



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

**IDENTIFICACION DE POTENCIALES RESIDUOS DETERMINANTES DE
MAYOR AFINIDAD A CELULOSA EN UNA ENDOGLUCANASA MEDIANTE
ESTUDIOS DE DOCKING MOLECULAR**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL EN BIOTECNOLOGÍA
E INGENIERA CIVIL QUÍMICO

MARÍA TERESA BULDRINI OVIEDO

PROFESOR GUÍA:

ORIANA SALAZAR AGUIRRE

MIEMBROS DE LA COMISIÓN:

JOSÉ CRISTIAN SALGADO HERRERA
MARÍA ELENA LIENQUEO CONTRERAS

SANTIAGO DE CHILE
MARZO 2012

RESUMEN

El bioetanol es un combustible complementario a las gasolinas, producido mediante fermentación de azúcares simples derivados de sacarosa, almidón o celulosa. Debido a su gran abundancia el material celulósico es particularmente atractivo como fuente de combustible. La hidrólisis de la celulosa puede realizarse utilizando métodos químicos (ácidos o bases) o enzimáticos. La hidrólisis enzimática genera productos más puros y consume menos energía; sin embargo, debido a su bajo rendimiento se requiere usar grandes cantidades de enzima, lo cual eleva considerablemente los costos de producción.

Las endoglucanasas son glicosil hidrolasas que rompen enlaces β -1,4 en polímeros de glucosa. Actúan sobre enlaces internos de la celulosa amorfa generando su despolimerización. La estructura de las celulasas incluye un módulo catalítico, unido a un dominio que actúa como centro de anclaje a la celulosa, por lo que se le ha denominado Módulo de Unión a Carbohidrato (CBM por su sigla en Inglés). Estudios previos han demostrado que una alta afinidad a celulosa se correlaciona con un aumento en la hidrólisis de ésta. En el presente trabajo se proponen mutaciones en el CBM de la endo β -1,4 glucanasa de *Trametes versicolor* (TvEG), que aumenten la adsorción de la enzima sobre celulosa, buscando de este modo aumentar la hidrólisis de ésta.

Se generó un modelo tridimensional del CBM de la endoglucanasa de *T. versicolor* mediante modelación comparativa. La estructura obtenida está formada por tres hojas beta antiparalelas unidas entre sí mediante dos puentes disulfuro. El modelo generado se utilizó para llevar a cabo estudios de Docking Molecular, para identificar regiones de interacción celulosa-CBM.

Usando los resultados de la simulación computacional de la interacción CBM-celulosa, se identificaron dos zonas de unión. La primera concuerda con la reportada previamente para CBM's de la familia I, mientras que no existen reportes en la literatura sobre interacción de la celulosa con la segunda zona identificada. A continuación se simuló mediante Docking Molecular, la interacción de celulosa con variantes de CBM que contenían diferentes aminoácidos en las posiciones clave de la segunda zona de interacción; los valores de energía libre de unión entregados por la simulación permitieron sugerir tres variantes de secuencia de CBM. Las variantes seleccionadas fueron: variante 1 (G7Y, F10W), variante 2 (G7Y, F10W, G12Q) y variante 3 (G7Y, G9Q, F10W, G12Q), donde G es glicina, Y tirosina, F fenilalanina, W triptófano y Q glutamina.

Se generó un sistema de expresión recombinante de TvEG mediante clonamiento del gen silvestre de la enzima en el vector pET22b(+) y transformación en *E. coli* BL21(DE3); la proteína se acumuló intracelularmente en forma activa. Las mutaciones planteadas fueron construidas exitosamente mediante mutagénesis sitio dirigida, utilizando la metodología de mutación por extensión de superposición. De esta forma, se deja planteado un sistema de expresión recombinante de las variantes construidas, de modo de en el futuro cuantificar experimentalmente la adsorción de los CBM mutados y validar los resultados del análisis realizado *in silico*.

Se concluye que la estrategia utilizada permitió simular apropiadamente la interacción CBM-celulosa. Los resultados permitieron proponer en base a los estudios de DM, una estrategia de doble anclaje del CBM a celulosa, que si bien es cierto debe ser comprobado experimentalmente, resulta ser novedoso. Además, permitió proponer variantes del CBM de TvEG que potencialmente se adsorberán con mayor afinidad sobre celulosa. Si bien estas mutaciones deben ser validadas experimentalmente, fueron planteadas a partir de un diseño racional lo que aumentaría la probabilidad que éstas presenten mayor afinidad a celulosa.

AGRADECIMIENTOS

Al finalizar esta importante etapa de mi vida, quisiera dar las gracias a todas las personas que me apoyaron, estuvieron conmigo en los buenos y malos momentos e hicieron posible que llegara a esta instancia.

En primer lugar quisiera agradecer a mi familia. A mis padres Beatriz y Horacio quienes siempre me apoyaron en las decisiones que tomé. Gracias por en todo momento animarme a seguir adelante. A mis hermanos Ruby, Karina y Horacio por estar ahí cada vez que los necesité. A mis tíos Celmira, Víctor, Armando y Mauricio, por su constante preocupación en cómo iba mi vida y mis asuntos académicos.

A Cristian, por acompañarme los últimos tres años y medios de mi carrera. Gracias por apoyarme cuando el tiempo escaseaba y las responsabilidades sobran. Gracias por siempre confiar en mí y darme el ánimo en los momentos difíciles.

También quiero agradecer el cariño y comprensión de mis amigos. A Vale y Jose por ser mis amigas desde los tiempos del colegio. A todas esas grandes personas que conocí y con quienes compartí los primeros años de Universidad. A mis amigos Elise, Erika, Sandra, Osvaldo, Nacho, Luchi, Heinz, Felipe T., Wesitos, Caro, Phillipe, en particular a quienes apoyaron ese loco proyecto que resultó ser *“Apoyo Escolar”*.

Mis amigos y futuros colegas. A Cristina, Nati J., Nati M., Claudia, Camilo, Cristian, Werner, Harris, Nacho, Aurelio. Gracias por compartir clases, informes, noches de estudio, buenos ratos y juntas. Gracias por este buen grupo que se formó.

A los profesores de mi comisión. A la profesora Oriana Salazar por su constante apoyo; por todas las dudas que me respondió en el día a día y por animarme a seguir adelante cuando las cosas no resultaban. Al profesor Cristian Salgado por mostrarme las muchas cosas que computacionalmente se pueden hacer y a la profesora María Elena Lienqueo por su buena disposición en el desarrollo de este trabajo.

Quisiera agradecer también a todas las personas y docentes que conocí en el CIBYB. A Nancy y Alexis, por ayudarme cada vez que lo necesité. A Paty, Anita, Icha, Jorge, Sergio, Cami, Coni, Dani S., Dani V., Gaby, Vida y Alicia, por hacer del laboratorio un lugar muy ameno para trabajar. Gracias por cada sobremesa compartida.

ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN	1
1.1 Descripción del proyecto y justificación	1
1.2 Estructura de la celulosa	3
1.3 Proceso de obtención de bioetanol	4
1.4 Hidrólisis enzimática de celulosa	4
1.4.1 Celulasas	5
1.4.2 Celulasas Fúngicas	6
1.4.3 Estructura y características de celulasas	7
1.4.4 Clasificación de módulos de unión a carbohidratos	8
1.5 Módulos de unión a carbohidrato de la familia I	8
1.6 Mecanismo de adsorción a celulosa	10
1.6.1 Modelos de adsorción	10
1.7 Métodos Computacionales utilizados en el diseño racional	12
1.7.1 Modelamiento de estructuras tridimensionales de proteínas por comparación	12
1.7.2 Docking molecular	12
1.8 Técnicas de Ingeniería de Proteínas	15
2. OBJETIVOS	16
3. MATERIALES Y MÉTODOS	17
3.1 Materiales	17
3.1.1 Microorganismos usados	17
3.1.2 Reactivos	17
3.1.3 Vectores usados	18
3.1.4 Enzimas de restricción	19
3.1.6 Herramientas computacionales	21
3.1.7 Síntesis de oligonucleótidos	21
3.1.8 Constructo molecular utilizado	21
3.2 Métodos	22
3.2.1 Metodología computacional	22
3.2.2 Metodología experimental	25
4. RESULTADOS Y DISCUSIÓN	33
4.1 Generación de un modelo tridimensional del módulo de unión a carbohidrato de TvEG	34

4.1.1	Alineamiento CBM de TvEG	34
4.1.2	Evaluación preliminar del modelo generado por Modeller	35
4.1.3	Evaluación del modelo generado y refinado con Modeller	38
4.1.4	Modelo tridimensional del módulo de unión a carbohidrato (CBM)	42
4.2	Estudio de la interacción CBM-celulosa mediante docking molecular	45
4.2.1	Conformaciones del complejo CBM-celulosa obtenidas a partir del análisis de Docking molecular	46
4.2.2	Energías de interacción del CBM de TvEG sobre celulosa	51
4.3	Identificación de mutaciones para mejorar la interacción CBM I de TvEG a celulosa	53
4.3.1	Planteamiento de las mutaciones	53
4.3.2	Análisis mediante docking molecular	55
4.3.3	Selección de mutaciones	59
4.3.4	Estructuras tridimensionales de las mutaciones seleccionadas	60
4.4	Diseño de la estrategia para la construcción de las variantes de CBM TvEG propuestas	64
4.4.1	Clonamiento del gen de la endoglucanasa de <i>T. versicolor</i> en el vector de expresión pET22b(+)	64
4.4.2	Estrategia desarrollada para la construcción de las variantes de TvEG	69
4.5	Construcción de las variantes de TvEG planteadas	72
4.5.1	Obtención de los fragmentos mutados	72
4.5.2	Ligación de los fragmentos mutados a pGem [®] -T Easy	72
4.6	Estudio de las condiciones de expresión de la endoglucanasa de <i>T. versicolor</i> en <i>E. coli</i> BL21 (DE3)	74
4.6.1	Análisis de expresión a distintas temperaturas de inducción	74
5.	CONCLUSIONES	76
6.	BIBLIOGRAFÍA	78
7.	ANEXOS	84
A.	Medios de cultivo y soluciones empleadas	84
	Medios de cultivo líquidos	84
	Medios de cultivo sólidos	84
	Soluciones Stock	84
B.	Secuencia gen (tveg) codificante para la endoglucanasa en estudio (TvEG)	85
C.	Secuencia gen codificante para las distintas variantes de endoglucanasa (TvEG)	85
D.	Ejemplos de códigos de configuración empleado por Vina y Autodock	88

Ej. Archivo configuración Vina:	88
Ej. Archivo de docking utilizado por Autodock:	88
E. Protocolo obtención DNA plasmidial con QIAprep® Spin Miniprep	89
F. Protocolo purificación con Qiagen	89
G. Análisis llevados a cabo para la proteína recombinante	90
Electroforesis de proteínas SDS-PAGE	90
Zimograma	91
Medición de actividad en medio líquido	91
H. Códigos empleados en el modelo por comparación (Modeller).....	92
Alineamiento secuencia objetivo con la plantilla	92
Generación del modelo	92
Refinamiento	93
I. Modelo utilizado como plantilla	93
J. Alineamientos al dominio de unión a carbohidrato	96

ÍNDICE DE FIGURAS

Figura 1. Representación gráfica de la estructura de la celulosa.....	3
Figura 2. Mecanismo de acción de celulasas.	6
Figura 3. Alineamiento CBM de la familia I.....	9
Figura 4. Comparación de la metodología empleada por Autodock y Vina.	14
Figura 5. Vector pGEM-T Easy.....	18
Figura 6. Vector pET-22b(+)	19
Figura 7. Esquema del procedimiento utilizado en la técnica de mutación por extensión de superposición	29
Figura 8. Pasos seguidos en la metodología computacional planteada.	33
Figura 9. Alineamiento CBM de la endoglucanasa de <i>T. versicolor</i> , con la plantilla utilizada (1azk).....	34
Figura 10. Perfil de DOPE score antes y con los refinamientos.	37
Figura 11. Gráfico de Ramachandran de la estructura final de la secuencia CBM de TvEG.	40
Figura 12. Modelo tridimensional para el CBM.	42
Figura 13. Puentes disulfuro presentes en el modelo tridimensional del CBM.....	43
Figura 14. Distancia de los aminoácidos aromáticos conservados en el CBM de TvEG	43
Figura 15. Comparación modelo generado con modelo de CBM existente.	44
Figura 16. Interacción entre CBM I de TvEG mediante Vina.....	46
Figura 17. Interacción entre CBM I de TvEG mediante Autodock	48
Figura 18. Orientación de los anillos aromáticos de las cadenas laterales hacia la celulosa	49
Figura 19. Energías de interacción calculadas con Autodock y Vina, para los cuatro casos en estudio.	51
Figura 20. CBM de TvEG unido a celulosa.....	52
Figura 21. Energías de afinidad obtenidas con Vina para las mutaciones de la primera estrategia planteada	55
Figura 22. Energías de afinidad obtenidas con Vina para las mutaciones de la segunda estrategia planteada	57
Figura 23. Estructuras tridimensionales de las mutaciones seleccionadas a implementar experimentalmente.	60
Figura 24. Gráfico de Ramachandran de las variantes planteadas.....	62
Figura 25. Pasos seguidos en la metodología experimental planteada.....	64
Figura 26. Diseño de partidores para la amplificación del gen de CBM de la endoglucanasa de <i>T. versicolor</i>	65
Figura 27. Electroforesis en gel de agarosa al 1% del gen de la endoglucanasa de <i>T. versicolor</i>	66
Figura 28. Electroforesis en gel de agarosa al 1% de la digestión de tveg_pGEM y del vector de expresión.....	67
Figura 29. Cuantificación del DNA correspondiente al gen de la endoglucanasa de <i>T. versicolor</i> y al vector de expresión linealizado.....	68
Figura 30. Verificación del clonamiento de Tveg a pET22b(+).	68
Figura 31. Estrategia desarrollada para la construcción de las variantes de TvEG.....	69

Figura 32. Diseño de partidores utilizados en la incorporación de las mutaciones planteadas para las 3 variantes de TvEG propuestas.	71
Figura 33. Electroforesis en gel de los fragmentos mutados.....	72
Figura 34. Verificación de la ligación de las variantes del gen de la endoglucanasa de <i>T. versicolor</i> al vector pGem [®] -T Easy.....	73
Figura 35. Gel de poliacrilamida al 10%. Expresión de la endoglucanasa de <i>T. versicolor</i> a tres temperaturas distintas.....	74
Figura 36. Zimograma de las fracciones solubles inducidas a diferentes temperaturas.	75
Figura 37. Zimograma tras la inducción a 37°C.	75

ÍNDICE DE TABLAS

Tabla 1. Resumen de propiedades físicas de materiales celulósicos.....	10
Tabla 2. Programas para docking molecular.....	13
Tabla 3. Reactivos utilizados en el desarrollo experimental.....	17
Tabla 4. Enzimas de restricción utilizadas.....	19
Tabla 5. Equipos utilizados en laboratorio.....	20
Tabla 6. Herramientas computacionales utilizadas.	21
Tabla 7. Mezcla PCR utilizada para la adición de sitios de restricción al gen de TvEG.	25
Tabla 8. Mezcla utilizada en la digestión del DNA plasmidial.	27
Tabla 9. Mezcla utilizada en el PCR de colonias.....	28
Tabla 10. Mezcla utilizada en la mezcla de PCR para la generación de los fragmentos A y B. ...	30
Tabla 11. Partidores y temperaturas utilizadas en la generación de los fragmentos mutados..	30
Tabla 12. Mezcla utilizada en la mezcla de PCR para la generación de la unión de A y B.	31
Tabla 13. Análisis estructural del primer modelo tridimensional obtenido para la secuencia del módulo de unión a celulosa (CBM).	36
Tabla 14. Análisis estructural del modelo tridimensional refinado obtenido para la secuencia del dominio de unión a celulosa (CBM).	38
Tabla 15. Tabla resumen del análisis de calidad de la estructura.....	41
Tabla 16. Mutaciones planteadas en los aminoácidos conservados en la familia I de CBM.	53
Tabla 17. Mutaciones planteadas en la segunda estrategia de mutación.....	54
Tabla 18. Energías de afinidad entregadas por Vina para las mutantes planteadas en la primera estrategia	56
Tabla 19. Energías de afinidad entregadas por Vina para las mutantes planteadas en la segunda estrategia	58
Tabla 20. Resumen de las tres variaciones de TvEG planteadas.....	59
Tabla 21. Variantes de CBM de TvEG planteadas. Se incluyen los aminoácidos mutados y el RMSD de las estructuras con respecto al CBM de TvEG nativo.....	61
Tabla 22. Volúmenes de los reactivos utilizados para la preparación del gel separador.	90
Tabla 23. Volúmenes de los reactivos utilizados para la preparación del gel concentrador.....	90
Tabla 24. Alinamientos encontrados para la secuencia aminoacídica del dominio de unión a carbohidrato de <i>T. versicolor</i>	96

LISTA DE ABREVIATURAS

Amp	Ampicilina
DNA	Ácido desoxirribonucleico
BSA	Seroalbúmina de bovino
CBM	Módulo de unión a carbohidrato
CD	Dominio catalítico
CIBYB	Centro de Ingeniería Bioquímica y Biotecnología
CMC	Carboximetil celulosa
dNTPs	Deoxiribonucleótidos trifosfato
DTT	Ditiotreitol
EDTA	Etilendiaminotetracetato ácido
HCl	Ácido clorhídrico
IPTG	Isopropil tio- β -D-galactósido
Kpb	Kilo pares de bases
kDa	Kilo Dalton
LB	Medio Luria-Bertani
O.D.	Densidad óptica
o.n.	Toda la noche (over nighth)
PAGE	Electroforesis en gel de poliacrilamida.
pb	Pares de bases
pH	Potencial hidrógeno
PCR	Reacción en cadena de la polimerasa
PM	Peso Molecular
p/v	razón peso/volumen
rpm	Revoluciones por minuto
SDS	Dodecil sulfato de sodio
TAE	Tampón Tris Acetato EDTA
TEMED	N, N, N', N'-tetrametiletilendiamina
Tm	Temperatura de apareamiento
Tris	Tris-(hidroximetil)-aminoetano
UV	Radiación ultravioleta
V	Voltio
X-GAL	5-bromo-4-cloro-3-indolil- β -D-galactopiranosido

1. INTRODUCCIÓN

La problemática de los combustibles, su agotamiento y alto costo, han incentivado a indagar sobre la obtención de energías alternativas como el biogás, biodiesel y bioetanol, a partir de la hidrólisis de residuos agroindustriales. Una alternativa de producción de biocombustibles corresponde al uso de biomasa lignocelulósica. Se estima que alrededor de 180 billones de toneladas de celulosa son producidas por plantas anualmente, por lo que constituye una de las fuentes de carbono renovables más importantes que hay sobre la tierra [Martínez-Anaya et al., 2008]. Sin embargo, una de las dificultades de utilizar estos residuos es que están formados por complejos lignina-celulosa y hemicelulosa los cuales son sumamente difíciles de degradar y sólo un grupo pequeño de organismos es capaz de llevarlo a cabo [Martínez et al., 2009].

En la actualidad, las enzimas se han convertido en herramientas fundamentales en la obtención industrial de productos de diversa índole; entre ellas, en la obtención de formas alternativas de combustibles [Polaina, 2004]. Las celulasas son las enzimas responsables de la hidrólisis de celulosa y se dividen en tres grandes grupos según su actividad enzimática: endoglucanasas o endo-1,4- β -glucanasas, exoglucanasas o celobiohidrolasas y β -glucosidasas. Las endoglucanasas, también llamadas carboximetilcelulasas, inician el ataque al azar en varios sitios internos de la región amorfa de la fibra de celulosa, abriendo nuevos sitios para el subsecuente ataque por las celobiohidrolasas que hidrolizan la celulosa cristalina [Martínez et al., 2009].

Se han llevado a cabo considerables investigaciones para mejorar la hidrólisis de materiales lignocelulósicos. Pretratamientos de la biomasa para remover la lignina y hemicelulosa pueden aumentar significativamente la hidrólisis de celulosa [Martínez et al., 2009]. Por otro lado, la optimización de la actividad de enzimas celulasas y de la acción puede también incrementar la hidrólisis [Sun y Cheng, 2002]. Los avances en la ingeniería de proteínas, sin duda han adquirido un rol fundamental en el desarrollo de mejoras enzimáticas, como la optimización de la actividad o termoestabilidad [Martínez et al., 2009].

Dentro de los métodos para desarrollar nuevas enzimas, una alternativa para mejorar las propiedades de éstas es el diseño racional. Si se cuenta con el gen y la estructura tridimensional de una enzima, se puede alterar la secuencia de aminoácidos y, por lo tanto, las propiedades catalíticas mediante un diseño racional seguido de mutagénesis sitio-dirigida [Peña-Montes y González-Saravia, 2008]. La mutagénesis sitio dirigida consiste en introducir una mutación específica a una secuencia conocida con el fin de conseguir una mejora en alguna propiedad de la proteína y con ello, por ejemplo, desarrollar biocatalizadores que se ajusten a condiciones de trabajo de procesos industriales [Hemsley et al., 1989].

1.1 Descripción del proyecto y justificación

El presente trabajo de título se enmarca dentro del proyecto *“Optimization of the lignocellulose treatment process for bioethanol production”* del programa de investigación *Domeyko-Energía* de las Académicas M. Elena Lienqueo y Oriana Salazar del Departamento de

Ingeniería Química y Biotecnología de la Universidad de Chile. En este proyecto se estudia la optimización del tratamiento de lignocelulosas con miras a la obtención de bioetanol.

La obtención de bioetanol a partir de residuos lignocelulósicos se inicia con un pretratamiento de la biomasa seguido de la hidrólisis de celulosa. De este modo se producen monosacáridos que posteriormente son fermentados en bioetanol. La etapa de hidrólisis consiste en la despolimerización de celulosa a glucosa y se lleva a cabo por medio de enzimas celulasas. Estas enzimas presentan un dominio de anclaje a carbohidrato (CBM) el cual promueve un acercamiento del dominio catalítico al sustrato, facilitando de este modo la hidrólisis de celulosa.

El trabajo de memoria se titula "Identificación de potenciales residuos determinantes de mayor afinidad a celulosa en una endoglucanasa mediante estudios de docking molecular". Tiene como objetivo proponer mutaciones en el dominio de unión a carbohidrato (CBM) de la endo β 1,4 glucanasa de *Trametes versicolor*, que aumenten la adsorción de la enzima sobre celulosa, y con ello mejorar la hidrólisis de ésta.

Se incorporó al estudio herramientas computacionales, que permitieron llevar a cabo análisis *in silico* de la estructura del dominio en estudio y su comportamiento de anclaje sobre celulosa. Esto permitió utilizar una estrategia de mutación racional, la cual involucra un menor uso de recursos (económicos y tiempo). Al implementar herramientas computacionales se facilitó el análisis de un amplio número de casos en estudio, lo que permitió evaluar un mayor número de mutaciones.

A partir de modelamiento por comparación se obtuvo un modelo de la estructura tridimensional del dominio de unión a celulosa de la endoglucanasa de *Trametes versicolor*, el cual se utilizó como base para el análisis de docking molecular. Esta metodología estudió la unión de la endoglucanasa de *T. versicolor* sobre celulosa, entregando información para identificar rasgos estructurales determinantes de la afinidad de la región CBM de la enzima a celulosa, con el objetivo de postular posibles cambios en la secuencia de aminoácidos que mejoren la adsorción.

Finalmente, dentro de las mutaciones analizadas computacionalmente, se escogieron 3 mutaciones a implementar de modo experimental. Estas mutaciones se construyeron mediante una estrategia de mutagénesis sitio dirigida basada en la técnica de mutación por extensión de superposición. Los genes mutados se clonaron en el vector pGEM[®]-T Easy.

Se determinaron también las condiciones de expresión soluble de la endoglucanasa de *Trametes versicolor* en *Escherichia coli*. De esta forma se dejan propuesta las condiciones de expresión de la enzima, para posteriormente analizar la endoglucanasa nativa y sus mutaciones, esperándose encontrar una mayor adsorción y actividad para las variaciones planteadas en el presente estudio.

1.2 Estructura de la celulosa

Los materiales lignocelulósicos son el principal componente de la biomasa, representando aproximadamente la mitad de la materia producida mediante fotosíntesis. Se compone principalmente por tres tipos de polímeros: celulosa, hemicelulosa y lignina, los cuales se encuentran fuertemente entrelazados y químicamente unidos mediante fuerzas no covalentes y covalentes [Pérez et al., 2002; Sánchez, 2009]. El porcentaje de cada uno de estos polímeros varía de planta en planta dependiendo de la especie. Suele encontrarse celulosa en mayor cantidad (50% p/p), y en menor cantidad hemicelulosa (25-30% p/p) y lignina (20-25% p/p) [Pérez et al., 2002].

La lignocelulosa o material lignocelulósico (residuos agrícolas, agro-industriales y forestales) puede presentar en su composición hasta 50% p/p de celulosa. Dado el alto porcentaje de celulosa presente en la biomasa, material lignocelulósico ha sido utilizado para la obtención del biopolímero y de sus derivados, tanto para la producción de papeles como compuestos de elevado valor comercial, tales como glucosa, etanol y otros [Mussatto et al., 2006; Pérez et al., 2002].

La celulosa es el principal componente de la pared celular en plantas [Lee, 1997] y consiste en un polímero lineal formado por unidades de D-glucosa, unidas entre sí mediante enlaces 1,4 β -glucosídicos, formando dímeros conocidos como celobiosa (ver Figura 1). Éstos dímeros a su vez forman largas cadenas unidas entre sí mediante puentes de hidrógeno y fuerzas de van der Waals [Béguin y Aubert, 1994; Pérez et al., 2002; Sánchez, 2009].

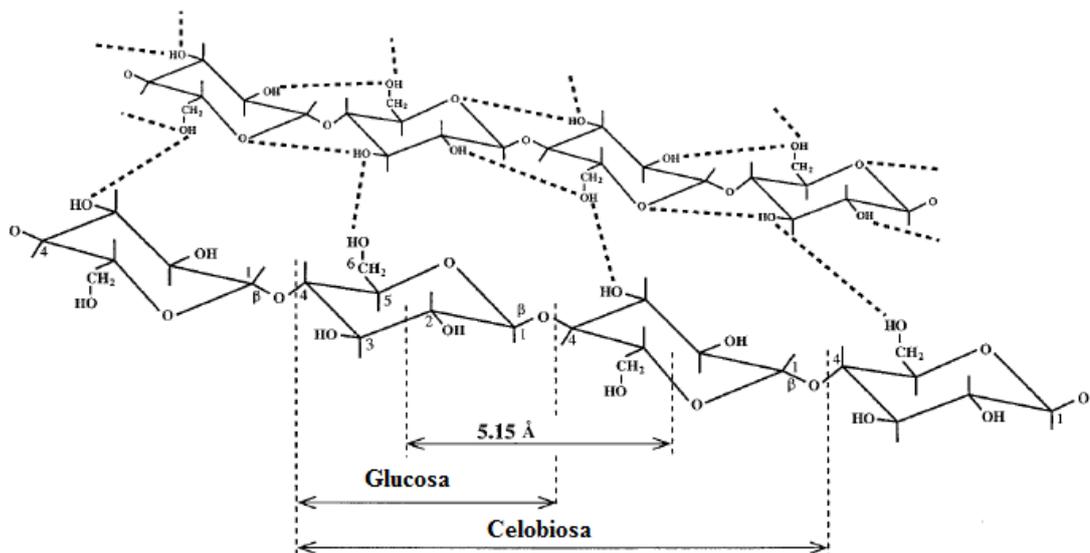


Figura 1. Representación gráfica de la estructura de la celulosa [Hilden y Johansson, 2004; Pandey, 1999].

En la naturaleza la celulosa generalmente se encuentra en un estado altamente organizado, conocido como fibra cristalina, acompañado de pequeñas cantidades de celulosa no organizada, en forma de celulosa amorfa. En esta conformación, la celulosa es más susceptible a ser degradada [Pérez et al., 2002].

Por otro lado, en los materiales lignocelulósicos, la celulosa se encuentra asociada con otras dos fracciones poliméricas, la hemicelulosa y la lignina. La hemicelulosa es un heteropolímero compuesto principalmente por los azúcares xilosa y arabinosa, aunque también presenta en su composición glucosa, manosa, galactosa y algunos ácidos glucurónicos. La lignina es una macromolécula polifenólica compuesta por diferentes unidades de compuestos fenólicos y aromáticos, y que tiene como función proporcionar rigidez a la pared celular del material [Mussatto et al., 2006; Pérez et al., 2002].

1.3 Proceso de obtención de bioetanol

En la actualidad la principal materia prima para la producción de bioetanol es la caña de azúcar, ya que su proceso resulta ser más sencillo y tiene un costo de producción bajo. De los 51 mil millones de litros de bioetanol producidos en el año 2006, la producción estadounidense, a base de maíz, y la brasileña, a base de caña, representaron el 70% del total de producción de bioetanol [BNDES y CGEE, 2008; García et al., 2011].

Dado la competencia con otros usos del suelo que conlleva la producción de bioetanol a partir de productos de origen agrícola, se hizo necesario buscar nuevas alternativas para su producción. En este ámbito, la biomasa lignocelulósica (paja de cereal, residuos forestales, madera) se proyecta como la materia prima del futuro para la producción de bioetanol, dada su alta disponibilidad al aprovechar materias consideradas desechos, menor emisión de agentes contaminantes y mayor eficiencia energética [Hahn-Hägerdal et al., 2006].

El proceso de producción de bioetanol a partir de biomasa lignocelulósica, y particularmente de residuos forestales, consiste en cuatro etapas: pretratamiento, hidrólisis, fermentación y destilación [Balat et al., 2008; Lee, 1997].

El pretratamiento es la primera etapa del proceso y tiene como objetivo alterar o remover compuestos o estructuras que impidan la hidrólisis [Balat et al., 2008]. Existen numerosos procedimientos para llevar a cabo el pretratamiento del material lignocelulósico, sin embargo, es posible agruparlos en pretratamientos físicos, químicos y biológicos [Martínez et al., 2009].

Posterior al pretratamiento se lleva a cabo la hidrólisis, que consiste en obtener los azúcares fermentables a partir de material lignocelulósico. Para esto, las largas cadenas de celulosa y ramificaciones de hemicelulosa son cortados hasta obtener monómeros fermentables. Este procedimiento, puede llevarse a cabo utilizando tanto métodos químicos (ácidos o bases) como biológicos (utilizando enzimas) [Sun y Cheng, 2002].

Finalmente se realiza la fermentación de los monómeros previamente liberados, que consiste en el crecimiento de microorganismos en azúcares de modo que éstos utilicen los monosacáridos como sustrato, produciendo etanol. El etanol producido procede a separarse, por ejemplo, mediante destilación [Balat et al., 2008; Chang et al., 1981].

1.4 Hidrólisis enzimática de celulosa

La hidrólisis de la celulosa puede ser realizada tanto por métodos químicos como son los tratamientos con ácidos o bases, como mediante procedimientos biológicos, utilizando

enzimas [Balat et al., 2008]. El pequeño tamaño de moléculas tales como ácido sulfúrico (hidrólisis ácida) e hidróxido de sodio (hidrólisis alcalina), les permite penetrar profundamente en la estructura de la celulosa; sin embargo estos métodos presentan desventajas, como poseer bajo rendimiento, alta formación de subproductos y requerir elevado consumo de energía (hidrólisis química requiere llevarse a cabo a temperaturas por sobre 120°C) [Mussatto et al., 2006].

La hidrólisis enzimática se ha vuelto una alternativa atractiva frente a la hidrólisis química. El elevado costo de producción de enzimas ha llevado a la incorporación de biotecnología moderna en el proceso. Esto ha permitido reducir los costos de producción y con ello el precio de las enzimas [Balat et al., 2008]. Por otro lado la hidrólisis enzimática puede producir productos más puros, y consumir menos energía. Los costos en la utilización enzimática son menores, ya que las enzimas requieren rangos de operación menos drásticos con respecto a las condiciones estándar de temperatura y presión (20 °C y 1 atm). Al utilizar hidrólisis enzimática además se evitan efectos de corrosión sobre los equipos utilizados en el proceso, debido a que se trabaja a pHs cercanos al neutro [Balat et al., 2008; Sun y Cheng, 2002].

Entre los factores que afectan la hidrólisis enzimática de celulosa, se incluye las características y accesibilidad del sustrato, actividad de las celulasas y las condiciones de operación bajo las cuales se lleva a cabo el procedimiento, sobretodo temperatura y pH [Sun y Cheng, 2002].

1.4.1 Celulasas

La hidrólisis biológica de la celulosa se lleva a cabo mediante la acción enzimática de celulasas, las cuales corresponden a enzimas glicosil hidrolasa, que son producidas por distintos microorganismos, incluyendo hongos, levaduras y bacterias [Lee, 1997].

Las celulasas pueden ser clasificadas en distintas familias, basándose en la similitud de las secuencias primarias de las glicosil hidrolasas. A partir de esto, se han encontrado enzimas con número EC distintos que pertenecen a la misma familia, como enzimas con especificidad de sustrato similares que pertenecen a familias distintas [Henrissat, 1991].

El proceso de degradación de la celulosa parece ser diferente para distintas especies, especialmente entre microorganismos aeróbicos y anaeróbicos. Microorganismos aeróbicos como *Trichoderma reesei* secretan una combinación de endoglucanasas y celobiohidrolasas, las cuales atacan su sustrato individualmente pero de forma sinérgica. Por otro lado, bacterias anaeróbicas como *Clostridium cellulovorans*, *Clostridium cellulolyticum*, *Clostridium josui*, y *Clostridium thermocellum*, producen complejos multienzimáticos extracelulares que tienen alta actividad degradando celulosa cristalina [Ohara et al., 2000].

La producción microbiana de celulasas está sujeta a diferentes mecanismos de regulación. A partir de los estudios realizados tanto en hongos como en bacterias, se ha propuesto un modelo general de regulación controlado principalmente por dos mecanismos. Por un lado, la inducción que se lleva a cabo por el sustrato natural celulosa, y por otro la represión por fuentes de carbono de bajo peso molecular como glucosa y celobiosa. Este modelo sugiere además, la existencia de un nivel basal de enzimas, que inician la hidrólisis del sustrato y

producen pequeños oligosacáridos que entran a la célula, actuando como inductores que encienden la transcripción de los genes [Ponce y Pérez, 2002].

1.4.2 Celulasas Fúngicas

El clásico sistema de celulasas de origen fúngico corresponde a un complejo de enzimático, el cual se encuentra compuesto por endoglucanasa, exoglucanasa y celobiasa. El efecto de este complejo es expresado por la acción sinérgica de estas enzimas sobre la celulosa [Lee, 1997].

Las endoglucanasas, también llamadas endocelulasas, son 1,4- β -D-glucan glucanohidrolasas que actúan de forma azarosa sobre las regiones de celulosa amorfa en el interior del polisacárido. Estas enzimas hidrolizan las cadenas de celulosa de manera aleatoria, obteniéndose una rápida disminución en el largo de éstas y en su grado de polimerización (ver Figura 2) [Hernández-Santoyo et al., 1999]. La configuración β -glucosa se conserva, obteniéndose como producto final una mezcla de glucosa y celobiosa [Martínez-Anaya et al., 2008].

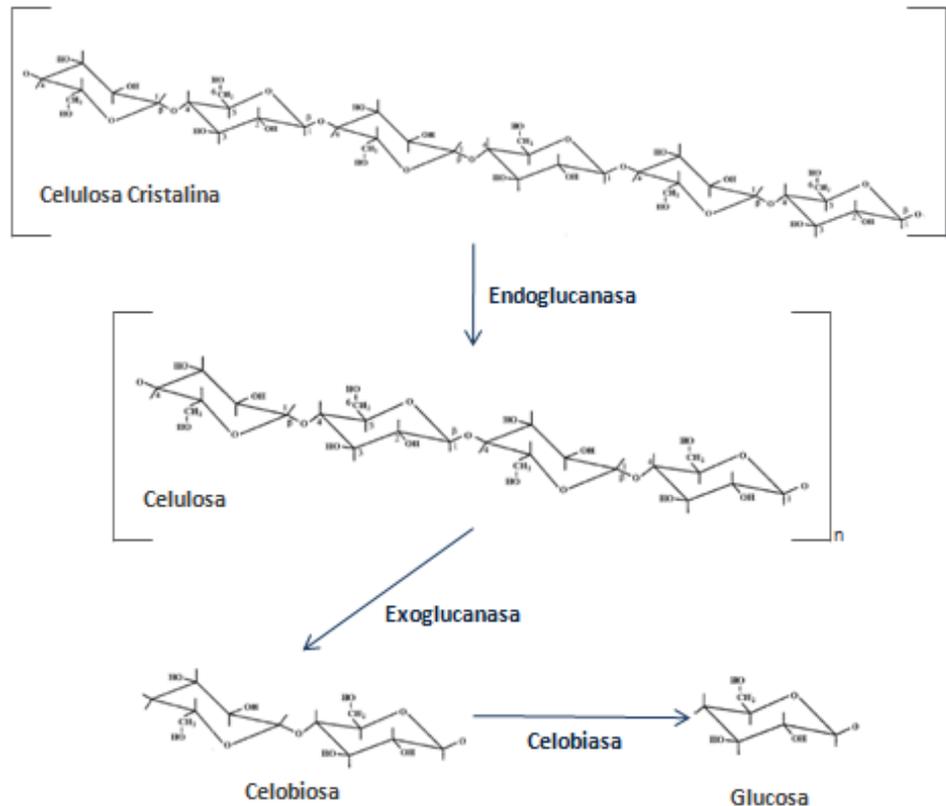


Figura 2. Mecanismo de acción de celulasas.

Las exoglucanasas son también conocidas como exocelulasas y actúan progresivamente en los extremos terminales del polímero liberando ya sea moléculas de glucosa (glucohidrolasas) o celobiosa (celobiohidrolasas) [Hernández-Santoyo et al., 1999; Martínez-Anaya et al., 2008].

Las celobiasas o β -glucosidasas son enzimas β -D-glucósido glucohidrolasas que se encargan de degradar la celobiosa a monómeros, liberando β -glucosa [Lee, 1997; Martínez-Anaya et al., 2008].

La mayoría de las endoglucanasas y exoglucanasas fúngicas poseen peso de 20 a 100 kDa, mientras que las celobiasas se encuentran entre 50 y 300 kDa. En general, celulasas fúngicas son óptimamente activas en pH 4 a 6, mientras que las celulasas bacterianas muestran mejor actividad a pH 6 a 7. Otra característica de las celulasas fúngicas es la presencia de glicosilaciones, las cuales son menos frecuente en celulasas de origen bacteriano [Bhat y Hazlewood, 2001].

1.4.3 Estructura y características de celulasas

Las celulasas presentan como estructura fundamental un dominio catalítico (CD) unido a un dominio de anclaje a carbohidrato (CBM). Estos dos dominios se encuentran conectados mediante una secuencia de unión (L) rica en prolina e hidroxiaminoácidos [Lynd et al., 2002; Rabinovich et al., 2002].

El dominio catalítico (CD) corresponde al dominio más grande dentro de la enzima. Se ha identificado como parte de la estructura donde se lleva a cabo la catálisis enzimática [Rabinovich et al., 2002].

En la mayoría de los casos el dominio catalítico y el dominio de unión en celulasas se encuentran separados por una secuencia de unión. Esta secuencia se compone por cadenas desordenadas, ricas en prolina, treonina y serina, así como tan bien por glicina y alanina. El largo de ésta varía entre 20 y 50 residuos, sin embargo, puede llegar alcanzar hasta 100 residuos. En muchas secuencias de unión es posible identificar entre 4 y 6 residuos repetidos los cuales pueden estar cargados o corresponder a residuos hidrofóbicos [Rabinovich et al., 2002].

La secuencia de unión provee una separación espacial entre el CD y el CBM, de modo de permitir sus funcionamientos de manera autónoma sobre la superficie del sustrato. Un rígido segmento glicosilado es el responsable de la separación espacial, mientras que un segmento flexible permite el funcionamiento independiente de ambos dominios [Rabinovich et al., 2002].

El dominio de unión (CBM) se puede encontrar tanto en el extremo N-terminal como en el extremo C-terminal de las celulasas [Linder y Teeri, 1997]. Un ejemplo de esto se observa en las celobiohidrolasas del hongo *Trichoderma reesei*, en las cuales la CBHI posee el dominio de unión en el extremo carboxilo, mientras que la CBHII posee el dominio de unión en el extremo amino de la enzima [Bray et al., 1996; Gilkes et al., 1991; Kraulis et al., 1989].

La función del módulo de unión a carbohidrato ha sido bastante debatida [Tavagnacco et al., 2011]. Se ha observado que al encontrarse el sitio activo sin estar acoplado a un CBM es activo en sustratos solubles, sin embargo presenta menor unión a sustrato y una leve disminución en la actividad enzimática [Linder et al., 1996; Tavagnacco et al., 2011]. En el caso de sustratos insolubles celulasas sin dominio de unión presentan una clara disminución en su

actividad [Linder y Teeri, 1997]. Se sugiere entonces que el CBM mejora la actividad enzimática sobre celulosa ya sea incrementando la concentración efectiva de la enzima en la superficie del sustrato o posiblemente promoviendo la solubilización de cadenas de azúcares de la superficie del polímero. De este modo se facilita la hidrólisis de celulosa a partir de un acercamiento del dominio catalítico al sustrato [Linder y Teeri, 1997; Lynd et al., 2002].

1.4.4 Clasificación de módulos de unión a carbohidratos

Así como existen familias de proteínas, existen agrupaciones más amplias conocidas como superfamilias, dentro de las cuales son clasificados los dominios de unión. Una de éstas es la superfamilia con dominios para la unión de carbohidratos, la cual presenta una estructura de forma beta en sandwich. Dentro de este grupo se consideran 5 familias [Lynd et al., 2002].

Mientras el número asignado a la familia aumenta, el número de aminoácidos conservados en ésta disminuye [Bray et al., 1996]. La gran mayoría de los dominios de unión presentes en celulasas de origen fúngico (incluyendo la endoglucanasa en estudio) se encuentran clasificadas en la familia I [Martínez-Anaya et al., 2008].

Todos los CBM de la familia I, consisten en 35 a 40 residuos aminoácidos. Su largo, como la variabilidad del extremo N-terminal de su secuencia, dependen de su localización en el extremo amino, carboxilo o al interior de la secuencia, ya que al ubicarse el CBM en el extremo n-terminal, incluye algunos péptidos señal, mientras que si se encuentra en el extremo c-terminal incluye péptidos de la secuencia de unión al dominio catalítico. Esta familia, sólo incluye CBM de origen eucariontes, principalmente hongos [Lynd et al., 2002].

La familia II, incluye CBM de varias enzimas producidas por bacterias aeróbicas y anaeróbicas. A parte de las celulasas y xylanases, CBM de la familia II pueden ser encontrados en otras enzimas relacionadas con la descomposición de hemicelulosa, como arabinosidasas y acetil esterases. Cabe notar que los CBM de la familia I y II, no sólo difieren en cuanto al organismo del cual son originarias, sino también presentan distinta especificidad de unión [Lynd et al., 2002].

Los CBM de la familia III, componen un gran número de bacterias aeróbicas y anaeróbicas con cadenas polipeptídicas de largo entre 131 y 172 residuos. Este dominio, por lo general se ubica en el extremo c-terminal de la proteína. La familia IV incluye CBM de bacterias aeróbicas, anaeróbicas y actinomicetos. Posee un tamaño similar al CBM de la familia III, sin embargo presenta una débil homología a éstos, ya que posee la habilidad de unirse a celulosa amorfa, no cristalina. La familia V, representa un único CBM de una endoglucanasa de la bacteria gram negativa *E. chrysanthemi*. Consiste en 63 residuos y se encuentra en el extremo c-terminal de la enzima [Lynd et al., 2002].

1.5 Módulos de unión a carbohidrato de la familia I

Los dominios de unión a carbohidratos que pertenecen a esta familia son pequeños péptidos rígidamente plegados, que presentan dos cisteínas conservadas involucradas en la formación de puentes disulfuros. Estos puentes disulfuro no sólo permiten un correcto plegamiento del

péptido, sino también estabilizan la estructura, otorgándole rigidez a la superficie de interacción a celulosa [Linder et al., 1996; Tavagnacco et al., 2011].

Poseen triptófanos y tirosinas conservados, los cuales generan una superficie hidrofóbica de unión a la celulosa (ver Figura 3) [Lynd et al., 2002; Rabinovich et al., 2002]. Estos aminoácidos se encuentran espaciados en intervalos regulares, semejantes al espaciamiento de los azúcares que forman la celulosa [Tavagnacco et al., 2011]. Además de los puentes disulfuros, la estructura es estabilizada por la presencia de tres hojas beta, las cuales permiten el correcto posicionamiento de los triptófanos y tirosinas sobre la celulosa [Linder et al., 1996].

		5		13		31	32																														
CBH I	<i>T. reesei</i>	-	T	Q	H	Y	G	Q	C	G	G	I	G	Y	S	G	P	T	V	C	A	S	G	T	T	C	Q	V	L	N	P	Y	Y	S	Q	C	L
CBH II	<i>A. bisporus</i>	-	Q	S	V	W	G	Q	C	G	G	N	G	W	T	G	P	T	T	C	A	S	G	S	T	C	V	K	Q	N	D	F	Y	S	Q	C	L
CBH I	<i>A. aculeatus</i>	N	V	A	L	Y	G	Q	C	G	G	Q	G	W	T	G	P	T	T	C	A	S	G	-	T	C	T	K	Q	N	D	Y	Y	S	Q	C	L
CBH I	<i>F. oxysporum</i>	G	S	V	Q	W	G	Q	C	G	G	Q	N	Y	S	G	P	T	T	C	K	S	P	F	T	C	K	K	I	N	D	F	Y	S	Q	C	Q
CBH II	<i>A. bisporus</i>	P	A	Q	M	W	G	Q	C	G	G	Q	G	W	T	G	P	T	A	C	Q	S	P	S	T	C	H	V	I	N	D	F	Y	S	Q	C	F
CBH I	<i>P. chrysosporium</i>	V	T	V	Q	W	G	Q	C	G	G	I	G	Y	T	G	S	T	T	C	A	S	P	Y	T	C	H	V	L	N	P	Y	Y	S	Q	C	Y
CBH I	<i>T. harzianum</i>	A	T	Q	H	Y	G	Q	C	G	G	T	G	W	T	G	P	T	R	C	A	S	G	Y	T	C	Q	V	L	N	P	F	Y	S	Q	C	L
CBH I	<i>T. viride</i>	P	T	Q	H	Y	G	Q	C	G	G	I	G	Y	S	G	P	T	V	C	A	S	G	S	T	C	Q	V	L	N	P	Y	Y	S	Q	C	L
CBH 58	<i>P. chrysosporium</i>	P	T	V	Q	W	G	Q	C	G	G	I	G	Y	S	G	S	T	T	C	A	S	P	Y	T	C	H	V	L	N	P	Y	Y	S	Q	C	Y
CBH-II	<i>T. reesei</i>	A	C	S	V	W	G	Q	C	G	G	Q	N	W	S	G	P	T	C	C	A	S	G	S	T	C	V	Y	S	N	D	Y	Y	S	Q	C	L
EG-1	<i>T. reesei</i>	C	T	Q	H	W	G	Q	C	G	G	I	G	Y	S	G	C	K	T	C	T	S	G	T	T	C	Q	Y	S	N	D	Y	Y	S	Q	C	L
EG II	<i>T. reesei</i>	-	Q	Q	V	W	G	Q	C	G	G	I	G	W	S	G	P	T	N	C	A	P	G	S	A	C	S	T	L	N	P	Y	Y	A	Q	C	I
EG type F	<i>F. oxysporum</i>	-	-	-	-	G	Q	C	G	G	N	G	W	T	G	A	T	T	C	A	S	G	L	K	C	E	K	I	N	D	W	Y	Y	Q	C	V	
EG 3	<i>H. insolens</i>	-	Q	G	A	W	Q	Q	C	G	G	V	G	F	S	G	S	T	S	C	V	S	G	Y	T	C	V	Y	L	N	D	W	Y	S	Q	C	Q
EG II	<i>T. longibrachiatum</i>	C	T	Q	H	W	G	Q	C	G	G	I	G	Y	T	G	C	K	T	C	T	S	G	T	T	C	Q	Y	G	N	D	Y	Y	S	Q	C	L
EG 4	<i>T. reesei</i>	P	T	Q	L	Y	G	Q	C	G	G	S	G	Y	S	G	P	T	R	C	A	P	P	A	T	C	S	T	L	N	P	Y	Y	A	Q	C	L

Figura 3. Alineamiento CBM de la familia I. CBH representa Exoglucanasa y EG representa endoglucanasas. Adaptación de [Nimlos et al., 2007].

El número de aminoácidos conservados y el espaciamiento entre ellos son factores críticos para un correcto anclaje sobre el sustrato. Al sustituir los aminoácidos conservados por otros como alanina o glicina se ha observado una disminución drástica en la adsorción a celulosa [Mattinen et al., 1997; Takashima et al., 2007].

Por otro lado estudios han encontrado configuraciones de aminoácidos conservados que presentan mejor adsorción que otras combinaciones de aminoácidos aromáticos. La configuración con mayor afinidad se presenta en el caso en que se ubica tirosina en la posición central (31) y triptófanos en ambos extremos (5 y 32). Se observa además que un aumento en la afinidad a celulosa se correlaciona con una alta actividad [Takashima et al., 2007].

1.6 Mecanismo de adsorción a celulosa

La formación del complejo enzima celulosa es necesaria para dar inicio a la hidrólisis del polímero. El dominio de unión CBM se adsorbe sobre los sitios accesibles contenidos en el sustrato, manteniéndose unido a éste, mediante fuerzas de interacción no covalentes [Lynd et al., 2002]. Dado esto, las propiedades estructurales de los distintos materiales celulósicos poseen una gran influencia en la adsorción y con ello en la reacción de hidrólisis de celulosa [Lee et al., 1982].

Dentro de las principales propiedades físicas de los materiales celulósicos utilizados en los modelos de adsorción se encuentra el índice de cristalinidad, el grado de polimerización y el grado de accesibilidad del material [Lee et al., 1982; Zhang y Lynd, 2004].

El índice de cristalinidad (Crl) representa la reactividad del sustrato, el cual es determinado a partir de la medición de patrones de difracción de rayos X [Zhang y Lynd, 2004]. Se ha reportado que la hidrólisis de celulosa mediada por celulasas fúngicas es 3 a 30 veces más rápida en celulosa amorfa que en celulosa cristalina [Lee et al., 1982; Lynd et al., 2002].

El grado de polimerización (DP) de los materiales celulósicos representa la abundancia relativa de enlaces β -glucosídicos del polímero. La cuantificación de éste se lleva a cabo mediante la disolución de celulosa y su posterior medición utilizando membranas o microscopía electrónica [Zhang y Lynd, 2004]. Se ha reportado que la solubilidad de celulosa disminuye drásticamente con el aumento del grado de polimerización de ésta, debido a los puentes de hidrógeno presentes intermolecularmente [Klemm et al., 1998; Miller, 1963; Pereira et al., 1988].

La accesibilidad (SSA) del material representa el área efectiva del polímero a la cual la enzima puede acceder para iniciar el ataque enzimático [Zhang y Lynd, 2004]. Algunos valores reportados para materiales celulósicos se presentan en la Tabla 1 a continuación.

Tabla 1. Resumen de propiedades físicas de materiales celulósicos.

Sustrato	Crl	SSA (m ² /g)	DP
Avicel	0,5 – 0,6	20	300
Algodón	0,81 – 0,95	nr	1000 – 3000
Papel filtrado	0 – 0,45	nr	750
Pulpa de madera	0,5 – 0,7	55 - 61	500 – 1500

Crl denota índice de cristalinidad, SSA accesibilidad, DP grado de polimerización. nr: no reportado.
Fuente: [Zhang y Lynd, 2004]

1.6.1 Modelos de adsorción

La descripción cuantitativa de la adsorción de celulasa a celulosa, generalmente incluye expresiones de concentraciones de complejos enzima-celulosa como función de variables relevantes para la adsorción en celulosa, que describen el estado del sistema. La mayoría de los modelos de adsorción incluyen parámetros como la cantidad total de celulosa y sustrato presente, y parámetros específicos de la enzima y el sustrato que influyen en la adsorción, como lo son la afinidad y capacidad, junto con variables químicas y físicas de la operación, como la temperatura y fuerzas iónicas participantes [Hong et al., 2007; Lynd et al., 2002].

La estimación de la concentración de complejo enzima-celulosa es usualmente determinada a partir de la diferencia entre la enzima total y la enzima no adsorbida, tal como se representa en la ecuación 1 [Lynd et al., 2002]:

$$[P_{AD}] = [P_T] - [P_L] \quad (Ec. 1)$$

Donde $[P_{AD}]$ representa la concentración de la enzima adsorbida a celulosa, $[P_T]$ corresponde a la concentración total de la enzima y $[P_L]$ es la concentración de enzima no adsorbida en la celulosa [Lynd et al., 2002].

La adsorción de celulasas es rápida comparada con el tiempo requerido para la hidrólisis de celulosa. En varios estudios se ha reportado que la adsorción enzimática alcanza estado estacionario a los 30 - 60 minutos [Lynd et al., 2002]. La descripción más común para la adsorción de celulasas es mediante isoterma de Langmuir, la cual considera que la adsorción puede ser descrita mediante una constante de equilibrio y una capacidad específica de adsorción [Zhang y Lynd, 2004]. La isoterma de Langmuir puede representarse:

$$[P_{AD}] = \frac{V_{ps}[S_0]}{K_{ps} + [S_0]} \quad (Ec. 2)$$

Donde $[S_0]$ representa la concentración de celulosa, K_{ps} (mg-celulosa/mL) la constante de saturación media y V_{ps} (mg-proteína/mL) la constante de máxima adsorción. Los parámetros de adsorción de la ecuación 2 se determinan a partir del análisis de una regresión no lineal de los datos experimentales [Lee et al., 1982; Zhang y Lynd, 2004]. La constante de saturación media (K_{ps}) es definida como la concentración de celulosa, a la mitad de la máxima adsorción [Lee et al., 1982].

1.7 Métodos Computacionales utilizados en el diseño racional

Los métodos computacionales utilizados en el diseño racional pueden ayudar a entender resultados experimentales y desarrollar nuevos estudios de ingeniería. Computacionalmente puede simularse la evolución dirigida y de este modo evaluar el efecto de reemplazar, agregar o eliminar uno o más aminoácidos de una proteína. La integración de datos computacionales focaliza el análisis, lo que a su vez permite llevar a cabo un estudio más dirigido y comprensivo [Lutz y Patrick, 2004].

A esta forma de recreación de estructuras complejas por medio de computadores se le ha dado el nombre de modelos *in silico* haciendo alusión a la estructura interna del computador que está hecho en un 90% de materiales de sílice. Cuando se escucha esta palabra se entiende que todo el proceso de simulación experimental ha ocurrido mediante utilización de modelos computacionales [Scior et al., 2007].

1.7.1 Modelamiento de estructuras tridimensionales de proteínas por comparación

El objetivo de los modelos de estructura de proteínas mediante comparación es construir un modelo tridimensional de una proteína a partir de una estructura conocida. La proteína a modelar recibe el nombre de *target*, mientras que las proteínas que se utilizan como modelo, son las *plantillas* [Sánchez y Sali, 2000]. Es posible llevar a cabo un modelo por comparación, ya que pequeños cambios en la secuencia de la proteína usualmente no provoca grandes cambios en la estructura tridimensional [Martí-Renom et al., 2003].

Tienen que darse dos condiciones para construir un buen modelo: primero, debe existir similitud entre el *target* y la plantilla (identidad de secuencias > 40%), y segundo, el correcto alineamiento entre el *target* y el la plantilla debe ser generado [Sánchez y Sali, 2000].

Los modelos por comparación, siguen 4 pasos: identificación de estructuras similares (plantillas), selección de platillas, alineamiento secuencia interés con la o las plantillas seleccionadas y construcción del modelo para la secuencia de interés [Sánchez y Sali, 2000]. Finalmente, se procede a evaluar el modelo (ángulos de torsión permitidos, colisiones estéricas, estereoquímica del modelo, etc.). En el caso de que el modelo generado no sea adecuado, se repite el procedimiento a partir del primer (búsqueda plantilla) o segundo (alineamiento *target*-plantilla) paso de modelamiento [Martí-Renom et al., 2003].

1.7.2 Docking molecular

El docking molecular es un método utilizado para predecir la orientación preferente de una molécula con respecto a otra cuando éstas forman un complejo estable [Jorgensen, 1991]. El objetivo del docking consiste en simular computacionalmente el proceso de reconocimiento molecular, de modo de identificar las estructuras involucradas en la unión receptor-ligando. El docking molecular permite encontrar la unión más probable entre el ligando y el receptor, es decir, la interacción que menos energía requiere. Permite además identificar el sitio idóneo de unión molecular, responsable del reconocimiento específico biológico. Dado esto, se han incorporado análisis de docking molecular en la industria farmacológica y en estudios de interacción enzima-sustrato [Kitchen et al., 2004; Scior et al., 2007].

Actualmente se han implementado diversos programas para llevar a cabo docking molecular (Tabla 2) [Smith y Sternberg, 2002]. Cada uno de ellos utiliza diversas representaciones tanto para el ligando como para el receptor, como también varía el algoritmo en que basan su análisis [Kitchen et al., 2004; Smith y Sternberg, 2002].

Tabla 2. Programas para docking molecular.

Programa	Laboratorio	Página web	Referencia
3D-Dock	Imperial Cancer Research Fund/Imperial college	www.bmm.icnet.uk/docking/	[Smith y Sternberg, 2003]
HEX	Aberdeen University	www.biochem.abdn.ac.uk/hex/	[Janin et al., 2003]
DOT	University of California San Diego	www.sdsc.edu/CCMS/DOT	[Ten Eyck et al., 1995]
BIGGER	Universidade Nova de Lisboa	www.dq.fct.unl.pt/bioin/chamera	[Palma et al., 2000]
MERL	University of Boston	Engpub1.bu.edu/bioinfo/MERL/software/decoys.html	[Camacho et al., 2000]
DOCK	University of California San Francisco	www.cmphram.ucsf.edu/kuntz/dock.html	[Meng et al., 1992]
AutoDock Vina	Scripps Institute	www.scripps.edu/pub/olson-web/download.html	[Trott y Olson, 2010]

Autodock es el software gratuito diseñado para realizar análisis de docking molecular, más citado [Chang et al., 2010; Scior et al., 2007]. Los archivos de entrada empleados para la simulación de Autodock y Vina poseen el mismo formato, lo cual facilita la utilización de ambos programas.

Tanto en el caso de Autodock como de Vina se utiliza una caja rectangular para definir el espacio de búsqueda. Esta caja puede estar centrada según la ubicación del receptor o ligando, mientras que el tamaño y la posición exacta de la caja pueden ser definidos manualmente por el usuario [Seeliger y de Groot, 2010].

Ambos programas requieren representar el receptor y ligando en un formato llamado *pdbqt*, el cual consiste en una versión modificada del formato *pdb* protein data bank [Berman et al., 2000]. Este nuevo formato incluye información sobre la carga atómica, tipo de átomos y, en el caso del ligando, información topológica (enlaces que pueden rotar). Una vez definido el lugar de unión y preparadas las moléculas, se inicia el proceso de docking [Seeliger y de Groot, 2010].

Autodock utiliza mapas de interacción, los cuales son calculados por el programa *autogrid*. La energía de interacción entre los átomos del ligando y del receptor es calculada en el espacio de unión, que se considera discreto a través de la asignación de una grilla [Seeliger y de Groot,

2010]. Si bien Vina también utiliza mapas de interacción, los calcula directamente al momento de llevar a cabo el docking, lo cual resulta más rápido y automatizado [Trott y Olson, 2010].

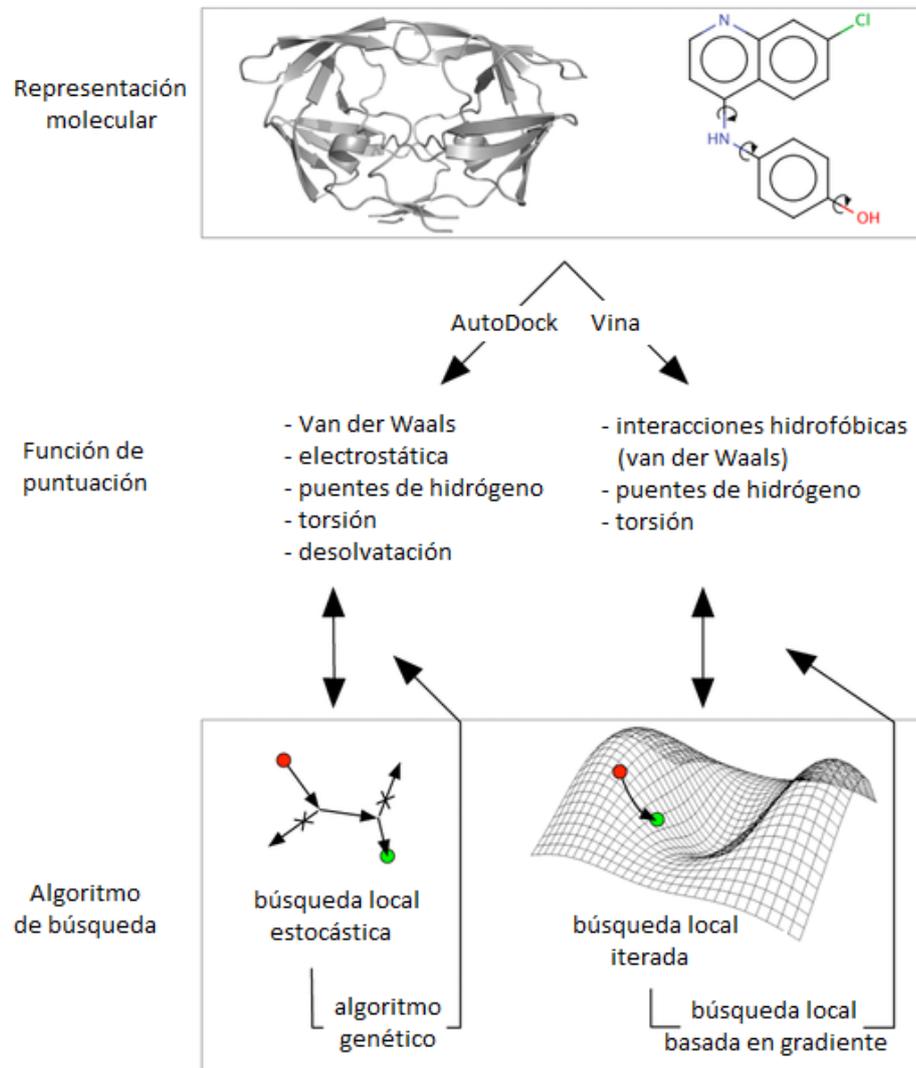


Figura 4. Comparación de la metodología empleada por Autodock y Vina. Ambos programas utilizan el mismo tipo de archivos de entrada para describir al receptor y al ligando. Las funciones de puntuación tienen parámetros similares, pero ponderados de manera distinta. Una diferencia clave entre los programas es el algoritmo de búsqueda local. Autodock utiliza búsqueda estocástica que genera conformaciones al azar. Vina en cambio calcula un gradiente mientras busca un óptimo local. Adaptado de [Chang et al., 2010].

Tanto Autodock como Vina operan de manera similar. Ambos emplean una función de puntuación ponderada empíricamente y un algoritmo de optimización global. Sin embargo las principales diferencias entre ambos programas radican en la función de búsqueda local y la parametrización de la función de puntuación (ver Figura 4) [Chang et al., 2010].

Por otro lado Vina funciona a una mayor rapidez que Autodock (dos órdenes de magnitud). Además se ha demostrado que se obtienen mejores resultados en cuanto al acoplamiento de proteínas y ligandos [Trott y Olson, 2010].

1.8 Técnicas de Ingeniería de Proteínas

Las proteínas pueden ser modificadas de su forma nativa y de este modo mejorar alguna característica de interés, como su actividad, termoestabilidad, o capacidad de unión al sustrato [Bornscheuer y Pohl, 2001]. Las mutaciones, se pueden introducir en la secuencia de nucleótidos en ubicaciones específicas mediante evolución dirigida, como a través de diseño racional.

La evolución dirigida se basa en la generación de una genoteca de mutantes (mutación aleatoria), los cuales posteriormente son expresados en un organismo determinado. Las proteínas producidas generalmente son analizadas en placas con múltiples pocillos y clasificadas en mutaciones deseadas y mutaciones negativas [Bornscheuer y Pohl, 2001; Peña-Montes y González-Saravia, 2008].

Si se cuenta con el gen y la estructura tridimensional de una enzima, se pueden diseñar mutaciones en base a la estructura de ésta, lo cual es conocido como diseño racional de proteínas. Las mutaciones diseñadas son incorporadas utilizando técnicas de mutagénesis sitio dirigida. El diseño racional, seguido de mutación sitio dirigida resulta ser una estrategia muy efectiva para la obtención de proteínas mejoradas [Bornscheuer y Pohl, 2001].

En la mutación sitio dirigida, la secuencia nucleotídica de fragmentos de ADN clonados puede ser modificada utilizando oligonucleótidos sintéticos y la técnica PCR [Carter, 1986]. Si bien son muchos los procedimientos reportados para introducir mutaciones mediante PCR, las dos técnicas más reconocidas, por su durabilidad y robustez, son mutagénesis por extensión de superposición y mutagénesis mediante la utilización de megapartidores [Sambrook y Russell, 2001].

En la mutagénesis por extensión de superposición, dos fragmentos que se superponen son amplificados independientemente mediante PCR. La mutación de interés se diseña en ambos fragmentos y dentro de la zona de superposición. Estos fragmentos son mezclados y en un PCR son extendidos usando dos partidores que se unen en los extremos de los fragmentos iniciales [Sambrook y Russell, 2001].

El diseño racional de enzimas, requiere un amplio conocimiento de la estructura y de las relaciones estructura y función, motivo por el cual, se han desarrollado técnicas computacionales que ayuden a estudiar las proteínas a mutar.

2. OBJETIVOS

2.1 Objetivos

2.1.1 Objetivo General

Proponer mutaciones en la Endo β 1,4 glucanasa de *Trametes versicolor*, que mejoren la adsorción sobre celulosa.

2.1.2 Objetivos específicos

1. Obtener un modelo tridimensional del dominio de unión a celulosa de la endoglucanasa de *Trametes versicolor*.
2. Estudiar la unión enzima – sustrato, utilizando técnicas de docking molecular.
3. Identificar rasgos estructurales determinantes de la afinidad de la región CBM de la enzima a celulosa, para postular posibles mutaciones que mejoren la adsorción.
4. Construir las mutaciones planteadas mediante mutagénesis sitio dirigida.

3. MATERIALES Y MÉTODOS

3.1 Materiales

3.1.1 Microorganismos usados

Para la clonación del gen de interés se utilizó *E. coli* DH5 α genotipo F- Φ 80dlac Δ M15 Δ (lacaya-argF) U169 recA1 endA1 hsdR178rk-,mk+) phoA supE44 λ - thi-1 gyrA96 relA1, mientras que para la expresión de la proteína recombinante se utilizó *E. coli* BL21(DE3) genotipo F- ompT hsdSB (rB-mB+) gal dcm (DE3).

3.1.2 Reactivos

Los reactivos utilizados y sus respectivos proveedores se detallan en la Tabla 3 a continuación.

Tabla 3. Reactivos utilizados en el desarrollo experimental.

Reactivos	Proveedor
Azul de bromofenol.	Amersham-
Rojo congo.	Pharmacia Biotech
Vent DNA Polimerasa, enzima de restricción NdeI.	Arquimed
Ampicilina.	BioLab
Extracto de levadura, agar, medio Luria-Bertani.	Calbiochem
Agarosa.	Difco
IPTG, Estándar de tamaño molecular de proteínas.	Fermelo
Partidores.	Fermentas
T4 DNA ligasa, estándar de peso molecular 1 kb, dNTPs, cepa <i>E. coli</i> DH5 α , BenchMark™ Protein ladder, EDTA, TEMED, High DNA Mass Ladder.	IDT
Cloruro de Sodio.	Invitrogen
CMC.	J.T. Baker
Sales, ácidos, bases, bromuro de etidio y solventes de grado analítico o de biología molecular.	Megazyme
pET-22b(+), <i>E. coli</i> BL21(DE3).	Merck
Kit Taq DNA polimerasa, kit pGEM®-T Easy, enzima de restricción XhoI, estándar de peso molecular 100 pb.	Novagen
Kit QIAprep Spin Miniprep (sistema de purificación de DNA plasmidial), kit QIAEX II (extracción DNA desde geles de agarosa), resina Ni-NTA Agarosa.	Promega
X-Gal, IPTG, BSA.	Qiagen
Acrilamida: Bisacrilamida 29:1, DTT, Tris, SDS, glicerol, Tritón X-100.	Sigma
	Winkler

3.1.3 Vectores usados

El vector de clonamiento pGEM-T Easy se utilizó para la amplificación de los productos de PCR desde DNA genómico. En su forma lineal posee un nucleótido de timina en sus extremos 3', que actúan como extremos cohesivos con los fragmentos de DNA amplificados por PCR, los cuales son generados con un nucleótido de adenina en los extremos 3' (Figura 5).

Para seleccionar los clones positivos, el sitio de clonamiento múltiple del vector interrumpe la secuencia codificante del péptido α de la enzima β -galactosidasa, la cual puede hidrolizar el sustrato X-gal, generando un precipitado de color azul. De esta forma, incluyendo este sustrato en el medio de cultivo, las colonias que adquieren coloración azul se identifican como negativas, y las blancas como positivas.

En la Figura 5 se muestra un esquema del vector de clonamiento pGEM[®]-T Easy, y sus principales puntos de referencia.

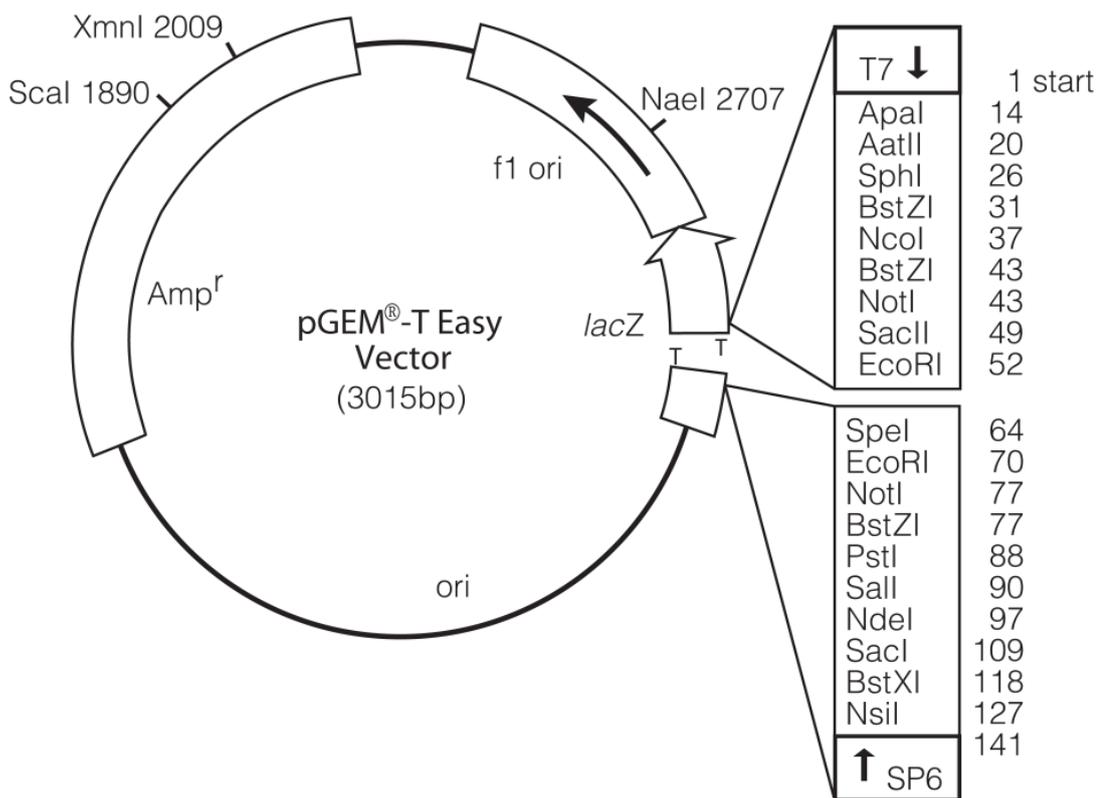


Figura 5. Vector pGEM-T Easy. F1 ori: origen de replicación de DNA de hebra simple derivado del bacteriófago F1. Amp^r: gen de resistencia a ampicilina; T7: promotor para T7 polimerasa; lacZ, región codificante para la β -galactosidasa interrumpida por el sitio de clonamiento. Promega.

Para la expresión de la Endoglucanasa recombinante en *E. coli* se utilizó el vector pET-22b(+), el que posee un sitio de clonamiento múltiple bajo el control del promotor T7 promoter.

Dentro del sitio de clonamiento múltiple el vector cuenta con la secuencia señal pelB. El fusionar el gen de interés a esta señal, permite una potencial localización periplasmática de la proteína. De manera adicional el vector agrega una cola de 6 histidinas al extremo carboxilo de la proteína recombinante, facilitando su posterior purificación. En la Figura 6 se presenta un esquema del vector pET-22b(+), utilizado en la expresión de la proteína recombinante.

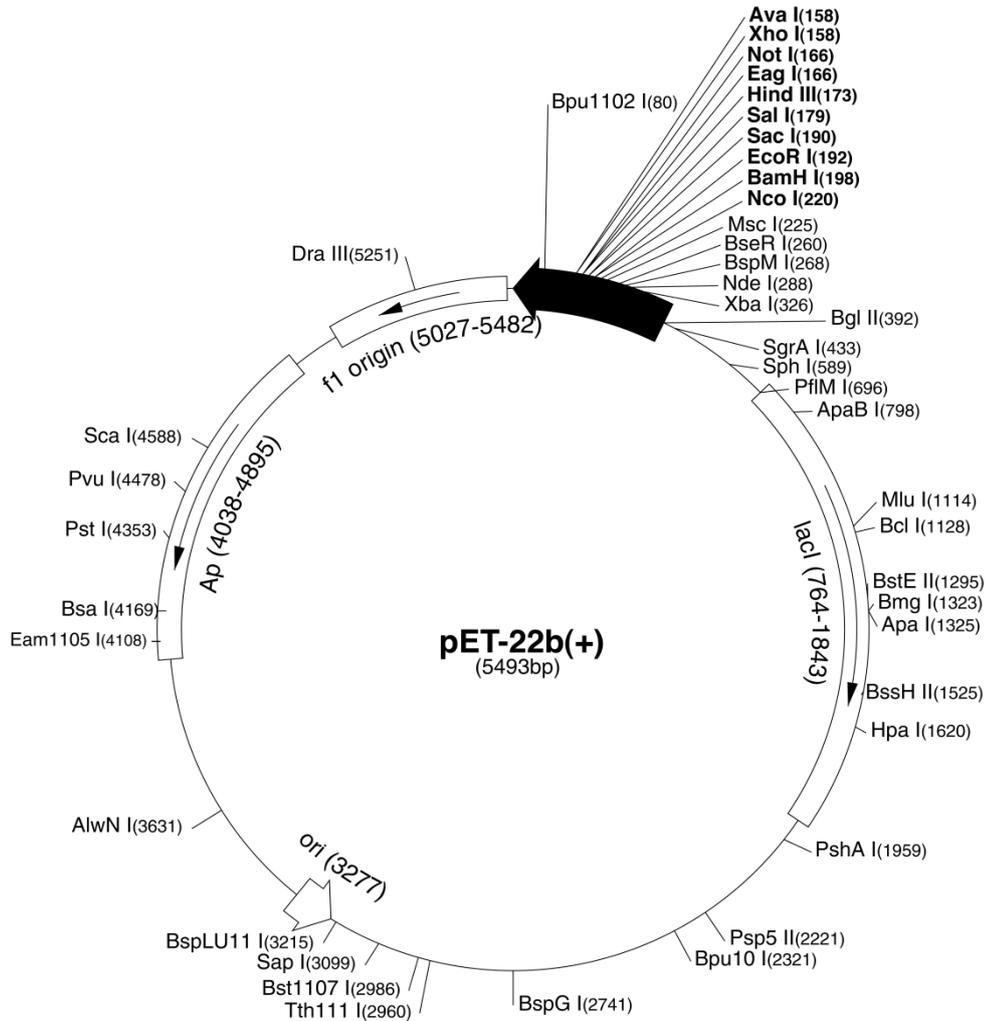


Figura 6. Vector pET-22b(+). F1 ori: origen de replicación de DNA de hebra simple derivado del bacteriófago F1; Ori pBR322(3277): origen de replicación bacteriano derivado de pBR322; LacI: gen que codifica el represor *lac*; Ap: secuencia codificante para *bla* contiene la región codificante para la β -galactosidasa interrumpida por el sitio de clonamiento. Novagen.

3.1.4 Enzimas de restricción

Las enzimas de corte a usadas en este trabajo corresponden a *Xho*I y *Nde*I. Ambas producen extremos de corte cohesivos.

Tabla 4. Enzimas de restricción utilizadas.

Enzima	Sitio de corte
<i>Xho</i> I	C / T C G A G
<i>Nde</i> I	CA / T A T G

/ indica el sitio de corte.

3.1.5 Equipos

En la Tabla 5, se presentan los principales equipos a utilizar, junto con su proveedor y modelo.

Tabla 5. Equipos utilizados en laboratorio.

Equipo	Proveedor
Agitador orbital MaxQ 4000	Barnstead, Lab-Line, USA
Autoclave	Orthmann, Chile
Camara Fotográfica Coolpix 4500	Nikon, Japon
Cámara horizontal Horizon-58	GIBCO BRL Life Technologies, Inc., USA
Fuente de poder Model 500	
Electroporador Cell-Porator Electroporationyst	Bio-Rad
Cámara vertical MiniProtean II	Laboratories Inc., USA
Fuente de poder Power Pac 1000	
Centrífuga Eppendorf 5804 R Rotor F34-63-8	Eppendorf AG, Alemania
Eppendorf Mastercycler Gradient	
Eppendorf Thermomixer confort	
Estufa	Gallenkamp, UK
Centrífuga IEC Modelo MP4R, Rotor 224	International Equipment Company, USA
Centrífuga preparativa RC-28S, Rotor GS-3 y F-24	DuPont, USA
Espectrofotómetro Genesys 6	Thermo Scientific, USA
Espectrofotómetro UV/visible ULtrospec 3000	Pharmacia Biotech Ltd, UK
Lector de placas de 96 pocillos Asys UVM 340	Biochrom Ltd., UK
Sistema de Purificación de Agua Milli-Q	Millipore, USA
Sonicador Microson ULtrasonic Cell Disruptor	Misonix, USA
Balanza Analítica Ohaus Analytical Plus	Ohaus, USA
Balanza Ohaus Precision Estándar	
Termociclador PTC-100	MJ Research, Inc., USA
Transiluminador UV	Vilber Loumat, Francia

3.1.6 Herramientas computacionales

En la Tabla 6 se presentan las herramientas computacionales a utilizar, su fuente o software utilizado, y el sitio web donde se encuentran disponibles.

Tabla 6. Herramientas computacionales utilizadas.

Herramienta: Fuente/software	Sitio web	Referencia
Modelo por comparación: Modeller 9v8	http://www.salilab.org/modeller/	[Sali y Blundell, 1994]
Visualización de moléculas: Pymol 0.99rc6	http://www.pymol.org/	[DeLano, 2002]
Docking molecular: Autodock 4.2	http://autodock.scripps.edu/	[Morris et al., 1998]
Docking molecular: Autodock Vina 1.1.2	http://vina.scripps.edu/	[Trott y Olson, 2010]
Preparación archivos docking molecular: ADT Tools 1.5.4	http://mglttools.scripps.edu/downloads	[Sanner, 1999]
Diseño de partidores: Oligo Calc	http://www.basic.northwestern.edu/biotools/oligocalc.html	[Kibbe, 2007]
Evaluación modelos 3D: VADAR	http://vadar.wishartlab.com/	[Willard et al., 2003]
Evaluación modelos 3D: Molprobit	http://molprobit.biochem.duke.edu/	[Chen et al., 2009]

3.1.7 Síntesis de oligonucleótidos

Los partidores diseñados para la amplificación y modificación de la endoglucanasa, fueron enviados a sintetizar a Integrated DNA Technologies, IDT.

3.1.8 Constructo molecular utilizado

Para el desarrollo de la metodología experimental se utilizó como material el constructo formado por el gen codificante para la endoglucanasa de *T. versicolor* clonado en el vector pGEM®-T Easy (tveg_pGEM). La secuencia del gen se incluye en el Anexo B.

3.2 Métodos

3.2.1 Metodología computacional

i. Obtención estructura tridimensional del módulo de unión a carbohidrato (CBM)

La estructura tridimensional del módulo de unión a celulosa (CBM) se obtuvo mediante la realización de un modelo por comparación. Se utilizó para esto la secuencia aminoacídica del dominio [Salinas et al., 2011]:

N-VWRQCGGTGFSGDTCASGSSCVAVNSYYSQCQ-C

y la metodología que se describe a continuación.

Selección de la plantilla

Se llevó a cabo una búsqueda de secuencias conocidas que presentan un buen alineamiento con la secuencia objetivo, con el fin de utilizarlas como plantillas para el modelo por comparación. Esta búsqueda se realizó en la base de datos del software MODELLER (pdb_95) [Eswar, Webb, et al., 2007] y en la base de datos PROTEIN DATA BANK [Bernstein et al., 1977], mediante búsqueda tipo Fasta.

Alineamiento plantilla y obtención del modelo

La obtención de la estructura del CBM se llevó a cabo mediante la utilización del software Modeller, el cual permite obtener modelos tridimensionales de proteínas por comparación con estructuras tridimensionales conocidas [Sánchez y Sali, 2000].

La plantilla utilizada para el modelamiento comparativo fue 1azk (Anexo I), la cual fue alineada con la secuencia objetivo. Posteriormente se generaron 200 modelos a partir de este alineamiento (ver códigos empleados en Anexo H). Utilizando el índice DOPE score como criterio de selección se escogió el modelo con menor índice.

Validación y refinamiento

El modelo previamente escogido fue validado y refinado. Para evaluar la validez y calidad de la estructura obtenida se utilizó el servidor en línea VADAR [Willard et al., 2003], con el objetivo de detectar problemas asociados a los residuos del péptido modelado. Una vez detectados los errores del modelo, se intentaron reparar con las herramienta *loop refinement* del software Modeller.

Se generaron 100 refinamientos y se seleccionó el mejor modelo, como aquel que presentó menor DOPE score.

Validación final del modelo tridimensional

La validación final del modelo se llevó a cabo mediante:

- Construcción de la gráfica del DOPE score para el modelo antes y después de ser refinado, comparando ambos perfiles.
- Validación mediante el servidor VADAR al modelo refinado. Comparación de los problemas presentes en el modelo, antes y después de llevar a cabo el refinamiento.
- Construcción de la gráfica de Ramachandram con el servidor MolProbity [Davis et al., 2004], para determinar si los residuos se ubican en zonas con o sin impedimento estérico.

ii. Evaluación de la unión del dominio-sustrato del CBM de TvEG con celulosa

Se evalúa la orientación preferente del péptido sobre el sustrato, mediante un análisis de docking molecular, llevado a cabo con Autodock y Vina. La preparación de archivo de entrada es casi la misma para ambos casos, salvo porque en el caso de Autodock se requiere calcular los mapas de búsqueda.

Preparación de los archivos del ligando (celulosa) y receptor (CBM)

Para la preparación de los archivos se utiliza el programa ADT Tools. En éste, se selecciona el archivo pdb correspondiente al receptor, en este caso el CBM de TvEG. Una vez abierto el archivo del receptor, se le agregan los hidrógenos polares y posteriormente las cargas parciales. Finalizado, se guarda el archivo en formato CBD.pdbqt. Para el caso en que se utiliza el receptor flexible, se seleccionan los residuos en los cuales se desea introducir movimiento, y se selecciona los enlaces en los cuales se permitirá rotación. Finalmente se salvan estos residuos (flexibles) en un archivo CBD_flexible.pdbqt y el resto de los residuos (rígidos) se salvan en otro archivo CBD_Rigido.pdbqt.

En el caso del ligando, se carga al programa y luego, se selecciona los enlaces en los que se permitirá rotación. Para el caso de la celulosa se define como rígida, debido a lo cual fue importante chequear que el número de rotámetros posibles se encuentre en cero. Una vez definida como rígida, se guarda en formato celulosa_rigida.pdbqt.

Selección del tamaño y ubicación de la caja

La caja empleada definirá los límites de búsqueda para la unión del ligando al péptido. Ésta debe ser suficientemente grande de modo de permitir la completa rotación del ligando dentro de ella. En base a esto, se define que las cajas empleadas serán de al menos 26Åx26Åx26Å (tamaño 26x26x26 con separación de 1 Å), el cual corresponde al tamaño mínimo para permitir la rotación del ligando.

Para el caso del docking ciego, la caja se definió centrada en el receptor, mientras que en docking forzado se desplazó el centro de la caja de manera de limitar el espacio de búsqueda al deseado.

Generación de los mapas de búsqueda (Autodock)

Si bien a Vina le basta con definir el ligando, receptor, tamaño y ubicación de la caja, en el caso de Autodock se requiere el cálculo de los mapas de búsqueda, lo cual se lleva a cabo mediante la ejecución del comando "Autogrid". Para ello, es necesario la generación de un archivo en formato .gpf (grid parameter file), el cual contiene la información asociada al ligando, receptor, tamaño y ubicación de la caja. Al ejecutar este archivo con Autogrid, se generan los mapas requeridos para la siguiente etapa de docking molecular.

En el caso de trabajar con un residuo flexible, los mapas calculados por Autogrid deben generarse a partir de la parte rígida del receptor (CBD_Rigido.pdbqt).

Docking molecular

Para Vina se genera un archivo de configuración que incluye: el ligando a utilizar (celulosa.pdbqt), el receptor (CBD.pdbqt), dimensiones y ubicación de la caja y parámetros de configuración (ver ejemplo en Anexo 0).

Para Autodock se selecciona el algoritmo genético Lamarkiano y se deja el resto los parámetros de simulación por defecto. Se genera un archivo .dpf (docking parameter file, ver ejemplo en Anexo 0), el cual es ejecutado con Autodock.

iii. Obtención estructura 3D de las mutaciones planteadas

El procedimiento de obtención de las estructuras tridimensionales de las mutaciones, se llevó a cabo mediante el mismo procedimiento empleado para el caso del CBM de TvEG, detallado previamente.

iv. Evaluación de unión del dominio-sustrato CBM mutado

La evaluación de unión de las distintas mutaciones (dominios mutados) con celulosa, se realizó mediante análisis de docking molecular, con el software Vina, utilizando para ello el procedimiento definido anteriormente.

3.2.2 Metodología experimental

i. Diseño de oligonucleótidos

Se diseñan partidores tanto *sense* como *antisense* para generar los genes de las enzimas mutadas, de manera que al utilizar la técnica de PCR, las mutaciones deseadas sean incorporadas en la secuencia.

Para el diseño de los partidores, se toma en cuenta la secuencia nucleotídica del gen del dominio de unión de la endoglucanasa nativa, de modo que los partidores tengan una secuencia complementaria de aproximadamente 10 nucleótidos (desde el extremo 3' del partidor). Además, se diseñan de modo tal de incluir en su secuencia las mutaciones deseadas, junto con sitios de restricción enzimática correspondientes. Estos sitios sirven para introducir el gen en el sitio de múltiple clonamiento del vector.

Otro punto a tener en cuenta para el diseño de los partidores, es que la temperatura de alineamiento (T_m) sea similar y que no exista alineamiento entre ellos. Para esto, se utiliza el programa Oligo Calc [Kibbe, 2007].

ii. Clonamiento del gen de la Endoglucanasa TvEG en pET22b

Adición de los sitios de restricción

Para la clonación de la endoglucanasa en el vector de expresión pET22b(+), se debió incorporar en los extremos del gen los sitios de reconocimiento de las enzimas de restricción. Se utilizó para ello el DNA plasmidial correspondiente a tveg-pGEM.

Se llevó a cabo un PCR a partir del DNA plasmidial, utilizando para ello partidores previamente diseñados de modo de amplificar el gen tveg y adicionar a sus extremos sitios de reconocimiento para enzimas del sitio de clonamiento múltiple del vector. Se agregó el sitio de *NdeI* en el extremo 5', mientras que en el extremo 3' se incorporó el sitio de la enzima *XhoI*.

La reacción de PCR se llevó a cabo con la enzima Deep Vent, utilizando para la mezcla un volumen final de 100 μL . La composición de la mezcla de PCR se detalla en la Tabla 7 a continuación.

Tabla 7. Mezcla PCR utilizada para la adición de sitios de restricción al gen de TvEG.

Reactivo	Concentración final
Buffer 10X	1X
Mix dNTPs	0,2 μM
Primer TVEG_PET_For	0,2 μM
Primer TVEG_PET_Rev	0,2 μM
Vent Pol	2U
DNA plasmidial	4 μL ¹
Volumen Total	100 μL

¹ Se utilizó 4 μL de DNA plasmidial de una extracción de DNA proveniente de 4 mL de cultivo.

El programa de amplificación utilizado fue:

- 1 ciclo a 95°C por 5 min.
- 25 ciclos a: 95°C por 30 seg, 64°C por 30 seg y 72°C por 30 seg.
- 1 ciclo a 72°C por 10 min.

Una vez llevada a cabo la reacción de PCR, se agregó a la mezcla 5U de la enzima Taq Polimerasa y se incubó 20 min a 72°C. El objetivo de esto fue agregar adeninas a los extremos 5' de la enzima amplificada, para su posterior ligación al vector de clonamiento pGEM®-T Easy.

Los productos de PCR de la reacción de amplificación del punto anterior fueron analizados en un gel de agarosa al 1% p/v por electroforesis. La banda de DNA interés fue cortada y colectada en un tubo Eppendorf de 1.5 mL. Posteriormente, se utilizó el procedimiento detallado en el Anexo F para la recuperación del DNA de interés.

Ligación Tveg en el vector pGEM®-T Easy

Para ligar los productos del PCR purificados desde el gel de agarosa al 1% p/v al vector pGEM®-T Easy, se llevó a cabo una mezcla de ligación de volumen final 10 µL, según la composición recomendada por el proveedor. La muestra se dejó ligando toda una noche, manteniéndola a 4°C. El producto de esta ligación se llamó Tveg_pGEM y se utilizó para transformar células de *E. coli* DH5α electrocompetentes.

Preparación de bacterias electrocompetente

Se inoculó 20 mL de medio LB con células stock de *Escherichia coli* DH5α, conservadas a -80 °C, y se incubó con agitación a 37 °C durante la noche. Se midió la densidad óptica (O.D.) del cultivo a 600 nm, y se calculó el volumen de inóculo necesario para 1 litro de cultivo con O.D₆₀₀ final igual a 0,05. A continuación, se creció durante 2 a 4 horas, hasta obtener una O.D₆₀₀ entre 0,6 y 0,8. Se enfrió durante 10 minutos, y luego se colectó las células centrifugando el cultivo a 5000 x g y a 4 °C, por 10 minutos. El pellet se lavó 2 veces con igual volumen de glicerol 10% v/v estéril, y se resuspendió en el glicerol residual del último lavado. Luego se alicuotó en volúmenes de 80 µL, para ser almacenado a -80 °C.

Transformación de DH5α con Tveg_pGEM

Se transformó mediante electroporación 20 µL de células electrocompetentes de *E. coli* DH5α con 1 µL mezcla de ligación. Las condiciones del electroporador fueron: 420 V, 330 µF, baja impedancia, tasa de carga rápida y tiempo máximo de 2,5 ms. Las bacterias transformadas se traspasaron a 1 mL de medio LB estéril y se incubaron a 37°C y a 200 rpm durante 1 h. Posteriormente se centrifugaron a 13.000 rpm durante 1 min y se descartó 900 µL del sobrenadante. El sedimento de células se resuspendió en el sobrenadante restante y se plaqueó sobre placas LB-Agar estériles suplementadas con Amp 100 µg/mL, X-Gal 50 µg/mL e IPTG 0,5 mM. Las placas se incubaron a 37°C durante toda la noche, y posteriormente se mantuvieron a 4°C.

Obtención y digestión del DNA plasmidial

Se seleccionan dos colonias blancas y con ellas se inoculan 4 mL de medio de cultivo LB-Amp estéril, el cual se incuba toda la noche a 37°C con agitación. Para la obtención del DNA plasmidial, se utiliza el sistema Qiagen, siguiendo los procedimientos del fabricante (ver Anexo E).

Una vez obtenida la construcción en vector pGEM®-T Easy se digiere para obtener el DNA codificante de la Endoglucanasa nativa con las enzimas de restricción *NdeI* y *XhoI*. El mismo procedimiento se lleva a cabo con el vector de expresión, de modo de ser digerido con las mismas enzimas que el vector de clonamiento.

Tabla 8. Mezcla utilizada en la digestión del DNA plasmidial.

Compuesto	Cantidad [μL]
DNA plasmidial ² (pGEM o pET22b)	250
<i>NdeI</i>	4
<i>XhoI</i>	4
Buffer D	50
dH ₂ O	192
Total	500

Para digerir el DNA plasmidial, se utilizaron las enzimas de restricción *NdeI* y *XhoI*. Se trabajó con un volumen de digestión de 500 μL (Tabla 8), de modo de contar con una alta concentración tanto de DNA inserto como de vector.

El resultado es visualizado en un gel de agarosa 1% p/v.

Los productos generados tras la digestión del DNA plasmidial (tanto Tveg_pGEM como pET22b), fueron analizados en un gel de agarosa al 1% p/v. Una vez llevado a cabo el gel de agarosa se utilizó el procedimiento detallado en el Anexo F para la recuperación del gen de interés.

Ligación del gen de la endoglucanasa al vector de expresión pET22b

Una vez purificados los fragmentos de DNA correspondientes al gen que codifica la enzima y al vector de expresión linealizado, se llevó a cabo un gel de agarosa al 1% p/v con el fin de estimar la concentración de DNA contenida en las muestras. Para llevar a cabo la cuantificación del DNA se utilizó el reactivo High DNA Mass Ladder (Promega), según las indicaciones del proveedor.

La reacción de ligación se realizó en volumen final de 10 μL. Se agregó el DNA previamente cuantificado de modo que la mezcla de ligación contuviese 120 ng de vector digerido y 80 ng de inserto. Además se agregó 1 μL de T4 DNA ligasa, 2 μL de tampón T4 DNA ligasa 5X y agua Milli-Q hasta completar los 10 μL. La mezcla de reacción se incubó a 16°C por 16 h.

² Se utilizó 250 μL de DNA plasmidial de una extracción de DNA proveniente de 20 mL de cultivo.

El plasmidio producto de la reacción de ligación se llamó Tveg_pET22b y se utilizó para transformar células de *E. coli* DH5 α electrocompetentes.

PCR de colonias

Este método se utilizó para corroborar la presencia del inserto en las colonias seleccionadas como clones positivos. En el caso de los clones que contenían el plasmidio Tveg_pET22b se seleccionaron las que lograron crecer en la placa. Estas colonias se transfirieron mediante una punta estéril a tubos Eppendorf con 30 μ L de agua Milli-Q estéril, incubándose a 100°C durante 10 min, luego se tomaron 5 μ L y se agregaron sobre 15 μ L de mezcla de PCR.

Tabla 9. Mezcla utilizada en el PCR de colonias.

Compuesto	Concentración final
Buffer 5X	1X
Mix dNTPs	0,4 μ M
Primer TVEG_PET_For	1 μ M
Primer TVEG_PET_Rev	1 μ M
Taq DNA polimerasa	1U
MgCl ₂	4 μ M
Volumen Total	10 μL

Para confirmar la presencia de inserto, se utilizaron los partidores correspondientes al fragmento clonado. Luego, los productos de PCR se analizaron mediante electroforesis en gel de agarosa al 1% p/v.

Se seleccionaron 2 colonias en las cuales se observó la presencia del inserto. Se dejó inóculo de éstas para posteriormente obtener su DNA plasmidial y enviar a secuenciar a Macrogen S.A. Las secuencias obtenidas se analizaron para verificar el correcto clonamiento de la endoglucanasa al vector de expresión.

iii. Mutagénesis sitio dirigida

La obtención de las mutantes planteadas se llevó a cabo mediante mutagénesis sitio dirigida, utilizando para esto la técnica de mutación por extensión de superposición (ver Figura 7) [Sambrook y Russell, 2001].

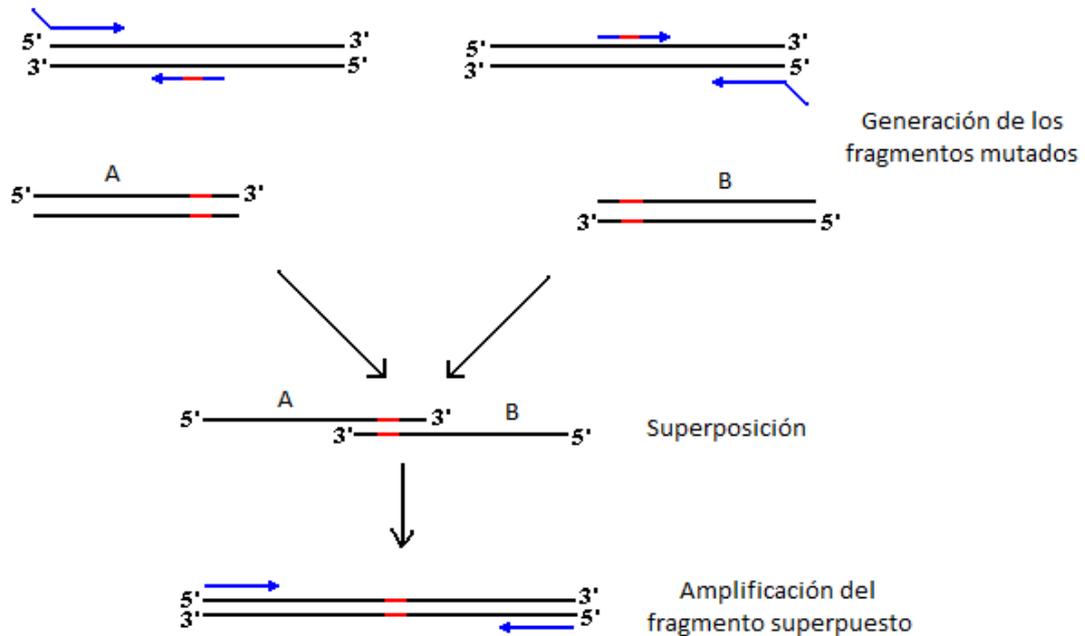


Figura 7. Esquema del procedimiento utilizado en la técnica de mutación por extensión de superposición. Mediante una primera reacción de PCR se generan los fragmentos mutados (denominados A y B). En una segunda reacción de PCR se alinean los fragmentos previamente generados y se amplifican.

PCR para la generación fragmentos mutados

La generación de los fragmentos mutados se llevó a cabo mediante la utilización de la técnica de PCR, usando para ésta partidores que incorporen las mutaciones en los aminoácidos deseados. Los fragmentos generados para la incorporación de la mutación 1 se denominaron A1 y B1, mientras que para las mutaciones 2 y 3 se nombraron A2, B2 y A3, B3 respectivamente.

En el caso de la mutación 1, se empleó como templado el constructo Tveg_pET22b previamente obtenido. En el caso de las mutaciones 2 y 3 se utilizó como templado el constructo generado para la mutación 1.

La reacción de PCR se llevó a cabo con la enzima Deep Vent, utilizando para la mezcla un volumen final de 200 μ L. La composición de la mezcla de PCR se detalla en la Tabla 10.

Tabla 10. Mezcla utilizada en la mezcla de PCR para la generación de los fragmentos A y B.

Compuesto	Concentración final
Buffer 10X	1X
Mix dNTPs	0,2 μ M
Primer Forward	0,2 μ M
Primer Reverse	0,2 μ M
Vent Pol	4U
DNA plasmidial	10 μ L ³
Volumen Total	200 μL

Los ciclos de amplificación para los distintos fragmentos correspondieron a:

- 1 ciclo a 95°C por 5 min.
- 25 ciclos a: 95°C por 30 seg, T^o alineamiento por 30 seg y 72°C por 30 seg.
- 1 ciclo a 72°C por 10 min.

Los partidores y temperaturas de alineamiento utilizadas para la generación de los distintos fragmentos se resumen en la Tabla 11 a continuación:

Mutación	Fragmento	Primer forward	Primer reverse	Temperatura de alineamiento
1	A1	pET_For	Mut1_Rev	62°C
	B1	Mut1_For	TVEG_PET_Rev	69°C
2	A2	pET_For	Mut2_Rev	64°C
	B2	Mut2_For	TVEG_PET_Rev	64°C
3	A3	pET_For	Mut3_Rev	62°C
	B3	Mut3_For	TVEG_PET_Rev	64°C

Tabla 11. Partidores y temperaturas utilizadas en la generación de los fragmentos mutados.

Los productos de PCR de la reacción de amplificación fueron analizados en un gel de agarosa al 1% p/v por electroforesis, para separar los fragmentos deseados. El tamaño esperado para los fragmentos amplificados corresponden a 556 pb y 1162 pb para A y B respectivamente.

Una vez llevado a cabo el gel de agarosa se utilizó el procedimiento detallado en el Anexo F para la recuperación de los fragmentos A y B.

PCR superposición de los fragmentos generados

Para la generación de las mutaciones fue necesario unir los fragmentos A y B, llevando a cabo un PCR que los superpusiese y posteriormente, adicionando partidores externos (complementarios a los extremos externos de cada fragmento) para la amplificación del resultado de la superposición.

La reacción de PCR se llevó a cabo con la enzima Deep Vent, utilizando para la mezcla un volumen final de 100 μ L. La composición de la mezcla de PCR se detalla en la Tabla 12.

³ Se utilizó 10 μ L de DNA plasmidial de una extracción de DNA proveniente de 4 mL de cultivo.

Tabla 12. Mezcla utilizada en la mezcla de PCR para la generación de la unión de A y B.

Compuesto	Concentración final
Buffer 10X	1X
Mix dNTPs	0,2 μ M
A	10 μ L
B	10 μ L
Vent Pol	2U
Volumen Total	100 μL

Los ciclos correspondieron a:

- 7 ciclos a: 95°C por 3 min y 69°C por 2 min.
- Pausa.
- 20 ciclos a: 95°C por 45 seg, 61°C por 30 seg y 72°C por 90 seg.
- 1 ciclo a 72°C por 5 min.

En la pausa se adicionó 1 μ L de los partidores externos (pET_For y TVEG_PET_Rev) de modo de amplificar el fragmento previamente superpuesto.

Una vez llevada a cabo la reacción de PCR, se agregó a la mezcla 1 μ L de la enzima Taq Polimerasa y se incubó 20 min a 72°C, para agregar adeninas a los extremos 5' de los fragmentos amplificados.

Los productos de PCR de la reacción de amplificación fueron analizados en un gel de agarosa al 1% p/v por electroforesis para separar el fragmento deseado (peso aproximado de 1,6 kb). Desde el gel de agarosa se cortó la banda de DNA interés desde el gel y se purificó según el protocolo descrito en el Anexo F.

Para ligar el fragmento purificado al vector pGEM®-T Easy, se llevó a cabo una mezcla de ligación de volumen final 10 μ L, con la composición recomendada por el proveedor. El producto de esta ligación se utilizó para transformar células de *E. coli* DH5 α electrocompetentes.

Dos colonias transformadas que presentaron coloración blanca (en placas LB-amp X-gal IPTG) se seleccionan e inoculan en 4 mL de medio de cultivo LB-amp estéril, el cual se incuba toda la noche a 37°C con agitación. Para la obtención del DNA plasmidial, se utiliza el sistema S.N.A.P.TM siguiendo los procedimientos del fabricante (ver Anexo E).

iv. Expresión y análisis de la endoglucanasa recombinante

Transformación de la enzima nativa a E. coli BL21

Del DNA plasmidial de la enzima nativa se tomó 1 μ L y se mezcló con 9 μ L de dH₂O estéril. De cada una de estas mezclas, se tomaron 1 μ L y se agregaron 20 μ L de células electrocompetentes de *E. coli* BL21. Las condiciones del electroporador fueron: 420 V, 330 μ F, baja impedancia, tasa de carga rápida y tiempo máximo de 2,5 ms.

Las bacterias transformadas se traspasaron a 1 mL de medio LB estéril y se incubaron a 37°C y a 200 rpm durante 1 h. Posteriormente se centrifugaron a 13.000 rpm durante 1 min y se descartó 900 µL del sobrenadante. El sedimento de células se resuspendió en el sobrenadante restante y se plaqueó sobre placas LB-Agar estériles suplementadas con Amp 100 µg/mL. Las placas se incubaron a 37°C durante toda la noche, y posteriormente se mantuvieron a 4°C.

Inducción de la síntesis de la endoglucanasa nativa

Para inducir la expresión del gen de la endoglucanasa nativa, en un 4 tubos de ensayo con 4 mL de LB más Amp, se colocó una colonia de las células *E. coli* BL21(DE3) previamente transformadas. Se incubó toda la noche a 37°C y 200 rpm.

Se midió absorbancia a 600 nm, utilizando como blanco medio TB estéril y se determinó el volumen requerido para inocular 200 mL de medio TB estéril, suplementado con Amp, a una D.O. final de 0,05. El cultivo se creció a 37°C y 200 rpm hasta alcanzar una D.O. de 0,6.

Una vez alcanzada la D.O. deseada, se agregó IPTG a una concentración final de 0,1 mM y se incubó por 24 h a 18°C y 200 rpm.

Obtención de la endoglucanasa

Tras alcanzar 24 h de inducción, el cultivo es centrifugado a 5.000 g por 5 min, de modo de colectar el pellet. El pellet obtenido se resuspendió en 5 mL de buffer acetato de sodio 50 mM (pH 5,5).

Como la endoglucanasa se clonó sin el péptido señal, se ubica en el citoplasma, por lo que para poder obtenerla es necesario romper las células. Para conseguir la lisis, se adiciona a las muestras 5 mg de lisozima y se incuba en hielo por 30 min. Luego de la incubación, las muestras son sonicadas, teniendo cuidado de mantenerlas permanentemente en hielo, y así evitar denaturación de la proteína debido al calentamiento de la muestra. Cada muestra se sonicó por 10 s, y luego se descansó 10 s adicionales. Este procedimiento se repitió 6 veces a cada muestra.

Una vez sonicadas las muestras fueron centrifugadas a 10.000 g por 20 min. El sobrenadante obtenido tras la centrifugación se traspasó a tubos eppendorf de 2 mL para su posterior análisis mediante electroforesis de proteínas (SDS-PAGE), zimograma y medición de actividad en medio líquido (Anexo G).

4. RESULTADOS Y DISCUSIÓN

El objetivo del presente estudio consiste en proponer mutaciones en la endo β 1,4 glucanasa de *Trametes versicolor* que mejoren la adsorción a celulosa. Para esto se llevo a cabo un diseño racional de las mutaciones, basado en análisis realizados *in silico*.

Ya que no se conocen cristales de celulosa completos y dado la secuencia (linker) que separa el dominio de unión con el sitio activo se puede considerar que estos tienen comportamiento independiente [Lee, 1997], pudiéndose analizar localmente lo ocurrido en el dominio de unión a carbohidrato. El análisis se inició con la obtención de un modelo tridimensional del CBM de TvEG, el cual se utilizó en los estudios posteriores de docking molecular (ver Figura 8).

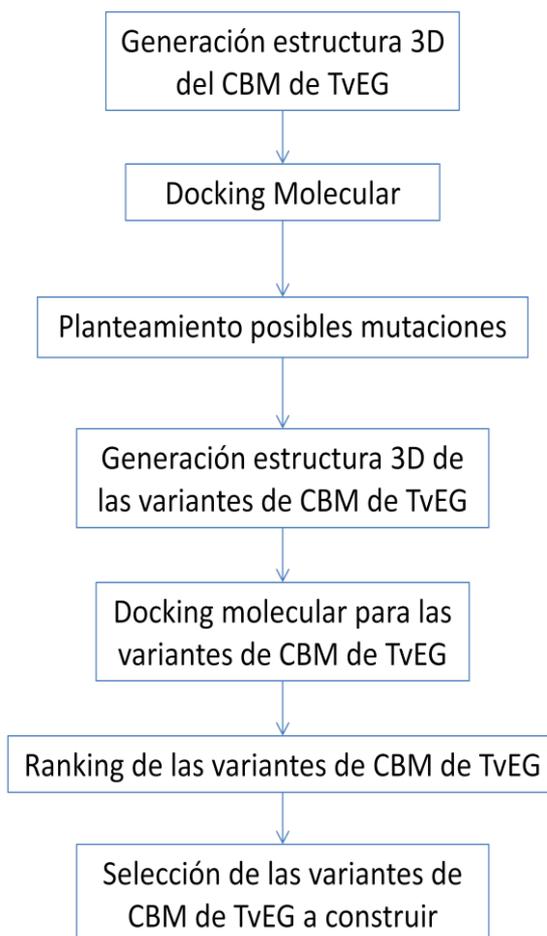


Figura 8. Pasos seguidos en la metodología computacional planteada.

A partir de la metodología computacional planteada se obtuvo el modelo tridimensional del CBM de la endoglucanasa *T. versicolor*, basado en modelamiento por homología. Este modelo fue evaluado y refinado. Posteriormente se utilizó como base para el análisis de docking molecular.

El análisis de docking molecular permitió estudiar la interacción del dominio y la celulosa, pudiéndose plantear posibles variantes de TvEG a realizar. Una vez planteada las variantes se generó un modelo tridimensional para cada una de ellas, el cual se utilizó para el análisis de docking molecular de éstas.

El docking molecular de cada una de las variantes permitió calcular la energía de interacción de éstas hacia la celulosa, a partir de la cual se generó un *ranking* de las posibles variantes a construir. Basados en la energía de interacciones calculadas se seleccionaron 3 variantes de TvEG.

A continuación se presentan cada una de las etapas anteriormente mencionadas con los resultados que se fueron obteniendo en cada una de ellas.

4.1 Generación de un modelo tridimensional del módulo de unión a carbohidrato de TvEG

Con el objetivo de poder estudiar la estructura del CBM de la endoglucanasa de *Trametes versicolor*, se obtuvo el modelo tridimensional de éste. Tanto para el modelo estructural, como para los análisis de docking molecular, se utilizó sólo el dominio de unión a carbohidratos.

4.1.1 Alineamiento CBM de TvEG

El modelamiento por comparación requiere la selección de una plantilla con estructura tridimensional conocida, sobre la cual se genera el modelo de la secuencia objetivo, en este caso, el CBM de la endoglucanasa de *Trametes versicolor*.

En la Figura 9 se muestra el alineamiento seleccionado para el CBM, el cual se utilizó como plantilla para la realización del modelo por homología. Los otros candidatos a plantilla descartados y sus alineamientos se encuentran en el Anexo J.

```

alineamiento                10      20      30
CBM TvEG                    ---VWRQCGGTGFSGDITTCASGSSCVAVNSYYSQCQ
1azk                        TQSHYGQCGGIGYSGPTVCASGTTCQVLNPNYASQCL
Aminoácidos conservados    ***** * ** * ***** * * * **

```

Figura 9. Alineamiento CBM de la endoglucanasa de *T. versicolor*, con la plantilla utilizada (1azk). Porcentaje de identidad 55,1%, porcentaje de similitud 62%, e-value: 0,2E-7. En amarillo se destacan las cisteínas conservadas y en rosado los aminoácidos aromáticos conservados en módulos de la familia I.

La molécula seleccionada como plantilla (1azk) corresponde al módulo de unión a celulosa de una celobiohidrolasa I del hongo *Trichoderma reesei* (Anexo H). Su módulo de unión pertenece a la familia I [Lynd et al., 2002; Mattinen et al., 1997] por lo que el largo de su secuencia (36 aminoácidos) es bastante parecido al largo de la secuencia del CBM en estudio (33 aminoácidos).

Dentro de los aminoácidos conservados entre estas dos secuencias podemos encontrar las cisteínas y aminoácidos aromáticos conservados en los módulos de la familia I. Estudios previos realizados al módulo de unión de la celobiohidrolasa I del hongo *Trichoderma reesei* han identificado que los aminoácidos conservados tienen funcionalidad clave en la unión del módulo con el carbohidrato. Las 4 cisteínas presentes son responsables del correcto plegamiento del péptido, y por consiguiente del correcto espaciado de los aminoácidos

aromáticos, los cuales se han identificado como los residuos que interactúan con los azúcares de la celulosa [Linder y Teeri, 1997; Lynd et al., 2002; Mattinen et al., 1997; Takashima et al., 2007; Tavagnacco et al., 2011].

El alineamiento entre las dos secuencias presenta un alto porcentaje tanto de identidad (55,1%) como de similitud (62%) con la secuencia objetivo, lo que sugiere que 1azk resulta ser una buena planilla para el modelamiento.

4.1.2 Evaluación preliminar del modelo generado por Modeller

A partir de la planilla seleccionada y mediante la utilización del software Modeller se obtuvo el modelo tridimensional del dominio de unión a carbohidratos de la endoglucanasa de *Trametes versicolor*. Es importante la evaluación del modelo, ya que ésta permite identificar los aminoácidos con posibles problemas estéricos, los cuales pueden ser corregidos mediante herramientas del software utilizado. En base a esto, se sometió el modelo generado a una evaluación de su área superficial accesible, volumen fraccional, ángulos phi, psi y omega.

La evaluación del modelo se llevó a cabo utilizando el servidor web VADAR [Willard et al., 2003], el cual señala posibles errores que pueden presentar los distintos residuos del modelo. Los errores que se pueden identificar con el servidor utilizado son:

A : Indica posible problema con ASA⁴ fraccional ($fASA > 1.0$).

V : Indica posible problema con volumen fraccional ($fV < 0.8$ o $fV > 1.2$).

P : Indica posible problema con ángulos Phi y Psi [Morris et al., 1992].

O : Indica posible problema con ángulo omega ($\omega > 170$ u $\omega < -170$)

C : Indica enlace peptídico cis ($-20 < \omega < 20$)

En la Tabla 13 se muestra el análisis de calidad de la primera estructura obtenida con Modeller, junto con la identificación de los residuos con posibles problemas.

⁴ ASA: Área superficial accesible

Tabla 13. Análisis estructural del primer modelo tridimensional obtenido para la secuencia del módulo de unión a celulosa (CBM). Los parámetros resaltados en rojo indican parámetros con problemas. Evaluación generada con VADAR.

RES. NUM.	RES. NAME	SCND STRUC	HBOND HBOND	BTURN BTURN	RES. ASA	FRAC. ASA	RES. VOL.	FRAC. VOL.	PHI PHI	PSI PSI	OMEGA OMEGA	PRBLM PRBLM
1	VAL	BCC C			152.9	0.89	137.3	1.01	360.0	-39.2	173.0	
2	TRP	BCB B			149.2	0.56	209.1	0.91	-134.1	72.2	-172.8	
3	ARG	BCB B	32, 32		129.5	0.53	196.8	1.12	-119.4	129.6	-175.7	
4	GLN	BCB B			36.5	0.19	153.7	1.11	-73.0	114.8	174.2	
5	CYS	BCB B	30		26.8	0.19	89.5	0.86	-90.9	-23.4	-177.5	
6	GLY	BCB B	30		26.0	0.29	57.7	0.92	179.6	144.6	-180.0	
7	GLY	CCC C	10	II	32.8	0.36	48.9	0.78	155.4	-177.9	-175.5	V
8	THR	CCC C	27	II	103.4	0.68	105.2	0.90	-66.6	140.8	176.8	
9	GLY	CCC C		II	84.2	0.93	49.8	0.79	77.8	18.1	-175.4	V
10	PHE	CCC C	7	II	49.5	0.22	159.4	0.81	-119.0	162.3	-179.3	
11	SER	CCC C			106.1	0.81	80.0	0.88	-128.4	-77.9	-173.0	
12	GLY	CCC C	14		50.5	0.56	51.5	0.82	169.6	21.0	-179.9	
13	ASP	CCC C			102.5	0.65	108.2	0.95	-66.0	-10.9	-173.0	
14	THR	CCC C	12		54.0	0.36	117.6	1.01	-43.5	92.8	175.5	
15	THR	CCC C			134.7	0.89	104.1	0.89	-98.2	69.2	-176.9	
16	CYS	CBC C			44.8	0.31	98.1	0.94	-92.3	171.8	-179.5	
17	ALA	CBC C	20		82.7	0.67	78.1	0.90	-118.6	163.7	177.9	
18	SER	CBC C			108.9	0.83	75.1	0.83	-77.0	179.6	-179.0	
19	GLY	CCC C			91.8	1.01	48.1	0.76	92.6	-30.2	177.3	A, V
20	SER	CCC C	17		38.4	0.29	97.0	1.07	-108.3	178.9	-173.3	
21	SER	BCB B	33		68.5	0.52	84.6	0.93	-130.8	-144.6	173.5	
22	CYS	BBB B			62.5	0.43	81.9	0.78	-152.8	89.0	-174.2	V
23	VAL	BBB B	31, 31		71.1	0.41	115.7	0.85	-114.8	156.5	-176.9	
24	ALA	CBC C			76.8	0.62	73.9	0.85	-72.8	146.7	174.1	
25	VAL	CCC C	29		87.2	0.51	120.1	0.88	-144.4	-122.9	179.5	
26	ASN	CCC C	28, 29	III	73.5	0.44	119.1	1.02	-71.1	166.2	-171.4	
27	SER	CCC C	8	III	100.0	0.76	80.8	0.89	-56.5	-25.6	178.8	
28	TYR	BCC C	26	III	120.8	0.50	161.8	0.85	-70.1	-44.4	-167.2	O
29	TYR	BBB B	26, 25	III	67.0	0.28	178.5	0.93	-132.9	128.1	160.6	O
30	SER	BBB B	6, 5		35.8	0.27	81.9	0.90	-98.1	115.8	-167.6	O
31	GLN	BBB B	23, 23		0.1	0.00	141.9	1.02	-133.0	165.5	-177.2	
32	CYS	BBB B	3, 3		1.8	0.01	96.7	0.93	-68.4	90.4	173.6	
33	GLN	BCB B	21		132.7	0.70	138.4	1.00	-89.2	360.0	360.0	

A : Indica posible problema con ASA fraccional (fASA > 1.0).
V : Indica posible problema con volumen fraccional (fV < 0.8 o fV > 1.2).
P : Indica posible problema con ángulos Phi y Psi [Morris et al., 1992].
O : Indica posible problema con ángulo omega (omega > 170 u omega < -170)
C : Indica enlace peptídico cis (-20 < omega < 20)

Los resultados entregados por el análisis mediante el servidor VADAR fueron claves para la determinación de los residuos que presentan problemas espaciales. De este modo, se identificó que los residuos GLY7, GLY9, GLY19 y CYS22 poseen problemas con su volumen fraccional (el residuo GLY19 además presenta problemas con su área superficial accesible), mientras que los residuos TYR28, TYR29 y SER30 presentan problemas de ángulo omega.

El volumen fraccional es calculado como el volumen ocupado por un residuo definido por su radio atómico y sus vecinos más cercanos [Willard et al., 2003]. Normalmente, si la proteína se plegó correctamente, sus residuos debiesen presentar volúmenes fraccionales cercanos a 1. En el caso de los residuos 7, 9, 19 y 22, este valor es menor a 0.8, lo cual indica que estos residuos se ubican en una región comprimida.

Por otro lado, el ángulo omega mide el giro entorno al enlace peptídico. En el caso de los residuos TYR28, TYR29 y SER30, éstos poseen su ángulo omega fuera de rango (valor es menor a 170 o mayor a -170), lo cual indica posibles problemas con la conformación *trans* de los enlaces.

Refinamiento con Modeller

Se llevó a cabo el refinamiento del modelo tridimensional obtenido para el CBM de la endoglucanasa de *T. versicolor* utilizando la herramienta de refinamiento de Modeller.

Modeller lleva a cabo refinamiento utilizando dinámica molecular con el algoritmo probabilístico SA (*simulated annealing*) [Sali y Blundell, 1994]. Una vez llevado a cabo el refinamiento con Modeller es posible generar el perfil DOPE⁵ del modelo antes y después de ser refinado. El perfil DOPE corresponde a una función de energía y permite identificar zonas del modelo con posibles errores [Shen y Sali, 2006].

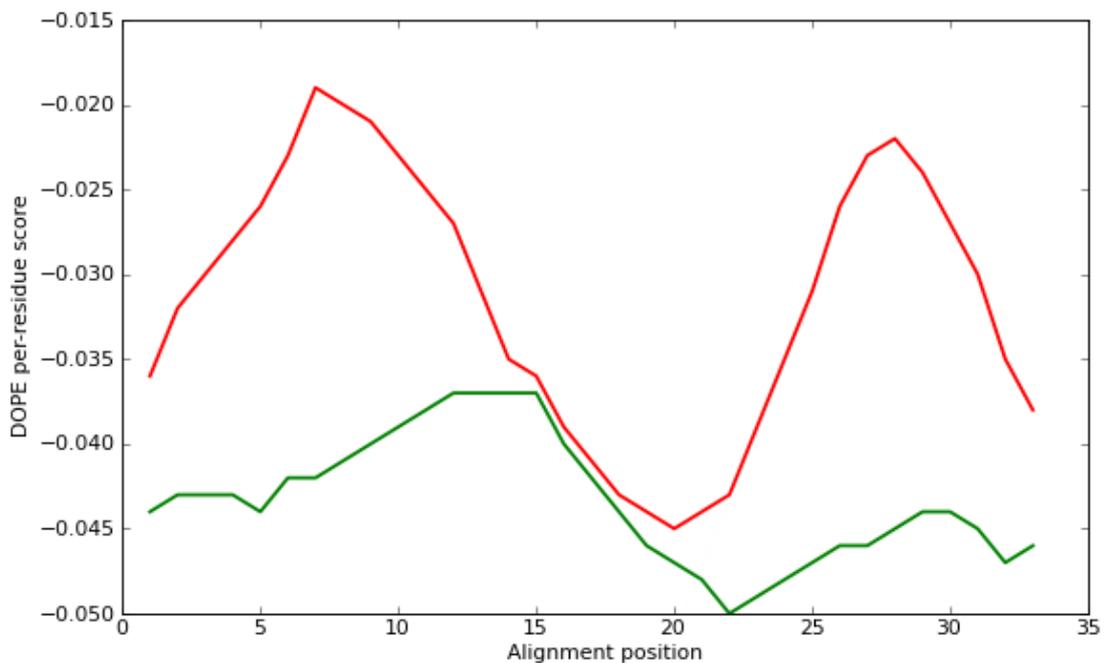


Figura 10. Perfil de DOPE score antes y con los refinamientos. En rojo se representa el perfil del modelo preliminar (antes de ser refinado), en verde el perfil del modelo tras ser refinado.

En la Figura 10 se observa que tras llevar a cabo el refinamiento, hubo una efectiva disminución en el DOPE score. Se decide entonces volver a evaluar el nuevo modelo para analizar la calidad de las nuevas estructuras y, de este modo, determinar si el refinamiento llevado a cabo tuvo un efecto positivo sobre la calidad del modelo.

⁵ DOPE: Discrete optimized protein energy

4.1.3 Evaluación del modelo generado y refinado con Modeller

El modelo estructural del CBM de la endoglucanasa de *T. versicolor* fue refinado con herramientas de Modeller. Luego del refinamiento es necesario llevar a cabo una nueva evaluación, a modo de determinar si los problemas detectados previamente consiguieron solucionarse.

En la Tabla 14 se muestra el análisis de calidad del modelo de CBM de TvEG después de ser refinado.

Tabla 14. Análisis estructural del modelo tridimensional refinado obtenido para la secuencia del dominio de unión a celulosa (CBM). Los parámetros resaltados en rojo indican parámetros con problemas.

RES. NUM.	RES. NAME	SCND STRUC	HBOND HBOND	BTURN BTURN	RES. ASA	FRAC. ASA	RES. VOL.	FRAC. VOL.	PHI PHI	PSI PSI	OMEGA OMEGA	PRBLM PRBLM
1	VAL	BCC C			156.5	0.91	130.6	0.96	360.0	-39.2	173.0	
2	TRP	BCB B			149.2	0.56	190.1	0.82	-134.1	72.2	-172.8	
3	ARG	BCB B 32,32			87.5	0.36	166.0	0.95	-119.4	129.6	-175.7	
4	GLN	BCB B			23.8	0.13	164.0	1.18	-73.0	114.8	174.2	
5	CYS	BCB B 30			47.0	0.33	86.8	0.83	-90.9	-23.4	-177.5	
6	GLY	BBC B 30			37.3	0.41	62.3	0.99	179.6	147.0	177.6	
7	GLY	BBC B			44.0	0.48	53.2	0.84	-171.5	-173.3	179.5	
8	THR	BBC B 27			125.1	0.83	118.6	1.02	-112.7	152.2	179.2	
9	GLY	CCC C			66.6	0.73	57.1	0.91	97.4	-179.6	179.8	
10	PHE	CBC C			214.6	0.95	171.1	0.87	-101.1	133.8	178.2	
11	SER	CBC C			85.6	0.65	85.4	0.94	177.0	171.3	177.7	
12	GLY	CBC C			69.3	0.76	49.5	0.79	-179.0	144.0	-179.7	v
13	ASP	CCC C			44.4	0.28	103.9	0.91	-88.6	-169.1	176.0	
14	THR	CCC C			97.2	0.64	97.2	0.83	-91.1	-148.2	170.7	
15	THR	CBC C			115.7	0.76	105.6	0.90	-87.3	160.0	177.0	
16	CYS	CBC C			53.0	0.37	99.3	0.95	-110.6	156.1	179.4	
17	ALA	CBC C			80.1	0.65	92.3	1.06	-114.1	157.8	177.3	
18	SER	CBC C			112.3	0.85	73.4	0.81	-88.2	164.5	178.6	
19	GLY	CCC C			87.2	0.96	48.7	0.77	108.6	-36.5	170.7	v
20	SER	CCC C			44.4	0.34	84.3	0.93	-97.2	-166.7	179.2	
21	SER	BBB B 33			66.5	0.50	84.1	0.93	-133.9	-144.6	173.5	
22	CYS	BBB B			60.7	0.42	87.2	0.83	-152.8	89.0	-174.2	
23	VAL	BBB B 31,31			71.1	0.41	115.2	0.85	-114.8	156.5	-176.9	
24	ALA	CBC C			76.8	0.62	75.2	0.86	-72.8	146.7	174.1	
25	VAL	CCC C 29			87.2	0.51	117.2	0.86	-144.4	-122.9	179.5	
26	ASN	CCC C 28,29		III	73.5	0.44	108.8	0.93	-71.1	166.2	-171.4	
27	SER	CCC C 8		III	101.1	0.77	86.6	0.96	-56.5	-25.6	178.8	
28	TYR	BCC C 26		III	132.3	0.55	174.9	0.91	-70.1	-44.4	-167.2	o
29	TYR	BBB B 26,25		III	66.0	0.27	177.3	0.93	-132.9	128.1	160.6	o
30	SER	BBB B 6,5			35.8	0.27	82.2	0.91	-98.1	115.8	-167.6	o
31	GLN	BBB B 23,23			0.1	0.00	147.4	1.06	-133.0	165.5	-177.2	
32	CYS	BBB B 3,3			1.7	0.01	107.2	1.03	-68.4	90.4	173.6	
33	GLN	BCB B 21			135.1	0.71	134.6	0.97	-89.2	360.0	360.0	

A : Indica posible problema con ASA fraccional (fASA > 1.0).

V : Indica posible problema con volumen fraccional (fV < 0.8 o fV > 1.2).

P : Indica posible problema con ángulos Phi y Psi [Morris et al., 1992].

O : Indica posible problema con ángulo omega (omega > 170 u omega < -170)

C : Indica enlace peptídico cis (-20 < omega < 20)

Como se puede observar en la Tabla 14, obtenida mediante VADAR, el análisis arrojado indica mejoras en la calidad de la estructura, específicamente en los residuos GLY7, GLY9, GLY19 y CYS22, lo cual es consistente con lo obtenido mediante el perfil de DOPE score. Pese a esto, se incorporó un nuevo problema en el residuo GLY12, el cual pasó a tener un volumen fraccional 0,79. Si bien este valor se encuentra fuera de rango (volumen fraccional >0,8), no se considera mayormente relevante, al encontrarse levemente (0,01) bajo el límite aceptable. Además el residuo GLY12 no se ubica en el sitio de unión a celulosa reportado previamente para otros módulos de la familia I.

Por otro lado, se mantuvo el problema de ángulos omega presentes en los residuos TYR28, TYR29 y SER30. Al comparar los valores antes y después del refinamiento, no se observa cambios en la magnitud de este ángulo, lo cual indicaría que el refinamiento no tuvo efectos sobre la conformación de los ángulos omega en el nuevo modelo planteado.

Una posible causa del problema presente en los ángulos omega de los residuos TYR28, TYR29 y SER30, se deben a un error arrastrado por la plantilla escogida, la cual presenta problema en varios de sus ángulos omega, incluyendo el de los residuos que alinean con el residuo TYR28, TYR29 y SER30 de la secuencia objetivo.

Para solucionar el problema presente en los ángulos omega, utilizando el software Swiss-pdb viewer [Guex y Peitsch, 1997; Kaplan y Littlejohn, 2001] se forzó la molécula para que sus residuos TYR28, TYR29 y SER30 tuviesen sus ángulos dentro del rango. Sin embargo, este cambio provocó un deterioro en la calidad de la estructura, apareciendo problemas en los residuos VAL1, GLY19, CYS22, VAL23, GLY31, CYS32 y GLY33, por lo que se descartó forzar el ángulo omega de residuos TYR28, TYR29 y SER30 a estar dentro del rango (± 170).

Finalmente, se procede a validar el modelo obtenido y refinado con Modeller, utilizando el servidor Molprobability. Se lleva a cabo un análisis sobre los ángulos phi psi permitidos, generando el gráfico de Ramachandran para el modelo evaluado. Este gráfico analiza el valor phi y psi de los residuos y a partir de éstos determina si los residuos se ubican en zonas con o sin impedimento estérico.

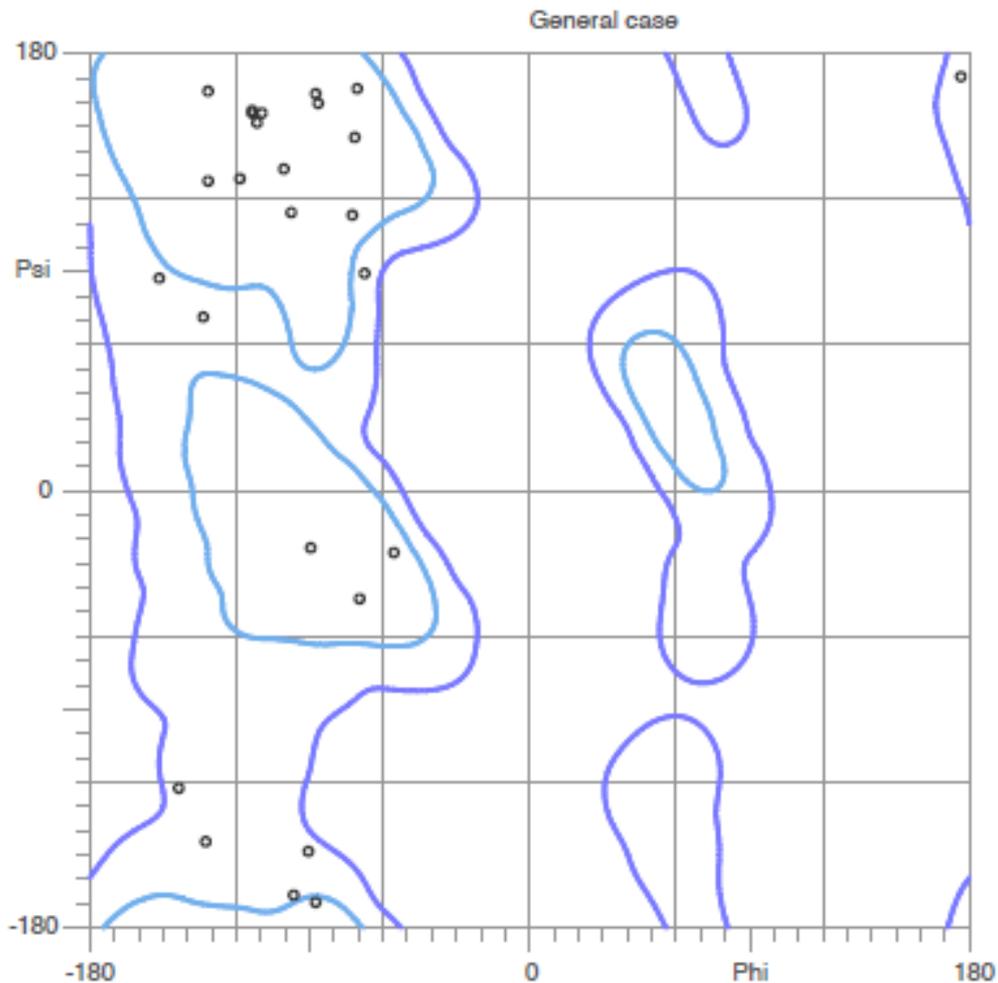


Figura 11. Gráfico de Ramachandran de la estructura final de la secuencia CBM de TvEG. Las regiones delimitadas por líneas celestes corresponden a zonas óptimas, mientras que aquellas delimitadas con una línea azul son zonas permitidas. Las regiones sin delimitación corresponden a zonas no permitidas. Gráfico obtenido con MolProbity.

En la Figura 11 se muestra el gráfico de Ramachandran para el modelo tridimensional del CBM de TvEG. Como se observa en la gráfica, la totalidad de los residuos (representados mediante un punto) se encuentran en áreas delimitadas con línea azul, lo cual indica ángulos permitidos. Por otro lado, un 74,19% de los residuos se ubican en zonas delimitadas mediante líneas celestes, las cuales representan las zonas óptimas de configuración dentro de la proteína, correspondiendo a conformaciones donde no hay impedimentos estéricos.

El gráfico de Ramachandran permite también aproximar la estructura tridimensional de la proteína, a partir de las combinaciones phi-psi presentes. En el caso del CBM de TvEG la mayoría de sus residuos presentan $-40 < \phi < -180$ y $60 < \psi < 180$ (cuadrante superior izquierdo de la gráfica), indicando una predominancia, en el péptido, de estructuras secundarias tipo hoja β , lo cual es concordante con antecedentes bibliográficos para este tipo de estructuras [Bray et al., 1996; Linder y Teeri, 1997; Lynd et al., 2002].

La Tabla 15 muestra el resumen del análisis de calidad para la estructura final del CBM de TvEG. Existe un 11,54% de rotámetros pobres (residuos 3, 26 y 31), es decir, aquellos residuos poseen problemas en la rotación de sus cadenas laterales sobre el esqueleto rígido polipeptídico. Por otro lado un 6,06% (14 y 30) de residuos presentan problemas en sus ángulos.

Tabla 15. Tabla resumen del análisis de calidad de la estructura (realizado con MolProbity).

Parámetro	CBM de TvEG	Meta
Rotámetros pobres	11,54% (3/26)	< 1%
Ramachandran outliers	0%	< 0.2%
Ramachandran favorecidos	74,19% (23/31)	> 98%
Desviaciones C β > 0.25 Å	0%	0 %
Residuos con malos enlaces	0%	0 %
Residuos con malos ángulos	6,06% (2/31)	< 0.1%

La mayoría de los problemas del modelo se encuentran ubicados en las hojas β . Esto sugiere que a pesar de los problemas presentes no se ve afectada la formación de la estructura secundaria. En el caso de que estos problemas se hubiesen presentado en el *loop* del péptido (problema más habitual en los modelos por comparación) podría haber un efecto en el posicionamiento y espaciado de las hojas β . De ser así podría verse afectado el posterior análisis de docking molecular, sin embargo en este caso los problemas en el modelo no parecen afectar la conformación final adquirida por el péptido.

Pese al refinamiento llevado a cabo aún existen problemas con el modelo obtenido. Si bien éstos son despreciados para los análisis desarrollados posteriormente, se aconseja sean corregidos a futuro, por ejemplo, utilizando técnicas de refinamiento basadas en simulaciones de dinámica molecular.

4.1.4 Modelo tridimensional del módulo de unión a carbohidrato (CBM)

En la Figura 12, se muestra el modelo de la estructura tridimensional del CBM de la endoglucanasa de *Trametes versicolor*, obtenido por comparación, utilizando como plantilla el módulo de unión a celulosa de una celobiohidrolasa I del hongo *Trichoderma reesei* (1azk).

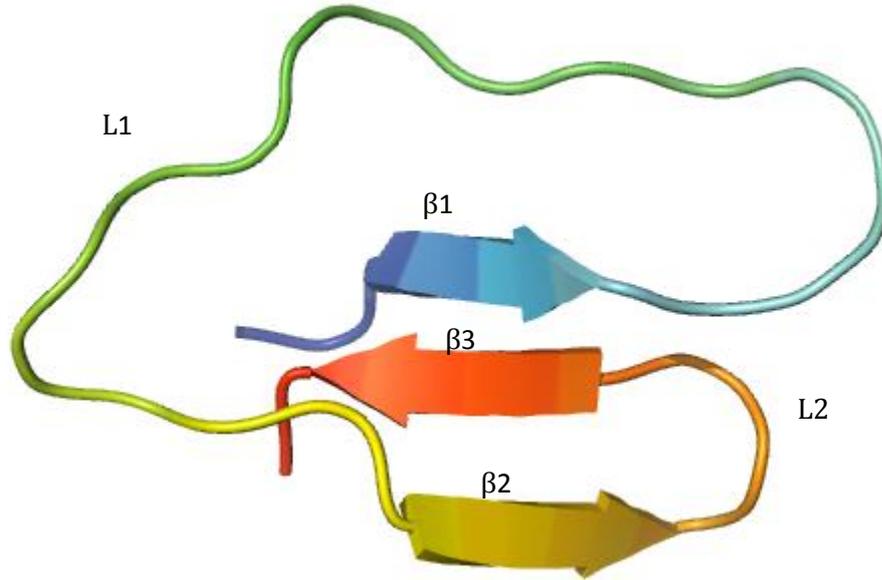


Figura 12. Modelo tridimensional para el CBM. En azul se indica el extremo amino y en rojo el extremo carboxilo. Imagen generada con Pymol.

Como se predijo con la gráfica de Ramachandran, el modelo obtenido para el CBM de TvEG presenta estructura secundaria, formada principalmente por hojas beta. El CBM de TvEG se conforma por tres hojas beta antiparalelas, $\beta 1$, $\beta 2$ y $\beta 3$, nombradas desde el extremo amino hacia el carboxilo (ver Figura 12). Éstas se encuentran conectadas entre sí mediante *loops*, presentándose uno de mayor tamaño (L1: entre las hojas $\beta 1$ y $\beta 2$) y uno de menor tamaño (L2: entre las hojas $\beta 2$ y $\beta 3$), formados por 18 y 3 residuos, respectivamente.

Este plegamiento se adopta gracias a la presencia de dos puentes disulfuro, formados entre las cisteínas 16 y 32, 4 y 22 (ver Figura 13), las cuales al quedar enfrentadas, establecen un enlace (de tipo covalente) entre los grupos $-SH$ de las cisteínas.

En otros CBM de la familia I se ha encontrado que los puentes disulfuro no sólo permiten un correcto plegamiento del péptido, sino también estabilizan y rigidizan la estructura [Linder y Teeri, 1997; Tavagnacco et al., 2011]. Las hojas beta, ayudan a la estabilidad de la molécula, y en conjunto con los puentes disulfuro, permiten el espaciamiento en intervalos regulares, de los residuos 2, 28 y 29, correspondientes a un triptófano y dos tirosinas respectivamente.

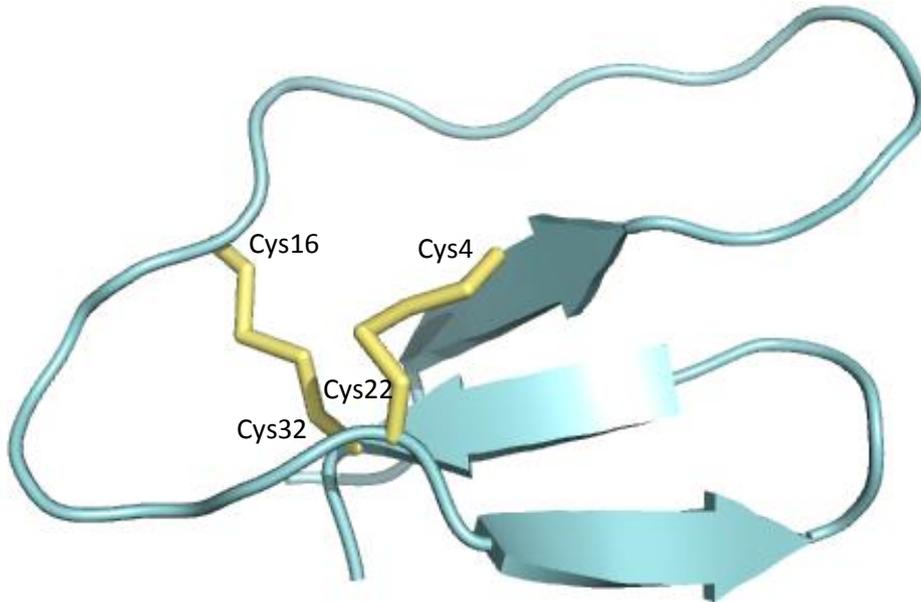


Figura 13. Puentes disulfuro presentes en el modelo tridimensional del CBM. En amarillo se indican las cisteínas que forman puentes disulfuro. Imagen generada con Pymol.

Los residuos 2, 28 y 29 corresponden a aminoácidos conservados en CBM de la familia I, los cuales se orientarían en un plano enfrentado a la celulosa [Linder y Teeri, 1997; Tavagnacco et al., 2011]. En la Figura 14 se muestra la orientación y distancia de los aromáticos de los residuos anteriormente mencionados.

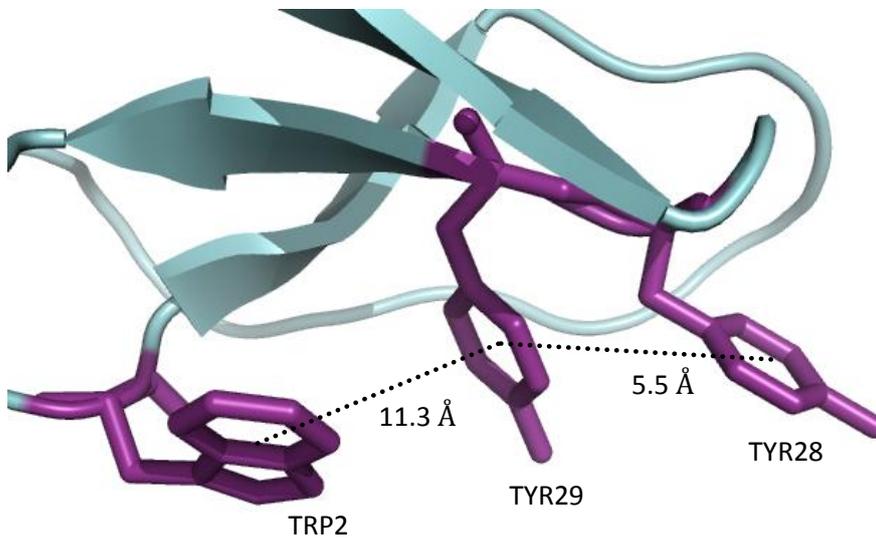


Figura 14. Distancia de los aminoácidos aromáticos conservados en el CBM de TvEG. En morado se destacan los residuos TRP2, TYR 28 y TYR 29 y en línea punteada la distancia entre estos. Imagen generada con Pymol.

Los residuos 28 y 29 se encuentran distanciados en un intervalo semejante al espaciamiento de dos azúcares que forman la celulosa, mientras que los residuos 2 y 29 se espacian en un intervalo correspondiente a tres azúcares.

Al comparar el modelo obtenido con el modelo de CBM I de una celobiohidrolasa de *Trichoderma reesei* [Beckham et al., 2010 ; Nimlos et al., 2007] se observa que los residuos 2, 28 y 29 no se orientan en un mismo plano, como sería lo esperado según el modelo previamente estudiado (Figura 15).

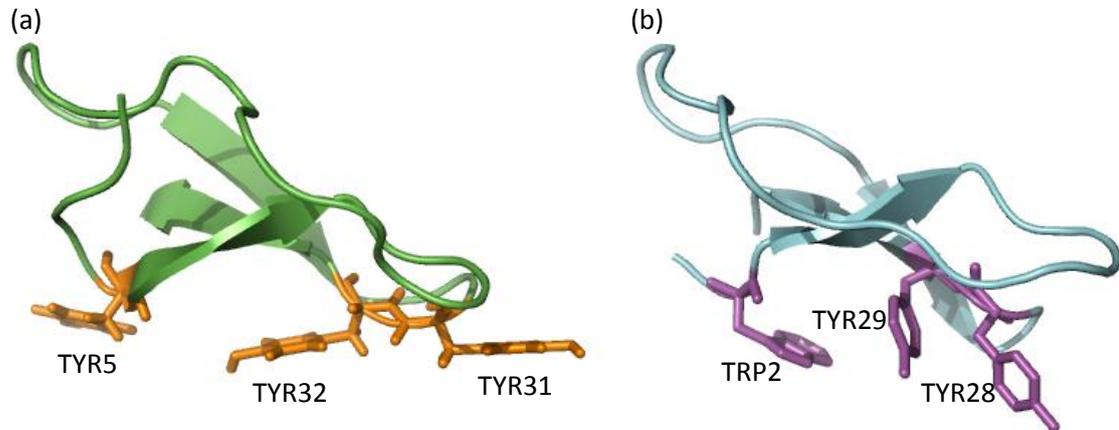


Figura 15. Comparación modelo generado con modelo de CBM existente. (a) Modelo existente para CBM I de *Trichoderma reesei*. (b) Modelo generado para CBM I de *Trametes versicolor*. En el caso del modelo previamente estudiado (a) se observan los aminoácidos conservados orientados en un mismo plano, mientras que en el modelo generado en este estudio (b) no ocurre lo mismo. Imagen generada con Pymol.

Un estudio previo (basado en dinámica molecular) sugiere que la celulosa induce un cambio conformacional sobre el dominio de unión, provocando la rotación de estos residuos [Nimlos et al., 2007]. El modelo generado en el presente estudio no fue sometido a dinámica molecular, lo que podría explicar la diferencia encontrada en el plano de orientación de sus residuos conservados.

Una vez obtenido el modelo para el CBM de TvEG se llevó a cabo análisis mediante docking molecular, con el fin de estudiar la conformación adquirida por el módulo en presencia de celulosa.

4.2 Estudio de la interacción CBM-celulosa mediante docking molecular

El modelo tridimensional del CBM de la endoglucanasa de *Trametes versicolor* obtenido previamente, se utilizó para llevar a cabo un análisis de docking molecular y de esta forma estudiar la unión del péptido a celulosa. Con este fin se utilizaron dos programas para llevar a cabo el análisis: Autodock y Vina. Para ambos se emplearon distintas condiciones o casos de estudio:

- Docking ciego: consiste en definir un espacio de búsqueda suficientemente grande de modo de permitir la unión del ligando en cualquier lugar del receptor.
- Docking forzado: consiste en definir un espacio de búsqueda limitado. En este caso, se limitó el espacio de búsqueda al área de interacción observada en otros dominios de unión de la familia I (entorno a los residuos TRP 2, TYR 28 y TYR 29), basando la elección en antecedentes de estudios previos [Beckham et al., 2010 ; Linder y Teeri, 1997; Mattinen et al., 1997; Tavagnacco et al., 2011].
- Receptor rígido: se define que el receptor no tendrá movilidad.
- Receptor flexible: se introduce flexibilidad a las cadenas laterales de los residuos TRP 2, TYR 28 y TYR 29.

La elección del tamaño del espacio de búsqueda se basó en que ésta fuese suficientemente grande de modo de permitir la rotación de la cadena de celulosa (utilizada como ligando) y de un tamaño tal que pudiese contener la totalidad de átomos del CBM en estudio (receptor). En base al primer criterio se definió que la caja de búsqueda fuese de al menos 26Å x 26Å x 26Å (caja de 26 x 26 x 26, con puntos espaciados cada 1Å). Para el caso del docking ciego, la caja de búsqueda se centró en el receptor, mientras que en el caso del docking forzado se centró de manera tal de limitar el espacio de búsqueda a la zona de interés.

Se llevaron a cabo 4 simulaciones, correspondientes a las combinaciones de las condiciones de docking descritas previamente:

- Docking ciego y receptor rígido.
- Docking forzado y receptor rígido.
- Docking ciego y receptor flexible.
- Docking forzado y receptor flexible.

Cada una de las simulaciones se realizó tanto en Autodock como en Vina. Ambos programas utilizan el mismo tipo de archivos de entrada para describir al receptor y al ligando, lo que facilita el uso de ambos programas y la comparación de sus resultados.

Las funciones de puntuación tienen parámetros similares, pero ponderados de manera distinta. Una diferencia clave entre los programas es el algoritmo de búsqueda local. Autodock utiliza búsqueda estocástica que genera conformaciones al azar. Vina en cambio calcula un gradiente mientras busca un óptimo local [Chang et al., 2010].

El objetivo de utilizar ambos programas para el análisis se basa en estudiar la interacción del módulo a celulosa a través de dos metodologías diferentes (distintos algoritmos de búsqueda). De este modo se busca llegar a resultados similares utilizando dos procedimientos independientes; al utilizar metodologías distintas se busca asegurar validez interna del modelo.

4.2.1 Conformaciones del complejo CBM-celulosa obtenidas a partir del análisis de Docking molecular

El análisis de docking molecular llevado a cabo con Vina permitió determinar la conformación y ubicación que adquiere la celulosa en el CBM de la endoglucanasa de *T. versicolor*, de manera rápida y precisa [Trott y Olson, 2010].

En la Figura 16 se muestran los resultados de las distintas conformaciones obtenidas con Vina para las simulaciones con los parámetros previamente descritos. Se observan dos zonas de interacción distintas; una identificada en el caso de llevar a cabo docking ciego considerando receptor rígido (a) y otra en el caso de forzar el espacio de búsqueda (b) o introducir flexibilidad al receptor (c) y (d).

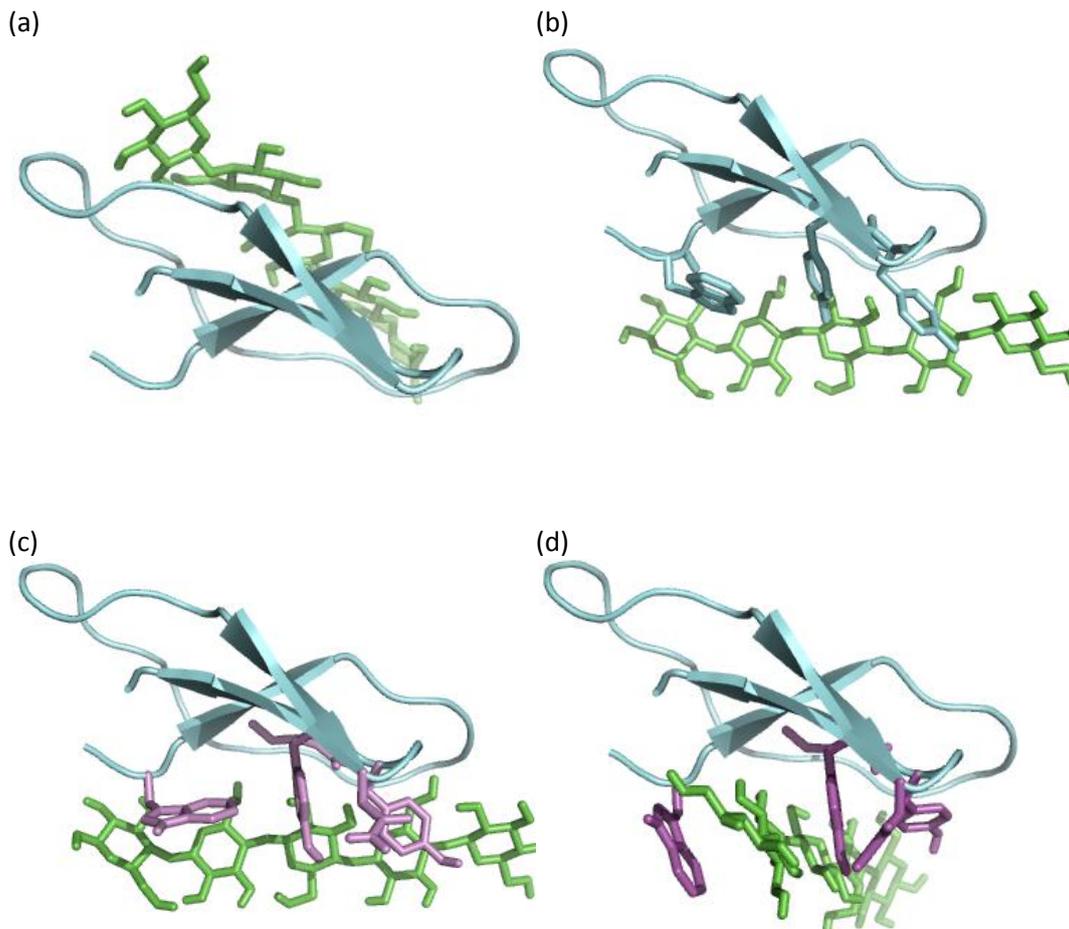


Figura 16. Interacción entre CBM I de TvEG mediante Vina. (a) Docking ciego y receptor rígido, (b) Docking forzado y receptor rígido, (c) Docking ciego y receptor flexible (TRP2, TYR28 y TYR29), (d) Docking forzado y receptor flexible (TRP2, TYR28 y TYR29). Imágenes generadas con Pymol.

A partir de la realización de un docking ciego, considerando al receptor rígido (ver caso (a) de la Figura 16), se obtuvo que la celulosa interactúa con el CBM de la endoglucanasa de *T. versicolor*, en un plano distinto al descrito previamente por otros autores para dominios de unión pertenecientes a la misma familia que el dominio en estudio (familia I) [Beckham et al., 2010 ; Linder y Teeri, 1997; Mattinen et al., 1997; Tavagnacco et al., 2011]. En base a esto, se lleva a cabo un nuevo análisis de docking, esta vez forzando el espacio de búsqueda a la zona de interacción esperada (ver caso (b)).

Al limitar el espacio de búsqueda, se obtiene que la celulosa se enfrenta al péptido, en el plano observado para otros dominios de unión. Si bien la celulosa se posiciona según lo esperado, no se observa una orientación de los grupos aromáticos de las cadenas laterales de los residuos TRP2, TYR28 y TYR29, debido a que estos mantienen la orientación obtenida en el modelo tridimensional generado. Para permitir que los grupos aromáticos se orienten, se incluyó flexibilidad al modelo y se llevaron a cabo otras dos simulaciones, considerando un docking ciego con receptor flexible (c) y docking forzado con receptor flexible (d).

En los casos en que se permitió movilidad en las cadenas laterales de los residuos TRP2, TYR 28 y TYR29, se consiguió una mejor orientación de los aromáticos sobre los anillos de azúcar de la celulosa. En el caso (c) basta incluir flexibilidad en las cadenas laterales de los residuos conservados para obtener una zona de interacción consistente con la reportada previamente por otros autores [Beckham et al., 2010 ; Linder y Teeri, 1997; Mattinen et al., 1997; Tavagnacco et al., 2011].

Se repiten las simulaciones recién planteadas utilizando el programa Autodock, a modo de comparar si se identifican las mismas zonas de interacción que las arrojadas tras el análisis con Vina.

En la Figura 17 se muestran los resultados de las distintas conformaciones obtenidas con Autodock para las simulaciones con los parámetros previamente descritos. Se representa la cadena de celulosa en naranja para diferenciarla de las simulaciones llevadas a cabo con Vina.

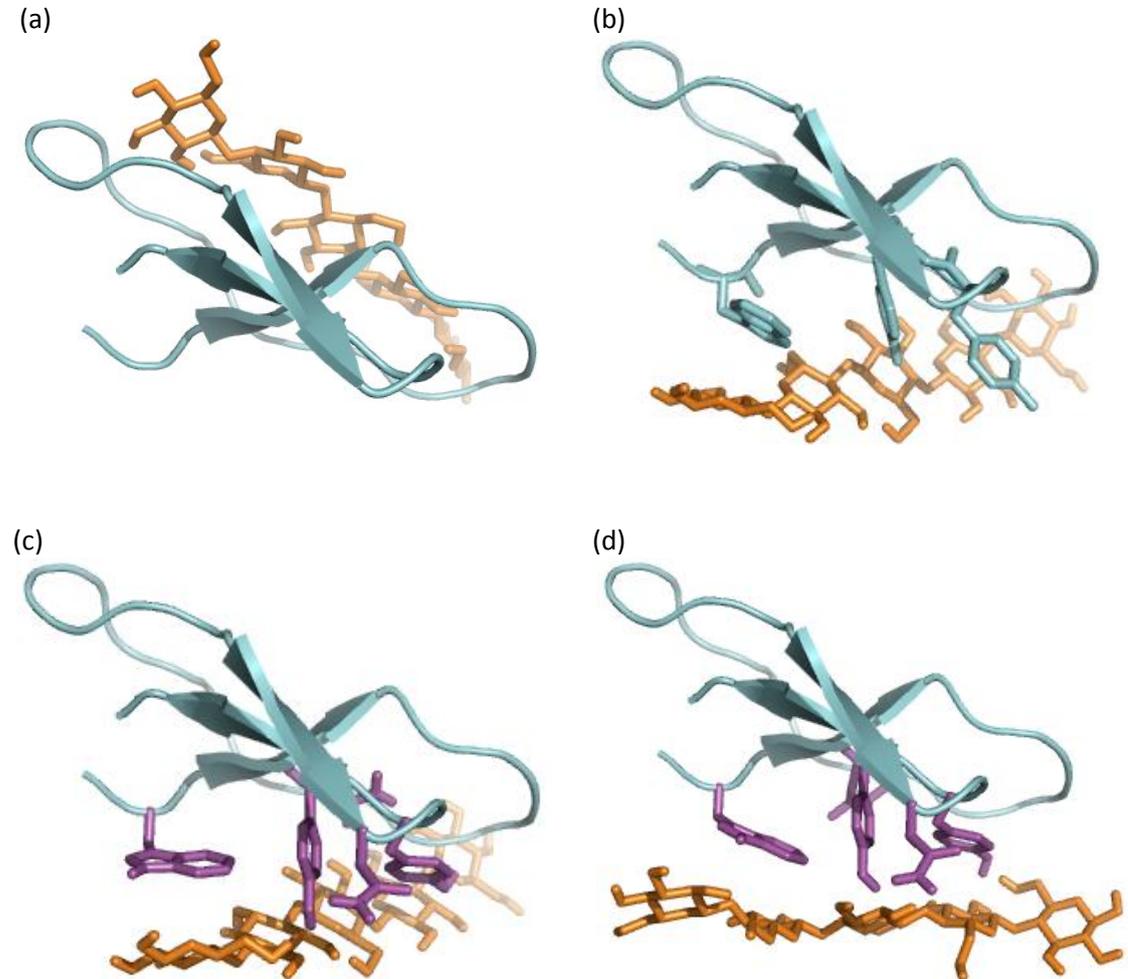


Figura 17. Interacción entre CBM I de TvEG mediante Autodock. (a) Docking ciego y receptor rígido, (b) Docking forzado y receptor rígido, (c) Docking ciego y receptor flexible (TRP2, TYR28 y TYR29), (d) Docking forzado y receptor flexible (TRP2, TYR28 y TYR29). Imágenes generadas con Pymol.

Al analizar los resultados de las conformaciones obtenidas con Autodock (Figura 17) vemos que éstas son consistentes con las encontradas al llevar a cabo el análisis de docking molecular con Vina, encontrándose conformaciones muy similares en ambos casos. Las zonas de interacción identificada en cada una de los casos coinciden con las zonas identificadas con Vina.

El docking ciego con receptor rígido (Figura 17, caso (a)) nuevamente sugiere que la interacción de la celulosa con el péptido ocurre en una zona distinta a la planteada por otros autores [Beckham et al., 2010 ; Linder y Teeri, 1997; Mattinen et al., 1997; Tavagnacco et al. 2011].

En la Figura 18 se muestra la orientación de las cadenas laterales de los aminoácidos conservados en los CBM de la familia I. Los casos analizados corresponden a una comparación de la orientación de los anillos aromáticos de los aminoácidos conservados tras llevar a cabo docking ciego (casos I y III) y docking forzado (casos II y IV), considerando para todos los casos flexibilidad en el receptor (cadenas laterales de los aminoácidos conservados). Los dos primeros casos (celulosa representada en verde) corresponden a resultados obtenidos mediante Vina mientras que los casos III y IV (celulosa representada en naranja) corresponden a resultados obtenidos con Autodock.

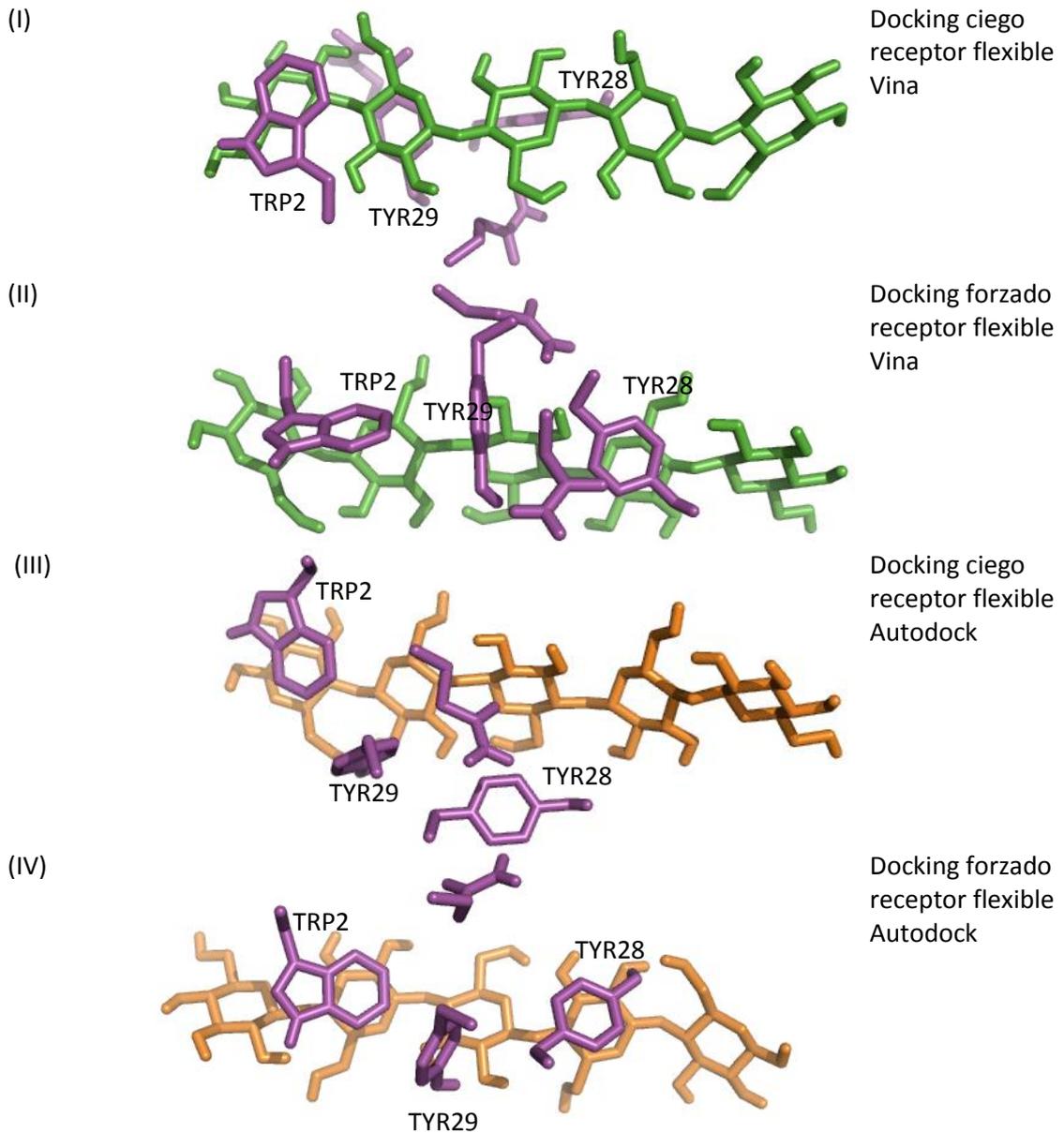


Figura 18. Orientación de los anillos aromáticos de las cadenas laterales hacia la celulosa. En todos los casos estudiados se consideraron flexibles las cadenas laterales de los aminoácidos aromáticos conservados (TRP2, TYR 28 y TYR 29). En los casos I y III se llevó a cabo docking ciego, mientras que en II y IV se llevó a cabo docking forzado. Se representa la cadena de celulosa en color verde para los análisis desarrollados con Vina y naranja para los análisis desarrollados con Autodock.

La Figura 18 representa la orientación de los residuos (conservados de la familia I) hacia celulosa. En todos los casos se observa que la periodicidad de los anillos de los grupos aromáticos es similar al de los azúcares de la celulosa. A pesar que el distanciamiento de los residuos coincide con el de los azúcares, los aminoácidos no siempre se orientan hacia la cadena de celulosa.

De los 4 casos analizados sólo en el primero la tirosina de la posición 29 se orientó frente a la celulosa. Los residuos triptófano 2 y tirosina 28 en cambio se orientaron de mejor manera a los azúcares de la celulosa.

Al comparar los resultados entregados por Vina y Autodock se observa una mayor consistencia en el caso del docking forzado (casos II y IV). Esto puede deberse a que en el caso de docking forzado el espacio de búsqueda es más pequeño, por lo que es más probable llegar a resultados similares.

Finalmente se concluye que si bien no se obtuvo una orientación total de los aromáticos como la planteada mediante análisis de dinámica molecular, la incorporación de flexibilidad al modelo de docking, resulta una buena aproximación a lo obtenido en las simulaciones de dinámica. El análisis mediante docking molecular presenta menor precisión, pero tiene la ventaja de ser más fácil de implementar y requerir menos recursos computacionales para ser llevado a cabo.

4.2.2 Energías de interacción del CBM de TvEG sobre celulosa

El análisis de docking molecular no sólo permite determinar la conformación que adquiere el complejo CBM – celulosa, sino también permite cuantificar la interacción mediante valores de energía.

Para los cuatro casos analizados anteriormente se exponen las energías de unión del CBM de TvEG a celulosa, tanto para las simulaciones realizadas con Vina como para las utilizadas con Autodock. El resumen de estas energías se grafica en la Figura 19.

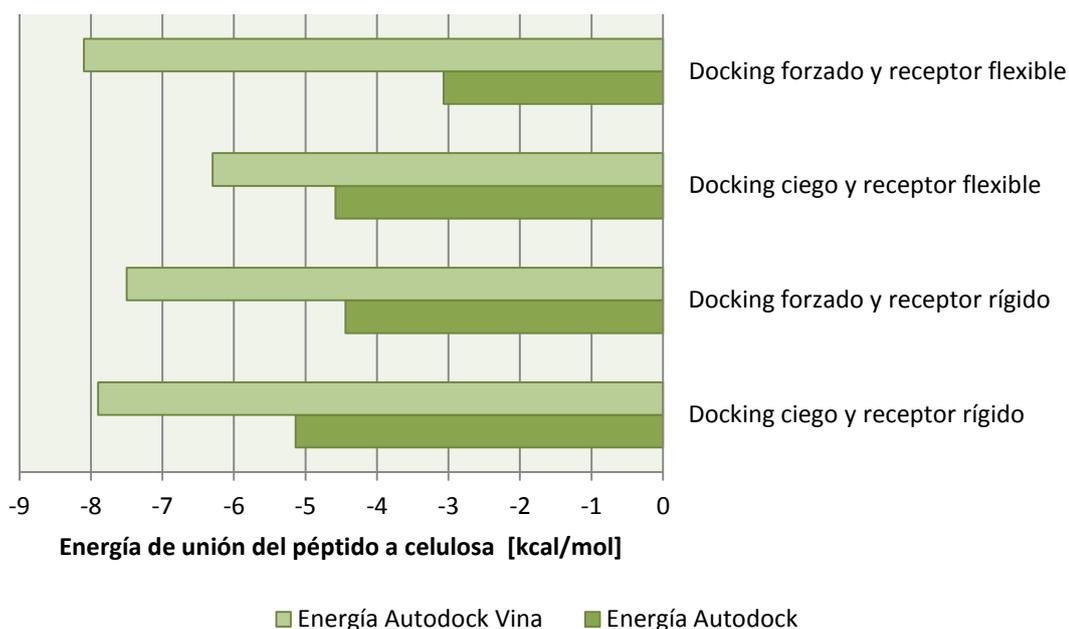


Figura 19. Energías de interacción calculadas con Autodock y Vina, para los cuatro casos en estudio.

Para el caso inicial de estudio (docking ciego y receptor rígido) la energía calculada fue de -7,9 kcal/mol con Vina y de -5,14 kcal/mol con Autodock. Al analizar la conformación obtenida se observa que la interacción de la celulosa con el péptido ocurre en una zona distinta a la planteada por otros autores.

En el docking forzado con receptor rígido se obtuvieron energías de -7,5 y -4,44 kcal/mol (calculadas con Vina y Autodock respectivamente). Si bien la interacción de la celulosa fue consistente con la reportada en la literatura, la magnitud de la energía fue menor a la obtenida en el caso anterior.

Por otro lado en los casos en que se consideró el receptor flexible no se observa una consistencia en las energías entregadas por ambos programas. Para el caso de llevar a cabo docking ciego es donde menos diferencia observamos entre ambas energías, sin embargo para Vina es el caso menos favorable (en relación a los otros estudiados) mientras que basados en Autodock se podría decir que es el segundo caso más favorable (en cuanto a energía).

En el caso en que se llevó a cabo docking forzado con receptor flexible es donde se presenta una mayor diferencia en las magnitudes de las energías entregadas por ambos programas. Resulta ser el caso más favorable si se analiza con Vina y el caso menos favorable si se analiza con Autodock.

Las diferencias en las magnitudes de energía calculadas por Autodock y Vina se deben a problemas en el cálculo realizado por Autodock cuando existen modelamientos difíciles, es decir, al utilizar un ligando muy grande o al permitir muchas rotaciones [Chang et al., 2010 ; Dhanik et al., 2011]. En el análisis llevado a cabo para el CBM de TeVG el ligando corresponde a una cadena de celulosa compuesta por 5 unidades de glucosa, lo que dificulta el cálculo realizado por Autodock y genera la diferencia de magnitud con la energía calculada con Vina. A partir de esto se puede inferir que al introducir flexibilidad al análisis y al utilizar un ligando de gran tamaño disminuye la correlación entre las energías calculadas por ambos programas.

En base a esto se analizan los dos casos en que se considera el receptor rígido, es decir con docking ciego y forzado. Para estos casos se identifican dos zonas de interacción distintas. La zona identificada tras llevar a cabo un docking ciego es distinta a la reportada previamente por otros autores. Debido a que esta zona no coincide con la documentada en otros estudios se designa “zona II”. La zona I en cambio representa la interacción descrita previamente para otros módulos de la familia I. Esta zona corresponde a la identificada al llevar a cabo un estudio de docking forzado (ver Figura 16).

La magnitud de energía de unión de la zona I es comparable con la energía de unión de la zona II, tanto para las energías calculadas por Vina (-7,5 y -7,9 kcal/mol para las zonas I y II respectivamente) como para las energías calculadas por Autodock (-4,44 y -5,14 kcal/mol para las zonas I y II respectivamente). Debido a esto se concluye que la unión en ambos sitios es igualmente favorable (en términos energéticos), a partir de lo cual se postula que el CBM de la endoglucanasa de *T. versicolor* posee dos zonas de interacción a celulosa; la zona I reportada previamente por otros autores y la zona II planteada en el presente estudio.

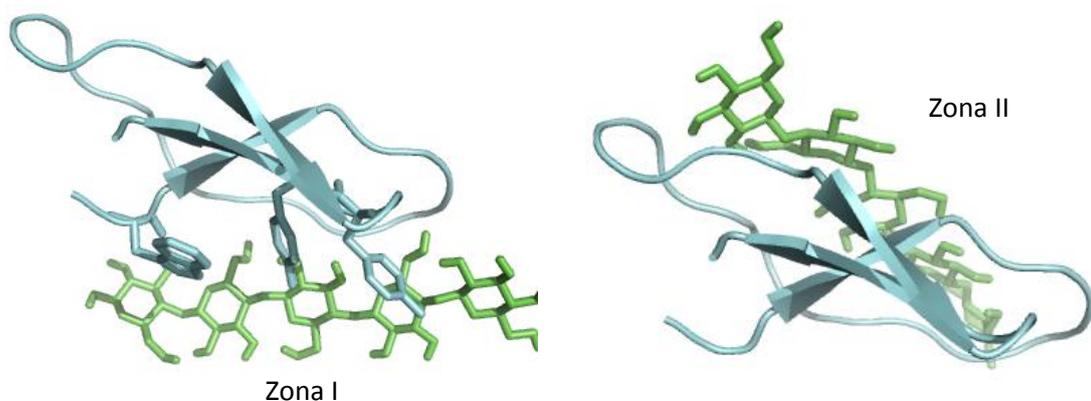


Figura 20. CBM de TvEG unido a celulosa. A la izquierda se muestra la zona I de interacción y a la derecha se muestra la zona II de interacción.

4.3 Identificación de mutaciones para mejorar la interacción CBM I de TvEG a celulosa

A partir del análisis previo de las zonas de interacción a celulosa, se plantearon un total de 24 mutaciones. Para cada una de ellas, se obtuvo un modelo tridimensional de estructura y se analizó mediante docking molecular, con el fin de seleccionar las mejores candidatas a llevar a cabo experimentalmente.

4.3.1 Planteamiento de las mutaciones

Estudios previos han determinado que la afinidad de CBM fúngicos a celulosa depende de la combinación y configuración de los aminoácidos aromáticos conservados [Beckham et al., 2010 ; Mattinen et al., 1997; Takashima et al., 2007].

Para el caso del CBM en estudio se han descrito con 4 aminoácidos aromáticos conservados: TRP2, PHE10, TYR28 y TYR29. Dado esto, en el presente estudio se planteó una primera estrategia de mutación, la cual consiste en mutar los aminoácidos aromáticos conservados y de esta forma estudiar el efecto de distintas configuraciones posibles. Estos aminoácidos se ubican en la zona I de interacción, previamente reportada como la zona de interacción de los CBM de la familia I a celulosa.

Se excluyó del análisis de planteamiento de mutaciones el efecto de la ubicación de fenilalanina, ya que previamente se ha concluido que su interacción con celulosa es menor y más inestable que la observada en el caso de la tirosina y triptófano [Takashima et al., 2007].

Se plantearon 10 mutaciones en las que se alterna la ubicación de triptófanos y tirosinas en las distintas ubicaciones de los aminoácidos conservados en la familia I de CBM, los cuales se han identificado previamente como los responsables de la interacción del módulo hacia celulosa.

En la Tabla 16 a continuación se exponen las mutaciones planteadas, destacando los aminoácidos involucrados y la mutación incorporada.

Tabla 16. Mutaciones planteadas en los aminoácidos conservados en la familia I de CBM. En rojo se destacan los aminoácidos mutados.

	Variante	Mutación incorporada
-	WFYY	Wild type (W2, F10, Y28, Y29)
1	W WWW	F10W, Y28W, Y29W
2	W WWY	F10W, Y28W,
3	W WYW	F10W, Y29W
4	W WYY	F10W
5	W YYY	F10Y
6	Y WWW	W2Y, F10W, Y28W, Y29W
7	Y WWY	W2Y, F10W, Y28W
8	Y WYW	W2Y, F10W, Y29W
9	Y WYY	W2Y, F10W
10	Y YYY	W2Y, F10Y

Como una segunda estrategia de mutación, se decidió incorporar mutaciones en la zona II de interacción a celulosa (ver Figura 20). Para ello se plantea reemplazar residuos de esa zona por aminoácidos aromáticos (triptófano y tirosina), junto con residuos que puedan participar en la formación de puentes de hidrógeno (como asparragina y glutamina). El principal objetivo de esto fue repetir el arreglo de aminoácidos presente en la zona I de interacción e incorporarlo en la zona II de interacción. Para esto, se planteó 14 nuevas mutaciones, las cuales se muestran a continuación en la Tabla 17:

Tabla 17. Mutaciones planteadas en la segunda estrategia de mutación. En rojo se destacan los aminoácidos mutados.

	Variante	Mutación incorporada
-	WGTGFG	Wild type (W2, G7, T8, G9, F10, G12)
11	WWTG WG	G7W
12	WY T N WG	G7Y, G9N, F10W
13	WY T Q WG	G7Y, G9Q, F10W
14	W G W G WG	T8W, F10W
15	W G Y G WG	T8Y, F10W
16	A G T G WG	W2A, F10W
17	WGTG WQ	F10W, G12Q
18	WY T Q WQ	G7Y, G9Q, F10W, G12Q
19	WY T N WQ	G7Y, G9N, F10W, G12Q
20	WY T G WQ	G7Y, F10W, G12Q
21	WY T Q WN	G7Y, G9Q, F10W, G12N
22	WY T Q WN	G7Y, G9Q, F10W, G12N
23	WY T G WN	G7Y, F10W, G12N
24	WY T G WG	G7Y, F10W

Una vez decididas las mutaciones a analizar *in silico*, se obtuvo la estructura tridimensional de éstas, para a continuación analizar su unión a celulosa.

4.3.2 Análisis mediante docking molecular

Se realizó el análisis de la afinidad de las distintas mutantes sobre celulosa, a partir de la evaluación de docking molecular. Este análisis se llevó a cabo utilizando el software Autodock Vina, debido a su rapidez, mejor precisión en sus predicciones y que no requiere el cálculo de mapas de búsqueda, ya que calcula su grilla automática y rápidamente [Trott y Olson, 2010]. Dado esto, Vina facilita el *screening* virtual de un gran número de estructuras.

Se llevó a cabo el análisis de docking molecular a las 10 primeras mutaciones planteadas (ver Tabla 16), correspondientes a las mutaciones localizadas en la zona I de interacción (descrita previamente por otros autores). En la Figura 21 a continuación se muestran las energías calculadas con Vina.

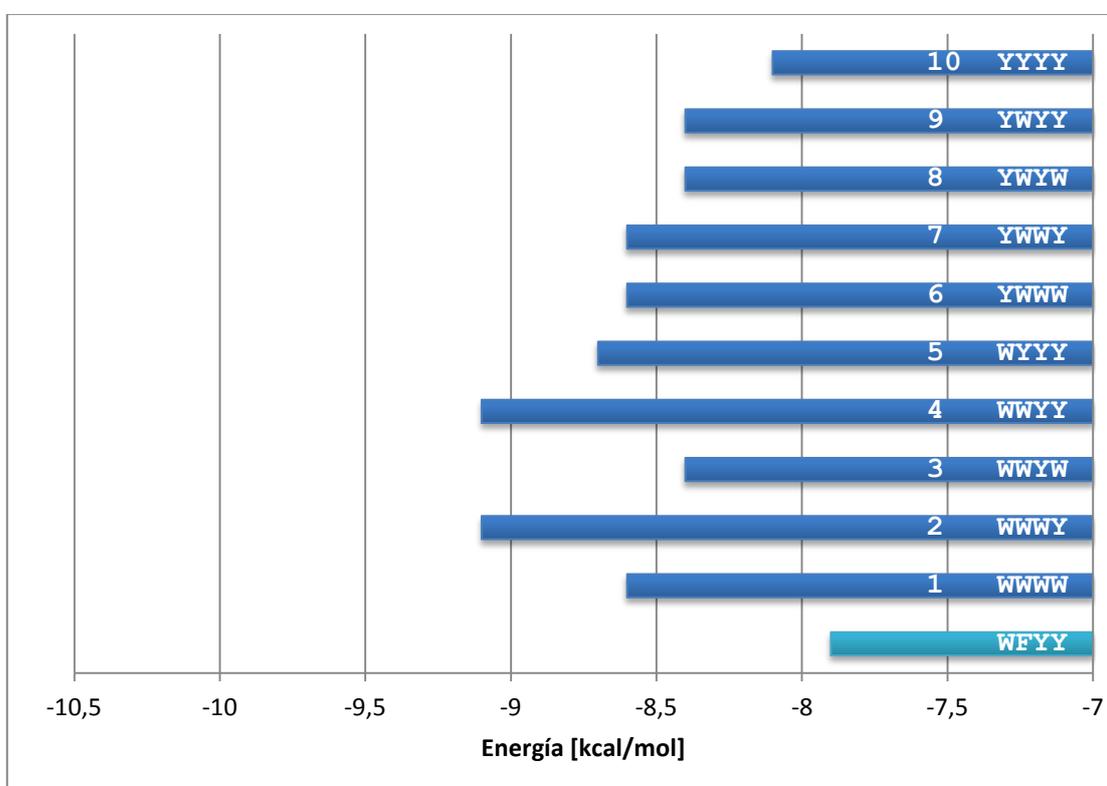


Figura 21. Energías de afinidad obtenidas con Vina para las mutaciones de la primera estrategia planteada (cambios en los aminoácidos conservados: W2, F10, Y28 e Y29). En celeste se muestra la energía de afinidad del CBM de TvEG nativo.

El error asociado al cálculo otorgado por Vina corresponde a 2,8 kcal/mol⁶ [Trott y Olson, 2010]. Las diferencias de energía bajo este valor no son estadísticamente significativas, motivo por el cual se considera que las mutaciones planteadas en la primera estrategia de mutación (ver Figura 21) no presentan mejoras considerables. A pesar de esto se encontró concordancia entre las energías calculadas y antecedentes bibliográficos previos, encontrándose una mejora

⁶ Este error fue calculado por el autor comparando una serie de datos experimentales y predicciones realizadas con Vina.

al sustituir la fenilalanina de la posición 7 por triptófano, [Beckham et al., 2010; Mattinen et al., 1997; Takashima et al., 2007].

En la Tabla 18 se indica la energía de unión previamente graficadas para la primera estrategia de mutación (cambios en los aminoácidos conservados: W2, F10, Y28 e Y29). En la columna de la derecha se incluyó la diferencia de energía de las mutaciones con respecto al CBM de TvEG (nativo).

Tabla 18. Energías de afinidad entregadas por Vina para las mutantes planteadas en la primera estrategia (cambios en los aminoácidos conservados: W2, F10, Y28 e Y29). En la columna de la derecha se indica la diferencia de energía de cada una de las mutaciones, con respecto al CBM de TvEG nativo. Se resalta en rojo las mutaciones con mayores mejoras de afinidad sobre celulosa.

	Variante	Energía de Afinidad [Kcal/mol]	Δ Energía de Afinidad [Kcal/mol]
-	WFYY	-7,9	-
1	WWWW	-8,6	-0,7
2	WWWY	-9,1	-1,2
3	WWYW	-8,4	-0,5
4	WYYW	-9,1	-1,2
5	WYYY	-8,7	-0,8
6	YWWW	-8,6	-0,7
7	YWWY	-8,6	-0,7
8	YWYW	-8,4	-0,5
9	YWYY	-8,4	-0,5
10	YYYY	-8,1	-0,2

El menor valor obtenido corresponde al caso en que los residuos THR2 y FEN7 fueron reemplazados por tirosinas, donde la diferencia de energía obtenida (con respecto al CBM nativo) sólo fue de un -0,2 kcal/mol. Junto con esto, se observó una baja afinidad del dominio de unión cuando éste posee tirosina en lugar del triptófano presente en la posición 2 (variantes 8 y 9).

Además se observó que la mejor afinidad se consiguió manteniendo las tirosinas de las posiciones 28 y 29, es decir, la mejor configuración obtenía corresponde a las variantes 2 y 4, a las cuales se incorporan 2 (F7W, Y28W) y 1 mutación respectivamente (F7W).

Si bien las energías obtenidas concuerdan con comportamientos presentados en los antecedentes bibliográficos, se procede a analizar las energías de unión obtenidas para las mutaciones de la segunda estrategia, con el objetivo de encontrar variantes que presenten mejoras estadísticamente significativas.

En el caso de la segunda estrategia de mutación, se utilizó como base la mejor mutación obtenida previamente, es decir la sustitución de la fenilalanina de la posición 7 por triptófano. Dado esto, todas las mutaciones planteadas (ver Tabla 17) cuentan con esta mutación incorporada.

En la Figura 22 se muestran las energías de interacción obtenidas con Vina para las mutaciones de la segunda estrategia de mutación. En esta estrategia se decidió mutar los aminoácidos ubicados en la zona II de interacción, correspondiente a la nueva zona identificada en este estudio (ver Figura 20).

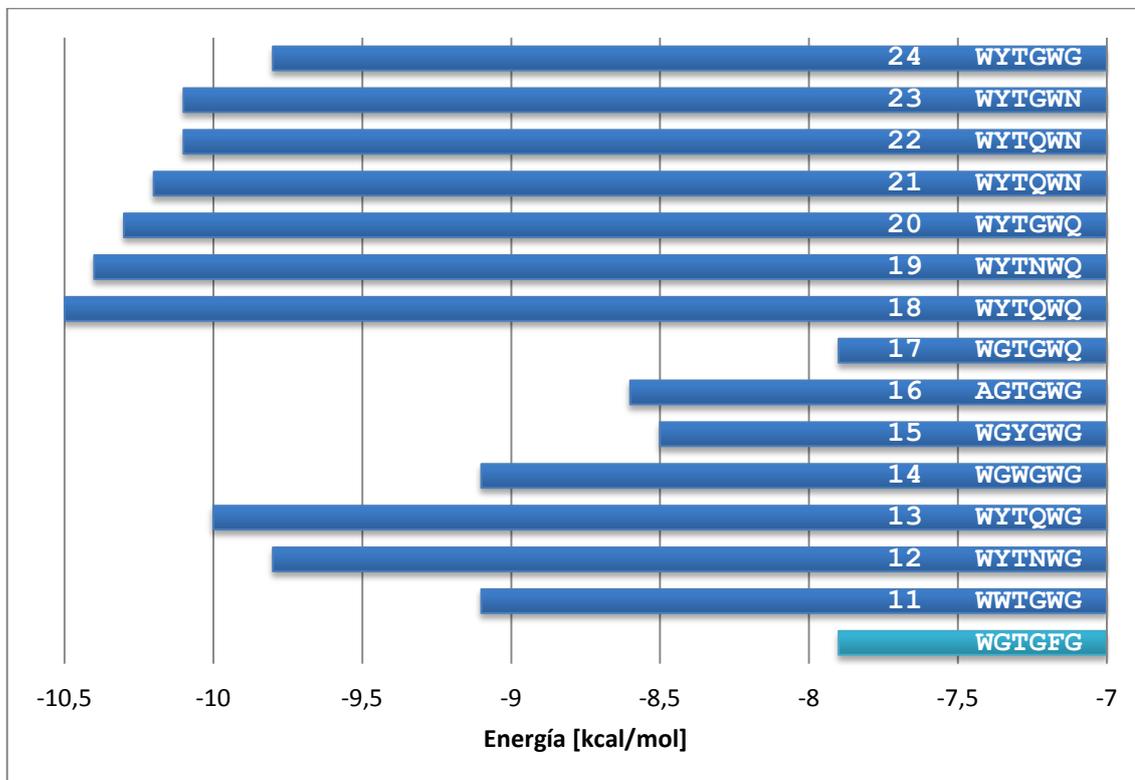


Figura 22. Energías de afinidad obtenidas con Vina para las mutaciones de la segunda estrategia planteada (cambios en los aminoácidos de la nueva zona de interacción identificada: W2, G7, T8, G9, F10, G12). En celeste se muestra la energía de afinidad del CBM de TvEG (nativo).

Al comparar la magnitud de estas energías con las obtenidas en la primera estrategia (Figura 21), se observa que la segunda estrategia generó mutaciones con energía de interacción mayor (energía promedio -9,6 kcal/mol) que las obtenidas previamente en la primera estrategia (energía promedio -8,6 kcal/mol).

En la segunda estrategia de mutación se planteó repetir el arreglo de aminoácidos presente en la zona I de interacción (ver Figura 20). El modelo de interacción fue planteado según antecedentes previos: aminoácidos aromáticos acompañados de aminoácidos que formen puentes de hidrógeno [Beckham et al., 2010 ; Mattinen et al., 1997; Takashima et al., 2007].

Se incorporó tirosina en las posiciones 7 y 8, sustituyendo glicina y treonina respectivamente. La selección de estas posiciones se basó en la cercanía en la cual estos aminoácidos quedan enfrentados a la cadena de celulosa, como resultado del análisis de docking molecular para docking ciego con receptor rígido (Figura 16, caso (a)).

En la Tabla 18 se indica la energía de unión previamente graficadas para la segunda estrategia de mutación (cambios en los aminoácidos de la nueva zona de interacción identificada: W2, G7, T8, G9, F10, G12). En la columna de la derecha se incluyó la diferencia de energía de las mutaciones con respecto al CBM de TVEG (nativo).

Tabla 19. Energías de afinidad entregadas por Vina para las mutantes planteadas en la segunda estrategia (cambios en los aminoácidos de la nueva zona de interacción identificada: W2, G7, T8, G9, F10, G12). En la columna de la derecha se indica la diferencia de energía de cada una de las mutaciones, con respecto al CBM nativo. Se resalta en rojo las mutaciones con mayores mejoras de afinidad sobre celulosa.

	Variante	Energía de Afinidad [Kcal/mol]	Δ Energía de Afinidad [Kcal/mol]
-	WGTGFG	-7,9	-
11	WWTGWG	-9,1	-1,2
12	WYTNWG	-9,8	-1,9
13	WYTQWG	-10	-2,1
14	WGWGWG	-9,1	-1,2
15	WGYGWG	-8,5	-0,6
16	AGTGWG	-8,6	-0,7
17	WGTGWQ	-7,9	0
18	WYTQWQ	-10,5	-2,6
19	WYTNWQ	-10,4	-2,5
20	WYTGWQ	-10,3	-2,4
21	WYTQWN	-10,2	-2,3
22	WYTQWN	-10,1	-2,2
23	WYTGWN	-10,1	-2,2
24	WYTGWG	-9,8	-1,9

En el caso en que se sustituyó la treonina de la posición 8 (variante 15, Tabla 19) se obtuvo una energía de interacción de -8,5 kcal/mol, mientras que en el caso en que no se sustituyó, se obtuvo una energía de -9,1 kcal/mol (variante 4, Tabla 18). La treonina es un aminoácido polar con cadena alifática pequeña; si esta cadena adquiere geometría adecuada podría formar puentes de hidrógeno con los átomos de la cadena principal. Lo anterior sugiere que el rol de treonina de la posición 8 es importante en la conformación adquirida por el CBM y por consiguiente, en la cercanía que éste adquiere con la celulosa.

En la Tabla 19 se observa que al incorporar la tirosina en la posición 7 (reemplazando una glicina), se obtuvo una energía de -9,8 kcal/mol (variante 24), mejor que en el caso en que no se realizó la sustitución (-9,1 kcal/mol, variante 4). La glicina, debido a que su cadena lateral consiste sólo en un hidrógeno, suele tener un rol importante en la conformación de la cadena polipeptídica, ya que puede adoptar ubicación impedida para otros aminoácidos. Dado esto, y pese a la mejora obtenida en la energía de interacción, es necesario analizar la nueva estructura adquirida tras la mutación planteada y determinar si ésta presenta problemas estéricos debido a la sustitución de la glicina.

Junto con la incorporación de un aminoácido aromático a la zona II de interacción, se planteó reemplazar, a uno y simultáneamente los aminoácidos adyacentes por aminoácidos que

puedan formar puentes de hidrógeno, y que de este modo, ayuden a la estabilización de la interacción. Para esto, se planteó distintas combinaciones de glutamina y asparragina en las posiciones 9 y 12 del CBM de TvEG.

Al reemplazar las glicinas presentes en estas posiciones por glutamina y asparragina, se obtuvieron mejoras en la energía de interacción obtenida. En el caso de llevar a cabo dos sustituciones simultáneas en las posiciones 9 y 12 (variantes 18, 19, 21 y 22, Tabla 19), se obtuvieron mejores resultados que al llevar a cabo sólo una sustitución (variantes 12, 13, 17, 20 y 23, Tabla 19). Por otro lado, la mejora obtenida, fue levemente mayor en el caso de incorporar glutamina en vez de asparragina.

Finalmente, y como se observa en la Tabla 19, la combinación energéticamente más favorable se lograría al reemplazar glicina de la posición 7 por tirosina, fenilalanina de la posición 10 por triptófano y glicinas de la posiciones 9 y 12 por glutamina (variante 18).

4.3.3 Selección de mutaciones

La zona de interacción descrita en la literatura corresponde a 3 aminoácidos aromáticos (tirosinas y triptófanos) espaciados de tal manera de coincidir con la periodicidad de los azúcares de la celulosa [Beckham et al., 2010; Linder y Teeri, 1997; Mattinen et al., 1997; Takashima et al., 2007].

En este trabajo los resultados de docking sugieren que existe una segunda zona de interacción, localizada entre los aminoácidos 7 y 12 del péptido. Se plantea entonces repetir el modelo de interacción del péptido, presente en la zona I de interacción, hacia la zona II de interacción, generando de este modo una estrategia de “doble anclaje” a celulosa.

De acuerdo a esto se decide plantear 3 variantes de TvEG a implementar para una posterior validación experimental. Las variantes de TvEG elegidas son mostradas en la Tabla 20.

Tabla 20. Resumen de las tres variaciones de TvEG planteadas. En morado se indican los aminoácidos a mutar en cada uno de los casos seleccionados. En la columna de la derecha se muestra la energía de interacción calculada con Vina.

Variante	Secuencia						Energía de Interacción [Kcal/mol]
	7	8	9	10	11	12	
Nativa	G	T	G	F	S	G	-7,9
Variante 1 WYTGWG	Y	T	G	W	S	G	-9,8
Variante 2 WYTGWQ	Y	T	G	W	S	Q	-10,3
Variante 3 WYTQWQ	Y	T	Q	W	S	Q	-10,5

El objetivo de la elección de éstas consistió en plantear mutaciones para validar experimentalmente lo postulado mediante el análisis de docking molecular. Se planteó evaluar el efecto de la incorporación de un aminoácido aromático, acompañado por aminoácidos (glutamina) que estabilicen su interacción con celulosa.

4.3.4 Estructuras tridimensionales de las mutaciones seleccionadas

Para llevar a cabo el análisis de docking molecular previamente expuesto fue necesario obtener un modelo tridimensional para cada una de las variantes planteadas. Para esto, se utilizó la metodología descrita en la sección 3.2.1, i.

El modelo tridimensional de las variantes seleccionadas se muestran en la Figura 23, en la cual se compara la estructura nativa del CBM de la endoglucanasa de *Trametes versicolor*, con los modelos tridimensionales obtenidos de las variantes. En el caso del péptido nativo (a), se indican los aminoácidos a sustituir en las mutaciones (glicinas de las posiciones 7, 9 y 12 y fenilalanina de la posición 10).

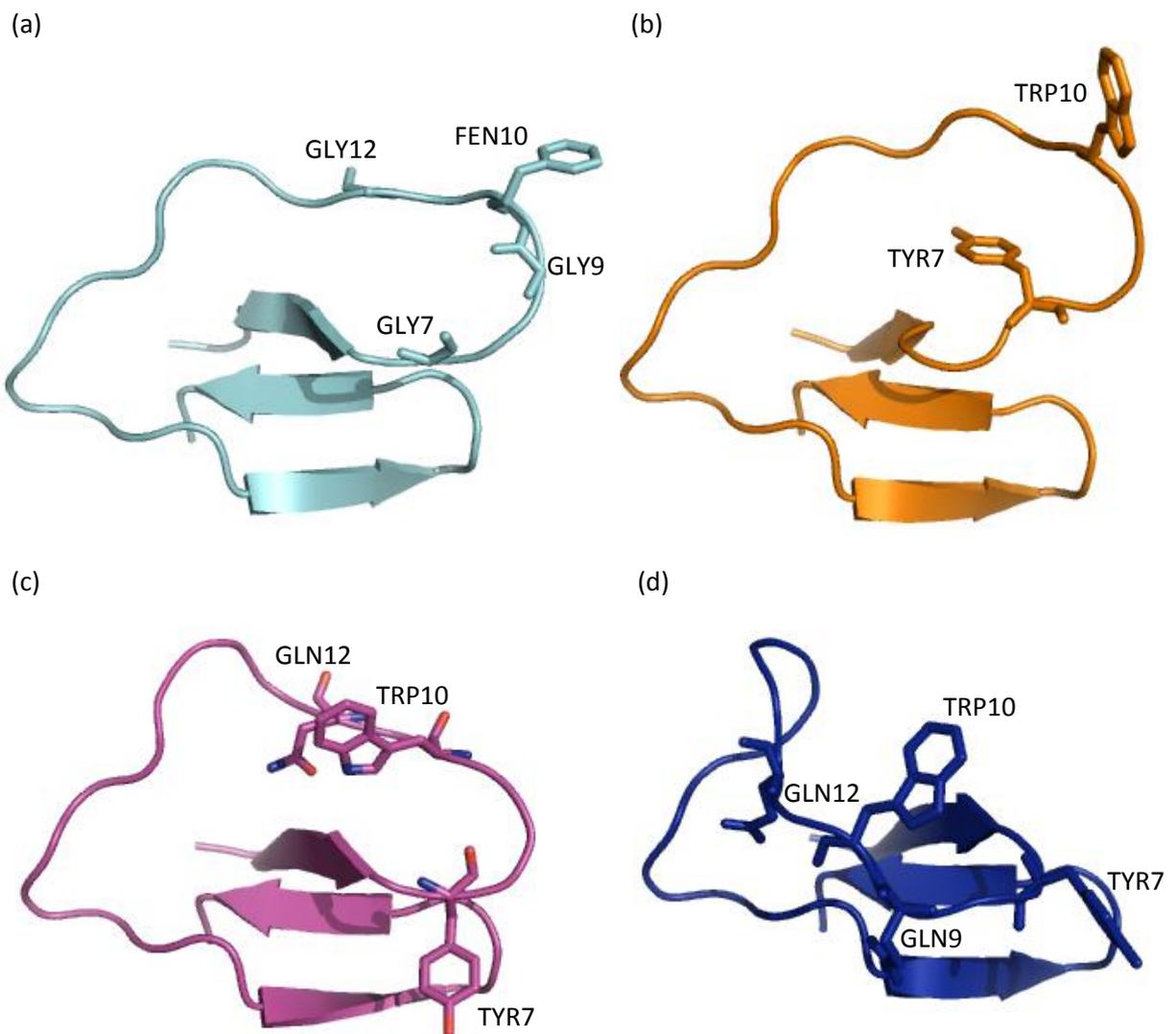


Figura 23. Estructuras tridimensionales de las mutaciones seleccionadas a implementar experimentalmente. (a) CBM de TvEG (nativa), (b) variante 1 (G7Y y F10W), (c) variante 2 (G7Y, F10W y G12Q) y (d) variante 3 (G7Y, G9Q, F10W y G12Q). En el caso del péptido nativo (a) se indican los aminoácidos a mutar. Imágenes generadas con Pymol.

La primera variante se representa en naranja. En este caso, fueron sustituidos dos aminoácidos, glicina de la posición 7 por tirosina y fenilalanina de la posición 10 por triptófano. Esta variante posee 2 mutaciones implementadas: G7Y y F10W.

La segunda variante, representada en púrpura, incorpora las dos sustituciones planteadas previamente en la mutación 1 y además incorpora una glutamina, reemplazando glicina de la posición 12. Esta variante posee 3 mutaciones implementadas: G7Y, F10W y G12Q.

La tercera variante, representada en azul, incorpora las sustituciones de la mutación 2 y además incorpora una segunda glutamina, reemplazando glicina de la posición 9. Esta variante posee 4 mutaciones implementadas: G7Y, G9Q, F10W y G12Q.

Al superponer la estructura de las distintas variantes con la estructura del CBM de TvEG nativo, se puede comparar la divergencia de sus estructuras. En términos cuantitativos es posible analizar sus RMSD⁷ y con ello comparar la divergencia de las estructuras.

En el caso de la variante 1 se obtuvo un RMSD de 1,3 Å, para la variante 2 un RMSD de 1,4 Å y para la variante 3 un RMSD de 2,4 Å, lo cual es consistente con el hecho de que a mayor número de mutaciones introducidas, más se afecta la estructura del péptido.

En la Tabla 21 se resumen las variaciones de CBM de TvEG propuestas a implementar experimentalmente. Estas variantes corresponden a las seleccionadas previamente tras el análisis de docking molecular llevado a cabo mediante el programa Vina.

Tabla 21. Resumen variantes de CBM de TvEG planteadas. Se incluyen los aminoácidos mutados y el RMSD de las estructuras con respecto al CBM de TvEG nativo.

	Mutaciones incorporadas	RMSD
Variante 1	G7Y y F10W	1,3 Å
Variante 2	G7Y, F10W y G12Q	1,4 Å
Variante 3	G7Y, G9Q, F10W y G12Q	2,4 Å

⁷ RMSD: Desviación de la media cuadrática.

Los modelos de las estructuras de las mutaciones, al igual que el de la enzima nativa, fueron evaluados, con el objetivo de determinar la calidad de los modelos obtenidos. La evaluación se llevó a cabo utilizando el servidor Molprobability. Se analizaron los ángulos phi y psi de los residuos, mediante un gráfico de Ramachandran, con el fin de determinar la calidad de los modelos.

En la Figura 24 se muestran los gráficos de Ramachandran de: (a) CBM de TvEG (nativo), (b) Variante 1, (c) Variante 2 y (d) Variante 3.

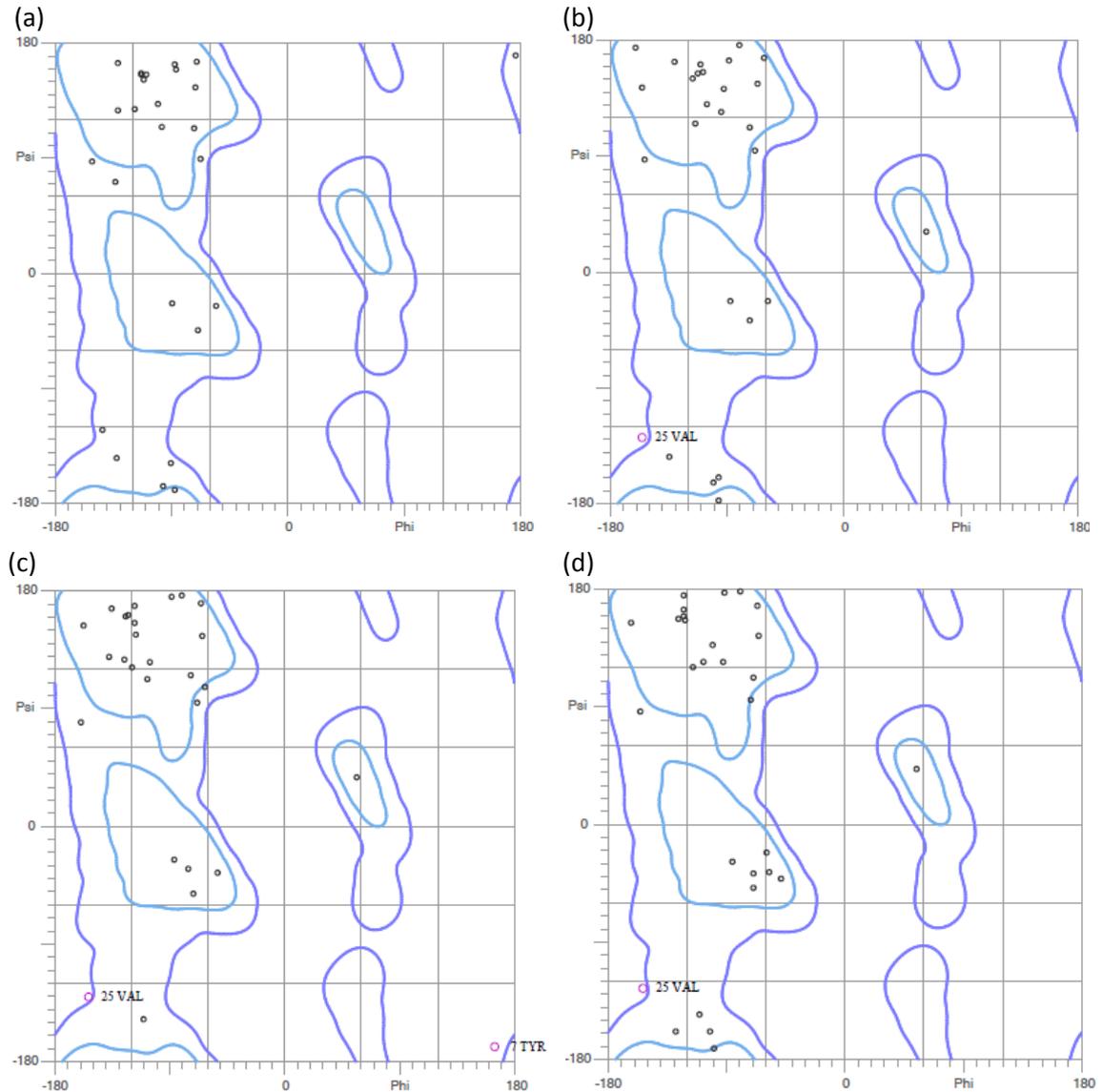


Figura 24. Gráfico de Ramachandran de: (a) la estructura final del CBM de endoglucanasa de *T. versicolor*, (b) estructura de la variante 1 (G7Y y F10W), (c) estructura de la variante 2 (G7Y, F10W y G12Q) y (d) estructura de la variante 3 (G7Y, G9Q, F10W y G12Q). Las regiones delimitadas por líneas celestes corresponden a zonas óptimas, mientras que aquellas delimitadas con una línea azul son zonas permitidas. Las regiones sin delimitación corresponden a zonas no permitidas. Análisis realizado con MolProbability.

Como se observa en los gráficos de la Figura 24 las estructuras obtenidas para las variantes presentan problemas en el residuo 25 (valina). La variante 2, además presenta problemas en el residuo 7 (tirosina), el cual corresponde a uno de los aminoácidos incorporados en las mutaciones planteadas. Estos residuos se encuentran fuera del área permitida de la proteína, lo que indica que probablemente poseen problemas de impedimento estérico.

En la posición 7 de la enzima nativa se encuentra una glicina, la cual se decidió sustituir por tirosina en las tres variantes planteadas. Al sugerir este cambio, era esperable que las estructuras presentasen problemas de impedimento estéricos, ya que se reemplazó un residuo de tamaño pequeño (glicina sólo posee un hidrógeno en su cadena lateral) por un residuo de un tamaño mucho mayor (tirosina posee un grupo fenólico en su cadena lateral). Sin embargo, este problema sólo se presentó en la variante 2, lo que quiere decir que las variantes 1 y 3 adquirieron una conformación que permite el acomodamiento del grupo fenólico de la tirosina, dentro de la cadena polipeptídica.

En la posición 25 se encuentra el residuo valina, el cual presenta problemas en sus ángulos ψ – ϕ en las 3 variantes planteadas. A pesar de que en el caso de la enzima nativa no se presentó problemas en este residuo, en las variantes se encuentra este residuo fuera del área de conformación permitida. En el caso de las variantes, se vio afectada la forma y ubicación del *loop* que une las hojas β 1 y β 2, generando una leve torsión en la hoja β 2 de las estructuras mutadas, lo que probablemente provocó un cambio en la ubicación de este residuo. Si bien este residuo se encuentra con problemas en sus ángulos permitidos, se ubica muy cercano al límite permitido (región delimitada en azul en la Figura 24), por lo que se permitirá trabajar con éste en esa posición.

Para el caso del péptido nativo, se obtuvo que un 74% de sus residuos se ubicaron en la zona favorecida de conformación. Al analizar lo ocurrido con las mutaciones se obtienen porcentajes de 77%, 83% y 80% para las variantes 1, 2 y 3 respectivamente. A partir de esto se concluye que si bien se generaron problemas puntuales con la valina de la posición 25 y tirosina de la posición 7 (en el caso de la mutación 2), en términos generales, y basado en el análisis de los ángulos ψ – ϕ , los modelos de péptidos mutados presentan una mejor calidad de estructura.

4.4 Diseño de la estrategia para la construcción de las variantes de CBM TvEG propuestas

Una vez seleccionadas las 3 variantes de CBM de TvEG a implementar experimentalmente, se procedió según los pasos detallados en la Figura 25.

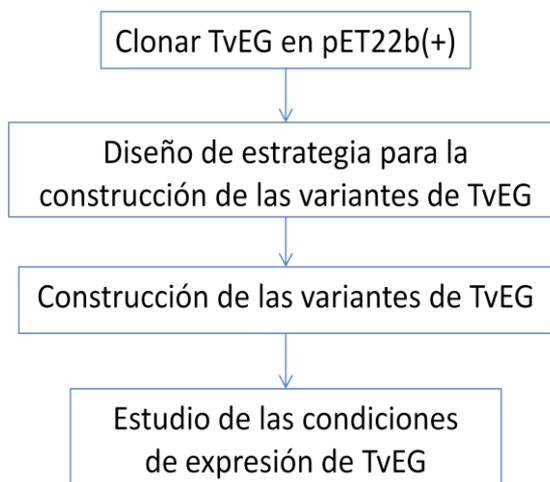


Figura 25. Pasos seguidos en la metodología experimental planteada.

Para la construcción de las distintas variantes fue necesario clonar el gen la enzima nativa (TvEG) en el vector de expresión pET22b(+), ya que se utilizó como base para la mutagénesis sitio dirigida. Una vez clonado el gen en el vector de expresión, se desarrolló una metodología para la construcción de las variantes planteadas. La construcción de las variantes se basó en reacciones de PCR, utilizando la metodología de mutación por extensión de superposición.

Finalmente, se procedió a realizar pruebas de expresión de la enzima nativa (TvEG), con el objetivo de evaluar la factibilidad de obtener la enzima soluble y activamente.

A continuación se presentan cada una de las etapas anteriormente mencionadas con los resultados que se fueron obteniendo en cada una de ellas.

4.4.1 Clonamiento del gen de la endoglucanasa de *T. versicolor* en el vector de expresión pET22b(+)

Una vez planteadas las mutaciones a implementar experimentalmente se procedió a clonar la endoglucanasa nativa al vector de expresión pET22b(+), para posteriormente utilizarla como templado en la construcción de las mutaciones.

La endoglucanasa de *Trametes versicolor* se encuentra clonada en el vector pGem[®]-T Easy. Para su amplificación y posterior ligación a pET22b(+) se requirió llevar a cabo el diseño de partidores.

Diseño de los partidores utilizados en la amplificación del gen de TvEG

Se diseñó un partidor *sense* y uno *antisense* para amplificación del gen la endoglucanasa de *Trametes versicolor* (ver Figura 26).

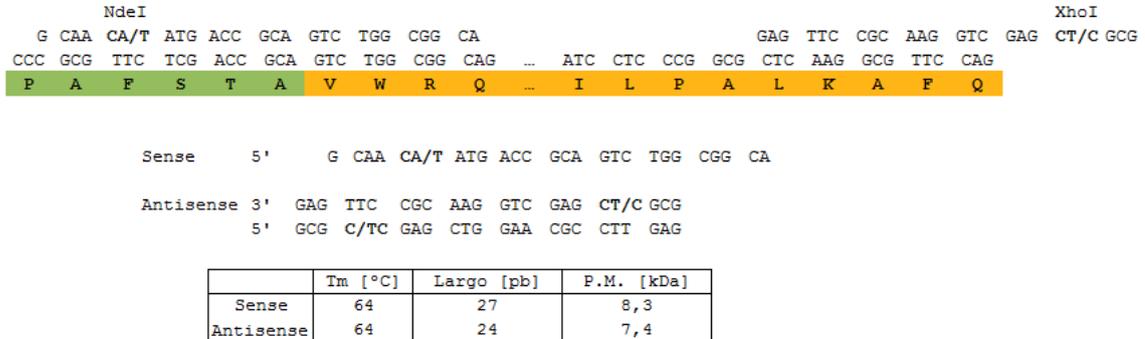


Figura 26. Diseño de partidores para la amplificación del CBM de la endoglucanasa de *T. versicolor*.

En el diseño del partidor *sense* se decidió incorporar 9 pb correspondientes a 3 aminoácidos conservados en los dominios de unión a carbohidrato de la familia I y no secuenciados al momento de obtener el gen de la enzima en estudio. Se consideró que lo más probable es que estos aminoácidos estén presentes también en este dominio y no hayan sido clonados previamente debido a algún error experimental en la obtención del gen. Dado esto, se incorporaron en la secuencia 3 aminoácidos conservados.

Se agregó también un sitio de reconocimiento para la enzima de restricción *XhoI* en el partidor *antisense* (región 3' de la secuencia codificante), mientras que en el partidor *sense* se incorporó un sitio de reconocimiento para la enzima de restricción *NdeI* (región 5' de la secuencia codificante). La elección de los sitios de reconocimiento a incorporar se basó en que estos estuviesen presentes en el vector de expresión, y que permitiesen clonar la enzima en un lugar tal que la producción de ésta tuviese unida en su extremo 3', una cola de histidina, para facilitar su purificación, y que la endoglucanasa se expresase en el compartimiento intracelular de la célula.

Amplificación y clonamiento de *tveg* en *pET22b(+)*

El gen de la endoglucanasa fue modificado con el fin de introducir los sitios de reconocimiento para las enzimas de restricción *NdeI* y *XhoI*, de manera de permitir su posterior clonamiento al vector de expresión *pET22b(+)*.

La incorporación se llevó a cabo mediante la técnica de PCR, para la cual se usó como templado el DNA plasmidial de la endoglucanasa nativa y como partidores *TVEG_PET_For* y *TVEG_PET_Rev*. El resultado del PCR llevado a cabo se visualizó en un gel de agarosa al 1% (Figura 27)

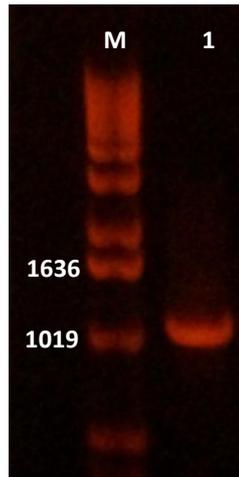


Figura 27. Electroforesis en gel de agarosa al 1% del gen de la endoglucanasa de *T. versicolor*, amplificado en la reacción de PCR. En esta reacción se utilizó como templado Tveg_pGEM y los partidores específicos para *T. versicolor*. M: marcador de peso molecular 1kb (Invitrogen).

En la Figura 27 se muestra el resultado de la amplificación del gen de la endoglucanasa. Como se observa en la figura, al utilizar una temperatura de amplificación (64°C) cercana a la teórica (66°C), se obtuvo una buena cantidad de DNA, lo que verifica el buen diseño de los partidores que amplifican el gen. Además al llevar a cabo la reacción de PCR y su posterior electroforesis en gel de agarosa, permite validar el correcto tamaño del gen (1179 pb).

Luego de generar el fragmento que contenía el gen de la endoglucanasa de *T. versicolor* modificada, fue ligado al vector de clonamiento pGem®-T Easy, para posteriormente electroporar *E. coli* DH5α. Estas bacterias fueron crecidas en placas LB-Agar con Amp, X-Gal e IPTG. De las colonias identificadas como positivas (blancas), se seleccionó una, la cual fue crecida para su posterior análisis y verificación de la correcta incorporación de los sitios de restricción.

La colonia identificada como positiva, se creció para luego extraerle el DNA plasmidial, el cual fue digerido con las enzimas de restricción *XhoI* y *NdeI*. En paralelo con la digestión del vector pGem®-T Easy que contenía ligada la endoglucanasa de *T. versicolor*, se digirió con las mismas enzimas de restricción el vector de expresión pET22b(+), con el fin de linealizarlo. Las digestiones realizadas para cada caso, las que se muestran en la Figura 28.

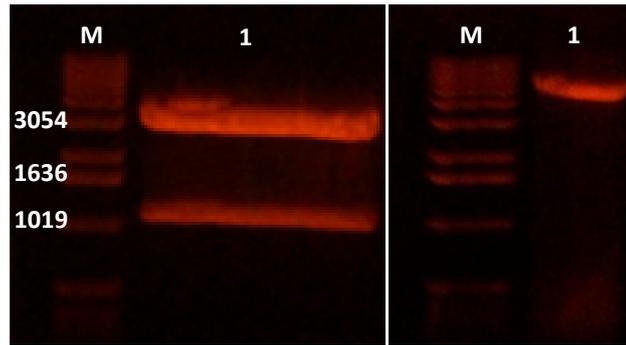


Figura 28. Electroforesis en gel de agarosa al 1% de la digestión de tveg_pGEM y del vector de expresión. A la izquierda, digestión del vector pGem[®]-T Easy que contiene el gen de la endoglucanasa de *T. versicolor*. A la derecha, digestión del vector de expresión pET22b(+). M: marcador de peso molecular 1kb (Invitrogen).

Se observan 2 bandas, donde la banda superior corresponde al vector de clonamiento pGem[®]-T Easy de 3015 pb y la inferior al gen de la endoglucanasa (1179 pb). A la derecha se observa 1 sola banda, correspondiente al vector de expresión pET22b(+) linealizado (5493 pb). Se purificaron las bandas correspondientes al gen de la endoglucanasa de *T. versicolor* y al vector pET22b(+) linealizado, para luego ligarlos.

Se realizó una electroforesis en gel de agarosa al 1% (Figura 29), con el fin de cuantificar el DNA purificado y de este modo determinar la composición de la mezcla de ligación a utilizar. En el segundo y tercer carril (de izquierda a derecha) se cargaron 1 μ L y 2 μ L de High mass ladder, el cual a partir de la intensidad de sus bandas permite estimar la cantidad de DNA contenida en un volumen dado. En este caso, se concluye que para el caso de la endoglucanasa se poseen 40 ng/ μ L y para el vector linealizado se poseen 30 ng/ μ L.

A modo de asegurar la ligación se emplean 120 ng de vector linealizado, y en base al largo del inserto, se decide agregar 80 ng de éste (para mantener una relación 3:1 de los nanogramos normalizados por el largo del segmento).

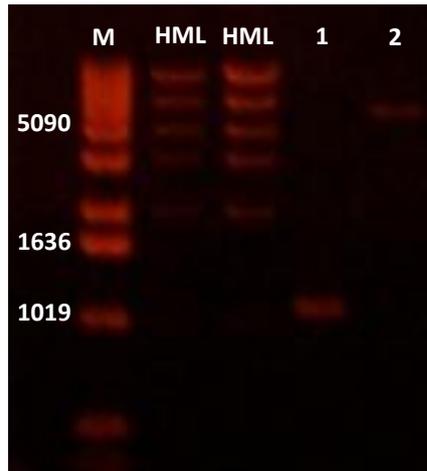


Figura 29. Cuantificación del DNA correspondiente al gen de la endoglucanasa de *T. versicolor* y al vector de expresión linealizado. M: marcador de peso molecular 1kb (Invitrogen). HML: High mass ladder (invitrogen). Carril 1: endoglucanasa de *T. versicolor*. Carril 2: pET22b(+) linealizado.

En base a la cuantificación anterior se llevó a cabo la ligación del gen de la endoglucanasa de *T. versicolor* al vector de expresión pET22b(+). La mezcla de ligación se usó para electroporar células *E. coli* DH5 α , las que fueron crecidas en placas LB Agar con Amp. A las 9 colonias que crecieron se les realizó PCR de colonias para verificar la presencia del inserto.

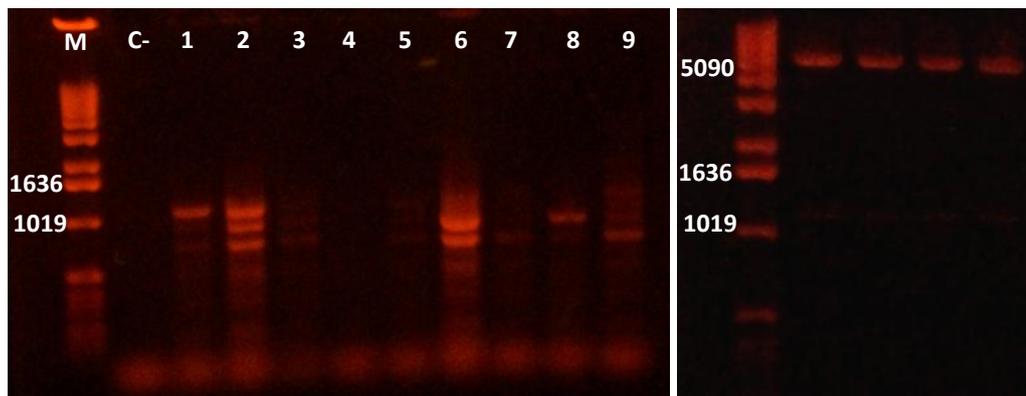


Figura 30. Verificación del clonamiento de Tveg a pET22b(+). A la izquierda, PCR de colonias. A la derecha digestión del DNA plasmidial de colonias positivas en el PCR. M: marcador de peso molecular 1kb (Invitrogen).

De las colonias analizadas mediante la técnica de PCR de colonias, se seleccionaron las correspondientes a las 1, 2, 5 y 9 como positivas. Éstas se crecieron para luego extraerles el DNA plasmidial, el cual se digirió con el fin de verificar el correcto clonamiento de la enzima. Como se observa en la Figura 30, las cuatro colonias seleccionadas tras el PCR de colonias, resultaron ser positivas.

Los plasmidios obtenidos se mandaron a secuenciar a Macrogen para verificar que las construcciones estuviesen correctas tanto en tamaño como en secuencia nucleotídica. Tras el análisis de las secuencias obtenidas, se concluye que se consiguió clonar con éxito la endoglucanasa de *Trametes versicolor*, en el plasmidio de expresión pET22b(+).

4.4.2 Estrategia desarrollada para la construcción de las variantes de TveG

La construcción de las mutaciones seleccionadas a partir del análisis *in silico* requirió desarrollar una estrategia para la construcción de éstas. Los criterios tomados en cuenta para el diseño de la estrategia fue minimizar el número de partidores requeridos, minimizar el número de reacciones de PCR para la construcción de las mutaciones y simultáneamente evitar el diseño de partidores que contuviesen sectores con muchas bases a mutar. En base a eso, se desarrolló la estrategia representada en la Figura 31.

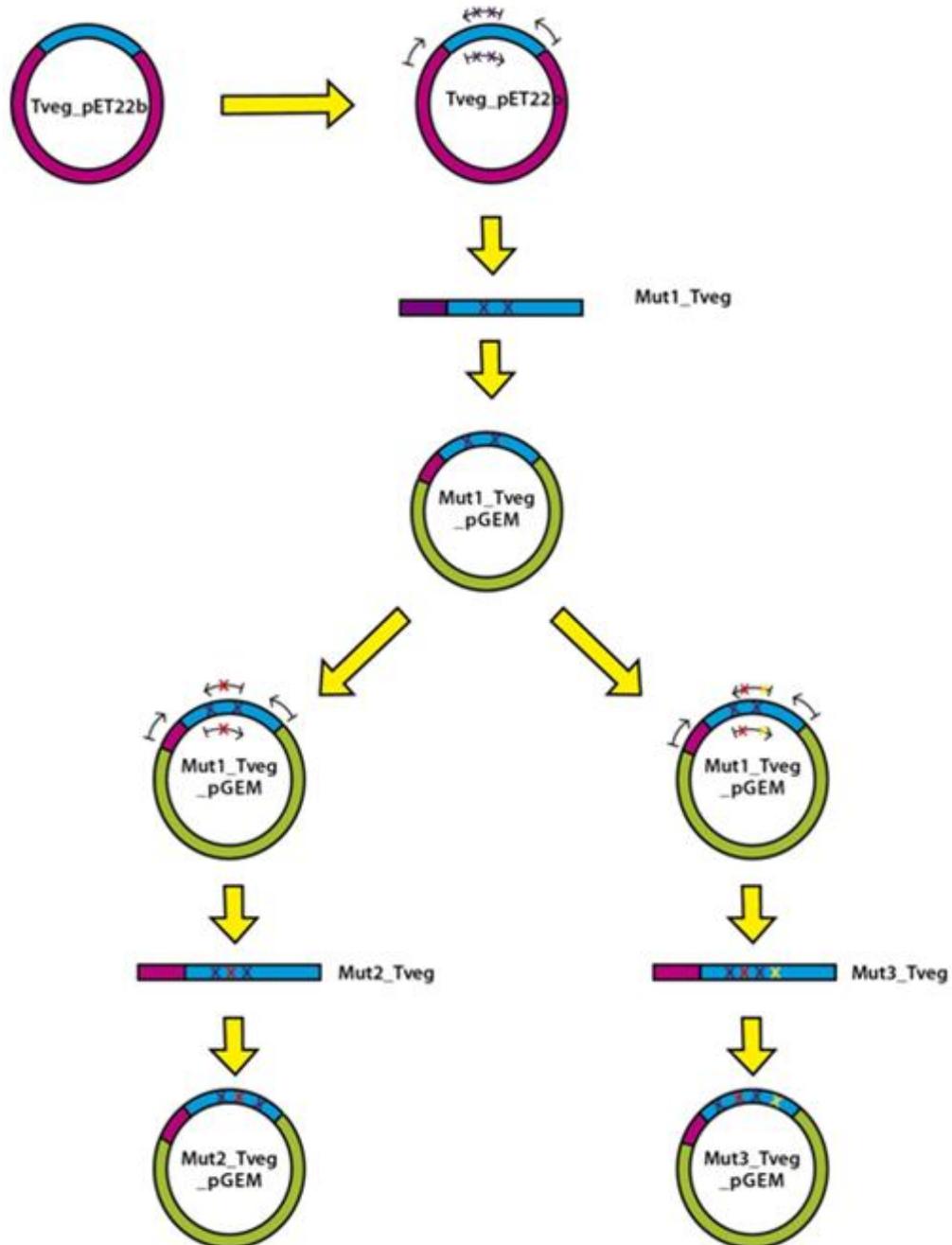


Figura 31. Estrategia desarrollada para la construcción de las variantes de TveG. La variante 1 se llevó a cabo utilizando como templado el gen de la enzima, previamente clonado en pET22b(+). A su vez, la variante 1 sirvió de templado para la generación de las mutaciones 2 y 3.

La estrategia desarrollada utiliza 3 pares de partidores mutagénicos, 1 par de partidores externos y 3 reacciones de PCR.

El partidor antisense utilizado como partidor externo corresponde al que amplifica desde el extremo terminal del gen de la endoglucanasa. El partidor sense empleado corresponde a un partidor auxiliar que alinea con pET22b(+) 500 pb río arriba promotor T7. La utilización de este partidor, en vez de aquel que amplifica desde el extremo inicial de la endoglucanasa, se debe a que las mutaciones planteadas se encuentran muy cerca (30 pb) del inicio del gen, lo cual generaría un fragmento mutado de tan sólo 30 pb, haciendo inviable su visualización en geles de agarosa y su posterior purificación.

La primera reacción de PCR utiliza como templado la enzima nativa clonada en pET22b(+). Para la amplificación emplea los partidores que introducen las mutaciones en el sector del gen que codifica para los aminoácidos 7 y 10. Para la extensión final del fragmento amplificado se utilizó los partidores externos. El fragmento generado mediante la mutación por extensión de superposición se denomina Mut1_Tveg, y consiste en 500 pb del vector pET22b(+) unido al gen que codifica para la endoglucanasa mutada. Este fragmento se liga al vector de clonamiento pGEM y el producto de esta ligación fue denominado Mut1_Tveg_pGEM.

Si bien el constructo generado incluye la amplificación de un fragmento del vector de expresión unido al gen de la endoglucanasa, esto no genera posteriores problemas para el clonamiento del gen en el vector de expresión, ya que la liberación del inserto se lleva a cabo mediante la digestión con enzimas de restricciones cuyos sitios de reconocimiento se encuentran en ambos extremos del gen de la endoglucanasa.

El constructo obtenido previamente (Mut1_Tveg_pGEM) se utiliza como templado para las reacciones de PCR de la variante 2 y 3. Para la construcción de la variante 2 se emplean los partidores que introducen las mutaciones en el sector del gen que codifica para el aminoácido 9. Análogamente, para el caso de la variante 3 se utilizan partidores que introducen mutaciones en los aminoácidos 9 y 12. Los fragmentos generados se denominaron Mut2_Tveg y Mut3_Tveg y se ligan al vector de clonamiento pGEM.

Con la estrategia planteada en las variantes 2 y 3 se incorporan 1 y 2 aminoácidos mutados con respecto a la variante 1, mientras que con respecto a la enzima nativa, se consigue incorporar 3 y 4 aminoácidos mutados.

La incorporación simultánea de los 4 aminoácidos mutados en un mismo par de partidores resulta muy difícil de implementar. La estrategia desarrollada, plantea una alternativa, en la que sin requerir un mayor número de partidores o reacciones de PCR se consigue construir las mutaciones planteadas. Esta estrategia tiene la ventaja de ser mucho más sencilla de implementar en términos experimentales, sin embargo tiene la desventaja de que las mutaciones deben realizarse en serie, lo cual requiere de más tiempo y cualquier problema presentado en la obtención de la primera mutación, afecta la construcción de las otras mutaciones.

Una vez desarrollada la estrategia para la obtención de las variantes, se procedió a diseñar los partidores que introducen las mutaciones. Para el diseño de éstos se tomó en cuenta la secuencia nativa de la enzima, de manera tal que los partidores tuviesen una secuencia complementaria de entre 30 y 36 pb con la enzima nativa.

En el diseño de los partidores (ver Figura 32) también se tomó en cuenta que la Tm de ambos fuese similar, se minimizó el alineamientos entre ellos y la formación de *hairpin* internos en un mismo partidor.

```

Sense 1      5'      CA CCA ATA TGG GTT ACC AGA GTT CTG TGA TGG ACA C
Antisense 1  5'      CA GGT AGT GTC TTG AGA CCA TTG GGT ATA ACC ACA C

Sense 2      5'      TGG CCG ACC AGA GTT CTG TGA TGG ACA CG
Antisense 2  5'      GC ACA GGT AGT GTC TTG AGA CCA GCC GGT

Sense 3      5'      CC GTC ACA CCA ATA TGG CCG ACC AGA CCA CTG TGA
Antisense 3  5'      AGT GTC ACC AGA CCA GCC GGT ATA ACC ACA CTG CC
    
```

	Tm [°C]	Largo [pb]	P.M. [kDa]	Hairpin	Complementary	Self-annealing
Sense 1	66	36	11,1	No	No	2
Antisense 1	66	36	11,9	Si	No	2
Sense 2	66	29	9	No	No	No
Antisense 2	66	29	9	No	No	No
Sense 3	69	35	10,7	Si	No	2
Antisense 3	69	35	10,7	Si	No	2

Figura 32. Diseño de partidores utilizados en la incorporación de las mutaciones planteadas para las 3 variantes de TvEG propuestas.

4.5 Construcción de las variantes de TvEG planteadas

4.5.1 Obtención de los fragmentos mutados

La construcción de las variantes planteadas se lleva a cabo con la implementación de la estrategia desarrollada. En la técnica de mutación por extensión de superposiciones, es necesario generar los fragmentos mutados que serán superpuestos. En este caso los fragmentos a superponer se nombraron A1, A2, A3 (de largo 556 pb) y B1, B2 y B3 (de largo 1162 pb).

Si bien la construcción de las variantes no se llevó a cabo en paralelo (ya que la variante 1 sirve de templado para la obtención de las otras dos), en la Figura 33 se muestra un resumen de los fragmentos generados.

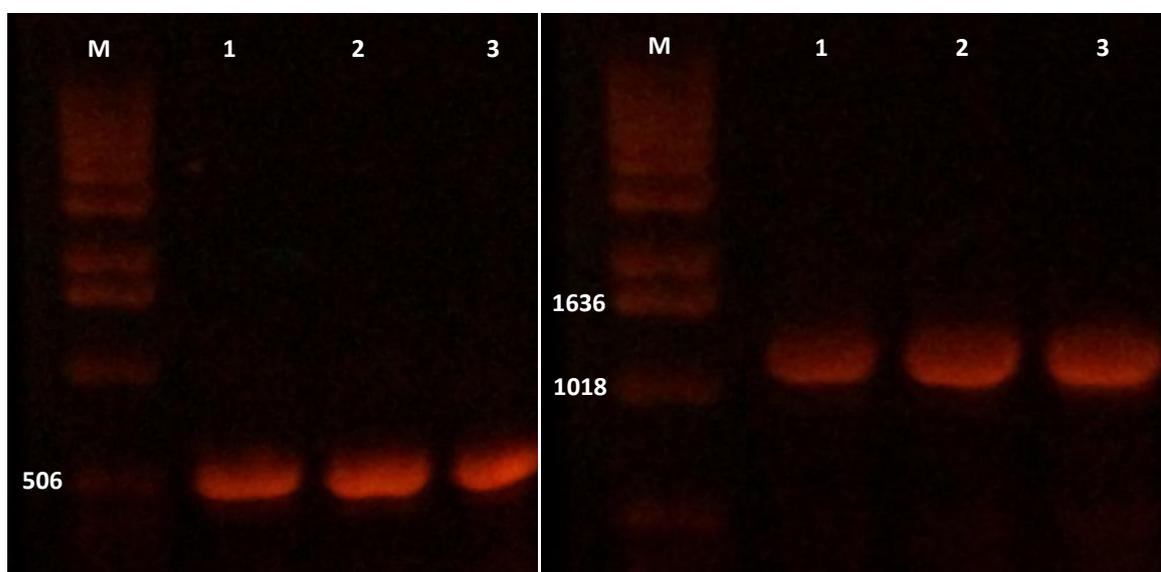


Figura 33. Electroforesis en gel de los fragmentos mutados. A la izquierda, se muestran los fragmentos A1, A2 y A3 generados. A la derecha se muestran los fragmentos B1, B2 y B3 generados. M: marcador de peso molecular 1kb (Invitrogen).

Como se observa en la Figura 33, los fragmentos obtenidos poseen tamaños según lo esperado (556 pb y 1162 pb). Dado esto, fueron superpuestos utilizando la técnica de PCR y utilizando partidores externos. Al producto de estas amplificaciones se les agregó extremos de adenina, para su posterior ligación a pGem[®]-T Easy.

4.5.2 Ligación de los fragmentos mutados a pGem[®]-T Easy

Luego de generar los fragmentos que contenían el gen de la endoglucanasa de *T. versicolor* mutado, fueron ligados al vector de clonamiento pGem[®]-T Easy, para posteriormente electroporar *E. coli* DH5 α . Estas bacterias fueron crecidas en placas LB-Agar con Amp, X-Gal e IPTG. De las colonias identificadas como positivas (aquellas blancas), se analizaron dos por cada mutación (mediante PCR de colonias), a modo de validar la presencia de la endoglucanasa mutada de *T. versicolor*.

De las colonias analizadas mediante la técnica de PCR de colonias, todas fueron identificadas como positivas, dado esto, se crecieron para luego extraerles el DNA plasmidial, el cual se digirió con el fin de verificar el correcto clonamiento de la enzima mutada. Como se observa en la Figura 34, las colonias 2, 3 y 5 seleccionadas tras el PCR de colonias, resultaron ser positivas.

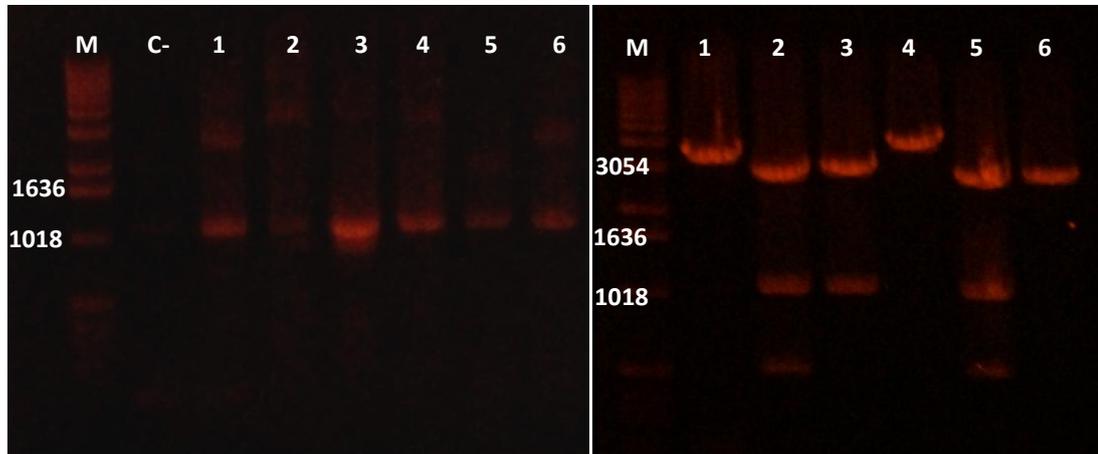


Figura 34. Verificación de la ligación de las variantes del gen de la endoglucanasa de *T. versicolor* al vector pGem[®]-T Easy. A la izquierda PCR de colonias. A la derecha digestión del gen de la endoglucanasa mutada clonada en el vector pGem[®]-T Easy. En ambos casos los carriles 1 y 2 corresponden a colonias de la mutación 1, los carriles 3 y 4 a colonias de la mutación 2 y los carriles 5 y 6 a colonias de la mutación 3. M: marcador de peso molecular 1kb (Invitrogen).

Los plasmidios obtenidos se mandaron a secuenciar a Macrogen para verificar que las construcciones estuviesen correctas tanto en tamaño como en secuencia nucleotídica. Tras el análisis de las secuencias obtenidas, se concluye que se consiguió clonar con éxito las tres mutaciones de endoglucanasa de *Trametes versicolor*, en el plasmidio de clonamiento pGem[®]-T Easy (ver secuencias Anexo C).

4.6 Estudio de las condiciones de expresión de la endoglucanasa de *T. versicolor* en *E. coli* BL21 (DE3)

Los sistemas bacterianos no suelen ser eficientes para expresar celulasas de hongos debido a que la formación de puentes disulfuro no se lleva a cabo y el plegamiento no es correcto, o porque las proteínas quedan atrapadas formando cuerpos de inclusión [Martínez-Anaya et al., 2008]. Si bien no se encuentra dentro de los objetivos del presente trabajo llevar a cabo la expresión recombinante de la endoglucanasa de *Trametes versicolor* en *E. coli*, se realizaron experimentos para evaluar el efecto de la temperatura en la producción soluble de la enzima, de modo de dejar propuesto condiciones de expresión para una futura producción de la endoglucanasa en *E. coli*.

4.6.1 Análisis de expresión a distintas temperaturas de inducción

La cepa escogida para el análisis de la expresión fue *E. coli* BL21(DE3), la cual se cultivó en el medio TB. La metodología de inducción utilizada corresponde a la detallada previamente en la sección 3.2.2, iv.

Una vez inducida la producción de la proteína, se recuperó la endoglucanasa desde el espacio intracelular, separando las fracciones solubles e insolubles y analizándolas mediante un gel de poliacrilamida de 10% (Figura 35).

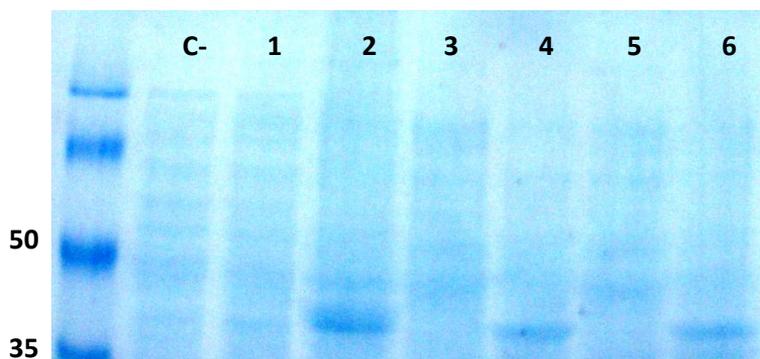


Figura 35. Gel de poliacrilamida al 10%. Expresión de la endoglucanasa de *T. versicolor* a tres temperaturas distintas. Carril 1: inducción a 37°C, fracción insoluble. Carril 2: inducción a 37°C, fracción soluble. Carril 3: inducción a 25°C, fracción insoluble. Carril 4: inducción a 25°C, fracción soluble. Carril 5: inducción a 18°C, fracción insoluble. Carril 6: inducción a 18°C, fracción soluble.

Como se puede ver en la Figura 35, se observa una banda correspondiente al tamaño esperado de la endoglucanasa (38 kDa). Esta se ve mucho más intensa en la fracción insoluble que en la soluble, lo cual sugiere la formación de cuerpos de inclusión. Por otro lado, se observa que a medida que se disminuye la temperatura de inducción, disminuye la cantidad de proteína insoluble. Esto puede deberse a que al inducir el cultivo a una temperatura menor, se esté produciendo menor cantidad de proteína, pero también puede deberse a que se produzca menos proteína insoluble y más proteína soluble. Dado que no se nota diferencia entre la intensidad de las bandas en la fracción soluble, se lleva a cabo un zimograma a modo de evaluar si existe actividad en las fracciones solubles.

Al analizar las bandas obtenidas en el zimograma se observa que la mayor actividad ocurre en los casos en que la inducción se lleva a cabo a 37°C y a 18°C. Al realizar la inducción a una menor temperatura se produce una menor cantidad de proteína total, sin embargo, la proteína soluble producida presenta mayor actividad, lo cual indicaría que si bien la proteína total disminuye, la proteína soluble aumenta.

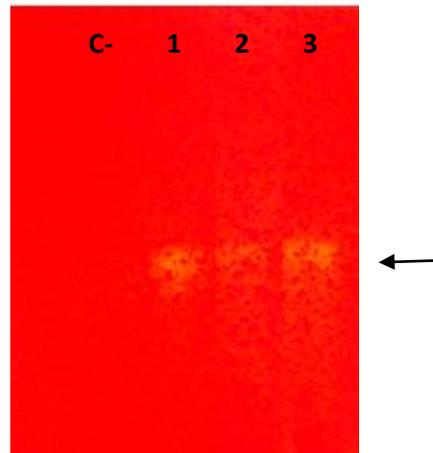


Figura 36. Zimograma de las fracciones solubles inducidas a diferentes temperaturas. Carril 1: fracción soluble obtenida tras la inducción a 37°C. Carril 2: fracción soluble obtenida tras la inducción a 25°C. Carril 3: fracción soluble obtenida tras la inducción a 18°C.

Junto con el zimograma a diferentes temperaturas, se llevó a cabo otro en que se comparó la actividad de la fracción soluble con la fracción insoluble, en una inducción a 37°C. Se obtuvo que pese a que la fracción insoluble había sido denaturada (con urea 8M), presentó una mayor actividad con respecto a la proteína producida solublemente. A partir de este resultado se propone que la cantidad de proteína insoluble es mucho mayor que la proteína soluble. Además, el hecho de que la fracción insoluble presente actividad, pese a las condiciones denaturantes a la cual fue sometida la muestra, indicaría que la proteína producida en cuerpos de inclusión es de fácil recuperación a una forma activa, ya que sólo con el paso de lavado del gel del zimograma, con triton X-100, se consiguió renaturarla. Se propone entonces, para una posterior expresión de la enzima, evaluar la posibilidad de producir grandes cantidades de proteína insoluble y renaturarla.

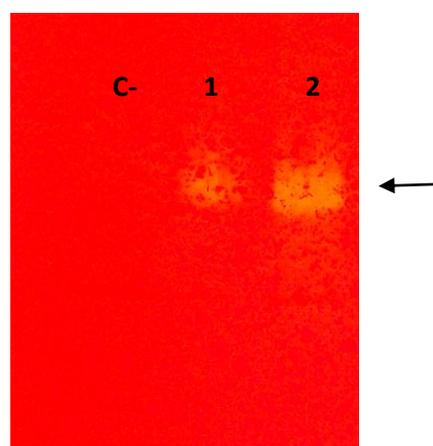


Figura 37. Zimograma tras la inducción a 37°C. Carril 1: fracción soluble. Carril 2: fracción insoluble.

5. CONCLUSIONES

Fue posible obtener un modelo de la estructura tridimensional del CBM de la endoglucanasa de *T. versicolor*. Este modelo se refinó utilizando herramientas de Modeller, sin embargo se encontraron problemas conformacionales en algunos residuos: los ARG3, ASN26 y GLN31 resultaron ser rotámetros pobres, mientras que THR14 y SER30 mostraron problemas en sus ángulos. Ninguno de estos residuos se ha identificado como aminoácidos conservados dentro de la familia I, responsables de la unión a carbohidrato. Además, los residuos con problemas se ubican en las hojas β y no parecen afectar la conformación adquirida por el péptido. Se considera entonces que el modelo obtenido es apropiado para los posteriores análisis de docking molecular.

El modelo tridimensional obtenido para el CBM presenta una estructura formada por tres hojas beta antiparalelas, las que se estabilizan por la formación de dos puentes disulfuro entre las cisteínas 16 y 32 y las cisteínas 4 y 22. De acuerdo al modelo esta conformación permite que los anillos aromáticos de los aminoácidos TRP2, TYR28 y TYR29 se ubiquen distanciados según el espaciamiento de los anillos de glucosa en la celulosa.

Los anillos aromáticos de los aminoácidos TRP2, TYR28 y TYR29 no se orientan en un mismo plano, sin embargo, al llevar a cabo docking molecular permitiendo la movilidad de las cadenas laterales de éstos, ocurre un acomodamiento de los residuos, de manera tal que se orientan sobre los azúcares de la celulosa. Por esto, se sugiere que la celulosa provoca cambios conformacionales sobre el CBM de TvEG y que estos pueden ser predichos utilizando docking molecular.

A partir del análisis de docking molecular se concluye que fue posible identificar dos zonas de interacción a celulosa presentes en el CBM de TvEG. El análisis de la energía de interacción sugiere que ambas son favorables, presentando magnitudes de -7,5 kcal/mol y -7,9 kcal/mol respectivamente. La primera zona de interacción corresponde a la zona de interacción descrita para otros dominios de unión a carbohidratos de la familia I, mientras que la segunda corresponde a una nueva zona de interacción no descrita previamente en la literatura.

Se plantearon 24 variantes, 10 para la zona ya descrita y 14 para la nueva zona identificada. Usando como criterio de selección la energía de unión entregada por el programa Vina, se comparó la interacción de las 24 variantes con la celulosa. Las variantes propuestas para la zona I no mejoran la interacción a celulosa. Se concluye que el arreglo de aminoácidos conservados, presentes en el CBM de TvEG silvestre es efectivamente, una combinación que optimizada naturalmente la unión del dominio a celulosa.

Se planteó introducir mutaciones en la zona II de interacción, copiando el arreglo de aminoácidos presentes en la zona I. Las variantes evaluada con cambios en la zona II presentaron un aumento en la energía de interacción a celulosa con respecto a la secuencia silvestre, lo que sugiere que sugiere que una zona con potencial de optimización. . De las 14 variantes planteadas en la zona II de interacción se seleccionaron 3 para la posterior validación experimental. Las variantes seleccionadas fueron: variante 1 (G7Y, F10W), variante 2 (G7Y, F10W, G12Q) y variante 3 (G7Y, G9Q, F10W, G12Q).

Con respecto al desarrollo experimental, fue posible clonar la endoglucanasa de *T. versicolor* en el vector de expresión pET22b(+). La expresión de TvEG se llevó a cabo en *E. coli* BL21 (DE3). Si bien los sistemas bacterianos no suelen ser eficientes para expresar celulasas de hongos, en el caso de TvEG se concluye que, bajo las condiciones de expresión estudiadas es posible producir la proteína de manera intracelular, soluble y activa. Ésto no sólo indica que es posible la validación experimental de las variantes propuestas sino que también plantea condiciones bajo las cuales es posible producir una enzima de origen fúngico (TvEG) en un sistema bacteriano.

Para la construcción de las variantes planteadas fue necesario diseñar una estrategia de mutación. La estrategia de mutagénesis sitio dirigida se basó en la técnica de mutación por extensión de superposición. Esta estrategia fue efectiva y los constructos de las distintas variantes se clonaron con éxito en el vector de expresión pGem[®]-T Easy sin incorporación de mutaciones adicionales no planificadas.

Mediante el estudio realizado se concluye que fue posible proponer mutaciones en la endo β 1,4 glucanasa de *Trametes versicolor* que potencialmente mejorarán su adsorción sobre celulosa. Si bien estas mutaciones deben ser validadas experimentalmente, fueron planteadas a partir de un diseño racional lo que aumentaría la probabilidad que éstas presenten mayor afinidad a celulosa. Se espera que estas variantes de la enzima TvEG mejoren no sólo su adsorción a celulosa, sino también el rendimiento de hidrólisis de ésta. Esto podría presentar una mejora en el proceso de producción de bioetanol, ya que al producir enzimas con mejores rendimientos podría disminuirse los costos para producir la hidrólisis de la celulosa.

6. BIBLIOGRAFÍA

BALAT, M., BALAT, H. y ÖZ, C. 2008. Progress in bioethanol processing. *Progress in Energy and Combustion Science*, vol. 34(5): 551-573.

BECKHAM, G.T., MATTHEWS, J.F., BOMBLE, Y.J., BU, L., ADNEY, W.S., HIMMEL, M.E., NIMLOS, M.R. y CROWLEY, M.F. 2010. Identification of amino acids responsible for processivity in a family 1 carbohydrate-binding module from a fungal cellulase. *The Journal of Physical Chemistry B*, vol. 114(3): 1447-1453.

BÉGUIN, P. y AUBERT, J.P. 1994. The biological degradation of cellulose. *FEMS Microbiology Reviews*, vol. 13(1): 25-58.

BERMAN, H.M., WESTBROOK, J., FENG, Z., GILLILAND, G., BHAT, T.N., WEISSIG, H., SHINDYALOV, I.N. y BOURNE, P.E. 2000. The protein data bank. *Nucleic Acids Research*, vol. 28(1): 235-242.

BERNSTEIN, F.C., KOETZLE, T.F., WILLIAMS, G.J.B., MEYER, E.F., BRICE, M.D., RODGERS, J.R., KENNARD, O., SHIMANOUCI, T. y TASUMI, M. 1977. The protein data bank: A computer-based archival file for macromolecular structures*. *Journal of molecular biology*, vol. 112(3): 535-542.

BHAT, M.K. y HAZLEWOOD, G.P. 2001. Enzymology and other characteristics of cellulases and xylanases. En M.R. BEDFORD y G.G. PARTRIDGE. *Enzymes in farm animal nutrition*. Oxon, UK: CABI Publishing. p. 11-60.

BNDES y CGEE. 2008. Bioetanol de caña de azúcar : energía para el desarrollo sostenible [online]. [Rio de Janeiro, Brasil]: BNDES. Disponible desde Internet:<<http://www.bioetanoldecanadeazucar.org/es/download/bioetanol.pdf>>.

BORNSCHEUER, U.T. y POHL, M. 2001. Improved biocatalysts by directed evolution and rational protein design. *Current opinion in chemical biology*, vol. 5(2): 137-143.

BRAY, M.R., JOHNSON, P.E., GILKES, N.R., MCINTOSH, L.P., KILBURN, D.G. y WARREN, R.A. 1996. Probing the role of tryptophan residues in a cellulose-binding domain by chemical modification. *Protein science*, vol. 5(11): 2311-2318.

CAMACHO, C.J., GATCHELL, D.W., KIMURA, S.R. y VAJDA, S. 2000. Scoring docked conformations generated by rigid-body protein-protein docking. *Proteins: Structure, Function, and Bioinformatics*, vol. 40(3): 525-537.

CARTER, P. 1986. Site-directed mutagenesis. *Biochemical Journal*, vol. 237(1): 1.

CHANG, M., CHOU, T. y TSAO, G. 1981. Structure, pretreatment and hydrolysis of cellulose. En *Bioenergy*. Springer Berlin / Heidelberg. vol. 20, p. 15-42.

CHANG, M.W., AYENI, C., BREUER, S. y TORBETT, B.E. 2010. Virtual Screening for HIV Protease Inhibitors: A Comparison of AutoDock 4 and Vina. *PloS one*, vol. 5(8): e11955.

CHEN, V.B., ARENDALL, W.B., HEADD, J.J., KEEDY, D.A., IMMORMINO, R.M., KAPRAL, G.J., MURRAY, L.W., RICHARDSON, J.S. y RICHARDSON, D.C. 2009. MolProbity: all-atom structure

validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, vol. 66(1): 12-21.

DAVIS, I.W., MURRAY, L.W., RICHARDSON, J.S. y RICHARDSON, D.C. 2004. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Research*, vol. 32(suppl 2): W615-W619.

DELANO, W.L. 2002. The PyMOL molecular graphics system.

DHANIK, A., MCMURRAY, J.S. y KAVRAKI, L. 2011. On modeling peptidomimetics in complex with the SH2 domain of Stat3. En *Proceedings de la 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '11)*. Boston, Massachusetts 2011, 3229-3332.

ESWAR, N., WEBB, B., MARTI-RENOM, M.A., MADHUSUDHAN, M.S., ERAMIAN, D., SHEN, M.Y., PIEPER, U. y SALI, A. 2007. Comparative protein structure modelling using Modeller. *Curr Protoc Protein Sci*, vol. 50(2.9):1-2.9.

GARCÍA, A.E., CARMONA, R.J., LIENQUEO, M.E. y SALAZAR, O. 2011. The current status of liquid biofuels in Chile. *Energy*, vol. 36(4): 2077-2084.

GILKES, N.R., HENRISSAT, B., KILBURN, D.G., MILLER JR, R.C. y WARREN, R.A. 1991. Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. *Microbiology and molecular biology reviews*, vol. 55(2): 303-315.

GUEX, N. y PEITSCH, M.C. 1997. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis*, vol. 18(15): 2714-2723.

HAHN-HÄGERDAL, B., GALBE, M., GORWA-GRAUSLUND, M.F., LIDÉN, G. y ZACCHI, G. 2006. Bio-ethanol - the fuel of tomorrow from the residues of today. *Trends in Biotechnology*, vol. 24(12): 549-556.

HEMSLEY, A., ARNHEIM, N., TONEY, M.D., CORTOPASSI, G. y GALAS, D.J. 1989. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Research*, vol. 17(16): 6545-6551.

HENRISSAT, B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal*, vol. 280(Pt 2): 309-316.

HERNÁNDEZ-SANTOYO, A., GARCÍA-HERNÁNDEZ, E. y RODRÍGUEZ-ROMERO, A. 1999. Celulosomas: sistemas multienzimáticos. *Journal of the Mexican Chemical Society*, vol. 43(3-4): 137-142.

HILDEN, L. y JOHANSSON, G. 2004. Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. *Biotechnology Letters*, vol. 26(22): 1683-1693.

HONG, J., YE, X. y ZHANG, Y.H.P. 2007. Quantitative Determination of Cellulose Accessibility to Cellulase Based on Adsorption of a Nonhydrolytic Fusion Protein Containing CBM and GFP with Its Applications. *Langmuir*, vol. 23(25): 12535-12540.

- JANIN, J., HENRICK, K., MOULT, J., EYCK, L.T., STERNBERG, M.J.E., VAJDA, S., VAKSER, I. y WODAK, S.J. 2003. CAPRI: a critical assessment of predicted interactions. *Proteins: Structure, Function, and Bioinformatics*, vol. 52(1): 2-9.
- JORGENSEN, W.L. 1991. Rusting of the Lock and Key Model for Protein-Ligand Binding. *Science*, vol. 254(5034): 954.
- KAPLAN, W. y LITTLEJOHN, T.G. 2001. Swiss-PDB Viewer (Deep View). *Briefings in bioinformatics*, vol. 2(2): 195.
- KIBBE, W.A. 2007. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Research*, vol. 35(suppl 2): W43-W46.
- KITCHEN, D.B., DECORNEZ, H., FURR, J.R. y BAJORATH, J. 2004. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature reviews Drug discovery*, vol. 3(11): 935-949.
- KLEMM, D., PHILPP, B., HEINZE, T., HEINZE, U. y WAGENKNECHT, W. 1998 *Comprehensive cellulose chemistry. Volume 1: Fundamentals and analytical methods.* Wiley-VCH Verlag GmbH. ISBN 3527294139.
- KRAULIS, P.J., CLORE, G.M., NILGES, M., JONES, T.A., PETTERSSON, G., KNOWLES, J. y GRONENBORN, A.M. 1989. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry*, vol. 28(18): 7241-7257.
- LEE, J. 1997. Biological conversion of lignocellulosic biomass to ethanol. *Journal of Biotechnology*, vol. 56(1): 1-24.
- LEE, S.B., SHIN, H.S., RYU, D.D.Y. y MANDELS, M. 1982. Adsorption of cellulase on cellulose: effect of physicochemical properties of cellulose on adsorption and rate of hydrolysis. *Biotechnology and Bioengineering*, vol. 24(10): 2137-2153.
- LINDER, M., SALOVUORI, I., RUOHONEN, L. y TEERI, T.T. 1996. Characterization of a double cellulose-binding domain. *Journal of Biological Chemistry*, vol. 271(35): 21268.
- LINDER, M. y TEERI, T.T. 1997. The roles and function of cellulose-binding domains. *Journal of Biotechnology*, vol. 57(1-3): 15-28.
- LUTZ, S. y PATRICK, W.M. 2004. Novel methods for directed evolution of enzymes: quality, not quantity. *Current opinion in biotechnology*, vol. 15(4): 291-297.
- LYND, L.R., WEIMER, P.J., VAN ZYL, W.H. y PRETORIUS, I.S. 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology and molecular biology reviews*, vol. 66(3): 506.
- MARTÍ-RENOM, M.A., STUART, A.C., FISER, A., SÁNCHEZ, R., MELO, F. y ŠALI, A. 2003. Comparative protein structure modeling of genes and genomes. *Annual review of biophysics and biomolecular structure*, vol. 29(1): 291.

- MARTÍNEZ-ANAYA, C., BALCÁZAR-LÓPEZ, E., DANTÁN-GONZÁLEZ, E. y FOLCH-MALLOL, J.L. 2008. Celulasas fúngicas: Aspectos biológicos y aplicaciones en la industria energética. *Revista Latinoamericana de Microbiología*, vol. 50(3-4): 119-131.
- MARTÍNEZ, C.C., VALVERDE, M.E. y LÓPEZ, O.P. 2009. Biocombustibles: biomasa lignocelulósica y procesos de producción. *Ide@s CONCYTEG*, vol. 4(54): 1246 - 1270.
- MATTINEN, M.L., KONTTELI, M., KEROVUO, J., DRAKENBERG, T., ANNILA, A., LINDER, M., REINIKAINEN, T. y LINDEBERG, G. 1997. Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. *Protein science*, vol. 6(2): 294-303.
- MENG, E.C., SHOICHET, B.K. y KUNTZ, I.D. 1992. Automated docking with grid-based energy evaluation. *Journal of computational chemistry*, vol. 13(4): 505-524.
- MILLER, G.L. 1963. Cellodextrins. *Methods Carbohydr Chem*, vol. 3: 134-139.
- MORRIS, A.L., MACARTHUR, M.W., HUTCHINSON, E.G. y THORNTON, J.M. 1992. Stereochemical quality of protein structure coordinates. *Proteins: Structure, Function, and Bioinformatics*, vol. 12(4): 345-364.
- MORRIS, G.M., GOODSSELL, D.S., HALLIDAY, R.S., HUEY, R., HART, W.E., BELEW, R.K. y OLSON, A.J. 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*, vol. 19(14): 1639-1662.
- MUSSATTO, S.I., DRAGONE, G., FERNANDES, M., ROCHA, G.J.M. y ROBERTO, I.C. 2006. Efecto de los tratamientos de hidrólisis ácida e hidrólisis alcalina en la estructura del bagazo de malta para liberación de fibras de celulosa. *Escola de Engenharia de Lorena*, vol. 22: 1-10.
- NIMLOS, M.R., MATTHEWS, J.F., CROWLEY, M.F., WALKER, R.C., CHUKKAPALLI, G., BRADY, J.W., ADNEY, W.S., CLEARY, J.M., ZHONG, L. y HIMMEL, M.E. 2007. Molecular modeling suggests induced fit of Family I carbohydrate-binding modules with a broken-chain cellulose surface. *Protein Engineering Design and Selection*, vol. 20(4): 179.
- OHARA, H., NOGUCHI, J., KARITA, S., KIMURA, T., SAKKA, K. y OHMIYA, K. 2000. Sequence of egV and Properties of EgV, a *Ruminococcus albus* Endoglucanase Containing a Dockerin Domain. *Bioscience, biotechnology, and biochemistry*, vol. 64(1): 80-88.
- PALMA, P.N., KRIPPAHL, L., WAMPLER, J.E. y MOURA, J.J.G. 2000. BiGGER: a new (soft) docking algorithm for predicting protein interactions. *Proteins: Structure, Function, and Bioinformatics*, vol. 39(4): 372-384.
- PANDEY, K.K. 1999. A study of chemical structure of soft and hardwood and wood polymers by FTIR spectroscopy. *Journal of Applied Polymer Science*, vol. 71(12): 1969-1975.
- PEÑA-MONTES, C. y GONZÁLEZ-SARAVIA, A.F. 2008. Evolución Dirigida en la Generación de Biocatalizadores: Biocatalizadores Hechos a Medida. *BioTecnología*, vol. 12(2): 4-23.
- PEREIRA, A.N., MOBEDSHAHI, M. y LADISCH, M.R. 1988. Preparation of cellodextrins. *Methods in Enzymology*, vol. 160: 26-38.

- PÉREZ, J.P., MUÑOZ-DORADO, J.M.-D., DE LA RUBIA, T.D.L.R. y MARTÍNEZ, J.M. 2002. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *International Microbiology*, vol. 5(2): 53-63.
- POLAINA, J. 2004. Estructura, función e ingeniería molecular de enzimas implicadas en la digestión de carbohidratos. *Mensaje Bioquímico*, vol. 28: 61-76.
- PONCE, T. y PÉREZ, O. 2002. Celulasas y Xilanasas en la industria. *Avance y Prospectiva*, vol. 21: 273-277.
- RABINOVICH, M.L., MELNICK, M.S. y BOLOBOVA, A.V. 2002. The structure and mechanism of action of cellulolytic enzymes. *Biochemistry (Moscow)*, vol. 67(8): 850-871.
- SALI, A. y BLUNDELL, T.L. 1994. Comparative protein modelling by satisfaction of spatial restraints. *Protein Structure by Distance Analysis*: 64-86.
- SALINAS, A., VEGA, M., LIENQUEO, M.E., GARCIA, A., CARMONA, R. y SALAZAR, O. 2011. Cloning of novel cellulases from cellulolytic fungi: Heterologous expression of a family 5 glycoside hydrolase from *Trametes versicolor* in *Pichia pastoris*. *Enzyme and Microbial Technology*, vol. 49(6-7): 485-491.
- SAMBROOK, J. y RUSSELL, D.W. 2001 *Molecular cloning: a laboratory manual*. New York: Cold Spring Harbor Laboratory Press. ISBN 978-0879695774
- SÁNCHEZ, C. 2009. Lignocellulosic residues: biodegradation and bioconversion by fungi. *Biotechnology Advances*, vol. 27(2): 185-194.
- SÁNCHEZ, R. y ŠALI, A. 2000. Comparative protein structure modeling. *Protein Structure Prediction: Methods and Protocols*, Humana Press, New Jersey.
- SANNER, M.F. 1999. Python: a programming language for software integration and development. *J. Mol. Graphics Mod*, vol. 17: 57-61.
- SCIOR, T., MORALES, E.M. y STEFANÓN, E.S. 2007. Los modelos In Silico, Una herramienta para el conocimiento farmacológico. *Elementos: Ciencia y cultura*, vol. 14(68): 45-48.
- SEELIGER, D. y DE GROOT, B.L. 2010. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *Journal of computer-aided molecular design*, vol. 24(5): 417-422.
- SHEN, M. y SALI, A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein science*, vol. 15(11): 2507-2524.
- SMITH, G.R. y STERNBERG, M.J.E. 2002. Prediction of protein-protein interactions by docking methods. *Current opinion in structural biology*, vol. 12(1): 28-35.
- SMITH, G.R. y STERNBERG, M.J.E. 2003. Evaluation of the 3D-Dock protein docking suite in rounds 1 and 2 of the CAPRI blind trial. *Proteins: Structure, Function, and Bioinformatics*, vol. 52(1): 74-79.
- SUN, Y. y CHENG, J. 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. *Bioresource Technology*, vol. 83: 1-11.

- TAKASHIMA, S., OHNO, M., HIDAKA, M., NAKAMURA, A., MASAKI, H. y UOZUMI, T. 2007. Correlation between cellulose binding and activity of cellulose-binding domain mutants of *Humicola grisea* cellobiohydrolase 1. *FEBS letters*, vol. 581(30): 5891-5896.
- TAVAGNACCO, L., MASON, P.E., SCHNUPF, U., PITICI, F., ZHONG, L., HIMMEL, M.E., CROWLEY, M., CESARO, A. y BRADY, J.W. 2011. Sugar-binding sites on the surface of the carbohydrate-binding module of CBH I from *Trichoderma reesei*. *Carbohydrate Research*, vol. 346(6): 839-846.
- TEN EYCK, L.F., MANDELL, J., ROBERTS, V.A. y PIQUE, M.E. 1995. Surveying molecular interactions with DOT. En *Proceedings de 1995 ACM*, 22.
- TROTT, O. y OLSON, A.J. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, vol. 31(2): 455-461.
- WILLARD, L., RANJAN, A., ZHANG, H., MONZAVI, H., BOYKO, R.F., SYKES, B.D. y WISHART, D.S. 2003. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Research*, vol. 31(13): 3316-3319.
- ZHANG, Y.H.P. y LYND, L.R. 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnology and Bioengineering*, vol. 88(7): 797-824.

7. ANEXOS

A. Medios de cultivo y soluciones empleadas

Medios de cultivo líquidos

Medio Luria-Bertani

El medio LB contiene por cada 1 L de solución: 10 g de triptona, 5 g de extracto de levadura y 5 g de cloruro de sodio. El volumen se llevó a 1 L con agua *Milli-Q* y el pH se ajustó a 7. Se esterilizó por autoclave y almacenó a temperatura ambiente.

Medio TB

El medio TB contiene por cada 1 L de solución: 12 g de triptona, 24 g de extracto de levadura y 4 mL de glicerol. El volumen se llevó a 900 mL con agua *Milli-Q*, se esterilizó por autoclave y se agregó 100 mL de fosfato de potasio estéril.

Medio SOB

El medio SOB contiene por cada 1 L: 20 g de triptona, 5 g de extracto de levadura y 0,5 g de cloruro de sodio. El volumen se llevó a 985 mL con agua *Milli-Q*. Se mezcló y se agregó 10 mL de una solución de cloruro de potasio 250 mM y se ajustó el pH a 7. Se esterilizó mediante autoclave. Antes de utilizar se agregó 5 mL de una solución estéril de cloruro de magnesio 2 M.

Medios de cultivo sólidos

Para la preparación de placas con medio sólido se agregó a los medios líquidos agar al 1,5% p/v. Los antibióticos se esterilizaron por filtración antes de ser agregados al medio de cultivo.

Soluciones Stock

IPTG 0,1 M

Se preparó una solución stock de IPTG 0,1 M en agua *Milli-Q* estéril. Se esterilizó por filtración mediante un filtro con un tamaño de poro de 0,2 μm (Sartorius, Hannover-Alemania). Se almacenó a 4°C.

X-Gal 50 mg/mL

Se preparó una solución stock de 5-bromo-4-cloro-3-indolil-D-galactosido (X-Gal) a una concentración de 50 mg/mL en N,N'- dimetil-formamida. Se almacenó a -20°C.

Ampicilina 100 mg/mL

Se preparó una solución stock de ampicilina a una concentración de 100 mg/mL en agua *Milli-Q* estéril. Se esterilizó por filtración utilizando un filtro estéril con un tamaño de poro de 0,2 μm . Se almacenó a -20°C en volúmenes pequeños.

B. Secuencia gen (tveg) codificante para la endoglucanasa en estudio (TvEG)

>tveg

```
GTCTGGCGGCAGTGTGGTGGCACCGGCTTCTCTGGTGACACTACCTGTGCCTCCGGCTCCAGCTGCGTT
GCAGTCAACTCATACTACTCGCAATGCCAGCCTGGCGCGTCGTCCCCACGTGACCGCGTCCGCGCCC
GGCCCTTCGGGATGCTCTGGCACGCGCACCAAGTTCAAGCTGTTCCGGTGTGAACGAGTCCGGCGCGGA
GTTCCGGGAACACCGTCATCCCGGGCGCGCTCGGCACGGACTACACCTGGCCGTCGCCCACCTCCATCG
ACTTCTTCATGGGCCAGGGCTTCAACACGTTCCGCATCCCGTTCCTGATGGAGCGCGTCAGCCCCCGGT
CGACGGGTGGCCTTACTGGCCCCGTTCAACAGCACCTACCTCGACGGTCTGAAGCAGACTGTTAGCTACA
TCACGGGCAAGGGTGGTTTTGCTATCATCGACCCGCACAATTTCATGATCTTCAACGGCGCGACTATCA
CGAGCACCAGCCAGTTCAGGCCTGGTGGCAGAAGCTTGCTGCTGAGTTCAAACCAACAACAACGTC
ATCTTCGACCTCATGAACGAGCCGACGACATCCCTGCGCAGACCGTCTTCCAGCTCATGCAAGCGGCC
GTCAACGGCGTGCGCGGAGCGGCGGACACCCAGCTCATCCTCGCCGAGGGCACGAGCTGGACTG
GGGCGTGGACGTGGACGACCTCGGGCAACAGCGACGCGTTCGGCGCGATCAAGGACCCGAACAACA
CATTGCTATCCAGATGCACCAGTACCTGGACTCGGACGGCTCGGGCACGTCCCCGATCTGCGTGTCCGA
CACGATCGGCGCGGAGCGCCTGCAGGCGGCGACGACGAGTGGCTGCAGCAGAAGGGCCTCAAGGGCTTC
CTCGGCGAGATCGGCACGGGCAACAACACGACGAGTGCCTGACCGCGCTGCAGGGCGCGCTCTGCGAGA
TGCAGCAGGCCGCGGGGACGTGGCTCGGCGCGCTCAGGTGGGCCACGGGCCCGTGGTGGGGAGACT
ATTACCAGAGCATCGAGCCTCCGAACGGCGATGCGATCTCTGCGATCCTCCCGGCGCTCAAGGCGTTCC
AG
```

>TvEG

```
VWRQCGGTGFSGD TTCASGSSCVAVNSYYSQCQPGASSPTSTASAPGPSG
CSGTRTKFKLFGVNESGA EFGNTVIPGALGTDYTWPSPTSIDFFMGQGFN
TFRI PFLMERVSPSTGGLTGPFNSTYLDGLKQTVSYITGKGGFAIIDPH
NFMIFNGATITSTSQFQAWWQKLAAEFKTNNNVIFDLMNEPHDIPAQTVF
QLMQAAVNGVRASGATTQLILAEGTSWTGAWTWTTSGNSDAFGAIKDPNN
NIAIQMHQYLDSDSGSTSPICVSDTIGAERLQAATQWLQQKGLKGFLEI
GTGNNTQCVTALQGALCEMQQAGGTWLGALRWATGPWWGDYYQSI EPPNG
DAISAILPALKAFQ*
```

C. Secuencia gen codificante para las distintas variantes de endoglucanasa (TvEG)

Donde:

NdeI: **CATATG**

XhoI: **CTCGAG**

Mut 1: GGT**TAT**ACCGGCT**GGT**CTGGTGAC

Mut2: GGT**TAT**ACCGGCT**GGT**CT**CAAG**AC

Mut3: GGT**TAT**AC**CAAT**GGTCT**CAAG**AC

>Mut1_tveg

CATATG ACCGCAGTCTGGCGGCAGTGTGGTTATACCGGCTGGTCTGGTGACACTACCTGTGCCTCCGGC
TCCAGCTGCGTTGCAGTCAACTCATACTACTCGCAATGCCAGCCTGGCGCGTCGTCCCCACGTCGACC
GCGTCCGCGCCCCGGCCCTTCGGGATGCTCTGGCACGCGCACCAAGTTCAAGCTGTTCCGGTGTGAACGA
GTCCGGCGCGGAGTTCGGGAACACCGTCATCCCGGGCGCGCTCGGCACGGACTACACCTGGCCGTGCGC
CCACCTCCATCGACTTCTTCATGGGCCAGGGCTTCAACACGTTCCGCATCCCGTTCCTGATGGAGCGCGT
CAGCCCCCGCTCGACGGGTGGCCTTACTGGCCCGTTCAACAGCACCTACCTCGACGGTCTGAAGCAGA
CTGTTAGCTACATCACGGGCAAGGGTGGTTTTGCTATCATCGACCCGCACAACCTTCATGATCTTCAACG
GCGCGACTATCACGAGCACCAGCCAGTTCAGGCCTGGTGGCAGAAGCTTGCTGCTGAGTTCAAAACC
AACAACAACGTCATCTTCGACCTCATGAACGAGCCGCACGACATCCCTGCGCAGACCGTCTTCCAGCTC
ATGCAAGCGGCCGTCAACGGCGTGCGCGAGCGGGCGGACCCAGCTCATCCTCGCCGAGGGCA
CGAGCTGGACTGGGGCGTGGACGTGGACGACCTCGGGCAACAGCGACGCGTTCGGCGCGATCAAGG
ACCCGAACAACAACATTGCTATCCAGATGCACCAGTACCTGGACTCGGACGGCTCGGGCACGTCCCCG
ATCTGCGTGTGCGACACGATCGGCGCGGAGCGCCTGCAGGCGGGCGACGCAGTGGCTGCAGCAGAAGG
GCCTCAAGGGCTTCTCGGCGAGATCGGCACGGGCAACAACACGCAGTGCCTGACCGCGCTGCAGGG
CGCGCTCTGCGAGATGCAGCAGGCCGGCGGGACGTGGCTCGGCGCGCTCAGGTGGGCCACGGGGCC
GTGGTGGGGAGACTATTACCAGAGCATCGAGCCTCCGAACGGCGATGCGATCTCTGCGATCCTCCCGG
CGCTCAAGGCGTTCAGCTCGAG

>Mut2_tveg

CATATG ACCGCAGTCTGGCGGCAGTGTGGTTATACCGGCTGGTCTCAA GACACTACCTGTGCCTCCGGC
TCCAGCTGCGTTGCAGTCAACTCATACTACTCGCAATGCCAGCCTGGCGCGTCGTCCCCACGTCGACC
GCGTCCGCGCCCCGGCCCTTCGGGATGCTCTGGCACGCGCACCAAGTTCAAGCTGTTCCGGTGTGAACGA
GTCCGGCGCGGAGTTCGGGAACACCGTCATCCCGGGCGCGCTCGGCACGGACTACACCTGGCCGTGCGC
CCACCTCCATCGACTTCTTCATGGGCCAGGGCTTCAACACGTTCCGCATCCCGTTCCTGATGGAGCGCGT
CAGCCCCCGCTCGACGGGTGGCCTTACTGGCCCGTTCAACAGCACCTACCTCGACGGTCTGAAGCAGA
CTGTTAGCTACATCACGGGCAAGGGTGGTTTTGCTATCATCGACCCGCACAACCTTCATGATCTTCAACG
GCGCGACTATCACGAGCACCAGCCAGTTCAGGCCTGGTGGCAGAAGCTTGCTGCTGAGTTCAAAACC
AACAACAACGTCATCTTCGACCTCATGAACGAGCCGCACGACATCCCTGCGCAGACCGTCTTCCAGCTC
ATGCAAGCGGCCGTCAACGGCGTGCGCGAGCGGGCGGACCCAGCTCATCCTCGCCGAGGGCA
CGAGCTGGACTGGGGCGTGGACGTGGACGACCTCGGGCAACAGCGACGCGTTCGGCGCGATCAAGG
ACCCGAACAACAACATTGCTATCCAGATGCACCAGTACCTGGACTCGGACGGCTCGGGCACGTCCCCG
ATCTGCGTGTGCGACACGATCGGCGCGGAGCGCCTGCAGGCGGGCGACGCAGTGGCTGCAGCAGAAGG
GCCTCAAGGGCTTCTCGGCGAGATCGGCACGGGCAACAACACGCAGTGCCTGACCGCGCTGCAGGG
CGCGCTCTGCGAGATGCAGCAGGCCGGCGGGACGTGGCTCGGCGCGCTCAGGTGGGCCACGGGGCC
GTGGTGGGGAGACTATTACCAGAGCATCGAGCCTCCGAACGGCGATGCGATCTCTGCGATCCTCCCGG
CGCTCAAGGCGTTCAGCTCGAG

>Mut3_tveg

CATATG ACCGCAGTCTGGCGGCAGTGTGGTTATACCAATGGTCTCAA GACACTACCTGTGCCTCCGGC
TCCAGCTGCGTTGCAGTCAACTCATACTACTCGCAATGCCAGCCTGGCGCGTCGTCCCCACGTCGACC
GCGTCCGCGCCCCGGCCCTTCGGGATGCTCTGGCACGCGCACCAAGTTCAAGCTGTTCCGGTGTGAACGA
GTCCGGCGCGGAGTTCGGGAACACCGTCATCCCGGGCGCGCTCGGCACGGACTACACCTGGCCGTGCGC

CCACCTCCATCGACTTCTTCATGGGCCAGGGCTTCAACACGTTCCGCATCCCGTTCCTGATGGAGCGCGT
CAGCCCGCCGTCGACGGGTGGCCTTACTGGCCCGTTCAACAGCACCTACCTCGACGGTCTGAAGCAGA
CTGTTAGCTACATCACGGGCAAGGGTGGTTTTGCTATCATCGACCCGCACAACCTTCATGATCTTCAACG
GCGCGACTATCACGAGCACCAGCCAGTTCAGGCCTGGTGGCAGAAGCTTGCTGCTGAGTTCAAACC
AACAAACAACGTCATCTTCGACCTCATGAACGAGCCGCACGACATCCCTGCGCAGACCGTCTTCCAGCTC
ATGCAAGCGGCCGTCAACGGCGTGCGCGGAGCGGCGGACCACCCAGCTCATCCTCGCCGAGGGCA
CGAGCTGGACTGGGGCGTGGACGTGGACGACCTCGGGCAACAGCGACGCGTTCGGCGCGATCAAGG
ACCCGAACAACAACATTGCTATCCAGATGCACCAGTACCTGGACTCGGACGGCTCGGGCACGTCCCCG
ATCTGCGTGTGGACACGATCGGCGCGGAGCGCCTGCAGGCGGCGACGCAAGTGGCTGCAGCAGAAGG
GCCTCAAGGGCTTCTCGGCGAGATCGGCACGGGCAACAACACGAGTGCCTGACCGCGCTGCAGGG
CGCGCTTGCAGATGCAGCAGGCCGGCGGGACGTGGCTCGGCGCGCTCAGGTGGGCCACGGGCCC
GTGGTGGGAGACTATTACCAGAGCATCGAGCCTCCGAACGGCGATGCGATCTCTGCGATCCTCCCGG
CGCTCAAGGCGTTCAGCTCGAG

>Mut1_TvEG

HMTAVWRQCGYTGWSGD TTCASGSSCVAVNSYYSQCQPGASSPTSTASAP
GPSGCSGTRTKFKLFGVNESGAIEFGNTV I PGALGTDYTWPSPTSIDFFMG
QGFNTFRIPFLMERVSPPSTGGLTGPFNSTYLDGLKQTVSYITGKGGFAI
IDPHNFMI FNGATITSTSQFQAWWQKLAAEFKTNNNVI FDLMNEPHDI PA
QTVFQLMQAAVNGVRASGATTQLILAEGETSWTGAWTWT TSGNSDAFGAIK
DPNNNIAIQMHQYLDSDGSGTSPICVSDTIGAERLQAATQWLQQKGLKGF
LGEIGTGNNTQCVTALQGALCEMQQAGGTWLGALRWATGPWWGDYYQSIE
PPNGDAISAILPALKAFQLE

>Mut2_TvEG

HMTAVWRQCGYTGWSQD TTCASGSSCVAVNSYYSQCQPGASSPTSTASAP
GPSGCSGTRTKFKLFGVNESGAIEFGNTV I PGALGTDYTWPSPTSIDFFMG
QGFNTFRIPFLMERVSPPSTGGLTGPFNSTYLDGLKQTVSYITGKGGFAI
IDPHNFMI FNGATITSTSQFQAWWQKLAAEFKTNNNVI FDLMNEPHDI PA
QTVFQLMQAAVNGVRASGATTQLILAEGETSWTGAWTWT TSGNSDAFGAIK
DPNNNIAIQMHQYLDSDGSGTSPICVSDTIGAERLQAATQWLQQKGLKGF
LGEIGTGNNTQCVTALQGALCEMQQAGGTWLGALRWATGPWWGDYYQSIE
PPNGDAISAILPALKAFQLE

>Mut3_TvEG

HMTAVWRQCGYTQWSQD TTCASGSSCVAVNSYYSQCQPGASSPTSTASAP
GPSGCSGTRTKFKLFGVNESGAIEFGNTV I PGALGTDYTWPSPTSIDFFMG
QGFNTFRIPFLMERVSPPSTGGLTGPFNSTYLDGLKQTVSYITGKGGFAI
IDPHNFMI FNGATITSTSQFQAWWQKLAAEFKTNNNVI FDLMNEPHDI PA
QTVFQLMQAAVNGVRASGATTQLILAEGETSWTGAWTWT TSGNSDAFGAIK
DPNNNIAIQMHQYLDSDGSGTSPICVSDTIGAERLQAATQWLQQKGLKGF
LGEIGTGNNTQCVTALQGALCEMQQAGGTWLGALRWATGPWWGDYYQSIE
PPNGDAISAILPALKAFQLE

D. Ejemplos de códigos de configuración empleado por Vina y Autodock

Ej. Archivo configuración Vina:

```
receptor = CBD.pdbqt

ligand = cellulose_rigid.pdbqt

center_x = 234.413
center_y = 25.014
center_z = -2.761

size_x = 26
size_y = 26
size_z = 26

exhaustiveness = 20

energy_range = 4

num_modes = 5

out= CBD_DC_R_cellulose_rigid.pdbqt
```

Ej. Archivo de docking utilizado por Autodock:

```
autodock_parameter_version 4.2 # used by autodock to validate parameter set
outlev 1 # diagnostic output level
intelec # calculate internal electrostatics
seed pid time # seeds for random generator
ligand_types C HD OA A N # atoms types in ligand
fld 7Y10W_ref_AD_DC_RB.maps.fld # grid_data_file
map 7Y10W_ref_AD_DC_RB.C.map # atom-specific affinity map
map 7Y10W_ref_AD_DC_RB.HD.map # atom-specific affinity map
map 7Y10W_ref_AD_DC_RB.OA.map # atom-specific affinity map
map 7Y10W_ref_AD_DC_RB.A.map # atom-specific affinity map
map 7Y10W_ref_AD_DC_RB.N.map # atom-specific affinity map
elecmap 7Y10W_ref_AD_DC_RB.e.map # electrostatics map
desolvmap 7Y10W_ref_AD_DC_RB.d.map # desolvation map
move cellulose_rigid.pdbqt # small molecule
flexres 7Y10W_ref_Flexi.pdbqt # file containing flexible residues
about 0.21 0.1201 -0.0466 # small molecule center
tran0 random # initial coordinates/A or random
axisangle0 random # initial orientation
dihe0 random # initial dihedrals (relative) or random
tstep 2.0 # translation step/A
qstep 50.0 # quaternion step/deg
dstep 50.0 # torsion step/deg
torsdof 0 # torsional degrees of freedom
rmstol 2.0 # cluster_tolerance/A
extrng 1000.0 # external grid energy
```

```

e0max 0.0 10000 # max initial energy; max number of retries
ga_pop_size 150 # number of individuals in population
ga_num_evals 2500000 # maximum number of energy evaluations
ga_num_generations 27000 # maximum number of generations
ga_elitism 1 # number of top individuals to survive to next generation
ga_mutation_rate 0.02 # rate of gene mutation
ga_crossover_rate 0.8 # rate of crossover
ga_window_size 10 #
ga_cauchy_alpha 0.0 # Alpha parameter of Cauchy distribution
ga_cauchy_beta 1.0 # Beta parameter Cauchy distribution
set_ga # set the above parameters for GA or LGA
sw_max_its 300 # iterations of Solis & Wets local search
sw_max_succ 4 # consecutive successes before changing rho
sw_max_fail 4 # consecutive failures before changing rho
sw_rho 1.0 # size of local search space to sample
sw_lb_rho 0.01 # lower bound on rho
ls_search_freq 0.06 # probability of performing local search on individual
set_psw1 # set the above pseudo-Solis & Wets parameters
unbound_model bound # state of unbound ligand
ga_run 10 # do this many hybrid GA-LS runs
analysis # perform a ranked cluster analysis

```

E. Protocolo obtención DNA plasmidial con QIAprep® Spin Miniprep

- Inocular una colonia en 4 mL de medio LB con antibiótico y crecer a 37 °C durante la noche.
- Centrifugar el cultivo de a 1 mL a la vez, descartando el sobrenadante (máx. 5 mL de cultivo) (1 minuto a 13000 rpm).
- Resuspender con la pipeta en 250 µL de Buffer P1.
- Agregar 250 µL de Buffer P2, y mezclar invirtiendo el tubo suavemente 4-6 veces.
- Agregar 350 µL de Buffer de N3 e invertir inmediatamente el tubo en forma suave para mezclar.
- Centrifugar 10 minutos a 13000 rpm.
- Recuperar el sobrenadante y depositar en una columna de spin.
- Centrifugar 1minuto a 13000 rpm.
- Descartar el líquido y lavar el DNA de la columna de spin con 750 µL de Buffer PE.
- Centrifugar 1minuto a 13000 rpm.
- Descarta el líquido y centrifugar nuevamente durante 1minuto para remover los residuos de buffer.
- Poner la columna en un tubo de 1.5 mL, agregar 50 µL de Buffer EB en el centro. Dejar reposar 1 minuto y centrifugar 1 minuto a 13000 rpm.

F. Protocolo purificación con Qiagen

- Cortar la banda del gel de interés y se deposita en un tubo eppendorf de 1,5 mL.
- Agregar 3 volúmenes del Buffer QX1 por cada 1 volumen del gel.
- Resuspender QIAEX II utilizando el vortex (por 30 s). Se agregan 10 µL de éste a la muestra a purificar.

- Incubar a 50 °C por 10 min, para solubilizar la agarosa. Cada 2 min se mezcla con el vortex, de modo de mantener el QIAEX II en suspensión.
- Centrifugar la muestra por 30 s. remover el sobrenadante cuidadosamente con una pipeta.
- Lavar el pellet con 500 µL de Buffer QX1.
- Lavar el pellet 2 veces con 500 µL de Buffer PE.
- Dejar secar el pellet de 10 a 15 min hasta que el pellet se ponga blanco.
- Para eluir el DNA, agregar 20 µL de 10 mM Tris-Cl pH 8,5 o agua. Resuspender utilizando el vortex.
- Incubar la muestra 5 min a temperatura ambiente.
- Centrifugar 30 s y cuidadosamente con una pipeta, traspasar el sobrenadante a un tubo limpio.

G. Análisis llevados a cabo para la proteína recombinante

Electroforesis de proteínas SDS-PAGE

Con el objetivo de estudiar la producción de la proteína recombinante en la fracción intracelular, se llevaron a cabo geles de SDS-poliacrilamida al 8%, compuesto por:

Tabla 22. Volúmenes de los reactivos utilizados para la preparación del gel separador.

Reactivo	Volumen
Acrilamida bisacrilamida 29:1	1,015 mL
Tampón Tris-HCl 1,5 M pH 8,8, SDS 0,4%	0,95 mL
Agua destilada	1,835 mL
PSA 10%	20 µL
TEMED	4 µL

Tabla 23. Volúmenes de los reactivos utilizados para la preparación del gel concentrador.

Reactivo	Volumen
Acrilamida bisacrilamida 29:1	0,17 mL
Tampón Tris-HCl 0,5 M pH 6,8, SDS 0,4%	0,25 mL
Agua destilada	0,58 mL
PSA 10%	7,5 µL
TEMED	1,25 µL

Las muestras se mezclaron en proporción 4:1 con tampón de carga 5X (Tris-HCl 60 mM, glicerol 25 % v/v, SDS 2 % p/v, 2-mercaptoetanol 14,4 mM y azul de bromofenol 0,1 % p/v pH 6,8), se calentaron a 100°C por 5 minutos y posteriormente se cargaron en gel de acrilamida. Las electroforesis se realizaron bajo voltaje constante de 200 V por 30 minutos. Se utilizó tampón de corrida Tris-glicina-SDS (Tris 25 mM, Glicina 192 mM y SDS 0,1 % p/v) y estándar de peso molecular preteñido de proteínas.

Posteriormente los geles se incubaron toda la noche en solución de tinción, agitando suavemente. Luego se realizó la destinción del gel, incubándolo en solución de destinción, con agitación suave por 1 a 2 horas, hasta que se apreció las bandas de las proteínas.

Zimograma

Se prepara un gel de poliacrilamida y adicionalmente se le agrega, al gel separador, CMC al 0,2%. Se usa un buffer de carga 5X (Tris-HCl 60 [mM], glicerol 25% v/v, SDS 2% p/v, 2-mercaptoetanol 14,4 mM y azul de bromofenol 0,1% p/v pH 6,8), con lo cual se mezclan 16 µL de muestra y 4 de éste, para luego cargar 10 µL en el pocillo. La corrida del gel se realiza de la misma manera que la electroforesis de proteínas, con la diferencia que la cámara se introduce en un recipiente con hielo para evitar la denaturación por calor. Finalizada la electroforesis, el gel se lava 1 h con Tritón X-100 al 2,5% v/v con agitación. Posteriormente se enjuaga con agua destilada por 5 min, para luego incubarlo por 1 h a 37°C con buffer acetato de sodio 50 mM pH 5,5. Para visualizar las bandas de actividad, el gel se tiñe con Rojo Congo al 1% p/v por 20 min, y luego es desteñido con NaCl 1 M por 5 min.

Medición de actividad en medio líquido

La actividad enzimática es medida utilizando el método DNS con CMC al 1% p/v en buffer acetato de sodio 50 mM pH 5,5. Para esto, se mezclan 400 µL de CMC, con 200 µL de muestra, y se incuba durante 2 h min a 37 °C. Pasado el tiempo de incubación, se toman 70 µL de cada reacción y se mezclan con 70 µL de solución DNS, y se incuba a 100 °C por 10 min. La mezcla obtenida, se enfría en hielo por 5 min, se toman 100 µL, se depositan en una microplaca y se lee la absorbancia a 550 nm. Si la absorbancia es mayor de 1,5 se diluye 1 en 2 en agua y se repite el procedimiento.

Soluciones para geles de acrilamida

Buffer de carga 5x para geles SDS-acrilamida (10 mL)

Tris-HCl 1 M (pH 6.8)	0,6 mL
Glicerol 50%	5 mL
SDS 10%	2 mL
Azul de bromofenol 1%	1 mL
H ₂ O	0,9 mL

Solución de tinción (1 L)

Coomassie	1 g
Metanol	450 mL
Ácido acético	100 mL
H ₂ O	450 mL

Solución de destinción (1 L)

Metanol	450 mL
Ácido acético	100 mL
H ₂ O	450 mL

H. Códigos empleados en el modelo por comparación (Modeller)

Alineamiento secuencia objetivo con la plantilla

```
from modeller import *

env = environ()
aln = alignment(env)
mdl = model(env, file='lazk', model_segment=('FIRST:A', 'LAST:A'))
aln.append_model(mdl, align_codes='lazk', atom_files='lazk.pdb')
aln.append(file='CBD.ali', align_codes='CBD')
aln.align2d()
aln.write(file='CBD-lazk.ali', alignment_format='PIR')
aln.write(file='CBD-lazk.pap', alignment_format='PAP')
```

Generación del modelo

```
from modeller import *
from modeller.automodel import *

env = environ()
a = automodel(env, alnfile='CBD-lazk.ali',
              knowns='lazk', sequence='CBD',
              assess_methods=(assess.DOPE, assess.GA341))
a.starting_model = 1
a.ending_model = 200
a.make()
```

Refinamiento

```
# Loop refinement of an existing model
from modeller import *
from modeller.automodel import *

log.verbose()
env = environ()

# directories for input atom files
env.io.atom_files_directory = './:../atom_files'

# Create a new class based on 'loopmodel' so that we can redefine
# select_loop_atoms (necessary)
class MyLoop(loopmodel):
    # This routine picks the residues to be refined by loop modeling
    def select_loop_atoms(self):
        # 10 residue insertion
        return selection(self.residue_range('1', '10'))

m = MyLoop(env,
            inimodel='CBD.pdb', # initial model of the target
            sequence='CBD')    # code of the target

m.loop.starting_model= 1      # index of the first loop model
m.loop.ending_model  = 10    # index of the last loop model
m.loop.md_level = refine.slow # loop refinement method; this yields
                              # models quickly but of low quality;
                              # use refine.slow for better models

m.make()
```

I. Modelo utilizado como plantilla

Código: 1azk



Molecula: Celobiohidrolasa I

Largo: 36 aminoácidos

Cadena: A

EC: 3.2.1.19

Organismo: Trichoderma reesei

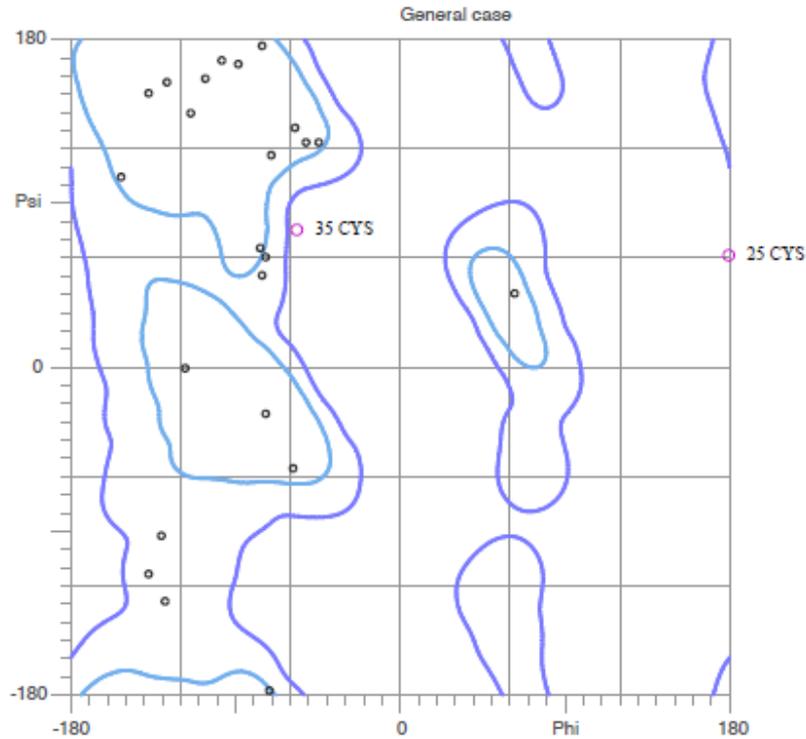
Detalle experimental: RMN en solución

Referencia: Mattinen, M. L., M. Kontteli, et al. (1997). "Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from Trichoderma reesei." *Protein science* 6(2): 294-303.

Evaluación obtenida con VADAR para la plantilla utilizada: 1azk:

RES. NUM.	RES. NAME	SCND STRUC	HBOND HBOND	BTURN BTURN	RES. ASA	FRAC. ASA	RES. VOL.	FRAC. VOL.	PHI PHI	PSI PSI	OMEGA OMEGA	PRBLM PRBLM

Chain A												
1	THR	BCC C			121.8	0.80	96.2	0.82	360.0	128.5	177.3	
2	GLN	BCC C			89.7	0.47	146.4	1.05	-74.9	61.1	-171.2	
3	SER	BCC C			42.3	0.32	89.4	0.99	-59.7	-55.6	-177.1	
4	HIS	BCC C			82.9	0.41	132.8	0.85	-115.2	140.8	-171.5	
5	TYR	CCC C 35A			153.9	0.64	166.7	0.87	62.9	41.9	-176.4	
6	GLY	BCB B 35A			2.5	0.03	68.1	1.08	-102.1	109.9	176.9	
7	GLN	BCB B			55.3	0.29	157.7	1.13	-52.4	124.4	173.4	
8	CYS	BCB B 33A			28.0	0.19	92.7	0.89	-118.5	-0.4	-175.7	
9	GLY	BCB B 33A			32.5	0.36	59.9	0.95	171.6	114.6	168.8	O
10	GLY	CCC C 13A		II	37.1	0.41	55.2	0.88	156.6	-160.3	-175.9	
11	ILE	CCC C 30A		II	117.8	0.60	149.0	0.92	-45.6	124.0	179.0	
12	GLY	CCC C		II	87.6	0.97	50.0	0.79	83.7	-13.1	-172.9	V
13	TYR	CCC C 10A		II	54.2	0.22	171.8	0.90	-89.7	167.4	166.7	O
14	SER	CCC C			103.9	0.79	83.8	0.92	-131.5	-92.4	172.2	
15	GLY	CCC C			52.5	0.58	50.1	0.80	-158.2	38.6	173.2	V
16	PRO	CCC C			90.6	0.59	107.0	0.93	-76.1	50.6	-171.9	
17	THR	CCC C			62.8	0.42	104.7	0.90	-77.7	66.6	178.8	
18	VAL	BCC C			113.0	0.66	122.4	0.90	-76.5	51.9	-176.1	
19	CYS	BBC C			31.4	0.22	96.2	0.92	-98.6	169.4	169.7	O
20	ALA	BBC C			44.9	0.36	80.9	0.93	-107.4	159.4	161.3	O
21	SER	CBC C			110.6	0.84	74.4	0.82	-72.9	-177.1	-165.3	O
22	GLY	CCC C			81.4	0.90	48.7	0.77	84.7	-57.0	153.6	V,O
23	THR	CCC C			36.5	0.24	105.0	0.90	-76.8	177.8	-159.4	O
24	THR	BCB B 36A			71.2	0.47	105.9	0.91	-129.1	-128.7	-176.6	
25	CYS	BCB B			65.3	0.45	92.3	0.88	179.8	62.5	-165.1	O
26	GLN	BCB B 34A, 34A			84.6	0.45	126.2	0.91	-128.8	157.8	-167.2	O
27	VAL	BCB B			133.5	0.78	119.9	0.88	-58.4	132.8	151.7	O
28	LEU	CCC C 32A			94.1	0.45	144.4	0.88	-138.1	-113.1	178.9	
29	ASN	CCC C 31A, 32A		III	64.8	0.39	105.5	0.90	-64.3	157.4	178.6	
30	PRO	CCC C 11A		III	113.3	0.73	115.7	1.01	-47.8	-17.4	172.6	
31	TYR	BCC C 29A		III	117.3	0.49	173.5	0.91	-74.2	-25.5	-154.9	O
32	ALA	BBB B 29A, 28A		III	6.6	0.05	82.6	0.95	-154.0	105.3	130.0	O
33	SER	BBB B 9A, 8A			36.0	0.27	82.8	0.91	-71.8	117.3	-160.2	O
34	GLN	BBB B 26A, 26A			17.5	0.09	139.7	1.00	-139.0	151.3	169.6	O
35	CYS	BBB B 6A, 5A			0.2	0.00	102.1	0.98	-58.0	76.3	153.6	O
36	LEU	BCB B 24A			76.4	0.37	160.1	0.98	-78.5	360.0	360.0	



Evaluación modelo CBM TvEG refinado y corregidos los ángulos omega

Utilizando el software Swiss-pdb viewer, se corrigieron manualmente los ángulos omegas que presentaron problemas (28, 29 y 30). Sin embargo, y como se observa a continuación, este cambio involucró pérdida de calidad del modelo, apareciendo problemas en los residuos 1, 19, 22, 23, 31, 32 y 33.

RES. NUM.	RES. NAME	SCND STRUC	HBOND HBOND	BTURN BTURN	RES. ASA	FRAC. ASA	RES. VOL.	FRAC. VOL.	PHI PHI	PSI PSI	OMEGA OMEGA	PRBLM PRBLM
1	VAL	CCC C			181.6	1.06	137.3	1.01	360.0	-39.2	173.0	A
2	TRP	CCC C			185.1	0.70	208.4	0.90	-134.1	72.2	-172.8	
3	ARG	CCC C			108.4	0.44	203.2	1.16	-119.4	129.6	-175.7	
4	GLN	CCC C			24.7	0.13	156.8	1.13	-73.0	114.8	174.2	
5	CYS	BCC C			47.7	0.33	96.6	0.92	-90.9	-23.4	-177.5	
6	GLY	BBC C 30			30.3	0.33	61.4	0.97	179.6	147.0	177.6	
7	GLY	BBC C			40.0	0.44	53.3	0.85	-171.5	-173.3	179.5	
8	THR	CBC C 27			128.2	0.85	120.1	1.03	-112.7	152.2	179.2	
9	GLY	CCC C			66.6	0.73	57.1	0.91	97.4	-179.6	179.8	
10	PHE	CBC C			214.6	0.95	171.1	0.87	-101.1	133.8	178.2	
11	SER	CBC C			84.7	0.64	85.3	0.94	177.0	171.3	177.7	
12	GLY	CBC C			69.3	0.76	52.6	0.83	-179.0	144.0	-179.7	
13	ASP	CCC C			44.4	0.28	107.9	0.95	-88.6	-169.1	176.0	
14	THR	CCC C			97.2	0.64	97.2	0.83	-91.1	-148.2	170.7	
15	THR	CBC C			115.7	0.76	105.6	0.90	-87.3	160.0	177.0	
16	CYS	CBC C			55.9	0.39	107.3	1.03	-110.6	156.1	179.4	
17	ALA	CBC C			82.0	0.66	95.4	1.09	-114.1	157.8	177.3	
18	SER	CBC C			112.3	0.85	73.4	0.81	-88.2	164.5	178.6	
19	GLY	CCC C			87.2	0.96	48.8	0.77	108.6	-36.5	170.7	V
20	SER	CCC C			73.6	0.56	98.4	1.09	-97.2	-166.7	179.2	
21	SER	CBB B			52.7	0.40	95.0	1.05	-133.9	-144.6	173.5	
22	CYS	CBB B 32			41.4	0.29	68.8	0.66	-152.8	89.0	-174.2	V
23	VAL	CBB B 31			38.3	0.22	70.5	0.52	-114.8	156.5	-176.9	V
24	ALA	CBC C			57.4	0.47	76.0	0.87	-72.8	146.7	174.1	

25	VAL	CCC	C	29		92.2	0.54	122.0	0.90	-144.4	-122.9	179.5	
26	ASN	CCC	C	28,29	III	74.5	0.45	118.4	1.01	-71.1	166.2	-171.4	
27	SER	CCC	C	8	III	96.3	0.73	86.8	0.96	-56.5	-25.6	-170.0	
28	TYR	CCC	C	26	III	147.6	0.61	179.6	0.94	-70.0	-44.4	170.0	
29	TYR	BBB	B	26,25	III	62.2	0.26	175.7	0.92	-132.9	128.1	-170.0	
30	SER	BBB	B	6		67.8	0.51	88.9	0.98	-98.1	115.7	-167.7	O
31	GLN	BBB	B	23		30.8	0.16	129.7	0.93	-132.9	165.5	-177.3	
32	CYS	CBB	B	22		47.3	0.33	66.9	0.64	-68.3	90.5	173.6	V
33	GLN	CCC	C			154.6	0.82	105.4	0.76	-89.2	360.0	360.0	V

J. Alineamientos al dominio de unión a carbohidrato

Tabla 24. Alineamientos encontrados para la secuencia aminoacídica del dominio de unión a carbohidrato de *T. versicolor*.

Alineamiento	Información
<pre> 1 10 20 33 . . . Query ---VWRQCGGTGFSGDTCASGSSCVAVNSYYSQCQ :::: :::: ::::: ::: ::::: Sbjct TQSHAGQCGGIGYSGPTVCASGTTTCQVLNPPYYSQCL 1 10 20 30 36 </pre>	<p>Código: 1AZ6</p> <p>e-value: 1.8E-8</p> <p>% identidad: 55</p> <p>% similitud: 62</p>
<pre> 1 10 20 33 . . . Query ---VWRQCGGTGFSGDTCASGSSCVAVNSYYSQCQ :::: :::: ::::: ::: ::::: Sbjct TQSHAGQCGGIGYSGPTVCASGTTTCQVLNPPYYSQCL 1 10 20 30 36 </pre>	<p>Código: 1AZH</p> <p>e-value: 1.8E-8</p> <p>% identidad: 55</p> <p>% similitud: 62</p>
<pre> 1 10 20 33 . . . Query ---VWRQCGGTGFSGDTCASGSSCVAVNSYYSQCQ :::: :::: ::::: ::: ::::: Sbjct TQSHYQCGGIGYSGPTVCASGTTTCQVLNPPAYSQCL 1 10 20 30 36 </pre>	<p>Código: 1AZJ</p> <p>e-value: 2.1E-7</p> <p>% identidad: 55</p> <p>% similitud: 62</p>
<pre> 1 10 20 33 . . . Query -----VWRQCGGTGF--SGDTCASGSSCVAVNSYYSQCQ :: ::-::: . :::::----- :::: Sbjct AQRCGDQARGAKCPNCLCCGKYGFCGSDAYCGAGS-C-----QSQCRGC 1 10 20 30 40 </pre>	<p>Código: 2LB7</p> <p>e-value: 1.0</p> <p>% identidad: 36</p>

	% similitud: 47
<pre> 1 10 20 33 . . . Query VWRQCGGTGFSGDTCASGSSCAVNSYYSQCQ : : : : : : : : : : Sbjct ---SCCPCCPSGCTKCASGCVCKGKTCDTSCCQ . . . 1 10 20 30 </pre>	<p>Código: 1M0G</p> <p>e-value: 8.8</p> <p>% identidad: 35</p> <p>% similitud: 41</p>
<pre> 1 10 20 33 . . . Query VWRQCGGTGFSGDTCASGSSCAVNSYYSQCQ . . . : : : : : : : : . . . Sbjct DPCECSKSG-----TCNCGSCTCTNCCKSK . . . 1 10 20 28 </pre>	<p>Código: 1M0J</p> <p>e-value: 8.3</p> <p>% identidad: 30</p> <p>% similitud: 62</p>
<pre> 1 10 20 33 . . . Query VWRQCGGTGFSGDTCASGSSCAVNSYYSQCQ : : : : : : : : : Sbjct -----GDCGCSGASSCNCGSGCSCNCGSK . . . 1 10 20 25 </pre>	<p>Código: 1T2Y</p> <p>e-value: 7.6</p> <p>% identidad: 31</p> <p>% similitud: 45</p>
<p>Legend: <i>Green</i> - identical residues <i>Orange</i> - similar residues <i>Red</i> - sequence mismatch <i>Blue</i> - deletion <i>Grey</i> - insertion</p>	