM71 transformada con pPIC9K_FC y pPIC9K_M1M2, inducidos con metanol por cuatro días a 28°C.	52
Figura 23. SDS-PAGE y Zimograma del sobrenadante del cuarto día de inducción de las colonias de <i>P. pastoris</i> transformadas.	53

ABREVIATURAS

cel7CDC: Secuencia genética codificante para el dominio catalítico de Cel7C.

Cel7CDC: Dominio catalítico de la proteína Cel7C.

CBH: Celobiohidrolasa

EG: Endoglucanasa

BGL: β -Glucosidasa

GH: Glicosil-hidrolasas

ADN: Ácido Desoxiribonucleico

ARN: Ácido Ribonucleico

CBD: Dominio de Unión a Carbohidratos

CMC: Carboximetilcelulosa

pNP-L: 4-nitrofenil- β -D-lactopiranosido

pNP-C: 4-nitrofenil- β -D-celobiosido

PCR: Reacción en cadena de la polimerasa

d-NTP: desoxirribonucleótidos trifosfato

LB: Luria Bertani

OD: Densidad Óptica

IPTG: Isopropil- β -D-1-tiogalactopiranosido

YPD: Yeast Extract-Peptone-Dextrose

BMGY: Medio complejo amortiguado con glicerol

BMMY: Medio complejo amortiguado con metanol

DTT: Ditioneitol

L, mL, uL: Unidades de Volumen

g, mg, ug, ng: Unidades de Masa

M, mM, uM: Unidades de Concentraciones Molares

pb, kb: Unidades de Tamaño Molecular en Pares de Bases Nitrogenadas

min, h: Unidades de Tiempo

°C: Grados Celsius, Unidad de Temperatura

K: Grados Kelvin, Unidad de Temperatura

V: Volts, Unidad de Potencial Eléctrico

rpm: revoluciones por minuto

UV: Ultravioleta

ANOVA: Análisis de Varianza

RESUMEN

Actualmente, el uso de enzimas celulasas, debido a su capacidad de degradar celulosa, ha crecido enormemente y ha sido reportado en industrias como la alimentaria, cervecera y vitivinícola, en la industria alimentaria animal, industrias textil y de detergentes, papelera, y de biocombustibles.

En la naturaleza estas enzimas son producidas por hongos y bacterias, y son principalmente secretadas al medio extracelular. Estas enzimas pertenecen a una familia compuesta de al menos tres grupos: exo-celobiohidrolasas (CBH), endo- β -1,4-glucanasa (EG) y β -glucosidasa (BG), las cuales forman parte de un complejo catalítico que actúa de manera sinérgica, en compañía de otras enzimas, degradando la biomasa. Debido al requerimiento de degradar celulosa por parte de diferentes industrias y ante las condiciones en las cuales se realiza este proceso, como la alta temperatura, el diseño de nuevas enzimas, realizando modificaciones sobre secuencias de proteínas mesófilas ya existentes sugiere ser una buena alternativa.

En un estudio anterior, con el fin de generar un método bioinformático predictivo para termoestabilización de proteínas, se analizó el efecto de la temperatura sobre la estructura de la exo-celobiohidrolasa I Cel7C del hongo *Phanerochaete chrysosporium* usando métodos de dinámica molecular. Como resultado de esto se propuso la introducción de tres puentes disulfuro, cada uno para una proteína Cel7CDC mutante diferente, los que deberían aumentar la rigidez de la estructura a altas temperaturas, y por lo tanto, su termoestabilidad.

Con el objetivo de corroborar experimentalmente el método de predicción bioinformática utilizado, en este trabajo se implementó uno de los tres puentes disulfuros propuestos para Cel7CDC. Para esto se realizó mutación sitio dirigida mediante la técnica de PCR con partidores que contenían la mutación en su interior, con objeto de generar dos mutaciones puntuales en el gen codificante. Con esto se reemplazó la fenilalanina 325 y la glicina 379 de la cadena aminoacídica de la proteína madura, por cisteínas. Las variantes del gen generadas fueron clonadas en los vectores pBluescript II SK (+) y fueron perpetuadas en *E. coli* TOP10, y también en el vector pPIC9k para ser expresadas en *P. pastoris* KM71. Se realizó la transformación y la selección de las variantes de *P. pastoris* se llevó a cabo por prototrofia en ausencia de histidina y luego con concentraciones crecientes de G418 según protocolos de selección del vector pPIC9k. La selección fue verificada mediante PCR de colonias y secuenciación. La expresión se llevó a cabo según el manual de expresión en *P. pastoris* con el vector, y tras ser analizadas mediante ensayos de actividad y zimogramas, no se encontró actividad asociada a ellas.

ABSTRACT

At present, the use of cellulose degrading enzymes, due to their degrading cellulose capacity, has hugely grown and has been reported in many industries including food and animal feed, brewing and wine production, textile, laundry and detergents industries and also in pulp and paper, and biofuel production.

In wildlife, cellulases are synthesized by fungi, bacteria, and are mostly secreted into the extracellular medium. These enzymes are a family composed of at least three groups: Endo- β -1,4-glucanases (EG), Exo-cellobiohydrolases (CBH) and β -glucosidases (BG), where all together are part of a catalytic complex acting synergistically, accompanied by other enzymes, degrading biomass. Due to the requirement of degrading cellulose by several industries and against the conditions this process is carried out, like the high temperatures needed, the design of new enzymes modifying sequences of mesophilic existing proteins, suggests to be a good option.

In a previous work, in order to generate a predictive bioinformatic method for protein thermostabilization, the temperature effect over the structure of the Exo-Cellobiohydrolase I protein, Cel7C, from the fungus *Phanerochaete chrysosporium*, was studied using molecular dynamic methods. Based on it, three disulfide bonds introduction were proposed, each for a single variant enzyme, which must increase the structure rigidity at high temperatures, and therefore, their thermostability.

To experimentally corroborate the bioinformatic prediction method proposed before, in this study, one of the three disulfide bonds was implemented for Cel7CDC. For this, two site directed mutations were made by PCR with primers that had the modification inside,

to generate punctual mutations into the gene sequence. Two aminoacids were replaced by Cysteines into the protein chain: 325 Phenylalanine and 379 Glycine. The different variants of the gene were cloned into the pBluescript II SK (+) and were perpetuated into *E. coli* TOP10 cells, and also in pPIC9k vector, for the expression in *P. pastoris* KM71. The transformation was made and the selection of *P. pastoris* variants was carried out by prototrophy in lack of histidine, and then with growing concentrations of G418, as it is mentioned in the pPIC9k vector expression protocol. The selection was verified by colony PCR and DNA sequencing. The proteins expression was carried out by multi-copy *Pichia* expression manual with the vector, and and after being analyzed through activity assays and zimograms, there was no associated activity found.

INTRODUCCIÓN

1. Rol de la Celulosa en la Industria.

La celulosa es el biopolímero más abundante y más renovable en la tierra utilizado como materia prima, y es el mayor residuo de la industria agrícola (Zhang y col., 2006). Se ha estimado que su producción anual a nivel mundial en células vegetales alcanza la cifra de aproximadamente $1,5 \cdot 10^{12}$ toneladas producidas mediante fotosíntesis (Sukumaran y col., 2005), implicando que una cantidad similar sea degradada anualmente.

La degradación de celulosa representa una parte importante dentro del ciclo del carbono en la biósfera. Naturalmente este proceso es llevado a cabo por microorganismos, entre ellos bacterias y hongos, tanto de manera aeróbica como anaeróbica, los cuales poseen enzimas que facilitan este proceso (Kuhad y col., 2011).

Actualmente, debido al uso de celulosa en diferentes industrias y la necesidad de su degradación, el uso de microorganismos implicados en este proceso o de las enzimas celulolíticas producidas por ellos se ha transformado en una herramienta de interés con aplicación industrial (Sukumaran y col., 2005).

En Chile, país con alto desarrollo en la industria forestal, gran parte de la madera obtenida es destinada a la producción de pulpa y papel (Figura 1). En el año 2015, la madera producida destinada a la producción de pulpa a través de procesos químicos creció en un 115,2%, reflejándose en un consumo equivalente al 36,5% del total de madera cosechado en trozas, y un porcentaje considerablemente menor es tratado para generar pulpa mediante procesos mecánicos (ODEPA – Gobierno de Chile, 2018).

Actualmente se suman a estos procesos químicos y mecánicos los tratamientos enzimáticos, incluyendo enzimas como xilanasas, lipasas, lacasas y celulasas (Kirk & Jeffries, 1996).

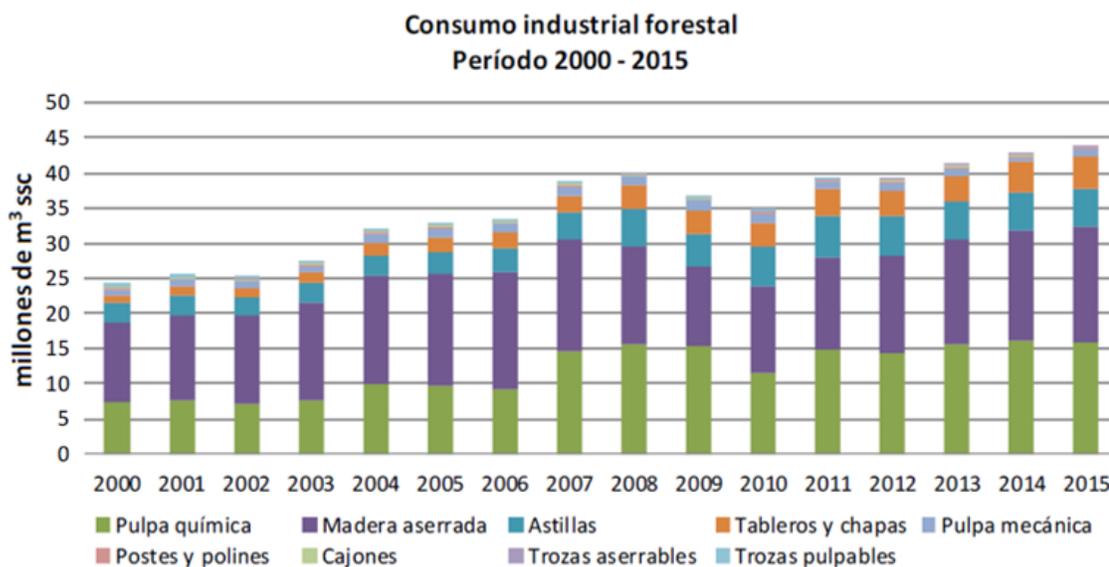


Figura 1. Consumo de biomasa por la industria forestal en Chile según producto, periodo 2000-2015. Ssc: sólidos sin corteza. Modificado de ODEPA – Gobierno de Chile, 2018.

Ante un uso más sustentable de los desechos provenientes de estas industrias, y ante una necesidad de optimizar el uso de energía disponible, es que actualmente estos están siendo utilizados por su alto contenido en azúcares, los cuales al ser fermentados por microorganismos permiten la generación de distintos productos de interés comercial e industrial. Dentro de los bioproductos de interés que son posibles de obtener a partir de la fermentación de azúcares liberados desde celulosa se encuentran aditivos para alimentos, ácidos orgánicos, glicerol, biocombustibles, etc. (Mussatto & Teixeira, 2010).

Utilizar la celulosa presente en los desechos lignocelulósicos de las industrias agrícolas y forestales para la obtención de biocombustibles es una forma eficiente de utilizar la energía. La celulosa es considerada energía solar almacenada como energía química,

con capacidad de ser convertida y utilizada como combustible (Palanna, 2009). Es por esto que el esfuerzo se ha centrado en el desarrollo de tecnologías asociadas a la generación de biocombustibles a partir de biomasa vegetal. Existen biocombustibles de primera, segunda, tercera y cuarta generación según la fuente de donde provenga dicha biomasa (Aro, 2016), siendo los de segunda generación los generados a partir de biomasa vegetal que no compiten con la producción de alimentos, sino que utiliza los desechos de la industria agrícola y forestal (Antizar-Ladislao & Turrion-Gomez, 2008).

En Chile, considerando la alta generación de desechos lignocelulósicos a partir de la agricultura y de la industria forestal, la producción de biocombustibles de segunda generación se plantea como una opción amigable con el medio ambiente ante la utilización de combustibles fósiles (Paneque y col., 2011), lo cual, al igual que en el resto del mundo, ha concentrado grandes esfuerzos en la investigación y desarrollo entorno a su generación.

Pero de manera previa a la generación de bioproductos, tanto de biocombustibles como de otros productos de interés a partir de celulosa es necesario disponer de azúcares simples los cuales sean fáciles de fermentar por los microorganismos, por lo cual, la degradación previa de los materiales celulósicos presentes en la biomasa, a glucosa, es esencial. (Sukumaran y col., 2005)

2. Síntesis y Estructura de la Celulosa.

La celulosa es el principal constituyente de las células vegetales. En gran parte de las plantas, se ha visto que el componente celulósico alcanza alrededor de un tercio de la masa total (Somerville, 2006), e incluso, en el caso del algodón, ésta llega a ocupar casi el 90% del peso total de la pared celular. Por el contrario, en células en crecimiento de

algunas plantas inferiores alcanza sólo un 1% (Preston, 1975). Por otro lado, su presencia ha sido descrita también en bacterias, algas y tunicados, siendo estos últimos, pertenecientes al reino animal (Saxena & Brown Jr, 2005).

Esta se encuentra embebida en una matriz de lignina y hemicelulosa, entre otros componentes, presente en la pared celular, cumpliendo una función principalmente estructural en la célula, resistiendo la turgencia ejercida desde su interior (Béguin & Aubert, 1994), y controlando así los procesos de crecimiento y división celular (Li y col., 2014). Al generarse nuevas células a partir de otras ya existentes, en los meristemas, estas comienzan a crecer según su citoplasma y sus organelos se van desarrollando. A lo largo de dicho periodo, la pared celular está continuamente siendo sintetizada en una zona cercana a la membrana plasmática, manteniéndose delgada y poco espesa para permitir el crecimiento celular, formándose así la pared celular primaria. Esta actúa controlando la forma que adoptará la célula y regulando su tasa de crecimiento, generando, por lo tanto, el mismo efecto sobre la planta en sí. Cuando la célula ya ha alcanzado la madurez, algunas células mantienen sólo este tipo de pared celular, mientras que en algunas células diferenciadas y especializadas, el depósito de celulosa continúa y la pared celular se engrosa y endurece debido a un proceso de lignificación, dando origen a la pared celular secundaria (Preston, 1975), la cual posee un mayor grado de polimerización de celulosa que la pared celular primaria (O'sullivan, 1997).

En cuanto a la estructura de la celulosa, ésta corresponde a un biopolímero formado por unidades repetitivas de glucosa unidas mediante enlaces β -1,4, los cuales se generan por la unión del grupo hidroxilo del carbono 1 en posición beta de una glucosa y del carbono 4 de la glucosa adyacente (Brett & Waldron, 1996). Dicho enlace genera una disposición entre las moléculas de glucosa en la que se encuentran giradas en 180° una

con respecto de la otra. Esta conformación da origen a la unidad base repetitiva de la cadena de celulosa: la celobiosa (Figura 2) (Béguin & Aubert, 1994).

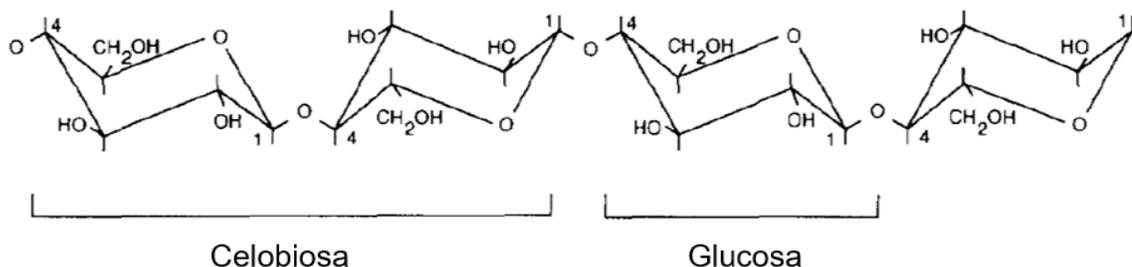


Figura 2. Representación estructural de la Celulosa. Celulosa con su unidad base repetitiva, Celobiosa, y el monosacárido del que se conforma, Glucosa. Modificado de Béguin & Aubert, 1994.

Las cadenas de celulosa a medida que se van sintetizando, se agrupan en la pared celular formando las microfibras, las cuales están generalmente compuestas de entre 25 a 36 cadenas de celulosa unidas entre sí mediante puente de hidrógeno, formando una estructura enrejada, conocida como estructura cristalina de la celulosa (Brett & Waldron, 1996). Cada microfibra contiene entre 500 a 14.000 moléculas de glucosa unidas a través de enlaces β -1,4 (Somerville, 2006).

Las fibras de celulosa, a medida que van siendo sintetizadas, van siendo depositadas en la pared celular, guiadas por los microtúbulos corticales, ubicándose paralelamente entre sí y generando una seguidilla de láminas de deposición, rotadas en diferentes ángulos con respecto a la lámina anterior (Barnett & Bonham, 2004). Esta red, en conjunto con la hemicelulosa, la cual son polisacáridos heterogéneos de diferentes monosacáridos, y la lignina, polímero fenólico con gran rol en prestar rigidez a las plantas, genera una estructura ideal para otorgar fuerza y rigidez a la célula ante situaciones de turgencia (Li y col., 2014), y representa una estructura altamente

recalcitrante a su degradación debido a la gran polimerización presente y al impedimento ejercido por la lignina.

3. Celulasas: Complejo Enzimático Degradador de Celulosa.

Dentro de los microorganismos aeróbicos degradadores de celulosa es posible encontrar hongos y bacterias los cuales se hallan, naturalmente, en todo lugar donde haya depósitos de restos celulósicos.

Debido a que la matriz en la que se encuentra embebida se compone, además, de lignina y hemicelulosa, el acceso a la celulosa para su degradación es dificultado, y para llegar a ella, algunos microorganismos, han evolutivamente desarrollado un complejo enzimático que actúa de manera cooperativa facilitando su alcance, como en el caso de *Phanerochaete chrysosporium* (Bumpus & Aust, 1987), un hongo de pudrición blanca. En éste el complejo degradador de celulosa actúa, primero, degradando la lignina mediante procesos oxidativos, y continúa hidrolizando celulosa y los distintos componentes de la hemicelulosa.

En específico, las enzimas que participan en la degradación de celulosa se clasifican en tres grupos: endoglucanasas (EG), celobiohidrolasas (CBH) y β -glucosidasas (BGL). Este complejo enzimático generalmente es secretado al medio extracelular o es expuesto en la membrana celular de algunos microorganismos celulolíticos, y actúa de manera sinérgica degradando la celulosa a diferentes niveles, teniendo, todas estas enzimas, especificidad por enlaces β -1,4-glicosídicos. Para *P. chrysosporium* y *Trichoderma koningii* se ha propuesto un modelo del proceso sinérgico que las reúne, el cual comienza por la acción de las EGs, las cuales rompen enlaces β -1,4-glicosídicos interiores en las cadenas de celulosa, dejando extremos libres, a través de los cuales las

CBHs pueden comenzar a degradar procesivamente la cadena de celulosa, liberando celobiosa. Posteriormente, las BGLs, hidrolizan celobiosa para prevenir la inhibición de las CBHs mediante su acumulación, liberando glucosa como producto de la reacción (Béguin & Aubert, 1994) (Figura 3). También se ha reportado, en *Trichoderma reesei*, sinergismo entre celobiohidrolasas de diferentes tipos, CBHI y CBHII, donde las primeras hidrolizarían las cadenas de celulosa a partir de los extremos reductores generados por las endoglucanasas, y las segundas hidrolizarían a partir del extremo libre opuesto (Nidetzky y col., 1994), no-reductor.

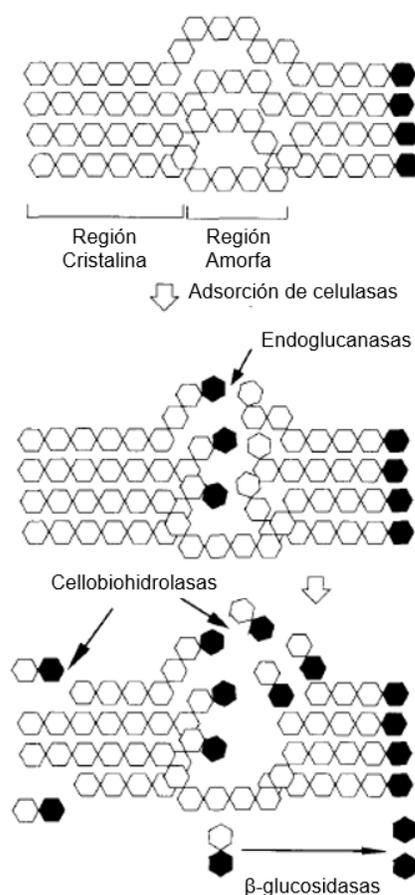


Figura 3. Sinergismo entre Endoglucanasas (EG), Celobiohidrolasas (CBH) y β-Glucosidasas (BGL) en un sistema fúngico de degradación de celulosa. Modificado de Béguin & Aubert, 1994.

3.1. Familia 7 de las Glicosil Hidrolasas.

La familia 7 de las GH (GH-7) está compuesta por EG y CBH, donde las últimas actúan hidrolizando la celulosa de manera procesiva a partir de extremos reductores generados por las EG (CAZy) y cuyo mecanismo de reacción es de retención. Su estructura está generalmente compuesta por cuatro dominios: un péptido señal para exportación extracelular, un módulo de unión a carbohidratos (CBD, por sus siglas en inglés Cellulose Binding Domain), un dominio catalítico, y un puente o linker de unión entre el CBD y el dominio catalítico, aunque pueden haber excepciones. Entre estas excepciones, existen algunas proteínas que combinan el orden o la presencia de dominios, con la restricción de que siempre están presentes el dominio catalítico y la señal de exportación. Existen enzimas que poseen CBD y otras que precinden de éste, y estos pueden provenir de distintas familias.

El dominio catalítico característico de las Celulasas pertenecientes a las GH-7 consiste en un sándwich de láminas β antiparalelas con loops superficiales formando un túnel cerrado donde se lleva a cabo la unión a celulosa. En este túnel, el sitio catalítico se encuentra en el extremo final, permitiendo la liberación de la molécula de Celobiosa y otros oligosacáridos, manteniéndose la proteína unida a la cadena de Celulosa (CAZylopedia, www.cazyclopedia.org).

Se ha visto que la escisión del CBD en celulasas provenientes de diferentes hongos genera efectos, desde no alterar la actividad de las enzimas, a provocar mejoras en la termoestabilidad de proteínas (Wang y cols., 2009), o, por el contrario, generar una reducción de su actividad (Takashima y col., 1998). Por otro lado, algunos estudios han demostrado el rol que cumple este dominio en la adsorción de la enzima sobre sustratos insolubles de celulosa, mediante la comparación de la actividad de distintas versiones

de una misma enzima, cuya única diferencia es la presencia o ausencia de CBD. Estos estudios han indicado que la reducción de la actividad visualizada en enzimas sin CBD, se debe a la disponibilidad y concentración del sustrato, donde ante una condición de alta hidratación, son menos capaces de unirse al sustrato que aquellas enzimas que poseen el dominio (Le Costaouëc y col., 2013).

La producción de estas enzimas por hongos es comúnmente inducida transcripcionalmente bajo la presencia de celulosa y controlada mediante un sistema de inducción-represión (Suto & Tomita, 2001), donde la celulosa y los oligosacáridos producto de su degradación participan como inductores, mientras que moléculas de glucosa y ocasionalmente celobiosa actúan como inhibidores (Mandels & Reese, 1965; Gruno y col., 2004).

3.1.1. Cel7C de *Phanerochaete chrysosporium*.

Phanerochaete chrysosporium es el hongo de pudrición blanca más estudiado debido a su capacidad de degradar lignocelulosa. Para hacer este proceso, *P. chrysosporium* secreta al medio enzimas que le permiten degradar la lignina, componente altamente recalcitrante de la pared celular, mediante procesos oxidativos, mientras que la hemicelulosa y la celulosa son hidrolizadas por celulasas y otras enzimas. Como resultado de esto, el hongo obtiene azúcares metabolizables a partir de la celulosa, los cuales ocupa como fuente de carbono. Dentro del set de enzimas secretado por *P. chrysosporium*, que permiten la descomposición de la celulosa, ha sido posible encontrar hasta ahora cinco endoglucanasas (Henriksson y col., 1999), al menos seis celobiohidrolasas (Muñoz y col., 2001) y dos β -glucosidasas (Covert y col., 1992a), que

actúan de manera sinérgica, y en conjunto con otras enzimas, como las xilanasas (Dobozi y col., 1992) y otras hemicelulasas (Wymelenberg y col., 2005).

Seis genes pertenecientes a la familia Cel7 han sido identificados en *P. chrysosporium*, mediante homología con el gen codificante para la proteína CBHI del hongo *Trichoderma reesei* (Muñoz y col., 2001), y también con respecto a su secuencia aminoacídica. Cel7C fue la primera enzima con actividad CBHI identificada en *P. chrysosporium*, cuyo gen codificante, además, es el mayormente conservado con la secuencia del gen *cbh1* de *T. reesei*. Mediante la determinación de proteínas extracelulares totales secretadas por este hongo, y la cuantificación de enzimas de interés, se ha encontrado que la expresión de Cel7C en presencia de celulosa corresponde al 3% del total versus un 12% de expresión correspondiente a otra CBHI, Cel7D (Uzcategui y col., 1991; Wymelenberg y col., 1993). Así también, mediante RT-PCR, se ha visto que la expresión de Cel7C a diferencia de la de Cel7D, no varía drásticamente al exponer el hongo a celulosa luego de haber sido reprimido de ésta. En las primeras horas de desrepresión, la expresión de Cel7C supera 35 veces a la de Cel7D, pero luego de cinco horas, Cel7D es altamente inducida, superando 3,5 veces a la expresión de Cel7C, la cual mostró tener un leve aumento a las 6 horas (Suzuki y col., 2009). A diferencia de este hongo, en *T. reesei* se ha reportado la expresión de un solo gen del tipo CBHI (Covert y col., 1992a).

Ante la abundancia de genes codificantes para CBHI, y con el fin de revelar el rol que cumplen ante diferentes contextos se ha estudiado la expresión de las diferentes isozimas en presencia de diferentes sustratos, mediante lo cual se determinó que la expresión de Cel7C y Cel7D es inducida en presencia de celulosa y es inhibida significativamente por glucosa, y como contraparte, en el caso de las enzimas Cel7A y Cel7B, se presume que tienen expresión constitutiva, ya que transcritos de los genes

que las codifican fueron encontrados tanto en crecimiento sobre medio con celulosa como con glucosa (Covert y col., 1992b).

Cel7C ha sido caracterizada en cuanto a su estructura y actividad. Mediante corte con Papaina se determinó que posee dos dominios, uno catalítico en el extremo N-terminal y otro correspondiente a un CBD en el extremo C-terminal, cuya disposición es similar a la enzima CBHI de *T. reesei* y diferente de la CBHII, que posee el CBD en el N-terminal. Además, se observó que la secuencia aminoacídica del CBD se encuentra muy conservada entre ambas especies (Uzcategui y col., 1991). En este mismo estudio se midió su actividad sobre sustratos comerciales derivados de celulosa como CMC y Avicel, y sobre 4-nitrofenil- β -D-lactopiranosido y 4-nitrofenil- β -D-celobiosido (*p*NP-L y *p*NP-C, respectivamente), cuya catálisis libera *p*-nitrofenol que al reaccionar con Na₂CO₃ adquiere un color amarillo, permitiendo así medir su concentración a través de espectrofotometría, y de esta manera se posibilita la cuantificación de la actividad de las enzimas. Se encontró que esta enzima presentaba actividad sobre todos estos sustratos, aunque débilmente en CMC. La actividad sobre CMC no es característica de las CBH, pero si lo es de las EG, lo que hizo que esta enzima fuera clasificada dentro de un grupo CBHI-EGI.

Además, en el mismo trabajo realizado por Uzcategui y colaboradores (1991), se observó el sinergismo entre exo-celobiohidrolasas I y II de *P. chrysosporium*, ante un antecedente sobre dicha relación entre enzimas de *T. reesei* (Henrissat y cols., 1985). Como resultado se confirmó que existe sinergismo entre las enzimas Cel7C o Cel7D (CBHI) con Cel6A (CBHII), y más aún, que el efecto sinérgico de acción entre ambas CBHI sobre Avicel es más fuerte.

La estructura de Cel7C ha sido predicha a partir de un modelo por homología con Cel7D (referencia), la cual a su vez fue modelada según la estructura de la enzima CBHI de *T. reesei* (referencia). La estructura tiene el plegamiento característico de las CBHI, donde dos sabanas β se enfrentan dejando una zona hidrofóbica entre sí, y cuyas hebras son continuadas por *loops*, los cuales se estabilizan mediante puentes disulfuro. Cel7C y Cel7D poseen 79,2% de identidad en la secuencia aminoacídica, y los residuos que actúan en el sitio catalítico de las diferentes enzimas Cel7 de *P. chrysosporium* se encuentran altamente conservados, lo que supone que el mecanismo de acción de estas enzimas también es el mismo.

Cel7C ha sido producida dos veces como proteína heteróloga en *E. coli*. La primera vez, su expresión dio como resultado una proteína inactiva en el medio intracelular y que necesitaba ser replegada para mostrar actividad, y como resultado de dicho proceso, la actividad que lograba recuperarse se presentaba solo ante sustratos artificiales (Howard, 1997). La segunda vez, y a partir de la conclusión generada en el primer estudio de Howard (1997), se logró generar una proteína, la cual era secretada al medio extracelular y presentó actividad sobre CMC y Avicel (Howard y col., 2003).

4. Celulasas en la Industria Biotecnológica.

Las celulasas provenientes de microorganismos, principalmente de hongos, se aplican en diversas industrias, donde existe una necesidad de degradación de biomasa. Su uso ha sido reportado en industrias como la de producción de pulpa y papel; de bioconversión de materiales celulósicos: para generación de biocombustibles, de alimentos para animales con alto contenido calórico, de rápida asimilación o mejorados nutricionalmente; en industrias cervecera y vitivinícola; generación de detergentes de

alta calidad basados en celulasas; en la industria alimenticia para facilitar la obtención de subproductos de interés como antioxidantes, pigmentos, aceites, etc.; entre otras (Kuhad y col., 2011).

Dentro de estas, durante décadas, las más utilizadas han sido las celulasas provenientes del hongo *Trichoderma reesei*, las cuales se encuentran comercialmente disponibles y sobre las cuales se han hecho diversas modificaciones para mejorar sus propiedades según los requerimientos de las industrias. También dentro de una búsqueda de alternativas a las enzimas provenientes de *T. reesei*, es que se han diseñado mixes de celulasas provenientes de otros hongos, como aquellos de los géneros *Penicillium*, *Acremonium* y *Chrysosporium*, los cuales han demostrado tener rendimientos más altos en comparación con sus pares provenientes de *T. reesei* (Gusakov, 2011), pero no han tenido mayor éxito en cuanto a su aplicación en la industria, debido a que en ésta área es muy importante la productividad del hongo, y es ahí también donde se han centrado las investigaciones, generando cepas de *T. reesei* con capacidad de producir grandes cantidades de enzimas (Martinez y col., 2008).

En industrias donde la materia prima celulósica corresponde a biomasa, la biodegradación de esta se ve limitada por diversos factores, entre las cuales se cuentan la cristalinidad de la celulosa, la superficie disponible y la cantidad de lignina que contenga. Por esta razón es necesario realizar pretramientos de diversos tipos (químico, mecánico, térmico, etc) (Hendriks & Zeeman, (2009). Éstos permiten que las enzimas alcancen la celulosa más fácilmente y puedan ejercer su función degradadora, aumentando la cantidad de celulosa hidrolizada y el rendimiento del producto final. El someter la biomasa a altas temperaturas corresponde a un tipo de pretratamiento, cuya aplicación hace que la celulosa se hinche y fraccione con mayor facilidad (Hendriks &

Zeeman, 2009). Para acompañar estos tratamientos, el uso de celulasas termoestables es ideal, como también para los procesos posteriores donde la celulosa disponible tras el tratamiento, es hidrolizada a altas temperaturas y por prolongados periodos para generar los productos de interés.

4.1. Celulasas Termoestables.

La termoestabilidad enzimática se refiere a la capacidad de una proteína de preservar su actividad aun cuando haya sido sometida a altas temperaturas, o cuando posee una temperatura de desplegamiento alta. Naturalmente es posible encontrarlas en organismos termófilos (temperatura óptima de crecimiento entre 50°C y 75°C) e hipertermófilos (temperatura de crecimiento desde los 75°C). Dichas enzimas poseen en sus estructuras mayor cantidad de puentes de hidrógenos, puentes disulfuro y enlaces salinos que la estabilizan o rigidizan, permitiéndoles soportar altas temperaturas. También pueden existir factores ambientales que promuevan su estabilidad (Kumar & Nussinov, 2001).

El uso de enzimas termoestables confiere una serie de ventajas al proceso de degradación de celulosa. Primero que todo, las reacciones a altas temperaturas favorecen la generación del producto final, necesitándose así menor cantidad de enzima, además también, por la baja pérdida de actividad de estas durante el procedimiento. Por otro lado, una alta temperatura inhibe el crecimiento de muchos microorganismos mesófilos que pudieran ser contaminantes, y aumenta la solubilidad de los materiales, como en el caso de la celulosa, permitiendo que las enzimas actúen con mayor facilidad sobre ellos (Ward & Moo-Young, 1988). También son una opción más ecoamigable, ya que estas proteínas suelen actuar en ausencia de metales pesados, por lo cual se

requiere menor adición de químicos para llevar a cabo las reacciones (Turner y col., 2007).

La obtención de este tipo de enzimas ha sido una gran puesta en valor de los microorganismos que las producen. Esto también se ha convertido en un desafío para la biología molecular, donde se ha propuesto obtener enzimas termoestables a partir de sus pares mesófilas, mediante procesos de evolución dirigida como de ingeniería de proteínas, imitando los mecanismos estructurales que confieren esta propiedad a las enzimas termoestables. Por ejemplo, para obtener celobiohidrolasas termoestables, ya que no es común encontrar este tipo de enzimas en organismos termófilos, pero son muy requeridas en industrias donde se vea implicada la degradación de lignocelulosa (Turner y col., 2007).

4.1.1. Microorganismos Termófilos.

Los microorganismos adaptados a vivir sometidos a altas temperaturas son la principal fuente de obtención de enzimas termoestables. Entre estos es posible encontrar bacterias y arqueas, resistiendo temperaturas sobre los 50°C, y solo unos pocos hongos creciendo entre los 50 y 55°C (Turner y col., 2007).

A pesar de las ventajas que estos organismos como productores naturales de enzimas termoestables, existe una serie de desventajas asociadas a su mantención y cultivo. Una de estas es que resulta muy caro replicar las condiciones ambientales para su correcto crecimiento y resulta difícil de realizar una producción a gran escala con estos microorganismos. Es por esto que hasta ahora, la manera más eficiente de ocupar estos organismos, es trabajar con los recursos genéticos que se puedan obtener a partir de éstos, produciendo las enzimas en sistemas heterólogos (Lasa & Berenguer, 1993). Se

ha reportado producción de celulasas termoestables provenientes de organismos termófilos en organismos heterólogos, las cuales han sido purificadas y caracterizadas, mostrando mantener su actividad luego de ser sometidas a temperaturas sobre los 85°C (Niehaus y col., 1999).

4.1.2. Ingeniería de Proteínas.

Ante las dificultades de trabajar con organismos termófilos y de la obtención de proteínas provenientes de estos, que actúen a las temperaturas específicas requeridas para diferentes procesos industriales, una buena alternativa es la ingeniería de proteínas para producir proteínas termoestables (Turner y col., 2007).

Tras análisis de la secuencia y de la estructura de enzimas termoestables homólogas a las que se busca termoestabilizar, se han determinado parámetros de flexibilidad y presencia de aminoácidos que cumplen una función estructural en estas enzimas, asociados a su termoestabilidad, en base a los cuales se deben diseñar las nuevas enzimas modificando otras mesófilas existentes que realicen la actividad esperada. Mediante técnicas de biología molecular como mutaciones sitio dirigida, es posible realizar la modificación de estas enzimas. La mayoría de las modificaciones consiste en agregar mayor cantidad de puentes de hidrógeno en el exterior de la proteína y generar nuevos enlaces salinos o puentes disulfuro a través de modificaciones aminoacídicas (Lehmann y col., 2000).

La técnica de mutación de sitio dirigida consiste en la generación de una mutación en un segmento de DNA a través de PCR (Figura 4), combinando un par de partidores que amplifiquen a los extremos del fragmento de DNA con el que se trabaja, y otro par mutagénico, el cual es complementario a una sección interna del fragmento de DNA,

contiene la mutación en su secuencia y es perfectamente complementario entre sí (Ho y col., 1989). Al combinar un partidor forward con uno reverse de cada par de partidores, se generan dos fragmentos, los cuales son complementarios en el extremo generado a partir del cebador mutagénico. Teniendo estos dos fragmentos, se procede a realizar la fusión de estos, y así obtener el segmento de DNA de interés con la mutación deseada.

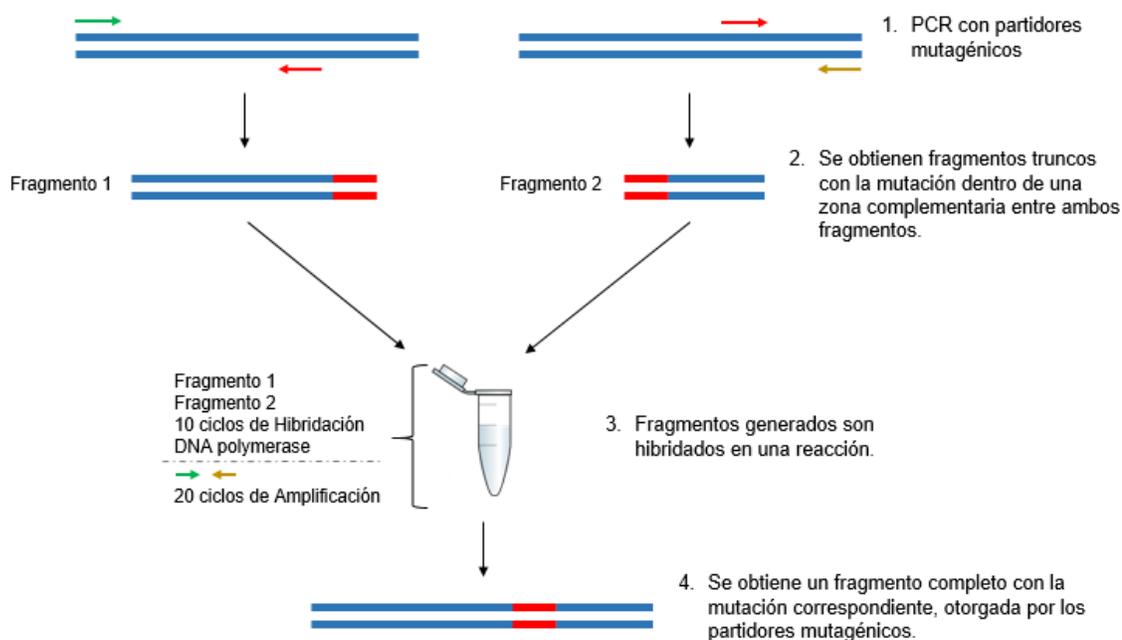


Figura 4. Metodología de mutagénesis sitio dirigida mediante PCR.

5. Antecedentes Para la Termoestabilización de Cel7C de *P. chrysosporium*.

Dentro del marco de la ingeniería de proteínas, en el año 2012 se realizó un trabajo en CeBiB (Jofré, 2012) el cual comprendía el modelamiento computacional de la enzima Cel7C de *P. chrysosporium* por homología y la simulación de su comportamiento tras ser sometida a diferentes temperaturas. Ésta metodología fue propuesta con el objetivo de poder contar con una herramienta de bajo costo y relativamente rápida respecto a los métodos tradicionales de diseño racional de enzimas.

El método de acción del modelo se validó previamente sobre las enzimas Cel7A de *Talaromyces emersonii* y Cel7B de *Melanocarpus albomyces*, cuyo modelo tridimensional y su termoestabilidad son conocidas. La estructura tridimensional de Cel7C fue obtenida con el molde de la enzima Cel7D de *P. chrysosporium*, disponible en las bases de datos. Éste modelo mostró tener una alta similitud con las enzimas previamente mencionadas al realizar un alineamiento, obteniéndose un RMS de 0,475 y 0,767 respectivamente. Para determinar las zonas de alta flexibilidad, candidatas a ser mutadas, se analizó su dinámica molecular a las temperaturas de 300, 350 y 400 K (26.85, 76.85 y 126.85°C) y se seleccionaron aquellos aminoácidos presentes en una región flexible que tuvieran un vecino modificable en otra región flexible enfrentada.

La propuesta generada a través de éste método, resultó en tres variantes de la enzima Cel7CDC de *P. chrysosporium*, donde cada una constaba de dos mutaciones puntuales, las cuales debieran generar un nuevo puente disulfuro en la estructura de la proteína, estabilizando zonas de alta flexibilidad (Figura 5), sin afectar a los aminoácidos presentes en el sitio activo de la proteína (Glu198, Asp200 y Glu203), las cuales es necesario validar de manera experimental, para así comprobar el método planteado.

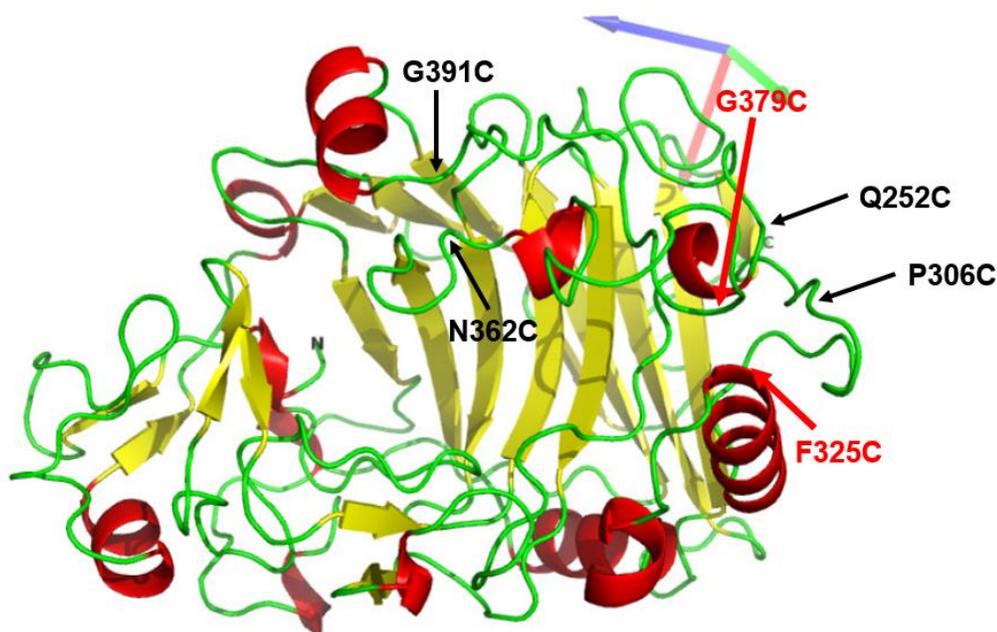


Figura 5. Representación en 3D de Cel7CDC y de las seis mutaciones propuestas para rigidizar la estructura de la proteína. Las mutaciones F325C y G379C (en rojo) corresponden a las que se llevarán a cabo en este trabajo. Modificado de Jofré, 2012.

6. Propuesta.

En este trabajo se pretende validar empíricamente el método de diseño racional de la termoestabilización de la enzima Cel7CDC del hongo *Phanerochaete chrysosporium* propuesto por Jofré.

Para la realización del presente trabajo, se escogió dentro de las tres variantes propuestas, aquella que predecía generar una mayor rigidización de su estructura en las regiones flexibles, y que además mostraba menor flexibilidad a lo largo de toda la estructura.

Se planificó que la validación experimental se llevaría a cabo mediante la implementación de dichas mutaciones, a través de mutación sitio dirigida sobre la secuencia de DNA codificante para el dominio catalítico de la proteína. Ésta luego se

expresaría en *Pichia pastoris* como sistema heterólogo de expresión y se procedería a evaluar su actividad y posteriormente, los parámetros de termoestabilidad.

7. Objetivos.

Objetivo General

El objetivo general que se pretende alcanzar dentro de este trabajo es generar las variantes mutadas de la enzima Cel7CDC para la posterior validación del método de predicción bioinformática para termoestabilización propuesto por Jofré (2012), a través de la implementación de dos mutaciones propuestas en dicho trabajo, sobre la secuencia codificante para el dominio catalítico de la proteína CBHI Cel7C del hongo *Phanerochaete chrysosporium*, en sitios puntuales para reemplazar aminoácidos por cisteínas en zonas flexibles de su estructura, y evaluar su actividad.

Para alcanzar este objetivo, los objetivos específicos con los cuales se trabajó fueron los siguientes:

1. Diseñar e implementar las mutaciones sobre el gen *cel7CDC* para realizar los cambios de aminoácidos F325C y G379C mediante mutagénesis sitio dirigida por PCR.
2. Clonar y expresar la proteína mutada en *Pichia pastoris* con el uso del sistema de expresión en multicopias con el uso del vector pPIC9k.
3. Evaluar las proteínas mutadas de acuerdo a su actividad y termoestabilidad, para determinar el efecto de las mutaciones sobre la enzima, a través de ensayos de actividad y zimogramas, utilizando *pNP-C* y *CMC* como sustratos respectivamente.

MATERIALES Y MÉTODOS

1. Material Biológico

Las cepas de *E. coli* y de *P. pastoris*, así como los plásmidos utilizados y generados en este trabajo se muestran en la siguiente tabla:

Tabla 1. Cepas de microorganismos, y plasmidios de clonación y de expresión utilizados en este trabajo.

Cepas o plásmidos	Descripción	Referencia
<i>E. coli</i> TOP10	Cepa de <i>E. coli</i> para clonación de plásmidos	Disponible en CeBiB ¹
<i>P. pastoris</i> KM71	Cepa de <i>P. pastoris</i> para expresión de proteínas recombinantes, con genotipo <i>his4, aox1::ARG4, arg4</i> y fenotipo Mut ^S His ⁻ Arg ⁺	Disponible en CeBiB ¹ Invitrogen
pBluescript II SK (+)	Vector estándar de clonación, Amp ^r	Disponible en CeBiB ¹
pPIC9k	Vector de expresión en <i>P. pastoris</i> , Amp ^r , G418 ^r , HIS4, posee factor α de <i>S. cerevisiae</i> bajo el control del promotor de la alcohol oxidasa 1 (AOX1)	Thermo Fischer
pPIC9k-Cel7C	Vector pPIC9K con la secuencia codificante para la proteína Cel7C completa.	Disponible en CeBiB ¹
<i>E. coli</i> TOP10/pPIC9k-Cel7C	Cepa de <i>E. coli</i> TOP10 portadora del plasmidio pPIC9k-Cel7C.	Disponible en CeBiB ¹
pBS_FC / M1 / M2 / M1M2	Vector pBluescript II SK (+) con la secuencia codificante para el dominio catalítico wild type de la proteína Cel7C, con la mutación M1 F325C, M2 G379C, o ambas mutaciones, respectivamente	Este trabajo
pPIC9k_FC / M1 / M2 / M1M2	Vector pPIC9k con la secuencia codificante para el dominio catalítico wild type de la proteína Cel7C, con la mutación M1 F325C, M2 G379C, o ambas mutaciones, respectivamente.	Este trabajo
<i>E. coli</i> TOP10/pBS_FC / M1 / M2 / M1M2	Cepa de <i>E. coli</i> TOP10 portadora de los plasmidios pBS_FC / M1 / M2 / M1M2, respectivamente.	Este trabajo
<i>E. coli</i> TOP10/pPIC9k_FC / M1 / M2 / M1M2	Cepa de <i>E. coli</i> TOP10 portadora de los plasmidios pPIC9k_FC / M1 / M2 / M1M2, respectivamente.	Este trabajo
<i>P. pastoris</i> KM71_FC / M1 / M2 / M1M2	Cepa de <i>P. pastoris</i> KM71 portadora de los fragmentos FC / M1 / M2 / M1M2, respectivamente.	Este trabajo

¹CeBiB, Centro de Biotecnología y Bioingeniería, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile.

2. Técnicas de Biología Molecular

2.1. PCR. Los diferentes PCR fueron realizados con la enzima X7 DNA Polymerase, producida en CeBiB. Los partidores utilizados se especifican en la Tabla 2. Estos, exceptuando el par 5' AOX1 y 3' AOX1, fueron diseñados en la plataforma virtual Primer design tool for Gibson Assembly (<http://pdtga.cofla.cl/>), creada por Diego Lagos, miembro de CeBiB, y fueron sintetizados por IDT (Integrated DNA Technologies). Una vez recibidos los partidores, estos fueron resuspendidos en la cantidad de agua recomendada por los proveedores, dejándolos en una concentración de 100 µM, de la cual se tomaba una alícuota para realizar un stock de uso habitual de 10 µM.

Tabla 2. Caracterización y descripción de los partidores utilizados para los diferentes PCR.

Partidor	Secuencia (5' → 3')	Tm (°C)	Tm Adj ¹ (°C)	Aplicación
Cel7C_FWD	AACCACCGCACGCTCACGTCG	60,2	67,2	Amplificación de secuencia codificante del dominio catalítico de Cel7C.
Cel7C_REV	CGTGGTGTGAGGTGCCGAA	58,3	65,2	
Cel7C_B_FC_F²	AAGACGGCGT G CGGCCGA	54,3	59,6	Mutación M1 F325C.
Cel7C_B_FC_R²	TCGCCG C ACGCCGTCTT	54,3	59,6	
Cel7C_B_GC_F²	ACCCGTCGACTCCCT T GTGTTG CCCG	65,9	73,1	Mutación M2 G379C.
Cel7C_B_GC_R²	CGGGCAACAC A GGGAGTCGA CGGGT	65,9	73,1	
Cel7C_Gib_FWD	GAGGCTGAAGCTTACGTAGAA AACCACCGCACGCTCAC	51,9*	57,2*	Amplificación de <i>cel7CDC</i> agregando extensiones complementarias a la secuencia del vector pPIC9k. Mantienen marco de lectura del factor α de secreción y terminador de traducción del vector.
Cel7C_Gib_REV	CATGTCTAAGGCGAATTAATT CGGCGTGGTGTGAGGTCGC	51,9*	57,2*	
5' AOX1	GACTGGTTCCAATTGACAAGC	55	-	Permiten comprobar la presencia de un inserto en el vector pPIC9k.
3' AOX1	GCAAATGGCATTCTGACATCC	55	-	

¹Tm Adj: Tm ajustada considerando las concentraciones de sales estándar presentes en el mix de PCR. ²Partidores utilizados para PCR de mutación sitio dirigida, complementarios entre sí, generan zona de hibridación entre dos fragmentos. *Tm estimada para zona complementaria a la secuencia de *cel7CDC*, no las extensiones. En rojo se muestra el nucleótido a mutar. En verde se muestra las secuencias diseñadas para generar regiones complementarias a la secuencia de pPIC9k para ligar el fragmento al vector mediante Gibson Assembly.

Las mutaciones M1, M2 y la doble mutante fueron realizadas a través de mutación sitio dirigida mediante PCR (Figura 4).

Se generaron dos fragmentos para realizar cada mutación. Ambos fragmentos poseen la mutación correspondiente dentro de una zona complementaria entre ambos fragmentos, mediante las cuales luego fueron apareados para generar la secuencia codificante para Cel7CDC con la mutación correspondiente. Las combinaciones de partidores y la descripción de los fragmentos generados se muestran en la Tabla 3. El ciclo de temperaturas utilizado para generarlos se encuentra especificado en la Tabla 4.

Tabla 3. Descripción de los fragmentos generados por los diferentes partidores utilizados en la de mutagénesis sitio dirigida mediante PCR.

Combinación de partidores	DNA Templado	Fragmento generado	Tamaño (bp)	Aplicación
Cel7C_FWD + Cel7C_REV	pPIC9k-Cel7C	FC	1251	Amplificación de la secuencia de <i>cel7CDC</i> .
Cel7C_FWD + Cel7C_B_FC_R	pPIC9k-Cel7C	F1	980	Mutación M1 F325C. Hibrida con F2 para generar la mutante M1 o F6 para generar la doble mutante M1M2.
Cel7C_REV + Cel7C_B_FC_F	pPIC9k-Cel7C y pBS_M2 respectivamente	F2 y F6	288	Mutación M1 F325C, para generar la mutante M1 o la doble mutante M1M2 respectivamente. Se hibrida con F1.
Cel7C_FWD + Cel7C_B_GC_R	pPIC9k-Cel7C y pBS_M1 respectivamente	F3 y F5	1145	Mutación M2 G379C para generar la mutante M2 o la doble mutante M1M2 respectivamente. Fragmento a hibridar con F4.
Cel7C_REV + Cel7C_B_GC_F	pPIC9k-Cel7C	F4	131	Mutación M2 G379C. Hibrida con F3 para generar la mutante M2 o F5 para generar la doble mutante M1M2.

Tabla 4. Programa de PCR mutación sitio dirigida.

¹PRIMERA FASE: Generación de fragmentos				
Etapa	T (°C)	Tiempo		
Denaturación Inicial	95	2 min		
Denaturación	95	40 s		
Alineamiento²	69	30 s	X27	
Elongación	72	80 s		
Elongación Final	72	5 min		
SEGUNDA FASE: Hibridación y Amplificación				
Etapa	T (°C)	Tiempo		
Hibridación de fragmentos (sin partidores)	Denaturación Inicial	94,5	2 min	
	Denaturación	94,5	40 s	
	Alineamiento	60	30 s	X10
	Elongación	72	80 s	
Agregar partidores	Elongación intermedia	72	40 s	
Amplificación fragmento completo mutado	Denaturación	95	40 s	
	Alineamiento	69	30 s	X20
	Elongación	72	80 s	
	Elongación Final	72	5 min	

¹Ciclo ocupado también al realizar PCR de colonias y Gibson Assembly. ²Temperatura de alineamiento para Gibson Assembly de 57°C.

Para visualizar el resultado de los diferentes PCR realizados, el producto obtenido fue sometido a electroforesis en gel de agarosa 1% (p/v) con GelRed® Nucleic Acid Gel Stain (Biotium) 0,02% (v/v) y posterior iluminación ultravioleta. Se utilizó marcador de peso molecular 1 kb DNA Ladder (NEB) para fragmentos de tamaño superior a 500 bp, y 50 bp DNA Ladder (NEB) para fragmentos de menor tamaño. La extracción y purificación de los amplicones de interés desde los geles de agarosa se realizó con el kit GeneJET Gel Extraction Kit (Thermo Fischer).

2.2. Ligación del DNA a los vectores. El vector pBluescript II SK (+) fue linealizado con la enzima de restricción *EcoRV* (NEB) en 1X NEBuffer™ 3.1 (NEB). Las ligaciones se llevaron a cabo con la enzima T4 DNA Ligase (Promega), incubando el vector con el inserto en una proporción de 1:3 respectivamente, a 4°C durante una noche.

La ligación del DNA al vector pPIC9K se llevó a cabo mediante Gibson Assembly. Para ligar el inserto al vector, éste fue linealizado con las enzimas de restricción *EcoRI* (NEB) y *NotI* (Fermentas) en Buffer O (10X) (Thermo Fischer), durante 3 horas a 37°C. Las ligaciones se llevaron a cabo mediante Gibson Assembly (Gibson y col., 2009), detallado en la sección dedicada a la técnica. Se realizó un PCR donde se agregaron extensiones complementarias al polilinker de pPIC9k, a las secuencias wild type y mutadas de *cel7CDC*. La reacción se llevó a cabo con los partidores *Cel7C_Gib_FWD* y *Cel7C_Gib_REV*, siguiendo el ciclo de amplificación señalado en la primera fase de la Tabla 4, pero con T_m de 57°C. El resultado de este PCR fue sometido a electroforesis en gel de agarosa al 1% (p/v) y fue visualizado bajo luz ultravioleta. La banda de interés fue purificada con el kit GeneJET Gel Extraction Kit (Thermo Fischer).

La reacción de ligación se llevó a cabo en un volumen total de 10 μ L con concentraciones equimolares de inserto y vector, durante una hora a 50°C. Se utilizaron las enzimas T5 Exonuclease (NEB), Phusion HF DNA Polymerase (NEB) y Taq DNA Ligase (NEB).

2.3. PCR de Colonias y Secuenciación de Plasmidios. Una vez transformadas y seleccionadas las colonias de *E. coli* TOP10 en medio sólido, se realizó PCR de colonias. Para obtener el DNA templado desde las células de *E. coli* se tomó un poco de cada colonia con una punta de pipeta y se puso en 15 μ L de agua. Las células fueron lisadas sometiéndolas a 100°C durante 10 min. Luego se centrifugó para separar los restos celulares del sobrenadante en el que se encontraba el DNA. Se tomó 2 μ L de sobrenadante para usar como DNA templado para la reacción. Ésta se llevó a cabo con el uso de los partidores *Cel7C_FWD* y *Cel7C_REV* en el caso de las colonias transformadas con los vectores derivados de pBluescript II SK (+), con las condiciones

especificadas en la primera fase de la Tabla 4, y con los partidores 5' AOX1 y 3' AOX1 en el caso de las colonias transformadas con los vectores derivados de pPIC9k, bajo las condiciones especificadas en Multi-Copy Pichia Expression Kit (Thermo Fisher). Los resultados de la amplificación se analizaron mediante electroforesis en gel de agarosa al 1% (p/v). Con el resultado de estos PCR se escogieron colonias recombinantes.

Los plasmidios de dichas colonias se extrajeron con GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) a partir de un cultivo líquido de 10 mL. Se cuantificó su concentración mediante espectrometría en MaestroNano (MaestroGen).

Se comprobó la presencia del fragmento de interés mediante estudio del patrón de digestión. El vector pBluescript II SK (+) y sus derivados fueron digeridos con la enzima *Sma*I (NEB) en 1X CutSmart® Buffer (NEB) a 25°C durante 3 horas, y el vector pPIC9k y sus derivados fueron digeridos con la enzima *Eco*RV (NEB) en 1X NEBuffer™ 3.1 (NEB) a 37°C durante 3 horas. Las secuenciaciones de los plasmidios se realizaron en MacroGen con partidores universales específicos para cada vector.

Las transformantes de *P. pastoris* KM71 con los vectores derivados de pPIC9k fueron analizadas mediante PCR de colonias con los partidores 5' AOX1 y 3' AOX1. El DNA genómico que se empleó como templado se obtuvo mediante lisis enzimática de la membrana celular con Lyticase de *Arthrobacter luteus* (Sigma-Aldrich). La solución de trabajo de esta enzima se preparó mediante disolución del polvo en Buffer Fosfato de potasio 0,1M, NaCl 2,5M y glicerol 50% (v/v), con una concentración final de 13 mg/mL (5 U/μL). Se incubó una colonia de *P. pastoris* diluida en 10 μL de agua libre de nucleasas, con 10 μL de la solución de Lyticase a 30°C durante 10 min. Luego se congeló la muestra a -80°C durante una noche. Se centrifugó la muestra para separar el DNA de

los restos celulares y se utilizó 2,5 µL del sobrenadante como templado. El PCR de colonias se realizó según lo descrito en Multi-Copy Pichia Expression Kit (Thermo Fisher).

También se realizó PCR desde el DNA genómico de los clones de *P. pastoris* KM71_FC y M1M2, con los partidores 5' AOX1 y 3' AOX1. El DNA genómico se obtuvo con el kit Wizard® Genomic DNA Purification Kit (Promega), según el protocolo de la sección Isolating Genomic DNA from Yeast. El producto de este PCR fue purificado con el kit GeneJET Gel Extraction Kit (Thermo Fischer), y se secuenció en Macrogen, para confirmar que el inserto se encontrara en marco de lectura con el ORF del factor α de secreción y que no existían otras mutaciones no deseadas.

Los resultados de las secuenciaciones fueron analizados realizando alineamientos en la herramienta Benchling (www.benchling.com) y en la herramienta bioinformática BioEdit versión 7.0.5.3 (Hall, 1999) para analizar con mayor detalle cuando fuera necesario.

3. Transformación y Selección de Clones Recombinantes

3.1. *Escherichia coli* TOP10. La transformación se llevó a cabo mediante transformación química. La selección de transformantes recombinantes, portadoras de vectores derivados de pBluescript II SK (+), se realizó sobre placas de LB (BD Difco™) con agar 1,5% (p/v), suplementadas con Ampicilina 100 µg/mL, X-Gal 40 µg/mL e IPTG 0,2 mM. Las placas inoculadas se incubaron durante una noche a 37°C. Se escogieron colonias blancas, las cuales fueron aisladas sobre placas con la misma composición. La selección de transformantes recombinantes, conteniendo vectores derivados de pPIC9k se realizó sobre placas de LB con agar 1,5% (p/v), suplementadas con Ampicilina 100

µg/mL, incubadas durante una noche a 37°C. Se realizó PCR de colonias con los partidores Cel7C_FWD y Cel7C_REV, y se escogieron colonias recombinantes.

Las diferentes versiones recombinantes de *E. coli* TOP10 escogidas fueron cultivadas en medio LB suplementado con Ampicilina 100 µg/mL durante una noche a 37°C con agitación de 250 rpm, y luego se procedió a extraer DNA plasmidial. Se guardó un stock de cada cepa en medio LB con glicerol 20% (v/v) a -80°C, a modo de reserva.

3.2. *Pichia pastoris* KM71. Las células electrocompetentes fueron preparadas según lo descrito en Multi-Copy *Pichia* Expression Kit (Thermo Fisher). Para proceder con la transformación, aproximadamente 20 µg de cada vector derivado de pPIC9k fueron digeridos con la enzima *SacI* (NEB) en 1X CutSmart® Buffer (NEB) durante una noche y luego fueron purificados con GeneJET Gel Extraction Kit (Thermo Fischer). La transformación se llevó a cabo mediante electroporación con el equipo Eppendorf Electroporator 2510 a 1500 V.

La primera fase de la selección de transformantes se realizó sobre placas de medio RDB sin histidina (sorbitol 1M, dextrosa 2% (p/v), YNB 1.34% (p/v), biotina 4 x 10⁻⁵% (p/v), aminoácidos 0.005% (p/v)) con agar al 2% (p/v), durante cuatro noches a 30°C. Las colonias que crecieron sobre este medio fueron suspendidas en agua estéril, y 50 µL de esta suspensión fueron sembrados sobre placas de YPD (extracto de levadura 1% (p/v), peptona 2% (p/v), dextrosa 2% (p/v), con agar al 2% (p/v)) suplementadas con concentraciones crecientes de G418 (Geneticin (G418) Sulfate, Santa Cruz Biotechnology): 1 mg/mL, 1,5 mg/mL, 2 mg/mL y 3 mg/mL de G418 en medio YPD. Aquellas colonias que crecieron en placas con 2 y 3 mg/mL de G418 se aislaron sobre placas con los mismos componentes y se analizaron mediante PCR de colonia.

4. Expresión de la Proteína Cel7C y sus Variedades Mutantes.

Se tomaron colonias aisladas desde placas con medio YPD suplementado con G418 en la concentración correspondiente, y se cultivaron durante una noche en medio YPD líquido a 28°C con agitación de 250 rpm. De estos cultivos se tomó la cantidad necesaria para obtener un cultivo con una A_{600} inicial de 0,01 en 250 mL de medio BMGY (extracto de levadura 1%, peptona 2%, Buffer fosfato de potasio pH 6,0 100 mM, YNB 1.34%, biotina $4 \times 10^{-5}\%$, glicerol 1%). Los cultivos se crecieron a 28°C con agitación de 250 rpm durante aproximadamente 20 h hasta que alcanzara una OD_{600} entre 2 y 6. Se centrifugó el cultivo y se descartó el sobrenadante. Los pellets fueron resuspendidos en 25 mL de medio BMMY (extracto de levadura 1%, peptona 2%, buffer fosfato de potasio pH 6,0 100 mM, YNB 1,34%, biotina $4 \times 10^{-5}\%$, metanol 0,5%) y se incubó a 28°C con agitación de 250 rpm, durante seis días. Se tomó 1mL de cada cultivo a tiempo cero y cada 24 h, y se agregó metanol a una concentración final de 0,5% cada día. Las alícuotas tomadas se centrifugaron a velocidad máxima (14.000 rpm) por 3 min, para separar el sobrenadante del pellet. Ambas fracciones fueron guardadas a -80°C.

5. Confirmación de la Presencia de la Secuencia de *cel7CDC* y *cel7C_M1M2*, y de su Expresión.

Se confirmó la presencia de los genes *cel7CDC* y *cel7C_M1M2* en el vector pPIC9k mediante secuenciación realizada en el proveedor de secuenciaciones MacroGen, usando los partidores universales 5' AOX1 y 3' AOX1.

Las proteínas en los sobrenadantes se concentraron en Centricón YM-10 (Millipore) de capacidad de 2 mL, tres veces, según instrucciones del proveedor, cuando fue necesario.

Las proteínas intracelulares fueron obtenidas mediante el protocolo indicado en Multi-Copy Pichia Expression Kit (Thermo Fisher).

5.1. SDS-PAGE. Se utilizaron geles de poliacrilamida al 12,5%. A 40 μ L de cada muestra se agregó una alícuota de 10 μ L de buffer de carga (Tris-HCl 60 mM, glicerol 25% (v/v), SDS 2% (p/v), 2-mercaptoetanol 14,4 mM, azul de bromofenol 0,1% (p/v) pH 6,8). La mezcla se incubó a 100°C durante 10 min. Se cargaron 15 μ L de las muestras en los geles, correspondientes cada uno a la cinética de las cuatro variantes generadas (*P. pastoris* KM71_ / FC / M1 / M2 / M1M2) y al control de *P. pastoris* KM71 transformada con el vector pPIC9k a lo largo de los seis días de inducción. Los geles fueron sometidos a electroforesis a 200 V durante una hora, y fueron revelados con tinción de plata y visualizados en transiluminador. Los pesos moleculares de las proteínas fueron determinados usando como referencia el marcador Pierce™ Prestained Protein MW Marker (Thermo Fisher).

5.2. Zimogramas. Se utilizaron geles de poliacrilamida al 12,5%, con CMC al 0,2% (p/v) como sustrato, presente en la fase separadora del gel. 40 μ L de sobrenadante concentrado y de una 1:10 de las proteínas intracelulares del cuarto día de inducción de *P. pastoris* KM71 transformada con el vector pPIC9k y de *P. pastoris* KM71_FC y M1M2 se mezclaron con 10 μ L de Buffer de carga semidesnaturante (Tris-HCl 125 mM, glicerol 20% (v/v), SDS 4% (p/v), azul de bromofenol 0,005% (p/v) pH 6,8). Se cargaron 15 μ L de las muestras en los geles correspondientes para cada fracción del cultivo, fue repetida la carga de las muestras en la misma secuencia para luego poder cortar el gel en dos, y fueron sometidos a electroforesis durante una hora a 200 V. El gel fue cortado verticalmente por la mitad, donde una mitad sería revelada con tinción de plata y la otra mitad seguiría el tratamiento para el zimograma. Las proteínas en los geles fueron

renaturadas durante una hora en una solución de Triton® X-100 (Merck Millipore) al 2,5% (v/v). Posteriormente los geles fueron lavados tres veces con agua durante 5 min y luego una vez con buffer acetato de sodio 50 mM pH 5,0. Fueron incubados durante una noche en buffer acetato de sodio 50 mM, pH 5,0, a 37°C, sin agitación. Se revelaron mediante tinción con Rojo Congo 0,1% (p/v) durante 15 min sin agitación y a temperatura ambiente, y se destiñeron mediante dos lavados con NaCl 1M durante 5 min cada uno. Todos los lavados se realizaron en agitación y a temperatura ambiente.

5.3. Ensayos de Actividad. Se realizaron en placas de 96 pocillos. Se agregaron 90 μ L de sustrato solución de pNP-C 1 mg/mL en buffer acetato de sodio 50 M pH 5,0, y 10 μ L de sobrenadante de cultivo de expresión. Las reacciones se hicieron en triplicado a 30°C, realizando las mediciones a las 3 y 6 h transcurridas, y luego se realizaron otras reacciones a 50°C con mediciones luego de 1, 1:30, 2 y 2:30 h. Las reacciones fueron detenidas con un volumen de reacción de Na₂CO₃ 2% (p/v) (100 μ L). Los blancos empleados fueron sustrato con medio de inoculación (BMMY) y sustrato con agua miliQ. Las mediciones se realizaron a 410 nm.

RESULTADOS

1. Diseño experimental de Mutagénesis de Sitio Dirigida.

Tras la propuesta realizada de manera bioinformática en el trabajo de Jofré (2012), se analizó y se diseñó una metodología para cumplir el objetivo de obtener el gen *cel7CDC* con las dos mutaciones que permitirían termoestabilizar su estructura gracias al cambio de aminoácidos por cisteínas que implican las mutaciones. El diseño experimental mediante el cual se realizó la mutagénesis de sitio dirigida para obtener los fragmentos mutados y los vectores derivados de pBluescript SK II (+), se estableció como se muestra en la Figura 6.

En la Figura 6 se puede ver que la generación de la variante del gen con las dos mutaciones es un proceso que consta de dos pasos de mutagénesis de sitio dirigida. En el primer paso se generaron las variantes M1 y M2 del gen *cel7CDC*, las cuales poseen una de las dos mutaciones a realizar, y luego en el segundo paso, a partir de cualquiera de las dos variantes del gen obtenidas, se generaría la mutante M1M2, que posee ambas mutaciones. Las mutaciones se integraron en la secuencia del gen *cel7CDC* a través de partidores mutagénicos. Para realizar la primera mutación se deben generar dos fragmentos parciales los cuales tienen una región de homología otorgada por dichos partidores. Estos fragmentos son hibridados en un ciclo de PCR de solapamiento, obteniéndose a partir de éste, la secuencia completa codificante para el dominio catalítico de Cel7C con una de las dos mutaciones correspondientes. Luego se procede a ligar el producto de esta reacción con el vector pBluescript II SK (+) previamente linealizado, generando los vectores pBS_M1 y pBS_M2. Estos vectores posteriormente

se utilizaron como plantilla para generar la variante doble mutante *cel7C*_M1M2 a través del segundo paso de mutagénesis de sitio dirigida, la cual también se ligó al vector mencionado.

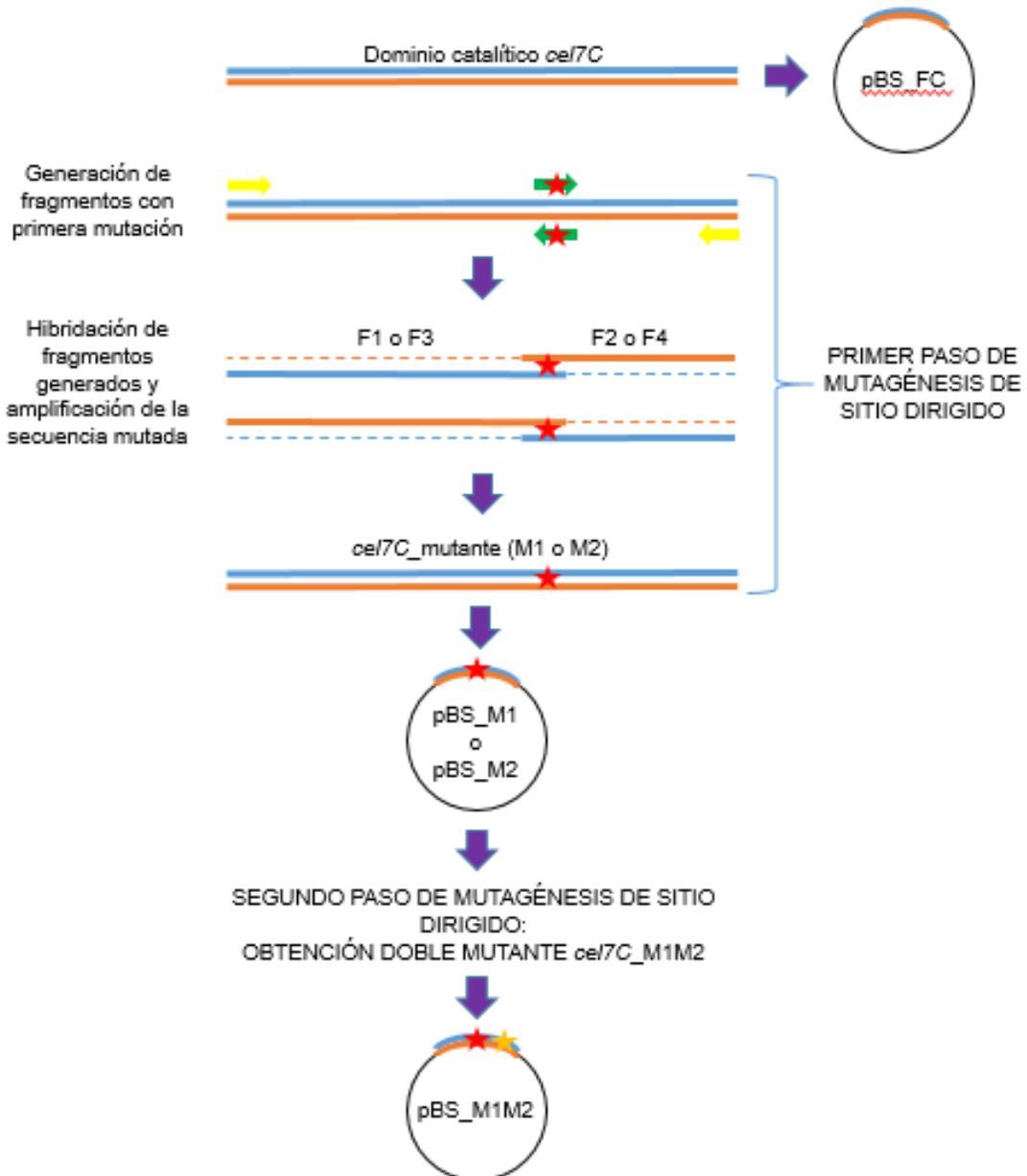


Figura 6. Diseño experimental para la obtención de vectores derivados de pBluescript II SK (+) para clonación de los insertos.

2. Implementación de mutaciones en la secuencia codificante para el dominio catalítico de Cel7C.

A partir del vector pPic9k-Cel7C (Tabla 1), que contenía la secuencia codificante del dominio catalítico de la proteína Cel7C silvestre, se generaron los fragmentos FC, F1, F2, F3 y F4 mediante PCR con el uso de los cebadores correspondientes (Tabla 2). El ciclo de temperaturas utilizado en esta etapa fue estandarizado mediante PCR con gradiente de temperaturas, a través del cual se determinó que 69°C es la temperatura óptima para llevarlos a cabo. En todos los casos se logró obtener una banda de amplificación de los tamaños esperados según lo que se muestra en la Tabla 3 (Figura 7).

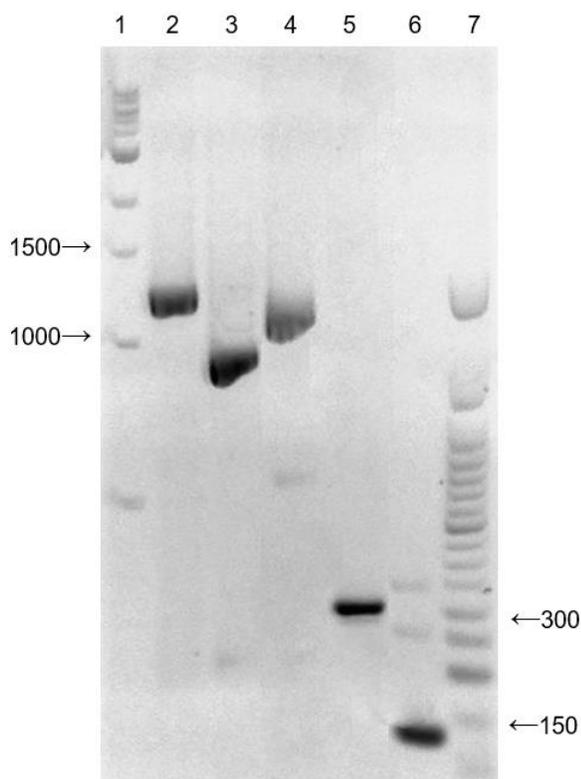


Figura 7. Amplificación de los fragmentos utilizados para realizar mutación sitio dirigida. Las reacciones fueron realizadas con una T_m de 69°C. Carril 1 y 7: Ladder 1kb y ladder 50bp respectivamente; 2: FC; 3:F1; 4:F3; 5:F2; 6:F4. El DNA templado utilizado corresponde al

vector pPic9k-Cel7C, y para obtener los diferentes fragmentos se utilizaron los partidores según se indica en la Tabla 3.

Una vez purificados los fragmentos, la unión de los fragmentos F1 y F2, y de F3 y F4, se llevó a cabo, generando de manera exitosa los fragmentos completos codificantes para el dominio catalítico de Cel7C con las mutaciones M1 y M2 respectivamente, los cuales poseen un tamaño de 1251 pb (Figura 8).

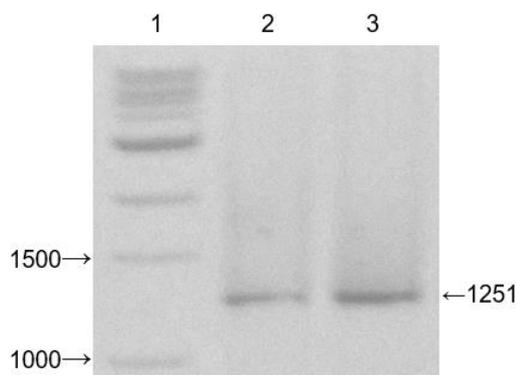


Figura 8. Fragmentos de *cel7C* mutados, purificados. Luego de realizar las hibridaciones de los fragmentos parciales, se obtuvieron fragmentos completos con una mutación cada uno correspondiendo el carril 2 a la mutante M1 y el carril 3 a la mutante M2. Carril 1: Ladder 1kb.

La variante *wild type* del gen y las mutadas generadas fueron ligadas al vector pBluescript II SK (+), formando los vectores pBS_FC, pBS_M1 y pBS_M2 (Tabla 1) respectivamente, con los cuales se transformó *E. coli* TOP10. Una vez obtenidas colonias transformantes en placas de LB con Ampicilina, IPTG y X-gal, se seleccionaron las blancas para ser analizadas mediante PCR de colonias (Figura 9A, B, C, D). Con los resultados obtenidos, aquellas colonias que presentaron un amplicón del tamaño esperado (1251 bp) se reconocieron como recombinantes (Figura 9E, F), y se realizaron aislamientos sobre placas de LB con Ampicilina, para así obtener colonias únicas.

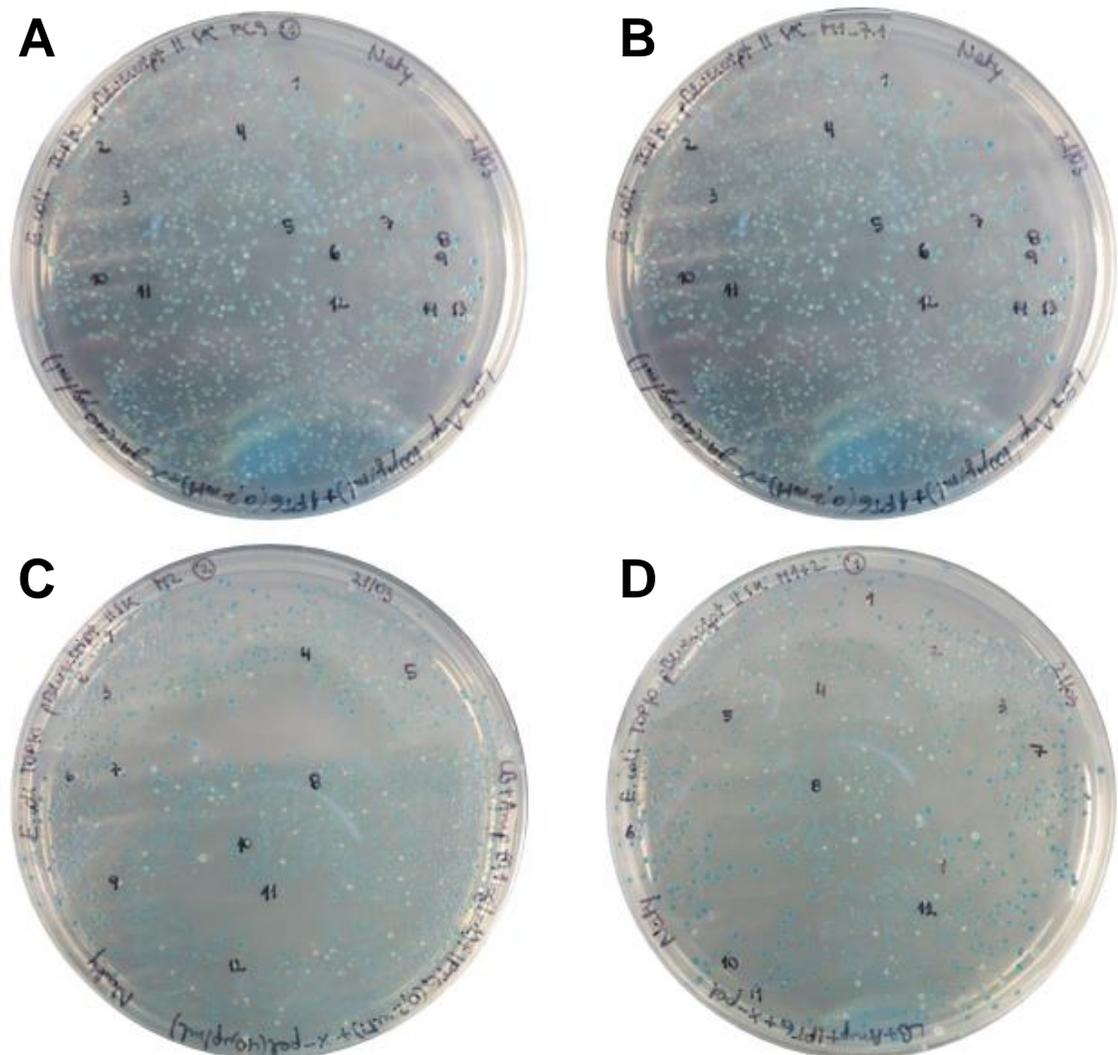


Figura 9. Resultados de la transformación de *E. coli* Top10 con vectores recombinantes pBluescript II SK (+) que contenían las versiones silvestre y mutadas de *cel7CDC*. Las transformantes fueron seleccionadas en medio LB con ampicilina, Xgal e IPTG. **A: Transformación con vector pBS_FC, **B:** Transformación con vector pBS_M1, **C:** Transformación con vector pBS_M2, **D:** Transformación con vector pBS_M1M2.**

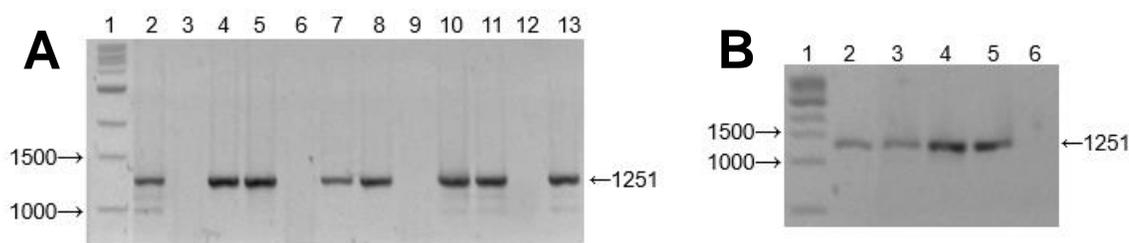


Figura 10. PCR de colonias de las transformaciones de *E. coli* TOP10 con vectores derivados de pBluescript II SK (+). Los PCR de colonias se realizaron con los partidores Cel7C_FWD y Cel7C_REV a 69°C. **A:** En los carriles 2-5: colonias transformadas con pBS_FC, 6-8: colonias transformadas con pBS_M1, 9-11: colonias transformadas con pBS_M2. **B:** PCR de colonias realizados con los mismos parámetros. En los carriles 2-4: colonias transformadas con pBS_M1M2. En ambos PCR el carril 1 corresponde al ladder 1kb. El control negativo sin DNA se encuentra en los carriles 12 (A) y 6 (B), y el control positivo vector pPIC9k-Cel7C se encuentra en los carriles 13 (A) y 5 (B).

Las colonias aisladas fueron crecidas en medio LBamp líquido para extraer el DNA plasmidial desde las células. Para corroborar la presencia del inserto de interés en los vectores, se analizó el patrón de digestión del DNA plasmidial de las colonias recombinantes con la enzima *Sma*I, comparando el resultado obtenido experimentalmente, con el patrón teórico obtenido in silico en ApE (software). En la Figura 11A se muestra la posición del sitio de corte de *Sma*I en el mapa de pBS_FC, obtenido a través del programa ApE. En la Figura 11B se muestra un análisis electroforético de los productos de corte de los vectores recombinantes con *Sma*I, mientras que en la Figura 11C se muestra el patrón teórico de corte de los vectores con *Sma*I. La enzima *Sma*I posee sólo un sitio de corte en el vector pBluescript II SK (+), lo cual genera la linearización del vector y se visualiza como una única banda de DNA a través de una electroforesis en gel de agarosa. La significativa diferencia de tamaños entre el vector vacío (2961 pb) y el vector con inserto (4212 pb) permite determinar su presencia o su ausencia en éste. Se comprobó que el tamaño del fragmento generado para cada vector, concordaba con el tamaño esperado, lo que sugiere que los fragmentos insertados en el vector corresponden al gen *cel7CDC* y sus variantes. Con

estos resultados se tomó la decisión de enviar los plasmidios a Macrogen para su secuenciación.

Una vez obtenidos los resultados de la secuenciación, y por lo tanto, la confirmación de la presencia de la primera mutación, se utilizó el vector pBS_M1 como DNA templado para realizar la segunda mutación sobre el gen *cel7CDC* y así obtener la variante con ambas mutaciones (M1M2), la cual, al igual que las demás secuencias, se insertó en el vector pBluescript II SK (+) y fue clonado en *E. coli* TOP10, siguiendo los pasos anteriormente mencionados para comprobar su presencia en el vector.

Los resultados de la secuenciación (Figura 12) confirmaron la presencia de los insertos en los vectores generados, encontrándose las secuencias íntegras, con las mutaciones específicas realizadas, y sin mutaciones adicionales. Un alineamiento de las secuencias completas de los variantes de *cel7CDC* se muestra en Anexo “Secuencias de vectores derivados de pBluescript II KS (+) clonados en *E. coli* TOP10”.

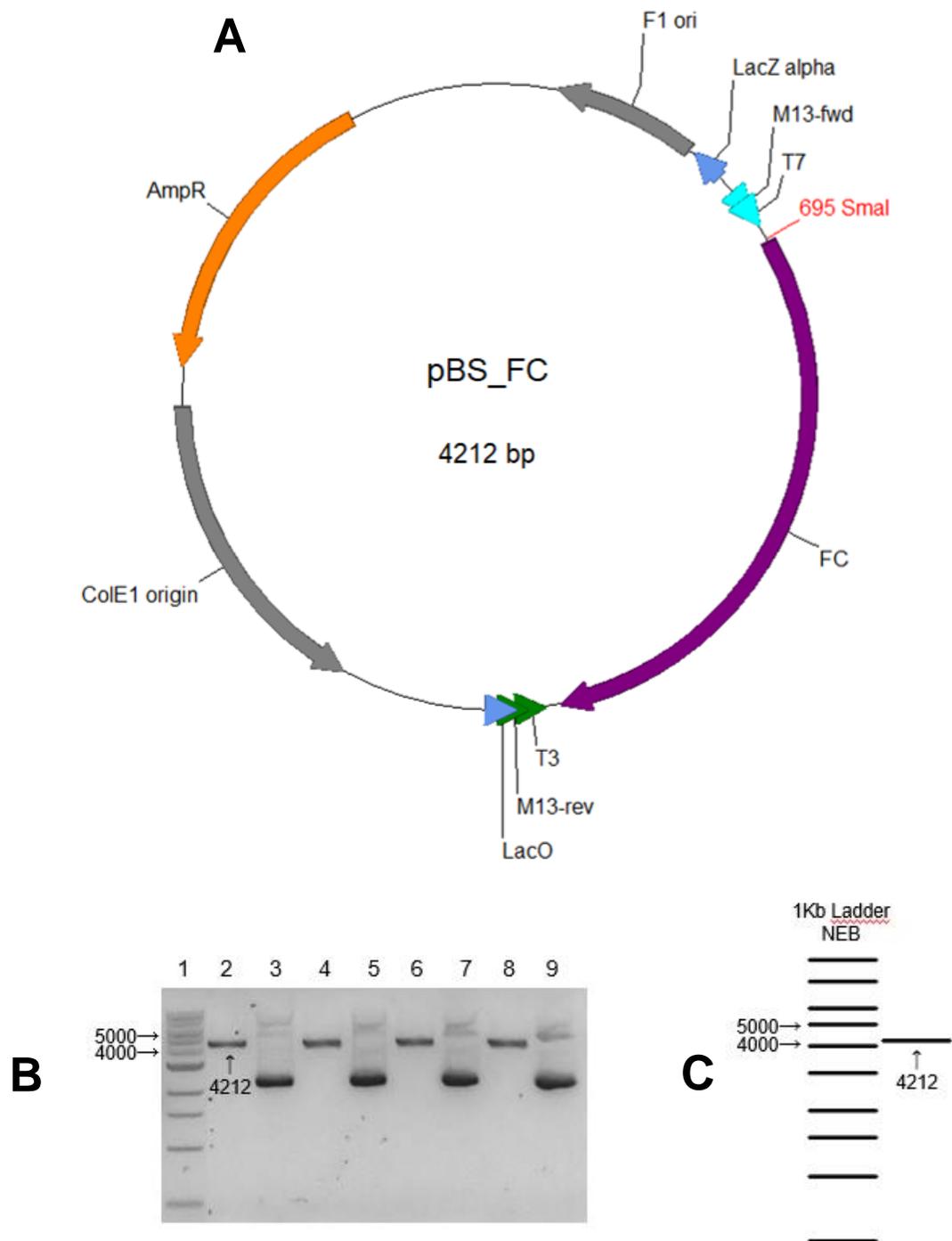


Figura 11. Análisis de restricción de vectores recombinantes derivados de pBluescript II SK (+) con la enzima *SmaI*. En **A**: mapa del vector pBS_FC. Se muestra el sitio de corte con la enzima de restricción *SmaI* en el gen *cel7CDC* (FC). Este sitio también debería estar presente en los vectores que contienen los insertos mutados de *cel7CDC*. **B**: análisis electroforético experimental de los productos de restricción con *SmaI*. Carril 1: Ladder 1Kb; carriles 2, 4, 6 y 8: vectores pBS_FC, pBS_M1, pBS_M2 y pBS_M1M2, respectivamente,

todos digeridos con *Sma*I, carriles **3, 5, 7 y 9**: vectores pBS_FC, pBS_M1, pBS_M2 y pBS_M1M2, respectivamente todos sin digerir. **C**: Digestión *in silico* del vector generada en el programa ApE.

Tabla 5. Colonias de *E. coli* TOP10 con los vectores derivados de pBluescript II KS (+), secuenciadas.

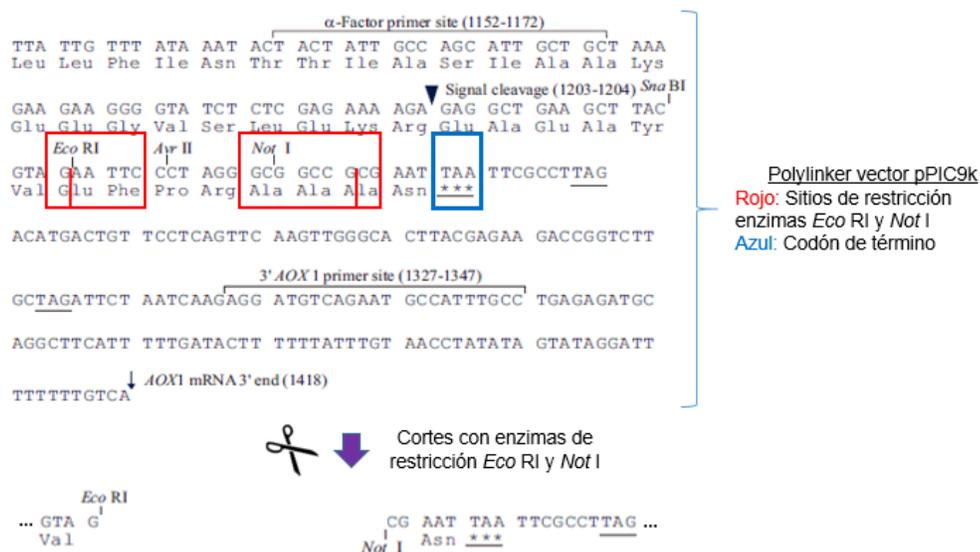
Nombre en Alineamiento	Colonias de <i>E. coli</i> TOP10
pBS_FC	Colonia 14
pBS_M1	Colonia 7.2
pBS_M2	Colonia 10
pBS_M1M2	Colonia 13



Figura 12. Extracto del alineamiento de secuenciaciones de plasmidios derivados de pBluescript II KS (+). Secuencia *cel7CDC*, codificante para el dominio catalítico de la proteína Cel7C wild type, es utilizada como templado para el alineamiento. Alineamiento muestra las zonas del gen donde se realizaron las mutaciones específicas. Los alineamientos completos en el Anexo "Secuencias de vectores derivados de pBluescript II KS (+) clonados en *E. coli* TOP10".

3. Generación de vectores de expresión derivados de pPIC9k.

Una vez confirmada la integridad de las secuencias de las cuatro variantes, se prosiguió con la generación de los vectores que permitirían la expresión de las proteínas en *P. pastoris*. Estos vectores fueron realizados ocupando como base el vector pPIC9k y los insertos fueron ligados utilizando la técnica de Gibson Assembly. Para realizar este procedimiento, se crearon los partidores que permitirían el solapamiento del inserto con la secuencia del vector en su polilinker. La estrategia de creación de los partidores específicos para esta ligación se detalla en la Figura 13.



Para mantener marco de lectura del gen *cel7CDC* (1251 pb) y que éste estuviera en marco con el Factor α de secreción y el primer terminador de traducción, se necesitaban dos bases nitrogenadas adicionales en el extremo 5' del gen y una en el extremo 3'.

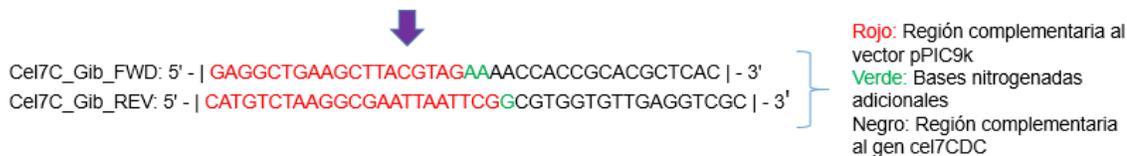


Figura 13. Diseño estratégico de los partidores para unir el gen *cel7CDC* y sus variedades al vector pPIC9k.

La confección de estos cebadores se pensó de manera que se mantuviera el marco de lectura con el factor α de secreción presente en el vector y con un codón de término, por lo tanto, al detectar su presencia se podría comprobar que la secuencia de la proteína se mantiene de manera adecuada para su expresión.

Las secuencias se amplificaron por PCR con los partidores *Cel7C_Gib_FWD* y *Cel7C_Gib_REV*, y los resultados de estas reacciones fueron purificados y se ligaron mediante Gibson Assembly al vector pPIC9k previamente linealizado. Con estas ligaciones se transformó *E. coli* TOP10 y se seleccionaron colonias transformantes en medio LB con ampicilina (Figura 14). Para identificar las colonias recombinantes se procedió a realizar un screening de las clonias mediante PCR usando el par de partidores

AOX1, los cuales amplifican una sección del vector adyacente al sitio de clonamiento múltiple, donde se encuentra el inserto (Tabla 2). En caso de ser recombinantes, se esperaba obtener un producto de amplificación de 1729 pb, correspondiente a la suma del tamaño del inserto de interés con secuencias propias del vector, mientras que en el caso de no ser recombinantes, se esperaba un amplificado de 493 pb. Los resultados (Figura 15) mostraron que 4 colonias contenían el vector pPIC9K con el gen silvestre *cel7CDC* (pPIC9k_FC), mientras que para las construcciones pPIC9K_M1, pPIC9K_M2 y pPIC9K_M1M2 se encontraron 2, 3 y 2 colonias transformantes recombinantes, respectivamente.

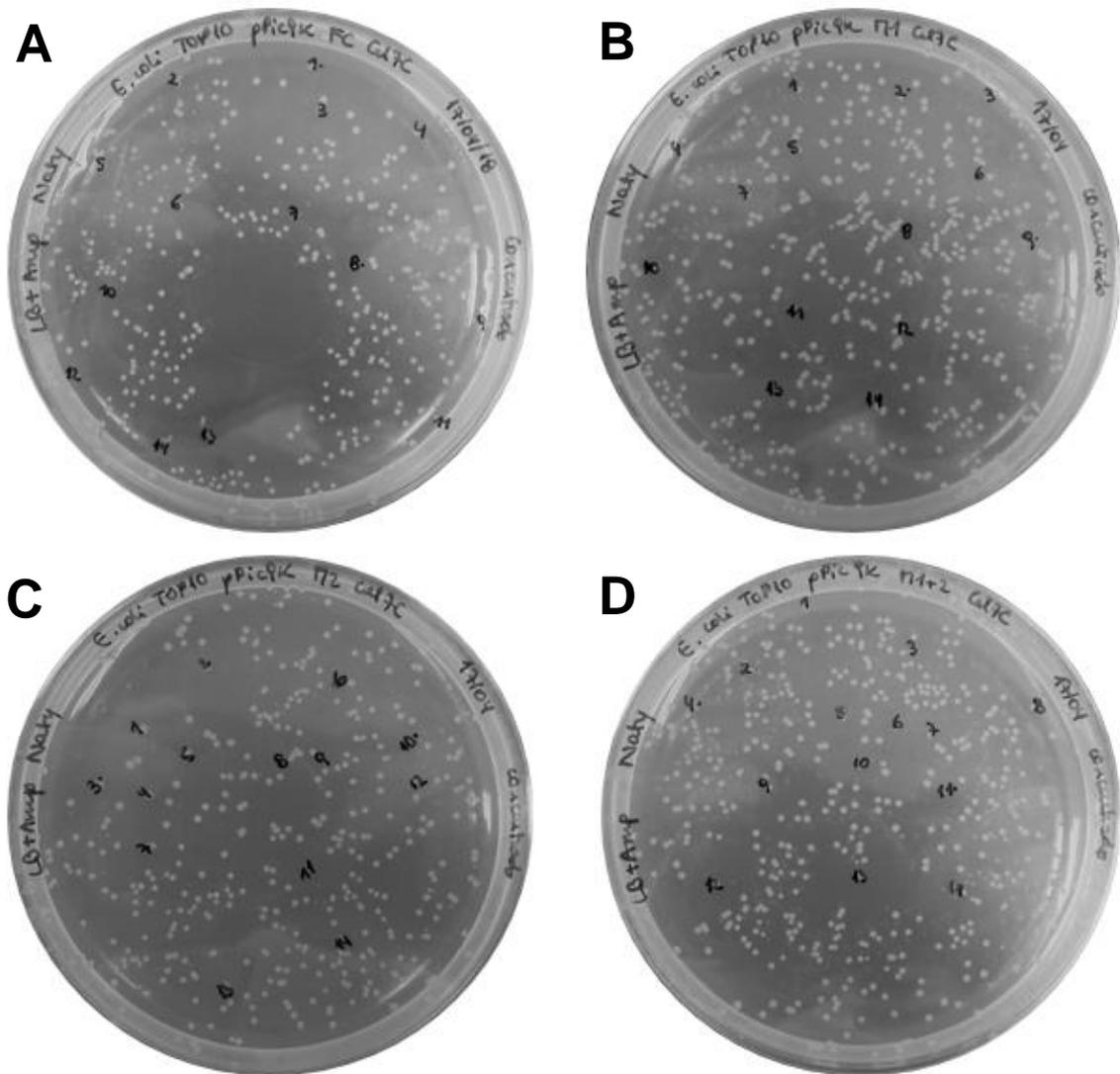


Figura 14. Resultados de la transformación de *E. coli* Top10 con vectores pPIC9K recombinantes que contenían las versiones silvestre y mutadas de *cel7CDC*. A: Transformación con vector pPIC9k_FC, **B:** Transformación con vector pPIC9k_M1, **C:** Transformación con vector pPIC9k_M2, **D:** Transformación con vector pPIC9k_M1M2.

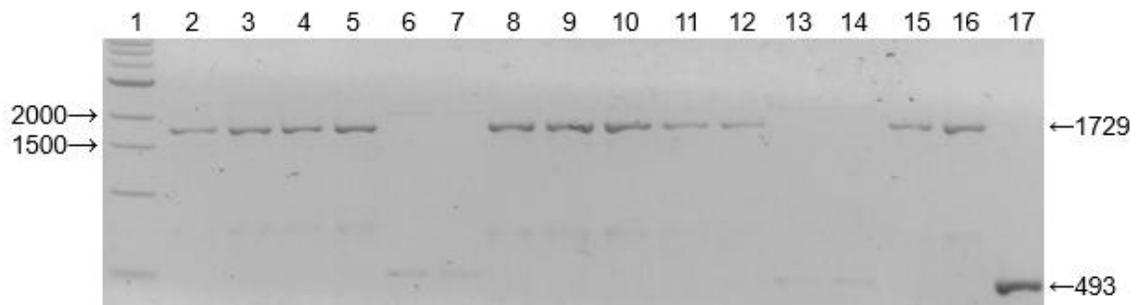


Figura 15. PCR de colonias de las transformaciones de *E. coli* TOP10 con vectores derivados de pPIC9k. Los PCR de colonias fueron realizados con el par de partidores AOX1,

carril 1: Ladder 1kb, **2-5**: colonias transformadas con pPIC9k_FC, **6-9**: colonias transformadas con pPIC9k_M1, **10-12**: colonias transformadas con pPIC9k_M2, **13-16**: colonias transformadas con pPIC9k_M1M2, **17**: control negativo, vector pPIC9k vacío.

Para comprobar la presencia del inserto de interés, el DNA plasmidial de las colonias recombinantes fue digerido con la enzima de restricción *EcoRV*, la cual corta los vectores que tienen los insertos en tres puntos, a diferencia del vector pPIC9k vacío, donde corta en dos puntos. El patrón de digestión obtenido se comparó con el patrón de digestión *in silico* generado con la herramienta ApE. En aquellos vectores en los que se obtuvo tres fragmentos de tamaños 5078, 4271 y 1163 pb respectivamente, se consideró la presencia del inserto de interés. Por otro lado, el patrón de digestión del vector pPIC9k vacío presentó solo dos fragmentos, uno de 5078 y otro de 4198 pb (Figura 16). Los vectores que sugerían tener los insertos fueron enviados a MacroGen para su secuenciación.

Mediante la secuenciación de los vectores derivados de pPIC9k, se confirmó la presencia de los insertos correspondientes, con las mutaciones respectivas y libres de mutaciones adicionales (Figura 17).

A través de esta secuenciación también se verificó la presencia de las secuencias contenidas en el par de cebadores Cel7C_Gib, lo cual demostró la eficiencia de la técnica de Gibson Assembly para realizar la ligación.

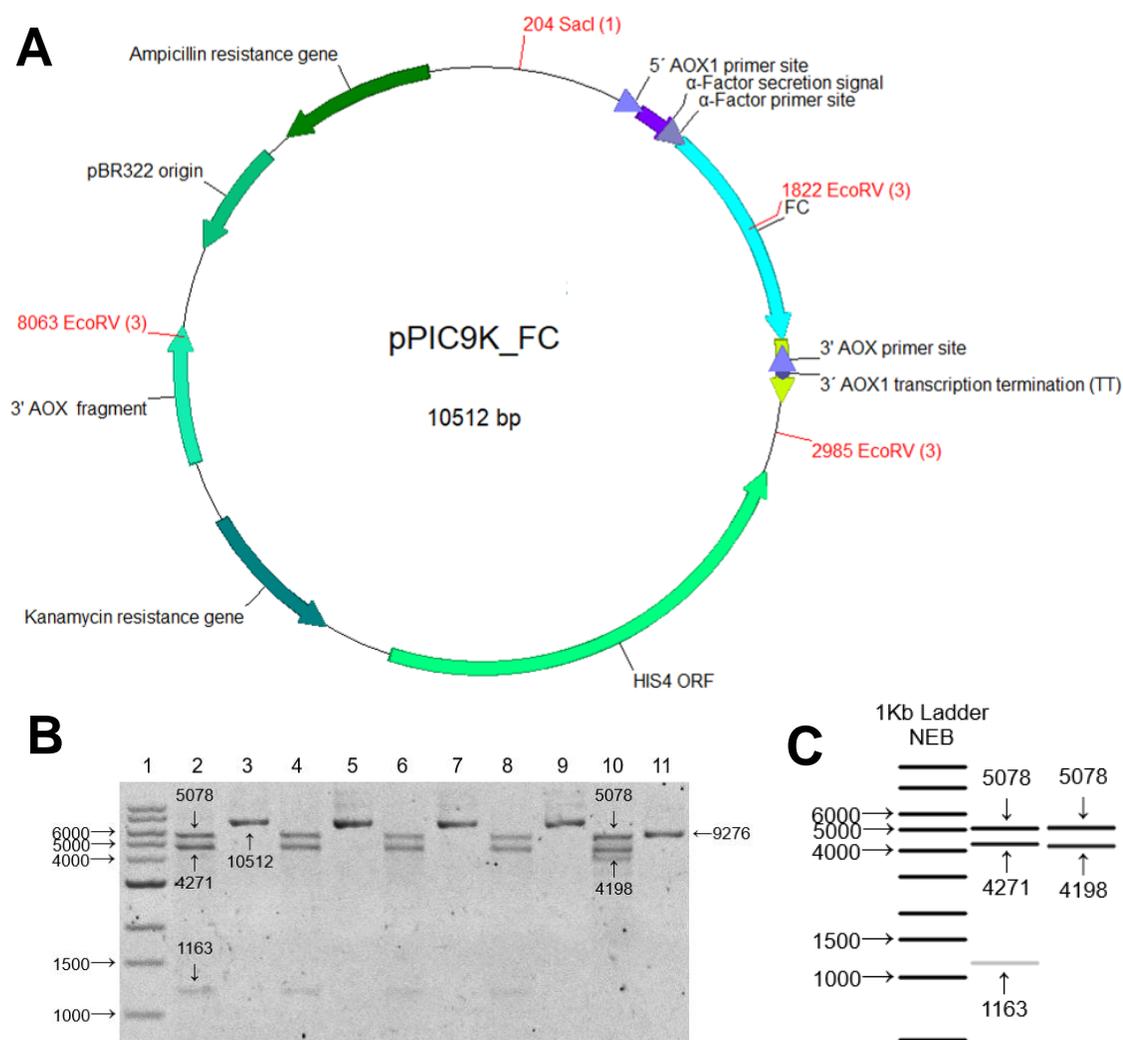


Figura 16. Análisis de restricción de vectores recombinantes derivados de pPIC9k con la enzima *EcoRV*. **A:** Mapa del vector pPIC9k_FC, muestra los sitios de corte con las enzimas de restricción *EcoRV* y *SacI*. Estos sitios también deberían estar presentes en los vectores que contienen los insertos mutados de *cel7CDC*. **B:** análisis electroforético experimental de los productos de restricción con *EcoRV*. Carril 1: Ladder 1Kb; carriles 2,4,6,8, 10: vectores pPIC9k_FC, pPIC9k_M1, pPIC9k_M2, pPIC9k_M1M2 y pPIC9k, respectivamente, todos digeridos; 3,5,7,9,11: vectores pPIC9k_FC, pPIC9k_M1, pPIC9k_M2, pPIC9k_M1M2 y pPIC9k, respectivamente, todos sin digerir; **C:** Digestión *in silico* del vector generada en el programa ApE.

Tabla 6. Colonias de *E. coli* TOP10 con los vectores derivados de pPIC9k, secuenciadas.

Nombre en Alineamiento	Colonias de <i>E. coli</i> TOP10
pPIC9k_FC	Colonia 8
pPIC9k_M1	Colonia 9
pPIC9k_M2	Colonia 3
pPIC9k_M1M2	Colonia 4

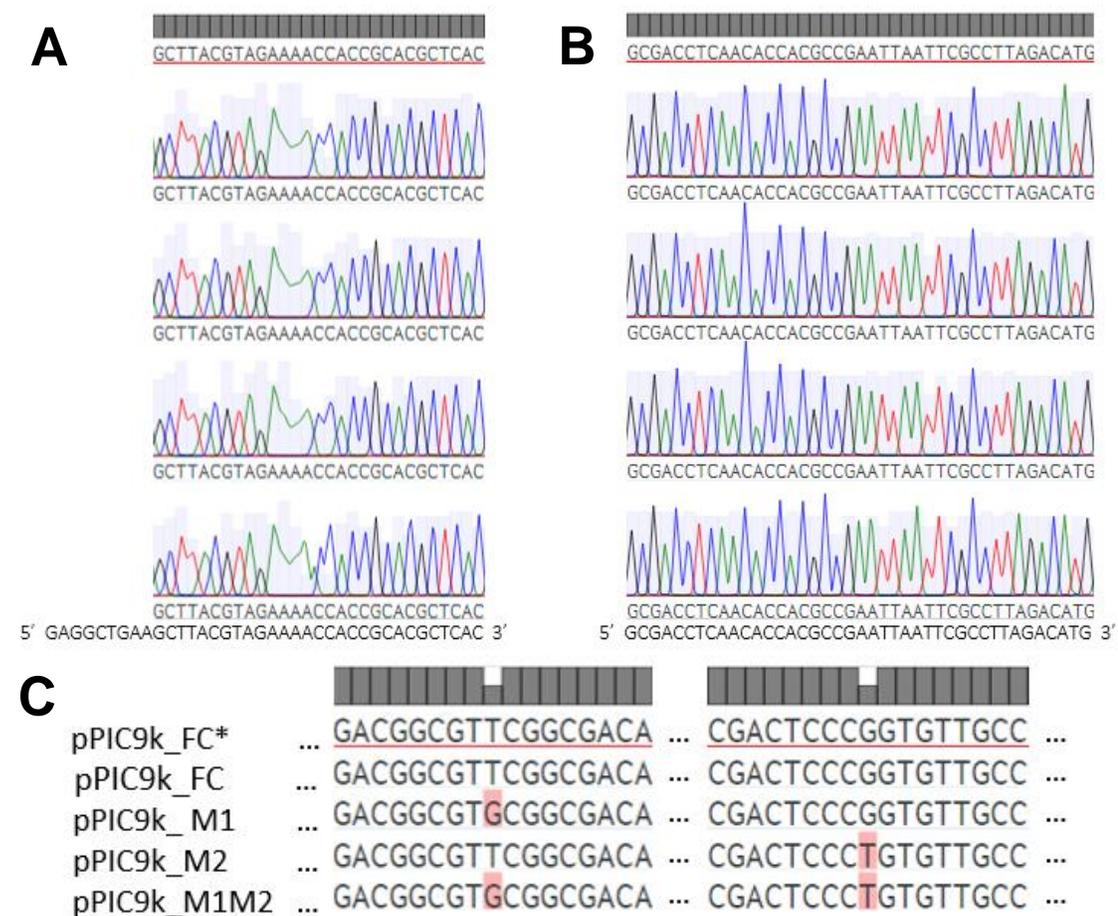


Figura 17. Detalle del alineamiento de secuenciaciones de plasmidios derivados de pPIC9k. *Se utilizó como plantilla para el alineamiento la secuencia a priori del vector con el inserto wild type para la secuencia codificante para el dominio catalítico de Cel7C. Alineamientos corresponden a las secuencias de los vectores de las colonias nombradas en la Tabla 6. **A:** Alineamiento de las secuenciaciones con el cebador Cel7C_Gib_FWD. **B:** Alineamiento de las secuencias con el reverso complementario del cebador Cel7C_Gib_REV. **C:** Alineamiento de las secuencias en zonas mutadas. Alineamientos completos se encuentran en Anexo “Secuencias de vectores derivados de pPIC9k clonados en *E. coli* TOP10.”

4. Expresión de las versiones silvestre y mutadas de Cel7CDC en *P. pastoris*.

Luego de obtener los vectores de expresión para *P. pastoris* con los insertos de interés y comprobar su integridad, se procedió a la transformación de esta levadura. Para esto, los vectores fueron linealizados mediante incubación con la enzima *SacI*, de manera de favorecer la inserción del vector en el genoma de esta levadura.

P. pastoris KM71 es una especie de ésta levadura desarrollada para ser ocupada como sistema heterólogo de expresión. La especie parental sobre la cual se desarrolló KM71 posee dos genes que se encargan del primer paso en el metabolismo del metanol por parte de la levadura, la oxidación. Estos son los genes *AOX1* y *AOX2*, de los cuales el primero es expresado bajo un promotor fuerte inducido por metanol como única fuente de carbono. Además también tiene una mutación en el gen *Arg4*, el cual provoca que la levadura sea auxótrofa para arginina.

En KM71, el gen *AOX1*, fue reemplazado por un constructo en el que el gen *Arg4* sin la mutación reemplaza desde el codón 16 hasta el 227 de dicho gen, quedando los extremos 5' y 3' del gen *AOX1* rodeando al gen *Arg4*, y de esta manera, le otorga prototrofia ante arginina. Así también se genera el sitio de recombinación homóloga 3' *AOX1*.

El vector pPIC9k posee la particularidad de integrarse en múltiples copias en el genoma de *P. pastoris* KM71. Se integra al genoma mediante recombinación homóloga con los sitios 3' *AOX1* el cual también se encuentra presente en su construcción. Al integrarse, este deja un nuevo sitio 3' *AOX1* disponible para una segunda inserción (Figura 18).

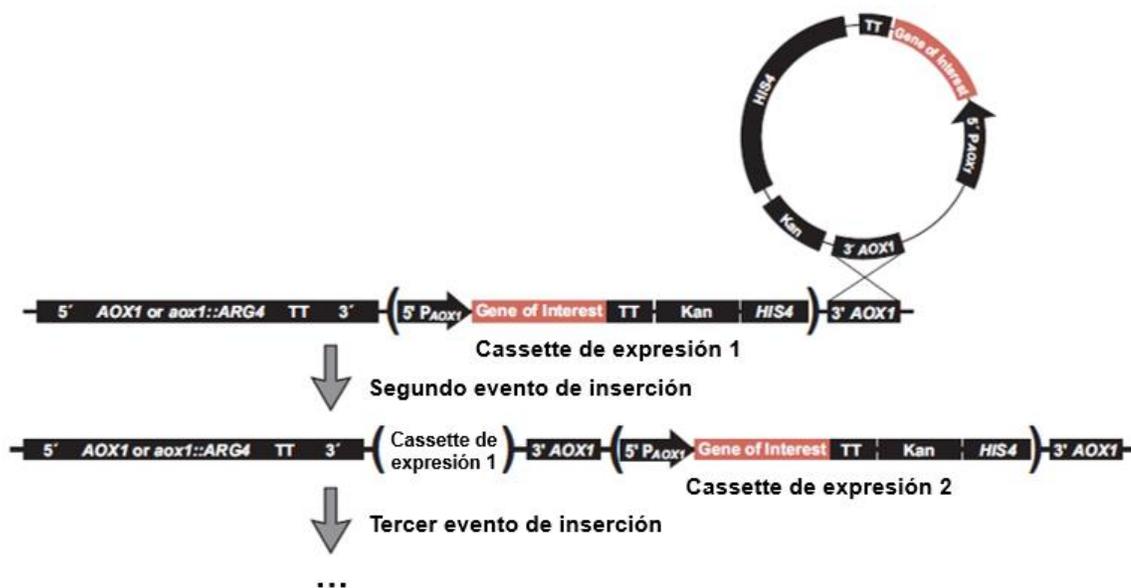


Figura 18. Inserción del vector pPIC9k en el genoma de *P. pastoris* por recombinación homóloga. El vector pPIC9k se inserta al genoma de *P. pastoris* por recombinación homóloga en los sitios 3' *AOX1* presente en ambos, dejando un nuevo sitio donde se puede producir una segunda inserción. Modificado de Expresión MultiCopia en *P. pastoris* (Invitrogen).

Pichia pastoris KM71 también es auxótrofa para histidina, y gracias a que el vector posee dentro de su construcción el gen *His4*, codificante para la última proteína que permite de la síntesis de histidina, es posible seleccionar las colonias transformantes mediante prototrofia con éste aminoácido.

El vector además posee dentro de su construcción el gen de resistencia a Kanamicina (*Kan*), el cual en *P. pastoris* otorga resistencia al antibiótico G418. Debido a que el número de veces que el vector se inserta en el genoma de la levadura conlleva a un mayor número de copias de éste gen, la resistencia al antibiótico es más alta según el número de eventos de inserción del vector que hayan ocurrido. Gracias a esto es posible tener una aproximación a la cantidad de veces que se insertó el vector en el genoma de la levadura, ya que esta se correlaciona con la concentración de G418 que puede resistir.

Es por esto que la selección de transformantes se realizó con una primera fase de selección mediante prototrofia, en un medio en ausencia de histidina, y luego se realizó una segunda fase de selección mediante resistencia a concentraciones crecientes de G418.

A partir de esto se escogieron colonias resistentes a concentraciones altas de G418 (2 y 3 mg/mL), lo que correspondería a aproximadamente 6 eventos de inserción del vector en el genoma de *P. pastoris* KM71 (Figura 19). Estas fueron analizadas mediante PCR de colonias, usando los partidores 5' AOX1 y 3' AOX1, y una colonia de *P. pastoris* KM71 transformada con el vector pPIC9k sin inserto como control negativo. Se aplicó un programa con arranque a alta temperatura para evitar amplificaciones inespecíficas. Las colonias que presentaron bandas del tamaño esperado (1729 pb) se seleccionaron como recombinantes. El control negativo se visualizó como una banda de amplificación de 493 pb (Figura 20).

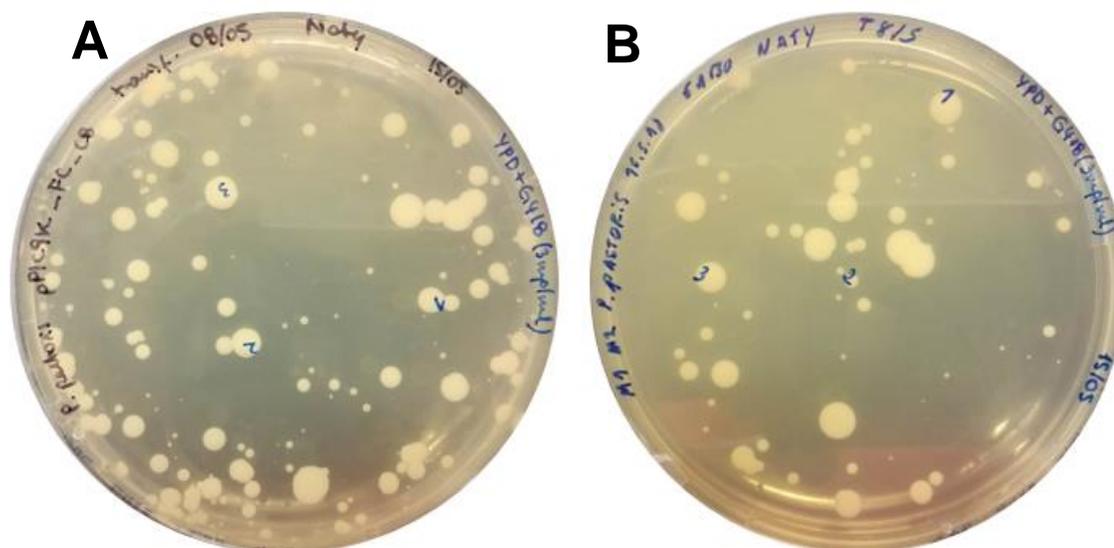


Figura 19. Resultados de la transformación de *P. pastoris* KM71 con vectores pPIC9K recombinantes que contenían las versiones silvestre y mutadas de cel7CDC. En A: Placas de YPD –G418 sembradas con *P. pastoris* transformada con el vector pPIC9k_FC, En B: Placas de YPD –G418 sembradas con *P. pastoris* transformada con el vector pPIC9k_M1M2.

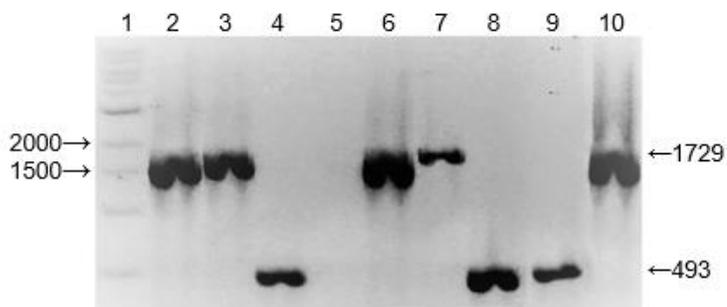


Figura 20. PCR de colonias de las transformaciones de *P. pastoris* KM71 con vectores derivados de pPIC9k. PCR de colonias con partidores AOX, carril 1: Ladder 1kb, 2-4: colonias transformadas con pPIC9k_FC, 6-8: colonias transformadas con pPIC9k_M1M2, 9: control positivo pPIC9k_FC, 10: control negativo, vector pPIC9k vacío.

A partir de las PCR realizadas, se purificó el producto obtenido y éste fue enviado a secuenciar a Macrogen, mediante lo cual, se verificó de la presencia del inserto de interés en el genoma de *P. pastoris* KM71 (Figura 21).

Tabla 7. Colonias de *P. pastoris* transformadas con los vectores derivados de pPIC9k, secuenciadas.

Nombre en Alineamiento	Colonias de <i>P. pastoris</i> KM71
pPIC9k_FC	Colonia 8
pPIC9k_M1M2	Colonia 4

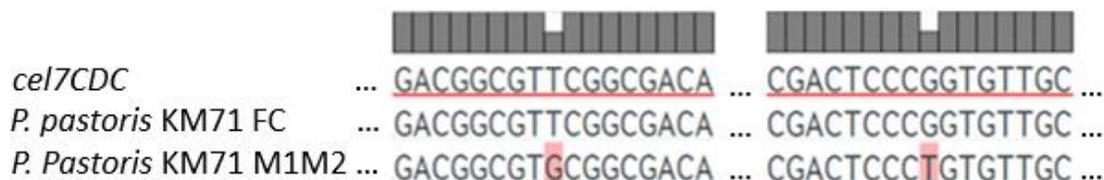


Figura 21. Detalle del alineamiento de la secuenciación de productos de PCR, proveniente de *P. pastoris* KM71 transformadas con vectores derivados de pPIC9k. Se utilizó como templado para el alineamiento la secuencia codificante para el dominio catalítico de Cel7C wild type. Se secuenciaron productos de PCR realizados con el par de partidores AOX 1. Alineamiento de las secuencias en zonas mutadas. Alineamiento de las colonias, de la región en buen estado se encuentra en el Anexo “Secuencias de colonias de *P. pastoris* transformadas con vectores derivados de pPIC9k”.

Desde las colonias resistentes a altas concentraciones de G418 y que fueron positivas en la PCR de colonias se seleccionó una colonia de *P. pastoris* KM71_ FC y otra de *P. pastoris* KM71_ M1M2, con las cuales se procedió a inducir la expresión protéica. El control negativo de la inducción, nuevamente fue *P. pastoris* KM71 transformada con el vector pPIC9k sin inserto. En este punto, debido a la falta de tiempo para continuar la investigación y a que hubo contaminación en las colonias transformadas con los vectores pPIC9k_M1 y pPIC9k_M2, el estudio de las líneas *P. pastoris* transformadas con dichos vectores fueron descartadas. Esta decisión fue tomada ya que su generación no afectaba al cumplimiento del objetivo general de esta investigación.

A través de los días de inducción se fueron recolectando muestras de las cuales se realizaron SDS-PAGE para analizar la cinética de expresión protéica a través de los seis días de la inducción. El peso molecular esperado del dominio catalítico de Cel7C es de 44,73 kDa. En las condiciones de inducción estudiadas se vió la expresión de una proteína con dicho tamaño a partir del segundo día de inducción. A través de esta

cinética, se determinó que en el cuarto día de inducción se obtiene mayor expresión de la proteína, que aún así no es alta (2 a 5 ng aproximadamente), por lo que las siguientes mediciones se realizaron éste día (Figura 22). En ese día se recolectó todo el cultivo, y se separó el sobrenadante de las células. El sobrenadante fue concentrado y fue utilizado para realizar los zimogramas y ensayos de actividad.

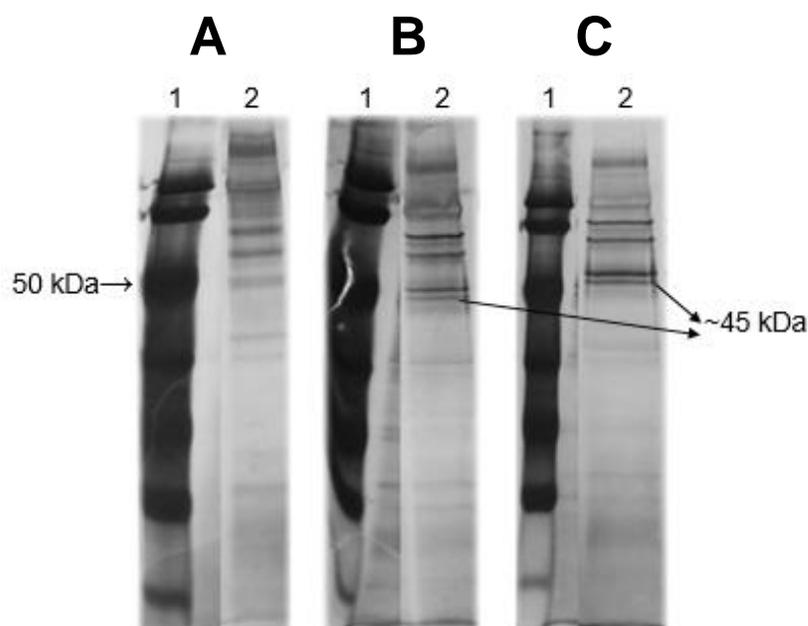


Figura 22. SDS-PAGE de sobrenadantes de cultivo de *P. pastoris* KM71 transformada con pPIC9K_FC y pPIC9K_M1M2, inducidos con metanol por cuatro días a 28°C. A: Inducción del control de inducción *P. pastoris* transformada con el vector pPIC9k vacío. B: Inducción de *P. pastoris* KM71_FC, C: Inducción *P. pastoris* KM71_M1M2. En todas las figuras, carril 1: Ladder Pierce™ Prestained Protein MW Marker, 2: Cuarto día de inducción de las muestras.

Se realizaron zimogramas utilizando como sustrato CMC. Los zimogramas son ensayos donde se permite observar la actividad de enzimas. Se realizan en geles de poliacrilamida en condiciones semidenaturantes y suplementados con el sustrato correspondiente. Luego se renaturan las proteínas y se deja reaccionar bajo las condiciones de temperatura y buffer correspondientes, lo cual posteriormente es revelado a través de tinción del gel. Los resultados (Figura 23) muestran una banda de

degradación del sustrato a la altura de 120 kDa en los clones generados y también en el control negativo de expresión. No se observa degradación por parte de alguna proteína de peso molecular cercano a los 45 kDa, el cual es el peso esperado para Cel7CDC, indicando que no existe producción de las enzimas de manera activa.

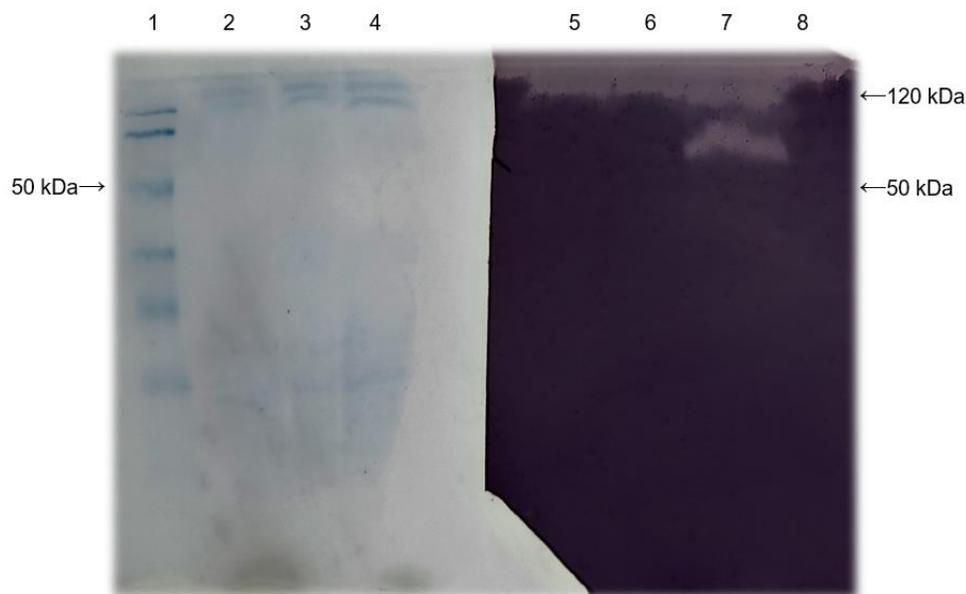


Figura 23. SDS-PAGE y Zimograma del sobrenadante del cuarto día de inducción de las colonias de *P. pastoris* transformadas. SDS-PAGE y Zimograma son imágenes especulares entre sí. Carriles **1,8**: Ladder Pierce™ Prestained Protein MW Marker, **2,7**: inducción de *P. pastoris* transformada con pPIC9k vacío. **3,6**: inducción de *P. pastoris* KM71_FC, **4,5**: inducción de *P. pastoris* KM71_M1M2.

Este resultado concuerda con el obtenido para el análisis de varianza (ANOVA) realizado sobre un ensayo de actividad en triplicado a 50°C, a tiempos de 60, 90 y 120 min, y sobre el ensayo de actividad en triplicado a 30°C, a tres y seis horas, en los cuales se utilizó pNP-C como sustrato. Tras detener la reacción con Na₂CO₃ 2% (p/v) las soluciones se tornaron amarillas y la concentración de p-Nitrofenol se midió mediante espectrofotometría. Resultó que las absorbancias detectadas no mostraron tener diferencias significativas que sugirieran la presencia de actividad celobiohidrolasa

atribuibles a la enzima wild type o a la enzima mutada al compararlas con el control negativo de expresión (Anexo 5) cultivado en el mismo medio e inducidos bajo las mismas condiciones.

Los ensayos de actividad y zimogramas realizados con las proteínas recuperadas del medio intracelular no mostraron mayor diferencia en cuanto a los realizados con las proteínas del medio extracelular (datos no mostrados).

DISCUSIÓN

En un contexto de creciente interés en la búsqueda de enzimas celulolíticas que sean activas y estables a altas temperaturas, se ha recurrido tanto a la utilización de proteínas provenientes de organismos termófilos, como también a la ingeniería de proteínas para generar proteínas mejoradas a partir de mesófilas ya existentes (Lehmann & Wyss, 2001; Ulmer, 1983).

En un estudio previo realizado en CeBiB, se planteó un modelo predictivo para termoestabilización mediante antecedentes bioinformáticos (Jofré, 2012). Para este estudio se seleccionó una proteína celobiohidrolasa de uno de los hongos de pudrición blanca mayormente estudiado, *Phanerochaete chysosporium*.

Se ha observado, en el genoma de éste hongo, la existencia de al menos seis genes codificantes para enzimas del tipo CBHI (Muñoz y col., 2001) y que al menos cuatro de éstos son expresados bajo diferentes condiciones (Wymelenberg y col., 2005). En este caso, la enzima seleccionada fue Cel7C. Esta enzima fue escogida por Jofré, ya que además de contar con una gran capacidad celulolítica, su transcripción es fuerte y específicamente inducida por celulosa, y es expresada y secretada en una alta cantidad en el organismo del cual proviene, por lo cual supone ser un buen ejemplar. Estas últimas características, dentro de un sistema de expresión heterólogo, no son de gran importancia, ya que su expresión depende netamente de las estrategias utilizadas por el sistema de expresión.

El diseño experimental, mediante el cual se propuso termoestabilizar esta proteína comprendía a la implementación de dos mutaciones en zonas flexibles de su estructura.

La propuesta de generar estas mutaciones se originó a partir de la predicción realizada en el trabajo de Jofré (2012), donde se estudió la dinámica molecular de la enzima ante diferentes temperaturas, las cuales simulaban temperaturas alcanzadas en procesos industriales (300, 350 y 400 K). A través de este estudio se identificó regiones flexibles en la proteína, que al mutar, se espera que rigidicen la estructura provocando un aumento en su estabilidad. La identificación de estas regiones se logró basándose en simulaciones generadas sobre un modelo de Cel7C, comparándolo con el comportamiento, también simulado, de estructuras cristalográficas de proteínas termoestables existentes. Jofré propuso realizar tres pares de mutaciones en Cel7C, reemplazando los aminoácidos existentes en las posiciones a mutar por cisteínas, con el fin de que se creara un puente disulfuro. Cada par de mutaciones fue propuesto para una variante diferente de Cel7C. Se propuso cambiar los aminoácidos Q252 y P306 para la variante A, los aminoácidos F325 y G379 para la variante B, y a los aminoácidos N362 y G391 en la variante C. Dentro de éstas, se escogió la variante B para realizar la validación del método ya que ésta era la única que mostraba reducir la flexibilidad de la estructura en las regiones que se simularon las mutaciones y su estructura resultaba ser más compacta que la enzima nativa a altas temperaturas.

En este trabajo, estas mutaciones se generaron mediante PCR de mutación sitio dirigida. Una vez realizadas las mutaciones, los fragmentos se integraron en el vector de clonación pBluescript II SK (+) y fueron perpetuados en *E. coli* TOP 10. Luego, a partir de los vectores, se generaron copias de los genes de interés mediante PCR, para posteriormente insertarlos en pPIC9k, un vector para expresión en *Pichia pastoris* regulado con metanol. Se analizó la secuencia de los productos de cada paso, mediante lo cual se comprobó la integridad, la presencia de las mutaciones específicas realizadas

y la ausencia de otras mutaciones espontáneas. En todos los casos las secuencias coincidieron con las esperadas, de manera que los resultados subsecuentes obtenidos son fieles a las secuencias de interés.

Anteriormente, la enzima Cel7C había sido expresada en *E. coli* utilizando el sistema pET de clonación y expresión (pET System manual), pero resultó expresarse de manera inactiva e intracelular (Howard, 1997). En ese trabajo fue necesario realizar un procedimiento para el plegamiento *in vitro* de la enzima, lo cual consistía en un protocolo tedioso del cual se obtuvo una enzima activa sobre 4-metilumbeliferonil- β -D-celobiosido/lactopiranosido (MUC y MUL respectivamente) y no sobre pNP-L, pNP-C, CMC ni Avicel. Más tarde, esta enzima se expresó fusionada a Glutathion S-transferasa (GST) provista por el vector pGEX-4T-3 en *E. coli* BL21(DE3) como sistema heterólogo de expresión, resultando expresarse de manera extracelular, y sin necesidad de un plegamiento posterior *in vitro*. La enzima resultó ser activa sobre los sustratos poliméricos CMC y Avicel (Howard y col., 2003).

En este trabajo se tomó la decisión de trabajar con *P. pastoris* como sistema heterólogo de expresión. Se escogió esta levadura por diversos motivos, dentro de estos está que su material genético es fácil de manipular, que posee una alta frecuencia de transformación con DNA, que también posee un alto nivel de expresión de proteínas tanto en el medio intracelular o extracelular y que realiza modificaciones postraduccionales, lo cual es muy importante en la expresión de proteínas de origen eucarionte (Cereghino & Cregg, 2000). Para la expresión de las proteínas en *P. pastoris*, se utilizó el vector pPIC9k, el cual es un vector diseñado específicamente para integrarse en el genoma de las cepas KM71 o GS115 de *P. pastoris* en múltiples copias, provocando así una alta expresión de la proteína de interés. Se describe que tras una

transformación de esta levadura con este vector se debe pasar por dos etapas de selección, las cuales permitirán dar con un clon portador de múltiples copias en el genoma. Primero se realiza una fase de selección por prototrofia en medio RDB sin histidina, y luego se seleccionan clones portadores de múltiples copias del vector insertas en el genoma, mediante cultivo de las colonias que crecieron en ausencia de histidina, sobre medio YPD suplementado con concentraciones crecientes de G418 (Geneticina). La resistencia de *P. pastoris* a altas concentraciones de este antibiótico (considerado alto desde 2mg/mL), está otorgada por una gran cantidad de copias del vector insertas en el genoma, por lo tanto se ocupa como un indicador.

Las colonias portadoras del fragmento *wild type* y del fragmento con las dos mutaciones que fueron escogidas, se analizaron mediante PCR de colonias a partir de DNA genómico de las levaduras y los productos de las PCR fueron secuenciados (Figura 21). Los análisis de estos resultados demostraron que ambos clones portaban las secuencias de interés en buenas condiciones, pero al momento de realizar la expresión de las enzimas, se observó que éstas no tenían actividad. Los zimogramas se realizaron utilizando carboximetilcelulosa (CMC) como sustrato debido a que en estudios anteriores se menciona que la enzima Cel7C posee actividad sobre este sustrato (Uzcategui y col., 1991). Estos ensayos fueron realizados con el fin de determinar si las proteínas observadas en el SDS-PAGE correspondían al dominio catalítico *wild type* y mutado de Cel7C, lo cual no pudo ser comprobado. Es posible que la diferencia de resultados entregados entre el SDS-PAGE donde se observó la producción de enzimas de un tamaño compatible con el esperado para la enzima de interés y el resultado del zimograma se deba a que la actividad de esta enzima sobre CMC como sustrato, no haya sido detectada, tanto por la cantidad de enzima producida, como por la

especificidad con el sustrato, como también porque la enzima detectada en el SDS-PAGE no sea Cel7C. El SDS-PAGE mostrado en la figura 22 corresponde a la primera inducción realizada, donde los pasos a seguir fueron exactamente los que se indican en el protocolo de expresión del manual de Expresión MultiCopia en *P. pastoris* (Invitrogen), luego en búsqueda de una optimización del método, el volumen y concentración de cultivos fueron modificados, manteniendo en todos los casos la temperatura de crecimiento de 28°C y una agitación constante de 250 rpm, lo cual puede haber afectado a los resultados posteriores obtenidos. Se necesita más tiempo para poder optimizar el método de expresión de la proteína, para así poder realizar análisis más confiables y replicar el método de manera trazable en el tiempo, para realizar comparaciones más directas.

Además, tras la realización del SDS-PAGE como una primera prueba de detección de la expresión de la proteína, su expresión también podría ser analizada y/o cuantificada en paralelo, realizando pruebas como Northern Blot o RT-Q-PCR, donde sería posible observar si el gen se está transcribiendo o no para ser traducido posteriormente. Realizar pruebas de detección de enzimas, como Inmunoblot, también sería posible de llevar a cabo. Existen anticuerpos monoclonales tanto para CBHI, CBHII como EG provenientes de *T. reesei* que han sido generados para detectar distintos dominios y partes de estas proteínas (Mischak y col., 1989; Aho y col., 1991), los cuales facilitarían la detección de celulasas con estructura homóloga a estas enzimas, como es el caso de Cel7C.

Los ensayos de actividad que se realizaron sobre los sobrenadantes recuperados, fueron realizados durante dos horas, realizando mediciones cada media hora, y durante seis horas, realizando las mediciones cada tres horas. La temperatura utilizada en el primer ensayo de actividad fue de 50°C y en el segundo ensayo de actividad fue de 30°C. En

ambos casos se utilizó *p*NP-C como sustrato. Previamente el ensayo estaba siendo probado utilizando *p*NP-L como sustrato, pero no se observaban cambios colorimétricos ni a través de las mediciones de absorbancia a través del tiempo, por lo cual en esta etapa se descartó como sustrato. Estos sustratos habían sido considerados porque previamente se había determinado que Cel7C tiene actividad sobre ellos (Uzcategui y col., 1991). A través de estos ensayos realizados no mostraron diferencias significativas de absorbancia a 405 nm respecto al blanco sin enzima, indicando ausencia de actividad celobiohidrolasa en los sobrenadantes de *P. pastoris* KM71_FC y KM71_M1M2. Estos resultados, en conjunto con los de SDS-PAGE y los zimogramas realizados, indican que existe la posibilidad de que la proteína Cel7C se esté produce en *Pichia pastoris*, pero de manera inactiva. Estos resultados no permiten validar el método propuesto por Jofré (2012) para la selección de residuos de aminoácido a mutar, pero si constituyen una aproximación para su validación. Es importante tener en cuenta que debido a las limitantes de tiempo en este trabajo de tesis, no fue posible proceder a realizar mejoras en el protocolo que permitieran obtener la enzima activa.

La optimización del método de inducción para la expresión de las proteínas por parte de *P. pastoris* es un asunto que debe llevarse a cabo destinando el tiempo necesario y la modificación de variables como la temperatura, para evaluar la sensibilidad a ésta en el proceso de expresión de las proteínas; la agitación del cultivo, trabajando con matraces con deflectores para, de esta manera, aumentar la oxigenación del cultivo; y el volumen del cultivo, tanto para encontrar una concentración óptima de células a inducir como para contar con el volumen suficiente de sobrenadante para realizar los estudios de expresión y de actividad de las enzimas evaluadas.

En un estudio anterior llevado a cabo en CeBiB se trabajó con la secuencia de Cel7C. En ese entonces, la secuencia de Cel7C con la cual se trabajó contaba con el CBD. En la expresión de la proteína generada en dicho trabajo se encontró una proteína que se expresaba en mayor cantidad el cuarto día de inducción y que correspondía al peso molecular esperado según su secuencia aminoacídica (Anexo 4). A diferencia de dicho trabajo, en esta investigación, la implementación de las mutaciones se realizó sobre la secuencia codificante al dominio catalítico de Cel7C, escindiendo el péptido señal de exportación de la secuencia nativa, el dominio de unión a carbohidratos y el linker que une ambos dominios, para así trabajar con el dominio responsable de la actividad celulolítica. Comparando los resultados de dicho trabajo con las investigaciones llevadas a cabo en el presente trabajo de título, se deduce que la presencia del CBD es importante para la funcionalidad de Cel7C.

Debido a la drástica modificación realizada sobre la estructura de la enzima al escindir el CBD, este factor resulta ser uno de los motivos aparentes por el cual no se logró observar su actividad. Se tomó la decisión de trabajar sobre el dominio catalítico de Cel7C ya que la predicción realizada en el estudio de Jofré fue hecha sobre esta parte de la enzima. Jofré se basó en el modelo tridimensional de Cel7D disponible en las bases de datos el cual corresponde al dominio catalítico de la enzima (Muñoz y col., 2001). Junto a esto, los sustratos con los que se midió la actividad celobiohidrolasa de la enzima no son sustratos de celulosa en estado cristalino, para los cuales la unión del CBD es específica, lo cual hacía suponer que no fuera requerido. La adición del CBD en este estudio podría afectar a la conformación de la enzima y comportarse diferente a la estructura modelada por Jofré, por lo cual la validación del método se pensó a modo de ser llevada a cabo de manera fiel a su trabajo.

Existe también la posibilidad de que se hayan generado puentes disulfuros entre pares de cisteínas diferentes a los propuestos por Jofré mediante las mutaciones, provocando un cambio conformacional en la estructura, lo cual se encuentra directamente relacionado con la actividad de la enzima. La posibilidad de que suceda un fenómeno tal es alta, ya que la secuencia con la que se trabajó contiene dieciocho cisteínas además de los dos reemplazos de aminoácidos realizados.

Otro factor que pudo haber afectado la expresión de la proteína y a su actividad, es la diferencia en la preferencia de codones entre organismos. Se ha visto que en *Pichia pastoris* esta diferencia afecta a la expresión de las proteínas de interés. Existen reportes en los que una optimización de codones a aquellos utilizados por *P. pastoris* ha aumentado significativamente el rendimiento en la expresión de algunas proteínas provenientes de otros organismos en comparación a la expresión de la secuencia wild type (Yu y col., 2013; Li y col., 2014). La utilización de Guanina y Citosina entre *P. pastoris* y hongo *P. chrysosporium* difiere en gran medida, teniendo *P. pastoris*, una menor utilización de estos nucleótidos en su genoma. Ante esto, tanto la corrección de esta fase del estudio, como la modificación de variables a estudiar para la expresión y la actividad de las proteínas, fueron destinadas a pruebas posteriores debido a limitaciones del tiempo destinado a la realización del Seminario de Título.

Junto a los factores descritos, a través de este estudio es posible percatarse de la limitación generada por el hecho de haber considerado sólo un clon de *P. pastoris* para evaluar la expresión y la actividad de la proteína wild type y de la proteína mutante. Se ha descrito que existe mucha variabilidad clonal en la expresión de proteínas recombinantes, tanto en bacterias como en levaduras (Almeida y cols., 2001; Cregg y cols., 1993). En ese sentido para tener mayor oportunidad de encontrar clones

productores de enzima activa, este estudio debió haber sido realizado con un número más grande de clones como muestra. Se recomienda hacer esto en futuras investigaciones, para de esta manera, encontrar un clon candidato con mejores características de expresión.

Por otro lado, para facilitar la manipulación de las enzimas, el estudio de su actividad y para simplificar su purificación para su caracterización, e incluso para ahorrar materiales y tiempo de investigación, sería bueno considerar agregar una etiqueta de afinidad como una cola de histidina a la secuencia de la enzima. De esta manera, una vez detectados los clones productores de las enzimas activas y con alto rendimiento de producción, sería posible trabajar directamente con ellos, descartando una segunda fase en el estudio para agregar la cola de histidina y una nueva selección de clones.

Actualmente se está desarrollando una línea investigativa en el CeBiB considerando las variables a mejorar a partir del trabajo descrito en éste documento. Para eliminar la variable de sensibilidad ante la temperatura en este estudio, se realizaron inducciones de los clones obtenidos en este trabajo para la producción de la proteína a temperaturas más bajas (20, 25 y 30°C). Como resultado, se ha obtenido que al expresar la proteína en los clones *P. pastoris* KM71_ FC y *P. pastoris* KM71_ M1M2 a una menor temperatura (20°C) es posible observar claramente la expresión de la proteína de un peso molecular de alrededor de 45 kDa, la cual no se encuentra en el clon de *P. pastoris* transformada con el vector vacío. Al igual que en este estudio, a pesar de ver su expresión mediante SDS-PAGE, al momento de medir su actividad sobre *p*NP-C, esta no fue detectada.

Actualmente los esfuerzos se centran en generar vectores con las mismas mutaciones realizadas en este trabajo, pero sobre la secuencia completa del gen *Cel7C*, es decir, la secuencia que posee el CBD, debido a que ya se comprobó que de este modo la proteína

se expresa activa. A modo de proyección se plantea el propósito de obtener clones productores de enzimas nativa y mutada activas con los cuales validar la propuesta de Jofré, quitando el foco de la producción de las enzimas con alto rendimiento.

CONCLUSIONES

En este estudio se realizaron dos mutaciones en la secuencia codificante para el dominio catalítico de la enzima Cel7C mediante mutagénesis por sitio dirigida mediante PCR. Estas fueron realizadas de manera exitosa, y con estas secuencias se generaron vectores derivados de pBluescript II SK (+) con los que se transformó *E. coli* TOP10, y de pPIC9k con los que se transformó también esta bacteria y la levadura *P. pastoris*. La integridad de las secuencias posterior a las transformaciones fue corroborada en cada paso a través de secuenciaciones.

A pesar de que se observó la expresión de una proteína de aproximadamente 45 kDa, como es esperado para la secuencia evaluada, no se encontró actividad celobiohidrolasa relacionada sobre *p*NP-C y CMC, existiendo una serie de factores que pudieran haber afectado, siendo el más probable, la ausencia de CBD en la proteína. Es por esto que esta investigación constituye un acercamiento a la validación del método de termoestabilización mediante predicción con herramientas bioinformáticas propuesto por Jofré (2012), proponiéndose también mejoras en los procedimientos a llevar a cabo para una correcta validación, y no corresponde a la validación en su totalidad, cumpliéndose así el objetivo del trabajo realizado.

BIBLIOGRAFÍA

- Aho, S., Olkkonen, V., Jalava, T., Paloheimo, M., BÜHLER, R., NIKU-PAAVOLA, M. L., ... & Korhola, M. (1991). Monoclonal antibodies against core and cellulose-binding domains of *Trichoderma reesei* cellobiohydrolases I and II and endoglucanase I. *European journal of biochemistry*, 200(3), 643-649.
- Almeida, M. S., Cabral, K. S., de Medeiros, L. N., Valente, A. P., Almeida, F. C., & Kurtenbach, E. (2001). cDNA cloning and heterologous expression of functional cysteine-rich antifungal protein Psd1 in the yeast *Pichia pastoris*. *Archives of biochemistry and biophysics*, 395(2), 199-207.
- Antizar-Ladislao, B., & Turrion-Gomez, J. L. (2008). Second-generation biofuels and local bioenergy systems. *Biofuels, Bioproducts and Biorefining*, 2(5), 455-469.
- Aro, E. M. (2016). From first generation biofuels to advanced solar biofuels. *Ambio*, 45(1), 24-31.
- Atalla, R. H., & Vanderhart, D. L. (1984). Native cellulose: a composite of two distinct crystalline forms. *Science*, 223(4633), 283-285.
- Barnett, J. R., & Bonham, V. A. (2004). Cellulose microfibril angle in the cell wall of wood fibres. *Biological reviews*, 79(2), 461-472.
- Béguin, P., & Aubert, J. P. (1994). The biological degradation of cellulose. *FEMS microbiology reviews*, 13(1), 25-58.
- Brett, C. T., & Waldron, K. W. (1996). *Physiology and biochemistry of plant cell walls* (Vol. 2). Springer Science & Business Media.
- Bumpus, J. A., & Aust, S. D. (1987). Biodegradation of environmental pollutants by the white rot fungus *Phanerochaete chrysosporium*: involvement of the lignin degrading system. *BioEssays*, 6(4), 166-170.
- Cereghino, J. L., & Cregg, J. M. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS microbiology reviews*, 24(1), 45-66.
- Covert, S. F., Bolduc, J., & Cullen, D. (1992a). Genomic organization of a cellulase gene family in *Phanerochaete chrysosporium*. *Current genetics*, 22(5), 407-413.
- Covert, S. F., Wymelenberg, A. V., & Cullen, D. (1992b). Structure, organization, and transcription of a cellobiohydrolase gene cluster from *Phanerochaete chrysosporium*. *Applied and environmental microbiology*, 58(7), 2168-2175.
- Cregg, J. M., Vedvick, T. S., & Raschke, W. C. (1993). Recent advances in the expression of foreign genes in *Pichia pastoris*. *Bio/technology*, 11(8), 905.
- Dobozi, M. S., Szakács, G., & Bruschi, C. V. (1992). Xylanase activity of *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.*, 58(11), 3466-3471.

Dragone, G., Fernandes, B. D., Vicente, A. A., & Teixeira, J. A. (2010). Third generation biofuels from microalgae. *Current research, technology and education topics in applied microbiology and microbial biotechnology*, 2, 1355-1366.

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison III, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, 6(5), 343.

Glycoside Hydrolase Family 7. Lugar de publicación: Carbohydrate-Active Enzymes (CAZy). <http://www.cazy.org/GH7.html>

Glycoside Hydrolase Family 7. Lugar de publicación: Carbohydrate-Active Enzymes (CAZypedia).

https://www.cazypedia.org/index.php/Glycoside_Hydrolase_Family_7#Three-dimensional_structures

Gruno, M., Välijamäe, P., Pettersson, G., & Johansson, G. (2004). Inhibition of the *Trichoderma reesei* cellulases by cellobiose is strongly dependent on the nature of the substrate. *Biotechnology and Bioengineering*, 86(5), 503-511.

Gusakov, A. V. (2011). Alternatives to *Trichoderma reesei* in biofuel production. *Trends in biotechnology*, 29(9), 419-425.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41:95-98.

Henriksson, G., Nutt, A., Henriksson, H., Pettersson, B., Ståhlberg, J., Johansson, G., & Pettersson, G. (1999). Endoglucanase 28 (Cel12A), a new *Phanerochaete chrysosporium* cellulase. *European journal of biochemistry*, 259(1-2), 88-95.

Hendriks, A. T. W. M., & Zeeman, G. (2009). Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresource technology*, 100(1), 10-18.

Henrissat, B., Driguez, H., Viet, C., & Schülein, M. (1985). Synergism of cellulases from *Trichoderma reesei* in the degradation of cellulose. *Bio/technology*, 3(8), 722.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, 77(1), 51-59

Howard, R. L. (1997). Cloning and Expression of *Phanerochaete Chrysosporium* Chb1-LIKE CDNAs in *Escherichia Coli* (Doctoral dissertation).

Howard, R. L., Masoko, P., & Abotsi, E. (2003). Enzyme activity of a *Phanerochaete chrysosporium* cellobiohydrolase (CBHI. 1) expressed as a heterologous protein from *Escherichia coli*. *African Journal of Biotechnology*, 2(9), 301-306.

Jofré, J. (2012). Mejoramiento de la termoestabilidad de enzimas mediante dinámica molecular y análisis de componentes principales. (Tesis de Magister). Universidad de Chile, Santiago, Chile.

Kirk, T. K., & Jeffries, T. W. (1996, March). Roles for microbial enzymes in pulp and paper processing. In ACS symposium series (Vol. 655, pp. 2-14). Washington, DC: American Chemical Society,[1974]-.

Kuhad, R. C., Gupta, R., & Singh, A. (2011). Microbial cellulases and their industrial applications. *Enzyme research*, 2011.

Kumar, S., & Nussinov, R. (2001). How do thermophilic proteins deal with heat?. *Cellular and Molecular Life Sciences CMLS*, 58(9), 1216-1233.

Lasa, I., & Berenguer, J. (1993). Thermophilic enzymes and their biotechnological potential. *Microbiologia*, 9(2), 77-89.

Le Costaouëc, T., Pakarinen, A., Várnai, A., Puranen, T., & Viikari, L. (2013). The role of carbohydrate binding module (CBM) at high substrate consistency: comparison of *Trichoderma reesei* and *Thermoascus aurantiacus* Cel7A (CBHI) and Cel5A (EGII). *Bioresource technology*, 143, 196-203.

Lehmann, M., Pasamontes, L., Lassen, S. F., & Wyss, M. (2000). The consensus concept for thermostability engineering of proteins. *Biochimica et Biophysica Acta (BBA)-protein structure and molecular enzymology*, 1543(2), 408-415.

Lehmann, M., & Wyss, M. (2001). Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Current opinion in biotechnology*, 12(4), 371-375.

Li, Y. M., Li, D. J., Xu, X. J., Cui, M., Zhen, H. H., & Wang, Q. (2014). Effect of codon optimization on expression levels of human cystatin C in *Pichia pastoris*. *Genet Mol Res*, 13(3), 4990-5000.

Li, S., Logan Bashline, L. L., & Gu, Y. (2014). Cellulose synthesis and its regulation. *The Arabidopsis book/American Society of Plant Biologists*, 12.

Lü, J., Sheahan, C., & Fu, P. (2011). Metabolic engineering of algae for fourth generation biofuels production. *Energy & Environmental Science*, 4(7), 2451-2466.

Mandels, M., & Reese, E. T. (1965). Inhibition of cellulases. *Annual review of Phytopathology*, 3(1), 85-102.

Martinez, D. et al. (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn *Hypocrea jecorina*). *Nat. Biotechnol.* 26, 553–560.

Mischak, H., Hofer, F., Messner, R., Weissinger, E., Hayn, M., Tomme, P., ... & Kubicek, C. P. (1989). Monoclonal antibodies against different domains of cellobiohydrolase I and II from *Trichoderma reesei*. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 990(1), 1-7.

Mueller, S. C., & Brown, R. M. (1980). Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. *The Journal of cell biology*, 84(2), 315-326.

Muñoz, I. G., Ubhayasekera, W., Henriksson, H., Szabó, I., Pettersson, G., Johansson, G., & Ståhlberg, J. (2001). Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Å resolution and homology models of the isozymes1. *Journal of molecular biology*, 314(5), 1097-1111.

Mussatto, S. I., & Teixeira, J. A. (2010). Lignocellulose as raw material in fermentation processes.

Naqvi, M., & Yan, J. (2015). First-Generation Biofuels. Handbook of Clean Energy Systems, 1-18.

Nidetzky, B., Steiner, W., Hayn, M., & Claeysens, M. (1994). Cellulose hydrolysis by the cellulases from *Trichoderma reesei*: a new model for synergistic interaction. Biochemical Journal, 298(3), 705-710.

Niehaus, F., Bertoldo, C., Kähler, M., & Antranikian, G. (1999). Extremophiles as a source of novel enzymes for industrial application. Applied microbiology and biotechnology, 51(6), 711-729.

Oficina de Estudios y políticas agrarias (ODEPA) – Gobierno de Chile. Barrera P., Daniel. (Enero, 2018). Celulosa chilena: Avances y perspectivas de su comercio exterior. Enero de 2018 [archivo PDF]. Recuperado de <https://www.odepa.gob.cl/publicaciones/articulos>

O'sullivan, A. C. (1997). Cellulose: the structure slowly unravels. Cellulose, 4(3), 173-207.

Palanna, O. G. (2009). Engineering chemistry. Tata McGraw-Hill Education.

Paneque, M., Román-Figueroa, C., Vázquez-Panizza, R., Arriaza, J. M., Morales, D., & Zulantay, M. (2011). Bioenergía en Chile. Organización de las Naciones Unidas para la Alimentación y la Agricultura. Facultad de Ciencias Agronómicas, Universidad de Chile.

"pET System manual" (PDF). Novagen.

Preston, R. D. (1975). X-ray analysis and the structure of the components of plant cell walls. Physics Reports, 21(4), 183-226.

Saxena, I. M., & Brown Jr, R. M. (2005). Cellulose biosynthesis: current views and evolving concepts. Annals of botany, 96(1), 9-21.

Sukumaran, R. K., Singhania, R. R., & Pandey, A. (2005). Microbial cellulases-production, applications and challenges.

Suto, M., & Tomita, F. (2001). Induction and catabolite repression mechanisms of cellulase in fungi. Journal of Bioscience and Bioengineering, 92(4), 305-311.

Somerville, C. (2006). Cellulose synthesis in higher plants. Annu. Rev. Cell Dev. Biol., 22, 53-78.

Takashima, S., Iikura, H., Nakamura, A., Hidaka, M., Masaki, H., & Uozumi, T. (1998). Isolation of the gene and characterization of the enzymatic properties of a major exoglucanase of *Humicola grisea* without a cellulose-binding domain. The Journal of Biochemistry, 124(4), 717-725.

Takashima, S., Ohno, M., Hidaka, M., Nakamura, A., Masaki, H., & Uozumi, T. (2007). Correlation between cellulose binding and activity of cellulose-binding domain mutants of *Humicola grisea* cellobiohydrolase 1. FEBS letters, 581(30), 5891-5896.

- Teeri, T. T. (1997). Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends in biotechnology*, 15(5), 160-167.
- Turner, P., Mamo, G., & Karlsson, E. N. (2007). Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microbial cell factories*, 6(1), 9.
- Ulmer, K. M. (1983). Protein engineering. *Science*, 219(4585), 666-671.
- Uzcategui, E., Ruiz, A., Montesino, R., Johansson, G., & Pettersson, G. (1991). The 1, 4- β -D-glucan cellobiohydrolases from *Phanerochaete chrysosporium*. I. A system of synergistically acting enzymes homologous to *Trichoderma reesei*. *Journal of biotechnology*, 19(2-3), 271-285.
- 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. *Bioresource technology*, 100(1), 345-349.
- Ward, O. P., & Moo-Young, M. (1988). Thermostable enzymes. *Biotechnology advances*, 6(1), 39-69.
- Wymelenberg, A. V., Covert, S., & Cullen, D. (1993). Identification of the gene encoding the major cellobiohydrolase of the white rot fungus *Phanerochaete chrysosporium*. *Applied and environmental microbiology*, 59(10), 3492-3494.
- Wymelenberg, A. V., Sabat, G., Martinez, D., Rajangam, A. S., Teeri, T. T., Gaskell, J., ... & Cullen, D. (2005). The *Phanerochaete chrysosporium* secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium. *Journal of biotechnology*, 118(1), 17-34.
- Yu, P., Yan, Y., Gu, Q., & Wang, X. (2013). Codon optimisation improves the expression of *Trichoderma viride* sp. endochitinase in *Pichia pastoris*. *Scientific reports*, 3, 3043.
- Zhang, Y. H. P., Himmel, M. E., & Mielenz, J. R. (2006). Outlook for cellulase improvement: screening and selection strategies. *Biotechnology advances*, 24(5), 452-481.

ANEXO

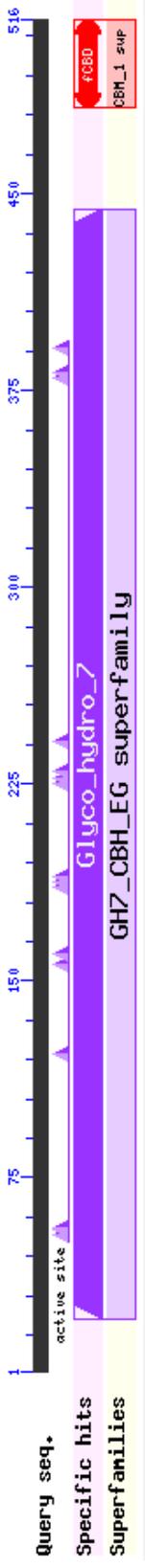
```

aaccaccgcacgctcacgtcgcagaagtgcacgaagagcggcggtgctcgaacctgaac
N H R T L T S Q K C T K S G G C S N L N
accaagatcgtgctcgacgcgaactggcgctgggtccactccactccggtacaccaat
T K I V L D A N W R W L H S T S G Y T N
tgctacactggcaaccagtgaggacgcaacgctctgccccgacggcaagacgtgtgctggcc
C Y T G N Q W D A T L C P D G K T C A A
aactgcgcgctcgacggcgccgactacaccggcacctacggcatcaccgcgagcggctcg
N C A L D G A D Y T G T Y G I T A S G S
tcgctcaagctgcagttcgtcactggctcgaacgctcggctccggtgtctacctcatggcg
S L K L Q F V T G S N V G S R V Y L M A
gacgacacgcactaccagatgttccagctcctgaaccaggagttcaccttcgacgtcgac
D D T H Y Q M F Q L L N Q E F T F D V D
atgtcgaacctcccctgcggtctcaacggcgccctctacctctctgcatggacgctgat
M S N L P C G L N G A L Y L S A M D A D
ggcggcatggccaagtaccgaccaacaaggctgggtgctaagtacgggtaccggctactgc
G M A K Y P T N K A G A K Y G T G Y C
gactcgcagtgcccgcgcacatcaagttcatcaatggcgaggctaacgttgagggctgg
D S Q C P R D I K F I N G E A N V E G W
aacgccaccagtgccaacgcgggtaccggtaactacgggtacgtgctgcactgagatggat
N A T S A N A G T G N Y G T C C T E M D
atctgggagggcaacaatgatgcccgtgcatacactccccacccttgaccaccaacgct
I W E A N N D A A A Y T P H P C T T N A
cagaccgctgctccggaagcgactgcaccgcgacaccggtctctgcgacgccagcggc
Q T R C S G S D C T R D T G L C D A D G
tgcgacttcaactccttccgcatggggcaccagaccttccctcggaagggcctgaccgctc
C D F N S F R M G D Q T F L G K G L T V
gacacctcgaagcccttaccgctcgtcaccagttcatcaccaacgacggcacctccgcg
D T S K P F T V V T Q F I T N D G T S A
ggcaccctcaccgagatccgctctctatgtccagaatgggaaggtcatccagaactcc
G T L T E I R R L Y V Q N G K V I Q N S
tccgtcaagatccccggcatcgaccgggtcaactcgatcaccgacaacttctgctcgcaa
S V K I P G I D P V N S I T D N F C S Q
cagaagacggcgttccggcgacaccaactacttcgcgacacgggtgggtctcaagcaggtc
Q K T A F G D T N Y F A Q H G G L K Q V
ggtagggcgttgcgcacgggcatgggtcctcgcgctctcgatctgggatgactacgccgcg
G E A L R T G M V L A L S I W D D Y A A
aacatgctctggctcgactccaactaccggacgaacaaggaccggcgtcgactcccgggtgtt
N M L W L D S N Y P T N K D P S T P G V
gcccgggcacctgcgcgaccacttccggcgctccccgcgacgatcgaggcgcagagcccc
A R G T C A T T S G V P A Q I E A Q S P
aacgcctacgttgtcttctcgaacatcaagttcggcgacctcaacaccacg
N A Y V V F S N I K F G D L N T T

```

SECUENCIA ORIGINAL	MUTACIÓN	REEMPLAZO AMINOACÍDICO
TTC	TGC	F → C
GGT	TGT	G → C

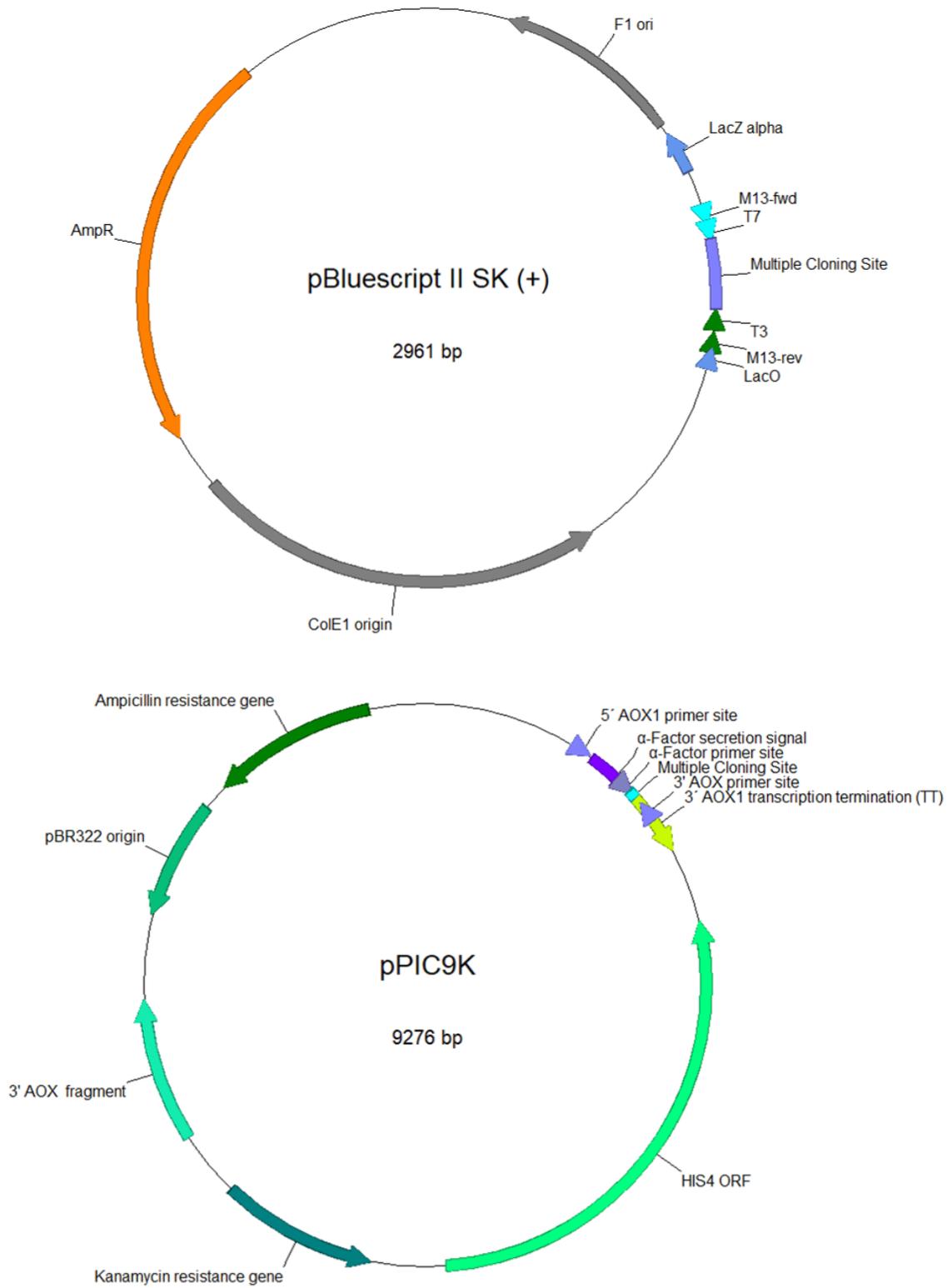
Anexo 1. Secuencia del Dominio Catalítico de Cel7C. Muestra sitios donde se realizan las mutaciones. En la tabla se muestran los cambios de nucleótidos (mutaciones) dentro de los codones correspondientes y el cambio de aminoácido que genera.



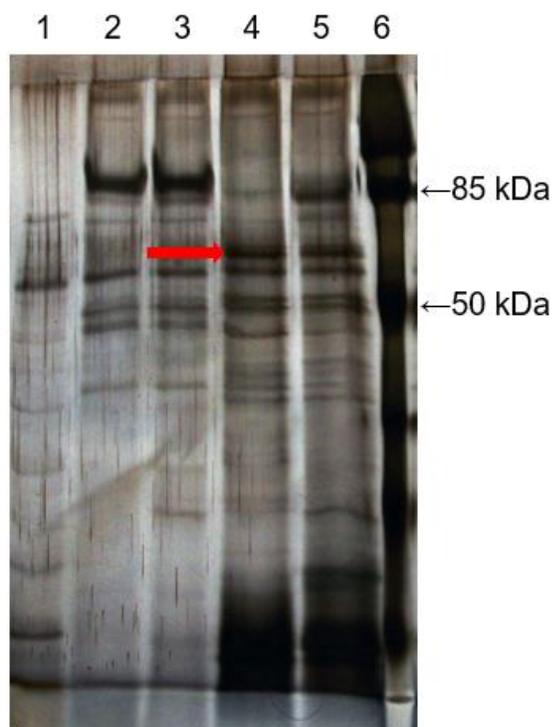
>AAB46373.1 exo-cellobiohydrolase I precursor [Phanerochaete chrysosporium]

MFR**TATLLAFTMAAMVFGQQVGTNTAE**NHRTLTLSQKCTKSGCSNLNTKIVLDANWRWLHSTSGYTNCYT
 GNQWDAFLCPDGKTCAAANCALDGADYTGTYGITASGSSKLQFVTGSNVGSRVYLMADDTHYQMFQLLNQ
 EFTFDVDMNSLPCGLNGALYLAMDADGGMAKYPTNKAGAKYGTGYCDSQCPRDIKFINGEANVEGWNAT
 SANAGTGNYGTCCTEMDIWEANNDAAYTPHPCTTNAQTRCSGSDCTRDTGLCDADGCDNFNFRMGDQTF
 LGKGLTVDTSKPFVVTQFITNDGTSAGTLTEIRRLYVQNGKVIQNSSVKIPGIDPVNSITDNFCSQQKI
 AFGDTNYFAQHGLKQVGEALRTGMVLALSIVDDYAAANMLWLDNSYPTNKDPSSTPGVARGTCATTSGVPA
 QIEAQSPNAYVVFNSIKFGDLNNTTYTGTVSSSVSSSHSSTSTSSSHSSSTPPTQPTGVTVP**PQWGQCGG**
IGYTGSTTCASPYTCHVLPNPPYSQCY

Anexo 2. Secuencia precursora de la enzima Cel7C completa e identificación de dominios conservados. En la secuencia se encuentra en verde el péptido señal de exportación, en gris el dominio catalítico y en rojo el dominio de unión a carbohidratos. La identificación de los dominios se observa en la gráfica.



Anexo 3. Mapas de los plasmidios pBluescript II SK (+) y pPIC9k.



Anexo 4. SDS-PAGE de inducción de Cel7C de un clon de *P. pastoris* generado en un estudio anterior. Se ve que al cuarto día de inducción hay una mayor expresión de la enzima de un

Anexo 5. Desviación y promedio de los datos obtenidos a través de los ensayos de actividad. Con estos datos se realizó el test de ANOVA.

TIEMPO (min)	DESVIACIÓN ESTÁNDAR			PROMEDIO		
	FC	M1M2	pPIC9k	FC	M1M2	pPIC9k
60	0,019	0,023	0,031	0,137	0,133	0,112
90	0,051	0,052	0,032	0,134	0,132	0,101
120	0,043	0,030	0,033	0,140	0,123	0,107

TIEMPO (h)	DESVIACIÓN ESTÁNDAR			PROMEDIO		
	FC	M1M2	pPIC9k	FC	M1M2	pPIC9k
3	0,025	0,021	0,040	0,113	0,102	0,174
6	0,027	0,028	0,062	0,283	0,254	0,435