

TRABAJO DE GRADO, LICENCIATURA EN BIOQUÍMICA:

Complementación de una cepa de *Saccharomyces cerevisiae* con genes de enzimas celulolíticas

Departamento de Bioquímica y Genómica Microbianas, Instituto de
Investigaciones Biológicas Clemente Estable

Belén Fernández

Diciembre 2017

Orientador: Dr. Francisco Noya

Tabla de contenido

RESUMEN.....	4
1. INTRODUCCIÓN.....	6
1.1. Bioetanol de primera generación.....	6
1.2. Bioetanol de segunda generación	7
1.3. Bioetanol de tercera generación	7
1.4. Estructura de la lignocelulosa	7
1.1.1. Celulosa.....	8
1.1.2. Hemicelulosa.....	9
1.1.3. Lignina.....	10
1.5. Enzimas degradadores de la lignocelulosa	10
1.2.1. Celulasas	10
1.2.2. Celulosoma.....	12
1.2.3. Hemicelulasas	13
1.6. Pretratamiento de la biomasa.....	13
1.3.1. Métodos físicos	14
1.3.2. Métodos químicos.....	14
1.3.3. Métodos fisicoquímicos	15
1.3.4. Métodos biológicos.....	16
1.7. Sacarificación	17
1.8. Fermentación	18
1.9. Tipos de procesos industriales	20
1.6.1. Hidrólisis y Fermentación por Separado.....	20
1.6.2. Sacarificación y Fermentación Simultánea	20
1.6.3. Sacarificación y Co-Fermentación	20
1.6.4. Bioproceso Consolidado.....	20
1.10. Ciclo de vida del alcohol lignocelulósico	21
2. HIPÓTESIS Y OBJETIVOS	22
2.1. Hipótesis	22
2.2. Objetivo general.....	22
2.3. Objetivos específicos.....	22
3. MATERIALES Y MÉTODOS	23
3.1. Obtención del vector de clonado pBSK	23
3.2. Extracción de ADN plasmídico.....	23
3.3. Electroforesis en gel de agarosa.....	25

3.4.	Vector de clonado: pBSK	26
3.5.	Genes a clonar: <i>gh1</i> y <i>cdt</i>	27
3.6.	Clonado de los genes <i>gh1</i> y <i>cdt</i> en pBSK	28
3.7.	Purificación del ADN del gel de agarosa	29
3.8.	Ligación	29
3.9.	Preparación de células competentes	30
3.10.	Transformación de células competentes	30
3.11.	PCR de colonia.....	32
3.12.	Secuenciación.....	35
3.13.	Electroforesis de proteínas por SDS- PAGE	35
4.	RESULTADOS	38
4.1.	Construcción del plásmido pBSK-inserto.....	38
4.2.	Screening de los clones positivos por PCR de colonia	38
4.3.	Ensayo de restricción para comprobar la presencia de clones positivos.....	39
4.4.	Análisis bioinformático.....	41
4.4.1.	Análisis aminoacídico de la secuencia	42
4.4.2.	Verificación de la identidad de la secuencia clonada	44
4.4.3.	Repetidos de la secuencia clonada.....	45
4.4.4.	Determinación del sitio de inserción del elemento transponible.....	46
4.5.	Expresión de las proteínas	47
5.	DISCUSIÓN	48
5.1.	Secuencias de Inserción de la familia IS4	48
5.2.	Repetidos de la secuencia de inserción clonada	51
5.3.	Sitio de inserción del elemento IS	51
5.4.	Contaminación del material de trabajo con IS4	52
5.5.	Ligación	52
5.6.	Expresión de proteínas en el extracto celular	53
6.	CONCLUSIONES Y PERSPECTIVAS	54
6.1.	Conclusiones	54
6.2.	Perspectivas	54
7.	BIBLIOGRAFÍA.....	55
8.	ANEXO.....	62

RESUMEN

En la actualidad, la creciente demanda energética lleva a la búsqueda de nuevas y novedosas fuentes de combustibles amigables con el medio ambiente, entre los que surge el bioetanol. Este biocombustible puede ser obtenido a partir de biomásas ricas en sacarosa, en almidón, o biomasa lignocelulósica. La biomasa lignocelulósica, como son los residuos agrícolas y forestales, está compuesta de tres unidades estructurales principales: lignina, celulosa y hemicelulosa. Actualmente la biomasa lignocelulósica es un compuesto co-generado como un desecho de las industrias forestal y agraria, el cual dado que no representa una fuente alimenticia, generalmente es incinerado obteniendo bajos niveles energéticos. Aunque esta biomasa es altamente recalcitrante, la celulosa y la hemicelulosa están compuestas principalmente por azúcares fermentables. Uno de los desafíos que enfrenta su aplicación es la obtención de microorganismos capaces de hidrolizar y fermentar todos los azúcares que componen dicha biomasa.

A la fecha no existen organismos que naturalmente sean capaces de degradar la biomasa lignocelulósica y fermentar los azúcares liberados a etanol. En este trabajo, se propone complementar una cepa de *Saccharomyces cerevisiae* especializada en la producción de etanol con genes que codifican para enzimas celulolíticas (endoglucanasas, celobiohidrolasas y/o beta-glucosidasas) con el fin de obtener una cepa de levadura capaz de llevar a cabo la sacarificación y fermentación de azúcares para la producción de bioetanol usando biomasa lignocelulósica de origen nacional.

Se ensaya el clonado de fragmentos sintetizados *in vitro* que codifican para dichas enzimas, marcadas con el epítipo FLAG en el vector pBluescript II SK (+). Se lleva a cabo la purificación del ADN plasmídico y ensayos de restricción para confirmar la presencia de los genes que codifican para las enzimas de interés en los mismos. Sin embargo, el producto de clonado obtenido es una contaminación proveniente de una secuencia de inserción de la familia IS4, subgrupo IS10. Dicha caracterización es realizada gracias al análisis bioinformático del producto de clonado que es mandado a secuenciar. Se cree que esto es consecuencia de una contaminación en la cepa bacteriana utilizada (*Escherichia coli* DH5 α).

Ensayos de expresión de proteínas en levaduras transformadas con genes celulolíticos tampoco arrojan resultados positivos. Se analiza la expresión de dichas proteínas en el extracto celular, no pudiéndose evidenciar expresión diferencial en las distintas cepas. Este resultado se adjudica a la posible expresión de las proteínas al medio extracelular, dado que las mismas presentan un péptido señal para dicha función.

A futuro se pretende alcanzar el objetivo de complementar una cepa de *S. cerevisiae* con genes que codifiquen para enzimas celulolíticas (para lo cual los experimentos de transformación deben ser realizados con huéspedes sin contaminaciones genómicas

provenientes de transposones); realizar ensayos de Western Blot, utilizando anticuerpos anti-FLAG; y análisis de actividad enzimática en medios sintéticos con sustratos específicos para la identificación de celulasas.

1. INTRODUCCIÓN

La demanda energética ha aumentado desde 302EJ/año en 1980 a 491 EJ/año en 2006, y se espera que se duplique para el año 2050. Se ha anticipado que la reserva de combustibles fósiles resultará agotada en los próximos 40 a 50 años dada la creciente demanda energética. Además, los combustibles fósiles dan lugar a la emisión de gases de efecto invernadero, contribuyendo al calentamiento global y dan lugar a la emisión de óxido de sulfuro el cual no solo es carcinogénico sino que contribuye a la lluvia ácida [1]. Esto lleva a la búsqueda de nuevas y novedosas fuentes de combustibles amigables con el medio ambiente, entre las que surge el bioetanol. En la actualidad muchos países, como es el caso China, Brasil, países de la Unión Europea (UE) y particularmente Estados Unidos (EEUU), han incrementado su producción de bioetanol (ver Figura 1.1). Los biocombustibles, como es el caso del bioetanol, se definen como productos químicos ricos en energía generados a partir de procesos biológicos o que derivan de la conversión química de biomasa [2].

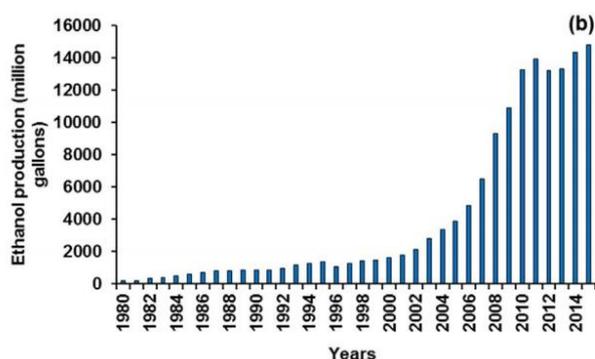


Figura 1.1. Producción de etanol por parte de los Estados Unidos en el período 1980–2015. Tomado de [3].

1.1. Bioetanol de primera generación

Para la síntesis de bioetanol, la materia prima de la cual se parte es hidrolizada en azúcares simples, los cuales posteriormente son fermentados a alcohol. La materia prima actualmente utilizada para llevar a cabo este proceso es el maíz (EEUU y China), caña de azúcar (Brasil), trigo y remolacha azucarera (Unión Europea) [4]. Si bien este es un proceso muy eficiente, presenta la desventaja de partir de materia prima que constituye una fuente alimenticia y su plantación con fines energéticos compite por la utilización del suelo.

1.2. Bioetanol de segunda generación

Otra alternativa actualmente abordada es la utilización de desechos agronómicos (rastrajo de maíz, paja y bagazo de caña de azúcar); forestales (biomasa leñosa como ramas y follaje) y municipales (desechos de comida, yerba) para la producción de bioetanol [5], compuestos principalmente por lignocelulosa. Esta es una propuesta muy atractiva dado que la biomasa lignocelulósica es la de mayor abundancia en la Tierra.

La disposición y composición de lignocelulosa dependerá de las condiciones climáticas de cada región [6]. Esto plantea el desafío de tener que desarrollar un sistema de tratamiento de la biomasa lignocelulósica que pueda abarcar la enorme heterogeneidad de sustratos sin comprometer el rendimiento. El proceso de conversión de la biomasa involucra cinco etapas principales: la correcta elección de la biomasa, su efectivo pretratamiento, la hidrólisis por parte de enzimas como son las celulasas y las hemicelulasas, la fermentación de los azúcares resultantes de dicha sacarificación y el posterior procesamiento del producto obtenido.

1.3. Bioetanol de tercera generación

Muchos biocombustibles de segunda generación requieren de grandes áreas de suelo húmedo para su producción [7]. Para superar este inconveniente surgen los biocombustibles de tercera generación, definidos como aquellos producidos a partir de organismos acuáticos (generalmente algas) [8]. Presentan una rápida tasa de crecimiento, la capacidad de fijación de gases de efecto invernadero y una alta capacidad de producción de lípidos [9]. Además, no compiten con los alimentos, y se pueden desarrollar en tierra no cultivable y agua salina. Actualmente se encuentra en estudio el acoplamiento del cultivo de microalgas para la producción de biocombustibles con el tratamiento de aguas residuales de manera de producir energía a la vez que se da la bioremediación del ambiente [9].

1.4. Estructura de la lignocelulosa

La lignocelulosa que da lugar a los biocombustibles de segunda generación presenta una composición típica de celulosa en un 30-45%, hemicelulosa en un 15-30% y lignina en un 12 a 25% [6]. La celulosa, la hemicelulosa y la lignina forman estructuras llamadas microfibrillas, que se organizan en macrofibrillas que median la estabilidad estructural en la pared celular de la planta. La celulosa compone la región central de la microfibrilla, encontrándose envuelta por la hemicelulosa y ésta a su vez por la lignina (ver Figura 1.2).

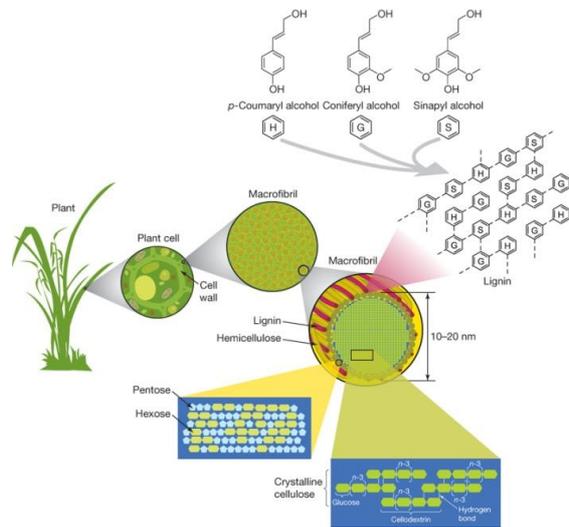


Figura 1.2. Estructura de la lignocelulosa, compuesta por celulosa, hemicelulosa y lignina organizadas en forma concéntrica. Tomado de [10].

1.1.1. Celulosa

La celulosa es un homopolímero lineal de β -D-glucopiranosas unidos por enlaces glucosídicos β -(1 \rightarrow 4) [11], donde cada cadena de glucano puede alcanzar una longitud de hasta 25.000 residuos [12]. Estas cadenas lineales se superponen interactuando mediante fuerzas de Van der Waals, dando lugar a las fibras de celulosa. Las moléculas en las fibras de celulosa cristalina se encuentran tan estrechamente empaquetadas que a través de las mismas no pueden permear las moléculas de agua [13]. Si bien la celulosa se caracteriza por presentar una estructura cristalina, se han encontrado regiones amorfas en las microfibras. La relación cristalinidad/regiones amorfas, depende de la fuente de la celulosa; siendo aquellas regiones amorfas más susceptibles a la biodegradación.

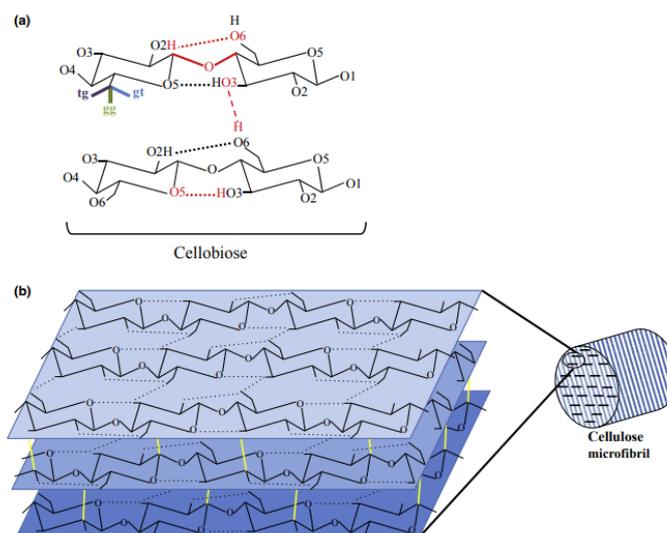


Figura 1.3. Estructura de la celulosa. (a) Estructura de la celobiosa, dímero formado por la unión β -(1,4) de dos glucosas rotadas 180° una respecto a la otra mediante enlace glucosídico. Las unidades de celobiosa forman largas cadenas con estructura de cinta plana, mantenida por los enlaces

de hidrógeno entre $O_3 - H / O_5$ y $O_2 - H / O_6$ (líneas rojas punteadas). A su vez estas cadenas se disponen en paralelo mediante enlaces de hidrógeno entre $O_6 - H / O_3$ formando láminas (b) Cada una de las láminas formadas por enlaces de hidrógeno complejos se apilan unas sobre otras por fuerzas de van der Waals (líneas amarillas). Tomado de [14].

1.1.2. Hemicelulosa

La hemicelulosa es un heteropolímero compuesto por pentosas, hexosas y ácidos urónicos (ácidos α -D-glucurónico, α -D-4-O-metilgalacturónico); mayoritariamente xilanos y los glucomananos [15].

El xilano se compone de unidades de D-xilosa unidas por enlaces glucosídicos β -(1,4), las cuales se encuentran abundantemente sustituidas por arabinosas. Los glucomananos contienen cadenas laterales de galactosa α -(1,6) ligadas, y las unidades de manosa se pueden sustituir con grupos acetato [16]. Los polímeros de β -manano se encuentran enlazados por uniones β -(1,4) y pueden presentar residuos de glucosa distribuidos aleatoriamente. El arabinano comprende una cadena principal L-arabinofuranosilo unidas por enlaces α -(1,5), el cual se encuentra generalmente sustituido por unidades de galactosa unidas por enlaces β -1,6 y cadenas laterales de L-arabinofuranosilo unidos por enlaces α -(1,3) [16].

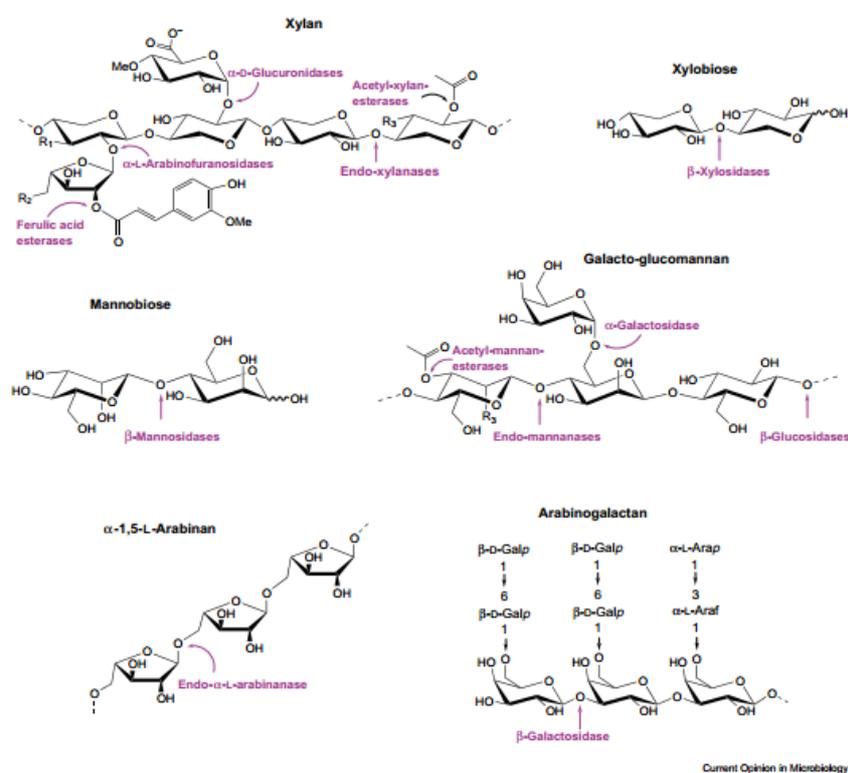


Figura 1.4. Componentes estructurales básicos encontrados en la hemicelulosa y las hemicelulasas responsables de su degradación. En la imagen se observa la estructura del xilano (compuesto por unidades de xilosa); xilobiosa; manobiosa (compuesta por unidades de manosa y glucosa); glucomananos; arabinanos. Se indican las enzimas responsables de la degradación de los diferentes componentes y los enlaces sobre los cuales se da la catálisis enzimática. Tomado de [16].

1.1.3. Lignina

La lignina por otra parte es un polímero amorfo formado principalmente por tres unidades fenólicas: el alcohol p-cumárico, el alcohol coniferílico y el alcohol sinapílico, los cuales polimerizan en una proporción que varía entre las diferentes plantas, tejidos de madera y capas de la pared celular [12] (ver Figura 1.5). Esta estructura altamente hidrofóbica hace de la lignocelulosa una estructura altamente recalcitrante, interfiriendo con el proceso de hidrólisis de la hemicelulosa y la celulosa. Su unión en forma irreversible a enzimas celulolíticas y bloquea el acceso de las celulasas a su sustrato [13].

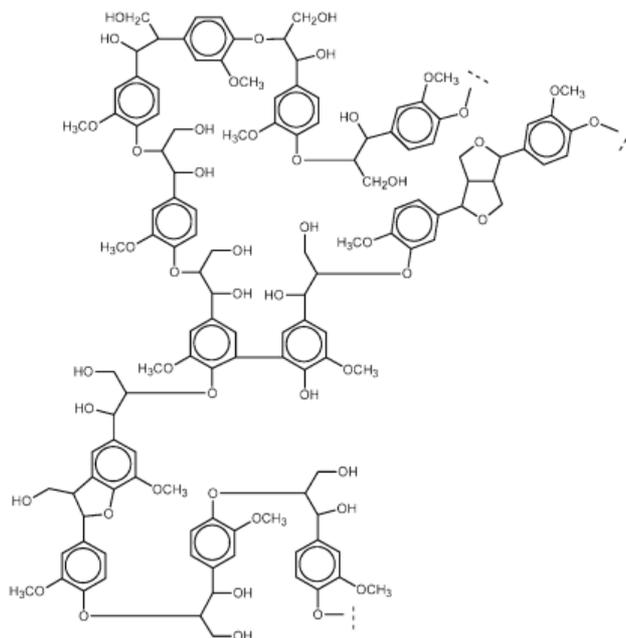


Figura 1.5. Estructura de la lignina. Es posible evidenciar su estructura amorfa y los diferentes enlaces entre las unidades de fenilpropano. Tomado de [17].

1.5. Enzimas degradadores de la lignocelulosa

A pesar de ser una estructura altamente recalcitrante, en la naturaleza existen microorganismos capaces de producir enzimas con actividad lignocelulósica. Entre estos microorganismos encontramos hongos, bacterias y algunos insectos. Los mismos expresan celulasas, hemicelulasas y lacasas, las cuales degradan eficientemente celulosa, hemicelulosa y lignina respectivamente. Se expresan además expansinas, “swollenins” y “loosenins”, las cuales generan rupturas en los puentes de hidrógeno que mantienen unidas las fibras de celulosa entre sí y con otros polisacáridos.

1.2.1. Celulasas

La hidrólisis enzimática de la celulosa involucra la acción conjunta de al menos tres tipos de enzimas:

1. endoglucanasas o 1,4- β -D-glucan-4-glucanohidrolasas (EC 3.2.1.4), encargados de la hidrólisis de enlaces glicosídicos en las regiones amorfas de la celulosa generando oligómeros de cadena larga y extremos no reductores en la cadena de celulosa [18].
2. exoglucanasas, de las cuales hay de dos tipos, las 1,4- β -D-glucan glucanohidrolasas (EC 3.2.1.74), conocidas como celodextrinasas y 1,4- β -D-glucan celobiohidrolasas (comúnmente llamadas celobiohidrolasas) (EC 3.2.1.91). Estas enzimas se encargan de degradar en forma unidireccional la celulosa, tanto a partir de sus extremos reductores como de sus extremos no reductores, liberando glucosa o celobiosa. Son capaces de hidrolizar las regiones cristalinas de la celulosa dado que atacan desde los extremos de la estructura de la fibra de celulosa.
3. β -glucosidasas o b-glucósido glucohidrolasas (EC 3.2.1.21), se encargan de la hidrólisis de celobiosa a su monómero constituyente: la glucosa.

Recientemente se ha identificado se ha identificado la presencia de otro tipo de celulasas, denominadas celulasas oxidativas, las cuales despolimerizan la celulosa por medio de reacciones de radicales libres [12].

La fuente de celulasas más comúnmente utilizada es *Trichoderma reesei*, dado que produce un sistema de celulasa extracelular, estable y eficiente [19]. Sin embargo este sistema de celulasa es deficiente en su actividad β -glucosidasa, generándose una hidrólisis incompleta de la celobiosa lo cual conduce a una inhibición por producto de las enzimas. Al momento de elegir una celulasa para su aplicación en la industria, se busca que la misma posea una alta actividad específica, sea termoestable, selectiva y presente una baja tasa de inhibición enzimática en presencia de glucosa o celobiosa y sinergismo entre las diferentes enzimas [19].

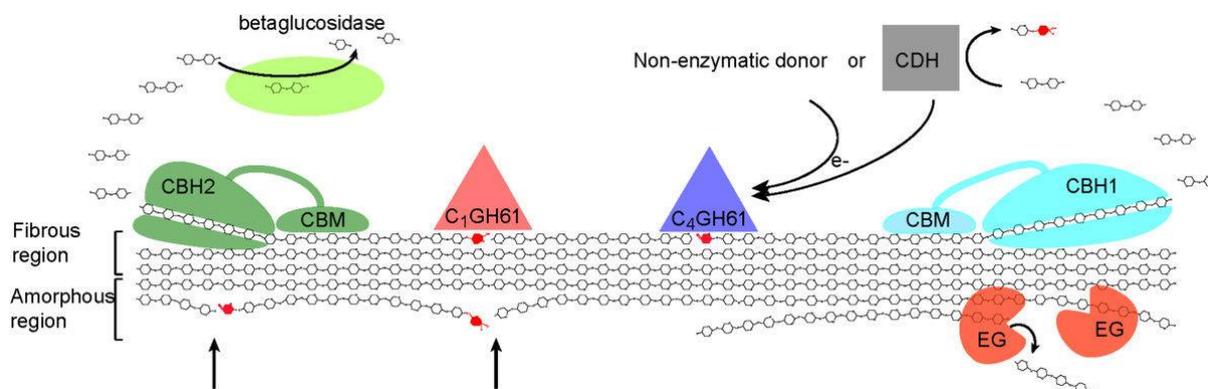


Figura 1.6. Degradación enzimática de la celulosa por parte de enzimas fúngicas. Se puede observar la catálisis enzimática de endoglucanasas (EG), β -glucosidasas, celobiohidrolasas (CBH1 y CBH2) distinción realizada dado que pueden degradar a partir del extremo reductor o no reductor de la celulosa, módulos de unión a carbohidratos (CBM) y celobiosa deshidrogenasas (CDH). Tomado de [20].

Las celulasas presentan una estructura modular, donde uno de los dominios es responsable de la catálisis enzimática mientras que el otro, conocido como módulo de unión a carbohidrato (CBM), aproxima el dominio catalítico a su sustrato. Ambos dominios se encuentran unidos por un péptido con una secuencia rica en los aminoácidos serina, prolina y treonina, lo suficientemente largo y flexible para permitir la correcta orientación y operación de ambos dominios [19].

1.2.2. Celulosoma

Las celulasas pueden ser secretadas al medio extracelular o pueden encontrarse asociadas en la superficie celular en una estructura conocida como celulosoma. El celulosoma es un complejo multienzimático localizado en la superficie celular, encargado de la degradación de la celulosa. Es producido frecuentemente por bacterias anaerobias, aunque se ha reportado que algunos hongos anaerobios también son capaces de secretar complejos tipo celulosoma [21]. Este complejo multienzimático se compone de una proteína de andamiaje sobre la cual se anclan cohesinas, las cuales establecen interacciones con las dockerinas. Cada dockerina en un lado se une a la cohesina y al otro lado a la subunidad enzimática mediante el dominio de unión a carbohidrato [15] (ver Figura 1.7). A su vez, las proteínas de andamiaje poseen sus propios dominios de unión a carbohidrato, permitiendo una eficiente captación de los oligosacáridos presentes en el medio.

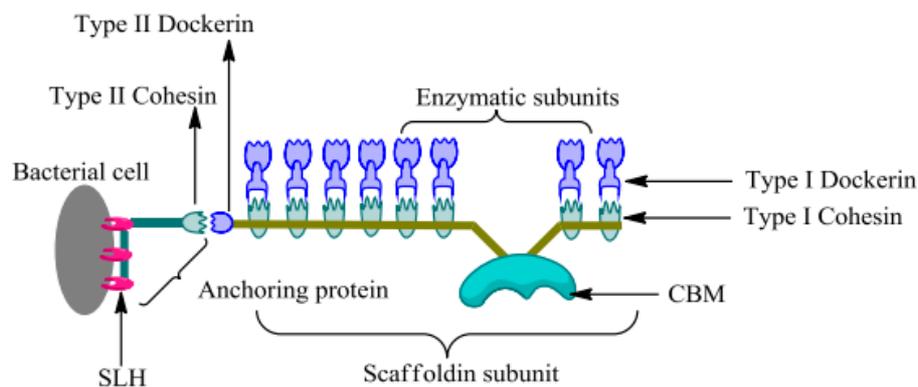


Figura 1.7. Esquema del celulosoma bacteriano. La unión del celulosoma a la pared bacteriana se da mediante proteínas de anclaje. El celulosoma presenta además módulos de unión a carbohidratos (CBM) que permiten la unión del celulosoma al sustrato glucosídico. Las subunidades enzimáticas se unen a las cohesinas mediante dockerinas. Tomado de [12].

La arquitectura del celulosoma proporciona múltiples ventajas: una adhesión directa y específica al sustrato de interés, permitiendo una competencia eficiente con otros microorganismos; y la proximidad de la célula a la celulosa asegura una captación celular eficiente de los oligosacáridos solubles, evitando su difusión en el espacio extracelular [16]. Desde el punto de vista enzimático, los celulosomas permiten una actividad concertada

óptima y el sinergismo de las celulasas, permite una óptima procesividad a lo largo de la fibra de celulosa y evita la competencia de las celulasas por el sitio de adsorción [17].

1.2.3. Hemicelulasas

La hemicelulosa es un heteropolímero, por lo que su efectiva degradación requiere la acción conjunta de muchas enzimas actuando en forma sinérgica, entre las que encontramos xilanasas, β -xilosidasas, β -mananasas, β -manosidasas, α -L-arabinofuranosidasa y α -D-glucuronidasas. Al igual que muchas celulasas, muchas hemicelulasas presentan una estructura modular, siendo uno de estos dominios un dominio catalítico y el restante puede resultar tanto un módulo de unión a carbohidrato como un módulo dockerina que media la unión mediante cohesinas-dockerina bien a la superficie de las células microbianas o a grandes complejos enzimáticos como el celulosoma. Los módulos catalíticos de las hemicelulasas se pueden clasificar en glicosil hidrolasas (GH), capaces de hidrolizar enlaces glicosídicos; o carbohidrato estererasas (CE), encargados de hidrolizar enlaces éster de grupos laterales de acetato o ácido ferúlico [16].

Las xilanasas (EC 3.2.1.8) hidrolizan el enlace β -1,4 de la cadena de xilano, liberando xilooligómeros los cuales serán posteriormente hidrolizados a xilosa por β -xilosidasas (EC 3.2.1.37), las cuales constituyen exoglicosidasas. La presencia del ácido 4-O-metil-glucurónico y las cadenas laterales de arabinofuranosa dificultan la unión de xilanasas e hidrólisis de xilano [25]. Hulbert et al. sin embargo, caracterizaron una xilanasas que necesita la sustitución del ácido 4-O-metil-glucurónico para el reconocimiento y la escisión del enlace β -1,4-xilosídico [26].

Las β -mananasas (EC 3.2.1.78) hidrolizan la hemicelulosa basada en manano, liberando β -1,4-mananoligómeros las cuales pueden ser adicionalmente hidrolizadas a manosa mediante β -manosidasas (EC 3.2.1.25) [16].

Por último, las L-arabinofuranosidasa degradan hemicelulasas que contienen arabinofuranósidos y α -D-glucuronidasas escinden el enlace α -1,2-glicosídico de la cadena lateral del ácido 4-O-metil-glucurónico que se encuentra en la cadena laterales de los xilanos [16].

1.6. Pretratamiento de la biomasa

Para que los azúcares de la celulosa y de la hemicelulosa estén disponibles para ser hidrolizados y posteriormente fermentados, la lignocelulosa debe sufrir un pretratamiento. El mismo tiene como objetivo general el separar los distintos componentes de la lignocelulosa [27]. Particularmente se busca un aumento de la superficie de contacto y porosidad de la celulosa y una disminución de su cristalinidad, así como una disminución de la lignina y de las

hemicelulosas [6]. Existen distintos pretratamientos de la biomasa, entre los que se enlistan los físicos, químicos, fisicoquímicos y biológicos. La elección del pretratamiento óptimo dependerá de la materia prima a utilizar, del costo económico y de su impacto ambiental.

Durante el pretratamiento se intenta prevenir la degradación completa de la lignina, ya que esto conduce a la formación de dos grupos de inhibidores de la fermentación: los furanos como el furfural y el 5-hidroximetilfurfural, y los derivados de fenoles como es el caso de los ácidos vanílicos y el ácido trans-p-cumárico [28].

Se considera un pretratamiento efectivo a aquel que a) da lugar a una fibra de celulosa reactiva al ataque enzimático; b) preserva la fracción de hemicelulosa y celulosa; c) no genera inhibidores enzimáticos o de microorganismos fermentadores debido a productos de degradación; d) resulta en una alta concentración de sólidos y una alta concentración de azúcares liberados en la fracción líquida; e) requiere un aporte energético bajo y posee un costo de materiales para la construcción de los reactores de pretratamiento bajo; f) consume poco o ningún producto químico y g) produce pocos residuos [29],[30]. La efectividad de un pretratamiento dependerá, entre otras cosas, de la biomasa a la cual se aplique.

1.3.1. Métodos físicos

Los métodos físicos tienen como principal objetivo la reducción de tamaño de las partículas de lignocelulosa y un aumento en la relación área/volumen. Como pretratamientos físicos se clasifican la sonificación, microondas, molienda y astillado. Estos métodos son costosos, por lo que es poco probable que sean usados a escala industrial [13].

1.3.2. Métodos químicos

Dentro de los pretratamientos químicos, el más ampliamente difundido es el uso de ácido sulfúrico, el cual puede realizarse con alta temperatura (180°C) por cortos períodos de tiempo (5 min), o a temperatura más baja (120°C) durante 30-90 min [31]. Este tiene la capacidad de hidrolizar la hemicelulosa a azúcares simples [32]. La principal desventaja de este pretratamiento es la producción de compuestos secundarios como el hidroximetilfurfural (HMF) que resultan tóxicos para los microorganismos [33]. Además, el tratamiento ácido genera corrosión en los reactores.

Por otra parte, encontramos los tratamientos con álcali, los cuales involucran el uso de hidróxidos de sodio, potasio, calcio y de amonio. Los mismos son llevados a cabo a menores presiones y temperaturas que los pretratamientos ácidos. Requiere de un paso de neutralización antes de la hidrólisis enzimática para eliminar la lignina y los inhibidores formados [13]. Causan la degradación de la lignina y el hinchamiento y la descristalización parcial de la celulosa [34], [35], [36].

Concluimos por tanto que el pretratamiento con ácido (que permite la eliminación de hemicelulosa) seguido de pretratamiento con álcali (permitiendo la eliminación de lignina) da como resultado celulosa de alta pureza.

Recientemente una nueva clase de solvente ha atraído mucha atención debido a su capacidad de disolver la lignina y los carbohidratos en forma simultánea: los líquidos iónicos o solventes verdes. Son térmicamente estables, amigables con el medio ambiente, reciclables y tienen baja volatilidad [37]. Consisten en una sal donde uno o ambos iones son grandes, y el catión tiene un bajo grado de simetría [38] (ver Figura 1.8). Estos factores tienden a reducir la energía reticular de la forma cristalina de la sal, y por lo tanto reducir el punto de fusión, haciendo del solvente verde un compuesto líquido a temperatura ambiente. El modo de acción de este solvente es mediante la disrupción de los enlaces entre celulosa, hemicelulosa y lignina, disolviendo estos compuestos sin la formación de inhibidores [39]. Sin embargo, a pesar de que este pretratamiento es menos costoso energéticamente, más fácil de operar y más amigable con el medio ambiente, todavía hay muchos obstáculos a superar antes de que sea utilizado a escala industrial [40], [41]. El costo de los líquidos iónicos es elevado, todavía se desconocen sus características toxicológicas y se desconoce sus requerimientos de regeneración.

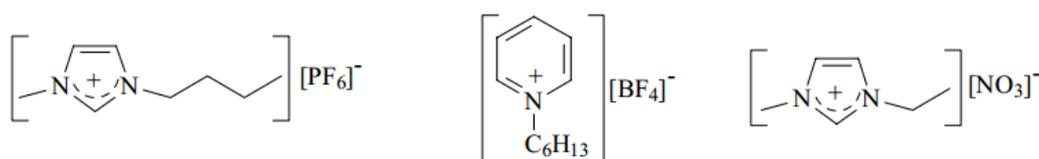


Figura 1.8. Ejemplos de líquidos iónicos simples a temperatura ambiente. Tomado de [38].

1.3.3. Métodos fisicoquímicos

Esta categoría comprende métodos que combinan un efecto tanto físico como químico. Este es el caso de la explosión de vapor, explosión de fibras de amonio, pretratamiento por agua caliente líquida, entre muchos otros.

Dentro de los métodos fisicoquímicos, el más comúnmente utilizado es el conocido como explosión de vapor. Este método se basa en la aplicación de vapor saturado a alta presión, y luego esta presión es rápidamente disminuida (causando una descompresión explosiva de los materiales). Esto genera la hidrólisis de la hemicelulosa y la transformación de la lignina, aumentando así la capacidad de hidrólisis de la celulosa. Típicamente se aplica una temperatura de 160-260°C durante unos segundos a varios minutos, y posteriormente se expone súbitamente el material a condiciones ambiente. Este brusco descenso de la presión dentro del reactor hace que las fibras se abran, rompiendo los enlaces inter e intra moleculares [42]. En algunos casos se genera un descenso del pH debido a la liberación de

ácidos orgánicos. De no ser así se puede adicionar ácido en forma extrínseca, lo cual favorece la hidrólisis y el posterior tratamiento enzimático [39]. El método de explosión de vapor ha sido aplicado a escala industrial en las plantas de Masonite [43]. A su vez, el Laboratorio Tecnológico del Uruguay (LATU) también cuenta con un equipo piloto capaz de llevar a cabo el tratamiento de explosión de vapor.

En el pretratamiento de explosión de fibras de amonio (AFEX, ammonia fiber explosion), la biomasa lignocelulósica es expuesta a amoníaco líquido a una alta temperatura y presión por un período de tiempo, y luego esta presión es disminuida súbitamente. En un proceso de AFEX típico, se agrega 1-2kg de amoníaco líquido/kg de biomasa seca, el cual se deja actuando a 90°C por alrededor de media hora [13]. Este tratamiento resulta en una descristalización de la celulosa y una despolimerización parcial de la hemicelulosa. Además genera una disrupción estructural que aumenta la superficie de contacto con el medio [44], [45]. Al igual que muchos pretratamientos alcalinos, no resulta muy eficiente si el porcentaje de lignina en la biomasa es relativamente alto como sucede en maderas [46]. Una de las grandes ventajas de este pretratamiento es que el amoníaco líquido puede ser recuperado y reutilizado. De no reutilizarse, el amoníaco debe ser tratado con precaución para evitar la contaminación del medio ambiente [29].

El pretratamiento de agua caliente líquida (LHW, liquid-hot water) por otra parte, utiliza altas presiones para mantener el agua en su estado líquido a altas temperaturas. Se ha reportado que LHW mejora la digestibilidad de la celulosa, la extracción de azúcares y la recuperación de pentosas [47]. Sumado a esto, durante el proceso no se generan compuestos inhibitorios y no es necesaria la reducción de tamaño de la biomasa dado que las partículas se disgregan durante el pretratamiento. Todo esto hace del mismo un prometedor candidato para trasladar a gran escala.

1.3.4. Métodos biológicos

Los pretratamientos biológicos pueden ser llevados a cabo utilizando microorganismos o un cocktail enzimático (lacasas y peroxidases). El objetivo de estos pretratamientos es la degradación de la lignina y la hemicelulosa, permaneciendo intacta la celulosa. Estos pretratamientos resultan técnicas prometedoras ya que no presentan requerimientos químicos, son amigables con el medio ambiente y no es necesario un suministro importante de energía. Sin embargo, dentro de sus desventajas encontramos que es un pretratamiento de larga duración, tardando de 10 a 14 días, se debe llevar un cuidadoso control de las condiciones de crecimiento de los microorganismos y son necesarias grandes superficies para llevar a cabo este pretratamiento. Además, la mayoría de los microorganismos lignolíticos no solo consumen lignina sino también celulosa y hemicelulosa.

Pretratamiento	Ventajas	Desventajas
Ácido	a) Solubilización de la hemicelulosa b) Alto rendimiento de glucosa	a) Formación de inhibidores b) Corrosión de los reactores
Alcali	a) Remoción de la lignina b) Baja formación de inhibidores	a) Alteración de la estructura de la lignina b) Alto costo de los catalizadores alcalinos
Líquidos iónicos	a) Hidrólisis de la hemicelulosa de la lignina b) Proceso menos demandante energeticamente	a) Costo de los solventes b) Toxicidad desconocida
Explosión de vapor	a) Hidrólisis de la hemicelulosa y transformación de la lignina b) Alto rendimiento de glucosa	a) Generación de compuestos tóxicos b) Degradación parcial de la hemicelulosa
AFEX	a) Baja formación de inhibidores b) Aumento de la superficie de contacto	a) Reciclaje del amonio b) Alteración de la estructura de la lignina c) Su efectividad depende del % lignina
LHW	a) No requiere de catalizadores b) No produce inhibidores	a) Grandes requerimientos energéticos y de agua b) Masa sólida remanente debe manejarse
Microorganismos	a) Amigables con el medio ambiente b) No presentan grandes requerimientos energéticos	a) Larga duración del proceso b) Ausencia de microorganismos lignolíticos que no consuman los azúcares necesarios para la fermentación

Figura 1.9. Tabla comparativa de ventajas y desventajas principales de distintos pretratamientos químicos, fisicoquímicos y biológicos tratados. En fondo amarillo se señalan los pretratamientos químicos; con fondo violeta se señalan los pretratamientos fisicoquímicos; y con fondo celeste se señala un pretratamiento biológico.

1.7. Sacarificación

La sacarificación es el proceso de conversión de un polisacárido complejo, como es el caso del almidón o la celulosa) en sus monosacáridos constituyentes. Este proceso puede ser llevado a cabo mediante agentes químicos como es el uso de ácido diluido, o agentes enzimáticos. Las diversas enzimas implicadas en la hidrólisis de los polisacáridos son obtenidas a nivel industrial de cepas modificadas genéticamente de microorganismos como *Trichoderma reesei* y *Aspergillus* sp.[48]. En un proceso de sacarificación estándar se adicionan 20mg de enzima/g de celulosa., aunque la carga enzimática ideal dependerá de la biomasa, la carga de sólidos y del pretratamiento utilizado [49]. Éstas enzimas son extremadamente costosas, encareciendo el costo de un proceso en el que sólo se degrada el 80% de la biomasa.

En la actualidad, se aplican diferentes técnicas con el fin de mejorar la eficiencia de las celulasas y aumentar los rendimientos de producción de los microorganismos que secretan estas enzimas. Las configuraciones industriales que dependen del uso de celulasas libres pueden resultar desfavorables económicamente debido al alto requerimiento, bajo reciclado y susceptibilidad a la inactivación de la enzima. La inmovilización de enzimas a un soporte las vuelve robustas y resistentes a un entorno heterogéneo. Además, evita la lixiviación molecular, permitiendo la reutilización de las enzimas, el conveniente manejo de las preparaciones enzimáticas y la fácil recuperación del producto.

Otra de las alternativas actualmente utilizadas es el diseño de enzimas. Se buscan enzimas con mayor actividad y termoestabilidad para que coincida con la óptima temperatura del proceso [50] a través de mutaciones específicas en su secuencia, basándose en el conocimiento de la relación secuencia aminoacídica/actividad enzimática [51]. Sin embargo esta relación no se encuentra del todo dilucidada, siendo muy difícil predecir cómo pequeños cambios moleculares afectan la función deseada [52]. Además muchas propiedades enzimáticas no están confinadas a uno o a unos pocos sitios de aminoácidos, sino que generalmente se distribuyen a lo largo de toda la proteína [53]. Por lo que otra posible estrategia para abordar este problema es la evolución dirigida. Para conseguir el rendimiento máximo del producto, las actividades enzimáticas de los componentes de la vía deben ser optimizadas y los niveles de expresión de varios genes de una vía metabólica necesitan ser regulados [54]. En este sentido, la estrategia de evolución dirigida ha demostrado ser eficaz para hacer frente a estos desafíos.

A su vez, Alper et al. demostraron que en lugar de modificar genéticamente las proteínas, los promotores pueden ser diseñados por medio de la evolución dirigida para lograr el nivel de expresión preciso y la regulación del control genético [55].

1.8. Fermentación

Una vez liberados los azúcares simples gracias al pretratamiento y la sacarificación, la biomasa lignocelulósica se encuentra en condiciones de ser fermentada para dar lugar a la producción de bioetanol. A pesar de ser el producto deseado, el etanol es un inhibidor del proceso de fermentación, interfiriendo con la membrana celular [56]. En los microorganismos respiratorios, esta inhibición es lineal con respecto a la concentración, pero en los anaerobios, la inhibición está exponencialmente correlacionada [57]. Esta inhibición es menor cuando el etanol es agregado externamente en relación a cuando se produce por la propia fermentación del organismo debido a su mayor concentración dentro de la célula durante la fermentación [58].

El microorganismo más utilizado para la obtención de etanol, tanto para la industria del vino y cerveza como para la industria de los biocombustibles es *Saccharomyces cerevisiae*, el cual presenta una temperatura óptima de crecimiento de 30°C y una alta tolerancia al etanol (hasta 10% v/v) e inhibidores químicos [59]. Este microorganismo es capaz de fermentar hexosas en condiciones anaerobias y a valores de pH=4, produciendo un rendimiento de etanol de hasta 12-17% m/v [60]. Sin embargo, la degradación de la hemicelulosa da lugar a la liberación de grandes cantidades de xilosa, la cual no puede ser metabolizada por *S. cerevisiae* [61]. Por lo que para que todos los azúcares presentes en la lignocelulosa puedan

ser metabolizados, es necesario la utilización de cepas de *S. cerevisiae* modificadas genéticamente.

También es posible la utilización de otros microorganismos como *Escherichia coli*, una bacteria mesófila, gram negativa, que tiene la capacidad de utilizar tanto pentosas como hexosas en la fermentación, aunque si azúcares como glucosa, sacarosa o fructosa se encuentran presentes se reprime la expresión de las enzimas necesarias para la utilización de xilosa, arabinosa, maltosa y lactosa. Además *E. coli* presenta un genoma sumamente plástico, siendo muy sencilla su modificación genética. Sin embargo presenta una tolerancia limitada al etanol y a los inhibidores, y posee un estrecho rango de crecimiento de pH y temperatura [60], limitando las condiciones de fermentación.

Zymomonas mobilis también es una bacteria gram negativa, cuyo rendimiento de etanol es de un 97%, incluso superior al de *Saccharomyces cerevisiae*. Presenta además una alta tolerancia al etanol (hasta 14% v/v) y una alta productividad del etanol (cinco veces más que la productividad volumétrica de *S. cerevisiae*). Entre sus desventajas se enlistan la incapacidad de fermentar la xilosa, la baja tolerancia a los inhibidores y que su óptimo crecimiento se da a pH neutro, cuando el proceso de fermentación es llevado a cabo a pH ácido debido a la liberación de ácidos durante el pretratamiento [60]. Ingram et al. describen una metodología para el mejoramiento de *E. coli* mediante la incorporación de la vía de producción de etanol de *Z. mobilis* [62], [63].

Entre otras posibles levaduras capaces de llevar a cabo la fermentación encontramos a *Pichia stipitis*, una levadura anaeróbica facultativa. La misma presenta un rendimiento de etanol del 82% y es capaz de fermentar la mayoría de los azúcares de material celulósico incluyendo glucosa, galactosa y celobiosa. Este microorganismo tiene la desventaja de ser intolerante al etanol en una concentración por encima de 40 g / L, no fermenta la xilosa a pH bajo, requiere condiciones micro-aerófilas para alcanzar el máximo rendimiento y reasimila el etanol formado.

Otra estrategia es la utilización de bacterias termófilas, como lo son *Clostridium thermocellum*, *Thermoanaerobacterium saccharolyticum* y *Thermoanaerobacter ethanolicus*. Estas bacterias anaerobias tienen la ventaja de presentar un máximo de crecimiento a temperaturas cercanas a los 70°C, lo cual las hace de gran utilidad para su uso en el Bioproceso consolidado. Si bien tienen la capacidad de fermentar una amplia variedad de azúcares, no son tolerantes a altas concentraciones de etanol.

1.9. Tipos de procesos industriales

1.6.1. Hidrólisis y Fermentación por Separado

El proceso más utilizado para la producción de biocombustibles es conocido como Hidrólisis y Fermentación por Separado (SHF por su sigla en inglés), donde la hidrólisis y la fermentación son llevadas a cabo en reactores diferentes [64]. Esta configuración permite que tanto la sacarificación como la fermentación se lleven a cabo en condiciones óptimas. Sin embargo, tiene la desventaja de que las enzimas celulolíticas se ven inhibidas por su producto final, disminuyendo la velocidad de hidrólisis a medida que se acumulan azúcares simples producto de la sacarificación [65].

1.6.2. Sacarificación y Fermentación Simultánea

Otra de las tecnologías para la conversión de biomasa es la Sacarificación y Fermentación Simultánea (SSF), en donde ambas etapas son llevadas a cabo en el mismo reactor. A diferencia de la SHF, no se observa una inhibición de las enzimas celulolíticas por sus productos ya que estos son fermentados a medida que se producen e implica una inversión menor dado que solo se requiere un reactor para que se lleve a cabo el proceso. A pesar de estas ventajas cabe recordar que la sacarificación y la fermentación son procesos llevados a cabo en condiciones óptimas diferentes, e implica un compromiso entre ambos procesos, aunque en algunos casos se han reportado mayores rendimientos de etanol bajo esta configuración [66].

La hidrólisis enzimática suele darse a 50°C, temperatura óptima de las enzimas implicadas. Por otra parte, la fermentación se realiza a temperaturas entre 37 y 42°C, dependiendo de la cepa utilizada [39].

1.6.3. Sacarificación y Co-Fermentación

La Sacarificación y Co-Fermentación (SSCF) se basa en la sacarificación y fermentación de la fracción sólida hidrolizada rica en glucosa en el mismo reactor que se fermenta la fracción líquida obtenida luego del pretratamiento rica en xilosa [39]. Este proceso ha sido reconocido como una de las opciones más convenientes para la producción de etanol a partir de biomasa lignocelulósica rica en xilosa [13] haciendo del mismo un proceso de suma importancia, dado que la falta de utilización de los azúcares derivados de la hemicelulosa es el mayor impedimento para la producción de los biocombustibles lignocelulósicos.

1.6.4. Bioproceso Consolidado

Una última alternativa es el Bioproceso Consolidado (CBP). Aquí el proceso no sólo es llevado a cabo en un único reactor, sino que un único microorganismo debe ser capaz de llevar a cabo

tanto la sacarificación como la fermentación de la biomasa [66]. Sin embargo, un microorganismo natural con capacidad celulolítica y que sea capaz de llevar a cabo la fermentación de los azúcares resultantes todavía no ha sido descrito. Estos microorganismos pueden buscarse en ambientes donde se lleve a cabo la degradación de la biomasa celulósica, como es el rumen bovino. También el desarrollo de microorganismos modificados genéticamente es una alternativa atractiva para el desarrollo de este proceso. Una aproximación de esto es la integración de enzimas capaces de conferir el fenotipo celulolítico a microorganismos fermentadores.

En este trabajo se plantea un abordaje de biología molecular, clonando genes de enzimas celulolíticas para complementar una cepa de *S. cerevisiae*, logrando obtener una cepa de levadura capaz de llevar a cabo la sacarificación y fermentación de azúcares para la producción de bioetanol usando biomasa lignocelulósica.

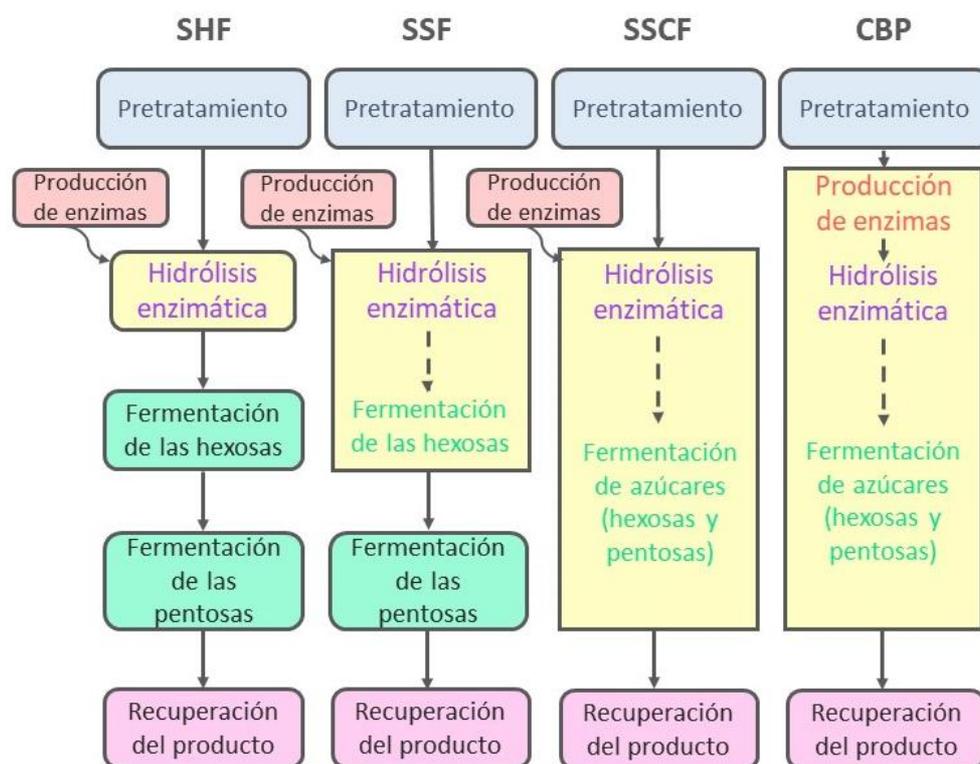


Figura 1.10. Esquema de los distintos bioprocesos disponibles para la conversión de la biomasa lignocelulósica. SHF: Hidrólisis y Fermentación por Separado; SSF: Sacarificación y Fermentación Simultánea; SSCF: Sacarificación y Co-Fermentación; CBP: Bioproceso Consolidado. Las flechas continuas indican el pasaje de un reactor al siguiente, las flechas discontinuas la sucesión entre distintas etapas dentro de un mismo reactor.

1.10. Ciclo de vida del alcohol lignocelulósico

Los análisis de ciclo de vida (LCA por su sigla en inglés: Life Cycle Assessment) permiten analizar y comparar cuantitativamente los impactos ambientales de cada sistema de generación de energía basado en biomasa [60]. Estos análisis se basan en estimaciones cuantitativas de impactos directos (agentes químicos contaminantes) e indirectos (emisiones

de gases de efecto invernadero, consumo de combustibles fósiles, consumo de agua) [67]. Análisis comparativos muestran que el uso de bioetanol de primera generación en motores de combustión interna, genera una reducción en el consumo de combustibles fósiles de un 30-50%, dicha reducción es de hasta un 90% para el caso de la sustitución por bioetanol de segunda generación [68].

2. HIPÓTESIS Y OBJETIVOS

2.1. Hipótesis

Es posible mejorar la producción de bioetanol a partir de biomasa lignocelulósica, dada la incorporación de genes que codifican para enzimas celulolíticas a microorganismos fermentadores.

2.2. Objetivo general

Otorgarle el fenotipo celulolítico a una cepa de *Saccharomyces cerevisiae*.

2.3. Objetivos específicos

- 1.** Clonar genes sintéticos de enzimas celulolíticas en vectores adecuados para su expresión en *S. cerevisiae*.
- 2.** Evidenciar la expresión de dichas enzimas tras la incorporación de las mismas en *S. cerevisiae*.

3. MATERIALES Y MÉTODOS

Se parte de la cepa DH5 α de *Escherichia coli*, la cual se crece en medio LB. La cepa DH5 α fue diseñada como una cepa de clonación, a la cual se le realizan múltiples mutaciones para permitir transformaciones de alta eficiencia. Ejemplos de esto son las mutaciones RecA1, una proteína involucrada en la recombinación homóloga, donde su mutación impide la incorporación del plásmido en el genoma; y endA1 que reduce la recombinación homóloga para un inserto más estable. En suma, esta cepa posee un buen rendimiento de plásmidos, insertos estables y puede albergar plásmidos de gran tamaño.

3.1. Obtención del vector de clonado pBSK

El vector pBluescript II SK (+) (pBSK) se encuentra en una cepa de *E. coli* almacenada a -80 °C. Para su obtención se toma con un ansa de las células almacenadas a -80°C, (sin permitir que las mismas se lleguen a descongelar pues es el estrés térmico podría lisarlas) y se rastrilla una placa de LB suplementada con ampicilina 50 μ g/ml. Se deja crecer durante 16 horas a 37°C. Una vez crecida esta placa se siembra un cultivo primario a partir de una colonia aislada y se incuba durante 16 horas a 37°C y 200 RPM en un medio líquido de iguales características. Posteriormente se realiza una extracción del ADN plasmídico a través de la técnica conocida como miniprep.

3.2. Extracción de ADN plasmídico

A partir de una colonia aislada de *E. coli* DH5 α – pBSK se realiza un cultivo primario en 5mL de LB suplementado con ampicilina 50 μ g/ml. Se incuba durante 16 horas a 37°C y 200 RPM. Posteriormente se centrifugan las células a 5000 RPM durante 90 segundos. Se descarta el sobrenadante y el *pellet* bacteriano se resuspende en 200 μ L de la solución A (anexo), una solución isotónica que impide la degradación del ADN por acción de las DNAsas. Esto se debe al agregado del EDTA, el cual quela los iones necesarios para la actividad de estas enzimas [69]. A esta solución se le adiciona también 1 μ L de RNasa 0,1 mg/mL para la degradación del ARN. Una vez resuspendidas las células se procede a generar la lisis celular y la desnaturalización del ADN a través del agregado de 200 μ L de solución B (anexo). Esta solución, compuesta de SDS y NaOH, establece un entorno alcalino para las células, induciendo su ruptura y la desnaturalización del ADN liberado. Se mezcla el tubo por inversión y se esperan 5 minutos. Por último, se agregan 200 μ L de la solución C que causa la precipitación de las proteínas y neutraliza el medio alcalino, permitiendo la renaturalización del ADN plasmídico. Se mezcla el tubo por inversión. Esta técnica permite la obtención

diferencial del ADN plasmídico basándose en sus diferencias de tamaño. El ADN plasmídico no solo es pequeño, sino que al desnaturalizar se mantiene concatenado por tratarse de ADN circular. Por lo tanto, al neutralizar la solución el plásmido renaturaliza fácilmente mientras que el ADN genómico es incapaz de encontrar regiones complementarias entre ambas hebras, formando precipitados con las proteínas. Esto permite su fácil separación del plásmido que se mantiene en solución. Para llevar a cabo esta separación se centrifuga durante 10 minutos a 12.000 g, recuperándose el sobrenadante. Con el objetivo de obtener ADN plasmídico de alta pureza, es posible agregar un volumen igual de fenol : cloroformo al sobrenadante obtenido. El eppendorf se centrifuga a 12.000 g por 2 minutos, transfiriéndose la fase acuosa donde se encuentra el ADN plasmídico en solución a un tubo estéril. Si bien este paso es opcional, su omisión puede resultar en ADN resistente al clivaje por enzimas de restricción [70]. Posteriormente, se precipita el ADN plasmídico con 600 μ L de isopropanol, se deja 5 minutos a temperatura ambiente y se centrifuga nuevamente durante 10 minutos a 12.000 g. Se descarta el sobrenadante y el ADN precipitado se lava para eliminar sales remanentes con 500 μ L de etanol 70% (V/V). Se centrifuga a 12.000 g durante 5 minutos y se descarta el sobrenadante. El pellet se seca en la estufa, aunque también es posible la evaporación del etanol remanente mediante centrifugación al vacío en SpeedVac. El pellet, una vez seco, se resuspende en 50 μ L de H₂O miliQ y se incuba a 60°C durante 15 minutos. El ADN plasmídico obtenido se guarda a -20°C.

Paralelamente se realiza la purificación del vector con el kit comercial PureLink® Quick Plasmid Miniprep Kits (Invitrogen), también basado en el método de lisis alcalina. Sin embargo, a diferencia de lo descrito anteriormente, la recuperación del ADN se da gracias a una columna de sílica. El lisado celular se aplica a la columna, uniéndose selectivamente el ADN plasmídico. Posteriormente los contaminantes se eliminan con un *buffer* de lavado mientras el vector se mantiene unido a la columna, y por último éste se eluye con el *buffer* TE (anexo) pudiéndose obtener hasta 40 μ g de ADN plasmídico.

Una vez finalizada la purificación del vector, el mismo se cuantifica basándose en su absorbancia a 260nm en el equipo NanoDrop 1000 Spectrophotometric (Thermo Scientific), un espectrofotómetro que requiere tan solo 1 μ L para la cuantificación de la muestra. La absorbancia del ADN a esta longitud de onda se debe a la presencia de grupos aromáticos en las bases nitrogenadas. A su vez, se corrobora que el ADN aislado presente una relación $A_{260}/A_{280} < 1.80$, indicativo de que se encuentra mayoritariamente libre de proteínas.

Con la finalidad de linealizar el plásmido, se digiere el vector pBSK con la enzima *BamHI* (Thermo Scientific) la cual presenta un único sitio de restricción y se incuba a 37 °C durante 1 hora. El producto de restricción se corre en un gel de agarosa al 0,8%, a fin de comprobar su pureza, integridad y tamaño.

La migración del ADN depende de si el mismo se encuentra en estado superenrollado, lineal o circular. El ADN superenrollado, altamente compacto, puede penetrar fácilmente en los poros del gel, migrando una mayor distancia que el ADN lineal y el ADN circular. Este último, ocupa un gran volumen, haciendo más dificultosa su migración. Dado que el marcador de peso molecular GeneRuler 1 Kb DNA Ladder (Thermo Scientific) se compone de fragmentos de ADN lineal, para determinar el tamaño del pBSK el mismo debe encontrarse en igualdad de condiciones, motivo por el cual el mismo debe ser linealizado.

3.3. Electroforesis en gel de agarosa

Con el fin de comprobar la integridad, pureza y tamaño del vector purificado se lleva a cabo una electroforesis en gel de agarosa al 0,8% en TAE 1X (anexo). El *buffer* Tris-Acético establece condiciones ligeramente alcalinas (pH=8-8.5) bajo las cuales el ADN está mejor protegido contra la hidrólisis. Por otra parte, el EDTA secuestra los iones divalentes, necesarios para la actividad de la DNAsa. De esta forma se mantiene la integridad del ADN mientras se ejecuta la electroforesis. Este *buffer* se puede sustituir por el TBE, donde el ácido acético es reemplazado por ácido bórico. TBE es un mejor amortiguador, observándose bandas de ADN más nítidas cuando los fragmentos son menores de 2 kb.

Se siembra una alícuota de 10 µL de la muestra a analizar y el peso molecular aproximado de los fragmentos obtenidos se determina por comparación con el marcador de peso molecular GeneRuler 1 Kb DNA Ladder (Thermo Scientific) (ver Figura 3.1), del cual se siembran 5 µL. Se aplica un campo eléctrico de 90 V por 45 minutos. La visualización del ADN es posible gracias al agregado de SYBR® Safe 10.000X (Invitrogen) al gel. Esto significa que se deberían agregar 5µL cada 50 mL de gel, sin embargo, con el agregado de 2,5 µL se obtiene una correcta visualización del gel y se expone a un transiluminador de luz azul para su revelado (Safe Imager™ 2.0, Invitrogen).

GeneRuler 1 kb DNA Ladder

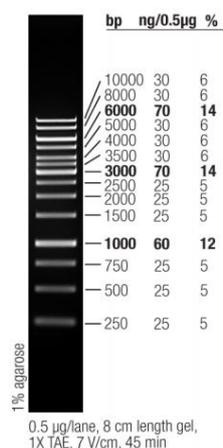


Figura 3.1. Marcador de peso molecular GeneRuler 1 Kb DNA Ladder (Thermo Scientific). Patrón de bandas obtenido al realizar una corrida electroforética a 7V/cm, durante 45 minutos en *buffer* TAE 1X, para un gel de agarosa al 1%.

3.4. Vector de clonado: pBSK

El vector pBluescript II SK (+) (pBSK) es un vector ampliamente utilizado para la transformación de *E. coli*. El mismo presenta como características fundamentales, un origen de replicación, una secuencia de resistencia al antibiótico ampicilina y un sitio de clonado múltiple (MCS). Este MCS tiene la particularidad de encontrarse dentro del gen que codifica para la subunidad α de la proteína denominada β -galactosidasa (LacZ α). (Figura 3.2). La β -galactosidasa es una proteína con la capacidad de ser dividida en dos péptidos: el péptido α , presente solamente en el plásmido para el caso de la cepa DH5 α de *E. coli*; y el péptido Ω , presente en el genoma. Ninguno de estos péptidos es activo en forma independiente, sin embargo, la β -galactosidasa se torna funcional cuando ambos péptidos (α y Ω) se encuentran presentes, proceso conocido como α complementación.

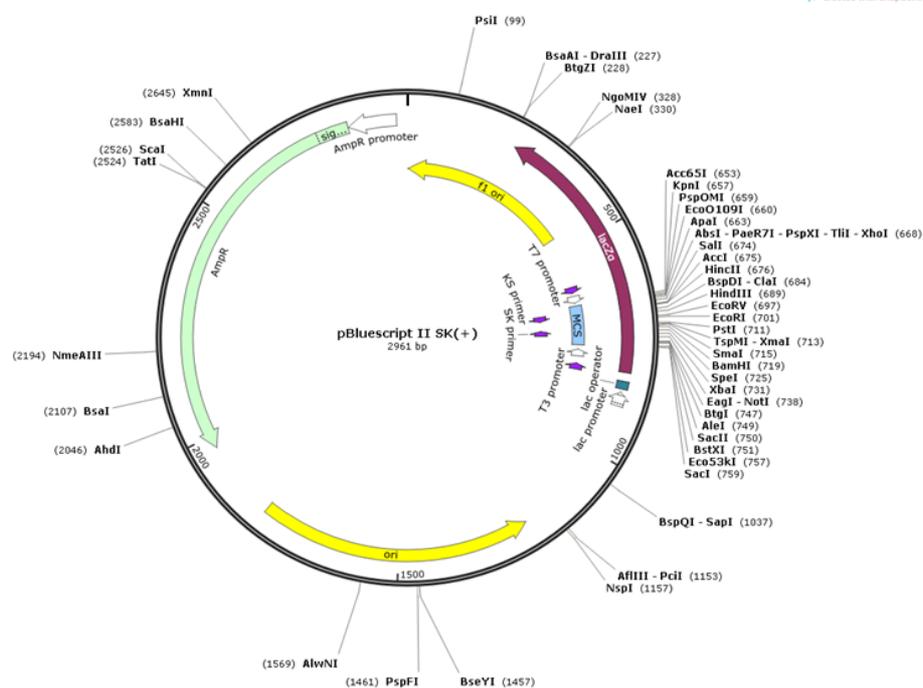


Figura 3.2. Mapa del vector de clonado pBSK II SK(+). Es posible evidenciar en el sitio de clonado múltiple del plásmido, los sitios de corte con las enzimas de restricción *XmaI* y *KpnI* en las posiciones 715 y 657 respectivamente, utilizados para clonar *gh1* y *cdt*.

LacZ cumple la función de gen reportero. Dicho gen requiere de un inductor para su expresión, que es la alolactosa, aunque generalmente se utilizan análogos como son el IPTG (isopropil- β -D-tiogalactósido). Esto se debe a que la β -galactosidasa hidroliza el sustrato conocido como X-Gal (5-bromo-4-cloro-3-indolil- β -D-galactopiranosido) dando lugar a

galactosa y 5-bromo-4-cloro-3-hidroindol. Este último se oxidada a 5,5'-dibromo-4,4'-dicloro-índigo, un compuesto azul insoluble que precipita, dando una coloración azul característica a las colonias. Esto nos permite verificar el correcto clonado del inserto de interés, mediante la pérdida de la α -complementación en aquellos plásmidos donde se encuentre presente el inserto. Esto se debe a que estas colonias no expresan una enzima β -galactosidasa funcional, pues el gen que codifica para su subunidad α se encuentra interrumpido.

Es importante que el vector otorgue una ventaja para la célula, en este caso: resistencia a la ampicilina, pues la replicación del plásmido es un proceso de alto costo y de no resultar en un beneficio para la bacteria, la misma terminará perdiendo el plásmido.

Este vector se utiliza en una primera instancia para obtener el gen de interés en un alto número de copias. Posteriormente, el inserto que contiene el gen de interés se clona en un vector de expresión de levadura, permitiendo la expresión de los genes de interés en *S. cerevisiae*.

3.5. Genes a clonar: *gh1* y *cdt*

gh1 es un gen de 1495 pb que codifica para una β -glucosidasa intracelular de 484 aminoácidos, capaz de hidrolizar el enlace β (1 \rightarrow 4) de la celobiosa para dar lugar a dos unidades de glucosa. Por otra parte, *cdt* es un gen de 1804 pb que codifica para un transportador de celodextrinas de 587 aminoácidos, capaz de interiorizar la celobiosa presente en el medio. Esto hace que ambas proteínas actúen en forma conjunta, ya que la β -glucosidasa codificada por *gh1* requiere del transportador de celodextrinas para llevar a cabo su función, y dicho transportador no posee sentido metabólico de no estar presente la β -glucosidasa. Ambos genes provienen del hongo celulolítico *Neurospora crassa* [71].

Sin embargo, *S. cerevisiae* salvaje es incapaz de asimilar la celobiosa pues carece tanto del transportador de celobiosa como de una β -glucosidasa intracelular capaz de su posterior degradación [72]. Es por esto que se decide complementar dicha cepa con enzimas del hongo celulolítico *Neurospora crassa*. Se construyen fragmentos de ADN sintéticos, lineales, conteniendo los genes *gh1* y *cdt* en forma independiente, recodificados para su óptima expresión en *S. cerevisiae*. A su vez se agrega un FLAG en el extremo de las secuencias para la posterior identificación de la proteína y los sitios de restricción *XmaI*, *BamHI* en el extremo 5' y *XhoI*, *Sall* y *KpnI* en el extremo 3' (ver Figura 3.3).

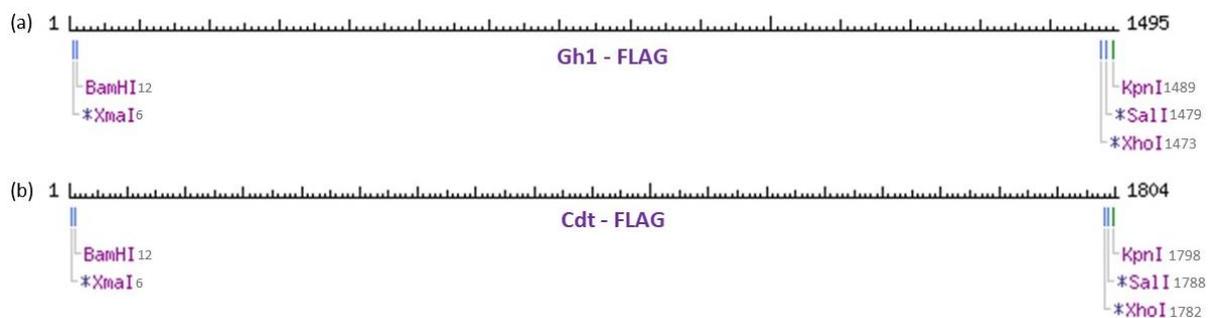


Figura 3.3. Esquema de las secuencias de ADN conteniendo (a) el gen *gh1* y (b) *cdt*. En los extremos de la secuencia se observan los sitios de restricción adicionados in vitro para el posterior clonado de los fragmentos.

3.6. Clonado de los genes *gh1* y *cdt* en pBSK

El pBSK y los insertos *gh1* y *cdt* se digieren con *XmaI* y *KpnI* de Thermo Scientific, cuyos sitios de restricción se adicionan intencionalmente al momento de sintetizarse los insertos. La digestión se realiza durante 2 horas a 37°C. Se elige esta temperatura de incubación pues *XmaI* es una enzima originaria de *Xanthomonas malvacearum* y *KpnI* fue originalmente aislada de *Klebsiella pneumoniae*; ambas Gammaproteobacterias con una temperatura óptima de crecimiento de 37°C [73][74][75]. Se decide llevar a cabo la digestión con éstas enzimas y no otras por ser las más externas en los insertos, permitiendo la posterior digestión del mismo con otras enzimas de restricción cuyos sitios de restricción se encuentran más internos en la secuencia. La digestión del inserto permite su linealización, y permite a su vez, la generación de extremos cohesivos. Esto permite la posterior inserción del inserto en el vector.

Tabla 3.1. Concentraciones y volúmenes de los reactivos utilizados para la doble digestión del pBSK con las enzimas *XmaI* y *KpnI*

Reactivos	Concentración en stock	Volumen (μL)
Buffer Cfr9I	10X	2
<i>XmaI</i> (Cfr9I)	10 U/μL	1
<i>KpnI</i>	10 U/μL	2
ADN	10 ng/μL	5
H₂O	-	10
V_f		20

Las condiciones óptimas para la doble digestión se obtienen a partir de la base de datos de Thermo Scientific donde se indica, para la combinación de dos enzimas, qué *buffer* utilizar y si alguna de estas enzimas se debe agregar en exceso como es el caso.

3.7. Purificación del ADN del gel de agarosa

Los productos de restricción se corren en un gel de agarosa 0,8% durante 45 minutos a 90 V. El gel se expone a un transiluminador de luz azul para su revelado y se recortan las bandas del tamaño esperado de acuerdo a lo que indica el marcador molecular GeneRuler 1 Kb DNA Ladder. Es importante destacar que el revelado debe realizarse con un transiluminador de luz azul y no con luz ultravioleta, pues ésta última introduce mutaciones en el ADN. De utilizarse luz ultravioleta se debe correr un carril extra para observar la localización de la banda de interés, para luego recortar el gel a esta altura, pero en el carril que no resultó irradiado.

La extracción del ADN del gel se lleva a cabo mediante el kit comercial QIAquick® Gel Extraction Kit capaz de purificar hasta 10 µg de ADN de entre 70 bp y 10 kb. Este kit permite la recuperación de hasta un 80% del ADN inicial gracias a la utilización de una suspensión de sílica. Esta unión se da en presencia de altas concentraciones de sales caotrópicas y un $\text{pH} \leq 7.5$. Luego mientras el ADN se mantiene unido a la columna, es posible lavar todas las impurezas como es el caso de sales, enzimas, agarosa, pigmentos, entre otros. Una vez finalizado este paso el ADN se eluye, lo cual ocurre en un pH entre 7.0 y 8.5 y con bajas concentraciones de sal [76]. Para esto es posible utilizar 50µL agua miliQ o el *buffer* EB, provisto por el kit (anexo).

3.8. Ligación

La ligación se lleva a cabo mediante la acción de la enzima T4 ADN ligasa (Thermo Scientific) en las condiciones especificadas (ver Tabla 3.2). Una vez mezclado el pBSK con el inserto se incuban ambas reacciones 16 horas a 14°C, de forma de generar los plásmidos pBSK-*gh1* y pBSK-*cdt* independientemente. También se ensaya la ligación incubando 1 hora a 25°C. Se utiliza una relación inserto : vector de 3:1, aunque posteriormente también se ensaya con una relación 2:1 y 6:1.

Tabla 3.2. Concentraciones y volúmenes de los reactivos para la reacción de ligación del gen *gh1* y *cdt* al vector pBSK

Reactivos	Concentración en stock	Volumen (µL)
<i>Buffer</i> T4	5X	6
T4 ligasa	1 U/µL	0,1
Inserto	10 ng/µL	20
pBSK	50 ng/µL	2
H ₂ O	-	1,9
V _f		30

Una vez finalizada la ligación, se incuba 10 minutos a 65°C con la finalidad de inactivar la ligasa. Está reportado que la inactivación de la ligasa aumenta la eficiencia de la posterior transformación [77].

La T4 ADN ligasa cataliza la formación de un enlace fosfodiéster entre el extremo 5' fosfato y el 3' hidroxilo terminal de ADN doble cadena. Además, esta enzima tiene la capacidad de reparar uniones de cadena simple de ADN, ARN o híbridos ADN/ARN.

3.9. Preparación de células competentes

Para la generación del estado de competencia se utiliza un método químico mediante el cual se generan poros en la membrana externa celular aumentando su permeabilidad al ADN. A partir de una colonia aislada de *E. coli* DH5 α , se inoculan 5 mL de LB sin antibiótico y se incuba durante 16 hs a 37 °C y 200 RPM. Posteriormente se realiza una dilución 1/100 en LB y se deja crecer hasta una $DO_{600nm} = 0,4$, permitiendo así que las células alcancen la fase exponencial. Las células se recuperan por centrifugación a 5.000 RPM y 4 °C durante 5 min en tubos estériles previamente enfriados en hielo. Se resuspende el *pellet* bacteriano en 20 mL de CaCl₂ 0,1 M frío por cada 50 mL de cultivo original. Se deja en hielo 20 min y se recuperan las células nuevamente por centrifugación a 5.000 RPM y 4 °C durante 5 min. Se descarta el sobrenadante y se resuspenden las células en 1 mL de CaCl₂ 0,1 M - glicerol frío por cada 50 mL de cultivo original. Se debe tener especial cuidado pues las células en esta etapa se encuentran en estado de competencia y son muy susceptibles, por lo que la resuspensión es llevada a cabo mediante pequeños golpeteos alternando con períodos de hielo. Las células competentes se almacenan a -80 °C en glicerol 25% (V/V) en alícuotas de 100 μ L.

3.10. Transformación de células competentes

Para introducir el ADN plasmídico en la cepa *E. coli* DH5 α , se utilizan 100 μ L de las células competentes, a las cuales se les agrega aproximadamente 50 ng de ADN y se deja en hielo 10 min. Se realiza un control negativo en el cual no se agrega ADN a las células, y un control positivo utilizando un plásmido cuantificado (1 ng del plásmido pBSK), para analizar la eficiencia de transformación. La eficiencia de transformación para el control positivo siempre es mayor, dado que este plásmido al no haber sido linealizado y posteriormente ligado, se encuentra en estado superenrollado, su volumen es menor y por tanto es más sencillo desde un punto de vista estérico que penetre a través de los poros que se inducen en las células competentes. Se utiliza el método de shock térmico, el cual favorece la formación de poros

en la membrana que facilitarán el ingreso del ADN. Con esta finalidad, la mezcla se somete a un golpe de calor a 42°C durante 90 segundos y rápidamente se coloca en hielo 2 minutos. Luego cada tubo se lleva a un volumen final de 1mL con LB y se incuba 1 h a 30 °C con agitación para permitir a las células asimilar el ADN y expresar la resistencia al antibiótico. Se concentran las células centrifugando 5 minutos a 2500 RPM y las mismas se resuspenden en 100 µL. Este volumen se plaquea en placas de LB suplementadas con ampicilina 50 µg/mL y X-gal 20 µg/mL, y se incuban a 37 °C durante 16 hs. La ampicilina vuelve selectivo al medio, dado que aquellas colonias que no presenten resistencia a la misma no podrán crecer, permitiéndonos discriminar aquellas colonias que fueron capaces de incorporar el plásmido. El X-gal por otro lado, hace al medio diferencial, permitiendo distinguir aquellas colonias que incorporaron el pBSK con el inserto de las que incorporaron el pBSK sin el inserto. Es de suma importancia que el tiempo de incubación no supere las 16 horas pues aquellas colonias que hayan incorporado el pBSK comienzan a secretar β-lactamasa al medio. Ésta enzima es responsable de la ruptura del anillo β-lactámico presente en la estructura de la ampicilina (Figura 3.4) y otros antibióticos β-lactámicos.

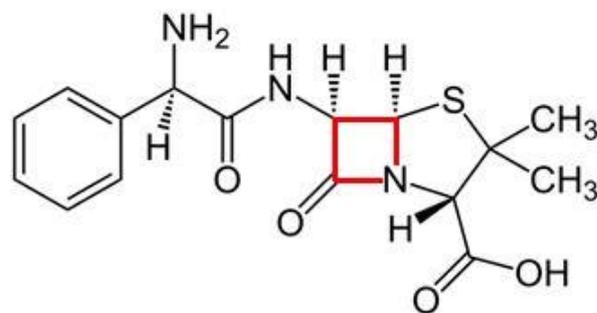


Figura 3.4. Estructura química de la ampicilina. Se destaca en rojo el anillo β-lactámico

La ampicilina actúa uniéndose a unas proteínas localizadas en la pared celular denominadas PBPs (Proteínas de Unión a Penicilina), lo cual inhibe la última etapa de síntesis de la pared celular [78].

La liberación de β-lactamasa hacia la zona circundante de la colonia que incorporó pBSK causa que la ampicilina sea degradada, permitiendo el crecimiento de colonias que no incorporaron el vector. A estas colonias se las conoce como “colonias satélites”, suelen ser diferenciables por su menor tamaño, pero llevan a confusión por ser blancas (dado que no hay plásmido incorporado).

Teniendo en cuenta las características del vector de clonado explicadas anteriormente, se seleccionaron las colonias blancas resultantes de la transformación, las cuales se espera tengan el plásmido pBSK-*gh1* y pBSK-*cdt*.

3.11. PCR de colonia

La presencia pBSK- *gh1* y pBSK- *cdt* en las colonias blancas se verifica por PCR de colonia.

La PCR (Reacción en Cadena de la Polimerasa) es una técnica *in vitro* que permite la amplificación enzimática de fragmentos específicos de ADN mediante la utilización de cebadores que hibridan con los extremos de la secuencia que se desea amplificar, el ADN molde, dNTPs y una ADN polimerasa que cataliza la reacción enzimática [79]. Consta de tres etapas. Un paso de calentamiento para separar las dos cadenas de ADN realizado a 95°C. A esto le sigue un paso de hibridación de los cebadores al ADN molde, cuya temperatura dependerá de la temperatura de *melting* (T_m) del cebador (el cual a su vez está condicionado por el %GC). Por último, se da la extensión, donde la ADN polimerasa copia el fragmento a amplificar a su temperatura óptima. Estos ciclos se repiten aproximadamente 30 veces, dando lugar a un aumento exponencial en la cantidad de ADN molde. Sin embargo, el aumento en el número de ciclos no conlleva a un aumento exponencial en el número de amplicones. En los últimos ciclos, la cantidad de enzima ya no es suficiente para extender todos los complejos cebador/molde presentes, resultando en una disminución de la eficiencia de la reacción de PCR [80]. A su vez, una vez se han completado todos los ciclos, se realiza una etapa de extensión final la cual es de mayor duración que la extensión que se dio anteriormente. Tiene como objetivo culminar la síntesis de aquellos fragmentos que no se haya amplificado en su totalidad.

La PCR de colonia presenta la particularidad de que la reacción de PCR no es llevada a cabo a partir de ADN aislado, sino que el material de partida es una colonia bacteriana. Se toma un pequeño volumen de la misma y se lo deposita en el tubo de PCR. Se debe tener particular precaución de que la cantidad de células de partida sean mínimas, de lo contrario la reacción de PCR se ve inhibida. En una primera instancia, se agrega la mitad del volumen final de reacción, sin la Taq, y se incuba 20 minutos a 95°C. Esto permite la lisis celular y la consecuente liberación del ADN. Posterior a este paso, se lleva a volumen final la reacción y se incluye la Taq. Ésta no se introduce desde un inicio para no generar un desgaste térmico y que no disminuya su eficiencia.

Como control positivo se utiliza un plásmido conteniendo un gen flanqueado por los mismos cebadores utilizados para el resto de las reacciones de PCR (pBSK-*cbh II* en este caso), mientras que para el control negativo se adiciona agua en lugar de ADN. El control positivo nos permite verificar que todos los componentes de la reacción se encuentren en correctas condiciones, y por tanto de no haber amplificación en las muestras de interés se debe tomar este resultado como negativo. Puede darse el caso de una PCR en la que no se observe amplificación de la muestra, pero el control positivo funcione, lo cual podría adjudicarse a inhibidores presentes en la muestra. Por otro lado, el control negativo nos permite determinar si se están dando reacciones inespecíficas. Esto implicaría que los cebadores

hibridan con secuencias parcial o total pero que no flanquean al gen de interés. Esto da lugar a otros productos de amplificación que se observan posteriormente en el gel de agarosa como bandas de distinto tamaño al deseado.

La reacción de PCR se realiza utilizando los cebadores T3 y T7 que flanquean el *multiple cloning site* (MCS) de pBSK donde se presume se insertó la secuencia de interés. Se utiliza el software <https://primer3plus.com/> para determinar la temperatura de *melting* de los cebadores. La temperatura utilizada en la fase de hibridación depende de dicho parámetro, siendo aproximadamente 5 unidades inferior a la T_m. Dado que la T_m de T7 es muy baja (ver Tabla 3.5), se decide utilizar una temperatura de hibridación de 45°C (ver Tabla 3.4). Una temperatura de hibridación muy baja da lugar a la unión inespecífica de los cebadores en otras regiones del ADN; mientras que una temperatura muy elevada impide la unión de los cebadores [81].

Tabla 3.3. Concentraciones y volúmenes de los reactivos utilizados en la reacción de amplificación.

Reactivos	Concentración en stock	Volumen (μL)
<i>Buffer</i>	10X	2
dNTP's	5 mM	0,8
Cebador T3	10 μM	0,4
Cebador T7	10 μM	0,4
Top Taq	5 U/μL	0,06
H ₂ O	-	16,34
V _f		20

Tabla 3.4. Condiciones de ciclado para la amplificación del gen *gh1* y *cdt*.

Etapas	Temperatura (°C)	Tiempo (minutos)	Número de Ciclos
Desnaturalización inicial	94	10	1
Desnaturalización	94	0.5	30
Hibridación	45	0.5	30
Extensión	72	1.5	30
Extensión final	72	10	1

Tabla 3.5. Secuencias y T_m de los cebadores utilizados.

Gen a amplificar	Cebadores	Secuencia Nucleotídica	T _m (°C)
<i>gh1</i>	T3 - Directo	5'-GCAATTAACCCTCACTAAAGG-3'	55,7
	T7 - Reverso	5'-TAATACGACTCACTATAGGG-3'	47,1
<i>cdt</i>	T3 - Directo	5'-GCAATTAACCCTCACTAAAGG-3'	55,7
	T7 - Reverso	5'-TAATACGACTCACTATAGGG-3'	47,1

En la figura 3.5 se resume la metodología de trabajo expuesta hasta el momento.

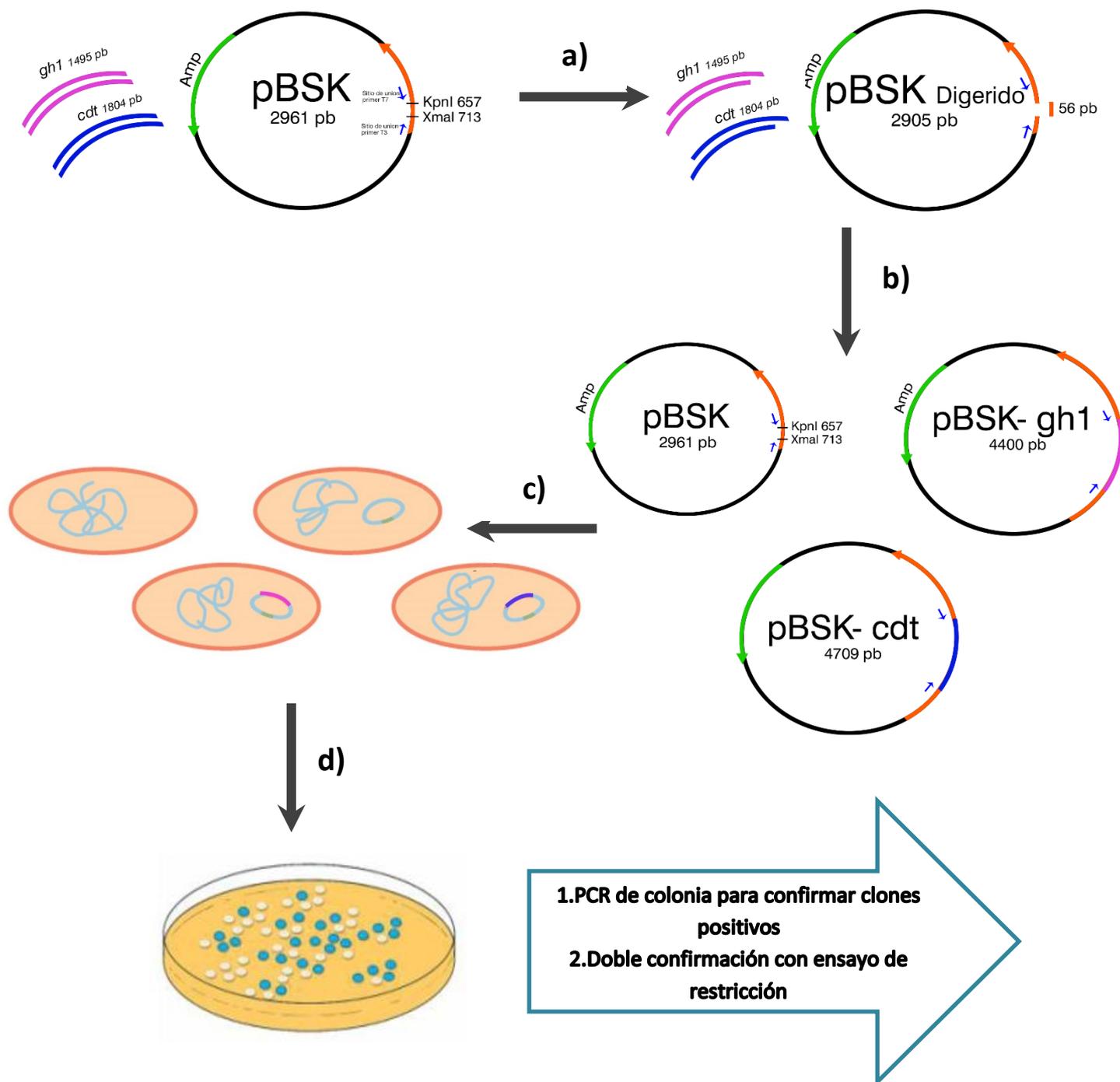


Figura 3.5. Resumen de la metodología de trabajo expuesta. a) Doble digestión con las enzimas de restricción *XmaI/KpnI*, permitiendo la linealización del pBSK y la formación de extremos cohesivos, tanto en el pBSK como en los fragmentos de ADN que contienen *gh1* y *cdt*. b) Ligación del pBSK con los insertos con la enzima T4 DNA Ligasa. También puede suceder que el pBSK se cierre sobre sí mismo, lo que otorgaría a las cepas transformadas resistencia a la ampicilina, pero no se vería interrumpido el *LacZ*, observándose las colonias de coloración azul. c) Transformación química de las células competentes de *E. coli* con el producto de ligación. d) Selección de colonias blancas como presuntos clones positivos y confirmación de los mismos.

3.12. Secuenciación

Con el fin de corroborar los presuntos clones positivos y cerciorarse que en ellos no hubiera ninguna mutación que pudiera afectar la expresión y actividad de la proteína, se secuencia el inserto. Se utiliza el servicio de secuenciación de Macrogen, Corea. Se secuencia en ambas direcciones utilizando los cebadores T3 y T7, con el fin de cubrir todo el gen.

El análisis de la secuencia obtenida se realiza mediante el algoritmo BLAST del programa de libre acceso NCBI (Centro Nacional para la Información Biotecnológica), donde se compara la secuencia clonada en el pBSK con la secuencia del inserto que en un comienzo se manda a sintetizar.

3.13. Electroforesis de proteínas por SDS- PAGE

Se analiza la expresión de proteínas clonadas en el vector pTEF1 en *S. cerevisiae* por otros integrantes del grupo de trabajo conteniendo genes codificantes para enzimas celulolíticas. Entre estos genes encontramos a una β – glucosidasa de *Thermobifida fusca* (BgICF), recodificada para su expresión en levaduras; una xilanasa de *Aspergillus niger* (XynB); y dos celobiohidrolasas: la celobiohidrolasa I de *Rasamsonia emersonii* (CbhI) la cual hidroliza desde el extremo reductor y la celobiohidrolasa II de *Myceliophthora thermophila* (CbhII) que hidroliza desde el extremo no reductor. Los genes de todas estas enzimas fueron modificados con la secuencia XYN, encargada de la excreción de las proteínas al medio extracelular.

Se realizan electroforesis en geles de poliacrilamida en condiciones desnaturizantes. Las mismas son proporcionadas por el dodecil sulfato de sodio (SDS), un detergente aniónico que desnaturiza las proteínas alterando su estructura secundaria y terciaria. Confiere además una carga negativa a las proteínas, proporcional a su masa, permitiendo su separación en un campo eléctrico según su peso molecular. Cuanto mayor sea la masa molecular de la proteína mayor será su carga negativa debido a la unión del SDS. Esto hará que su migración se vea impedida por su tamaño y por su carga.

Se utilizan geles discontinuos, compuestos por un gel de apilamiento y un gel de resolución (ver anexo). La polimerización de la acrilamida es posible gracias a la adición de persulfato de amonio (APS) y N,N,N',N'-tetrametiletlenodiamina (TEMED).

Para la realización del ensayo para cada una de las cepas analizadas se inocula una colonia aislada en 5mL de YPD líquido (ver anexo) y las células de dicho cultivo se concentran a las 16hs mediante centrifugación 2 minutos a 5000 RPM. Se re-suspenden los *pellets* bacterianos en un volumen tal de *buffer* de unión (ver anexo) que permite obtener una $Ab_{S620} = 2$, $Ab_{S620} = 5$ y $Ab_{S620} = 10$. De dicha solución se toman 25 μ L, a los cuales se agregan 25 μ L de *buffer* de carga (ver anexo) y se incuban 5 min a 100 °C en la campana de extracción. Se siembran 20 μ L

de muestra en el gel de poliacrilamida y se corre a 30 mA en *buffer* de corrida (ver anexo) durante 1 hora y 45 minutos.

Finalmente, los geles se teñen durante 2 hs con la solución de teñido (ver anexo) que contiene el colorante Azul de Coomassie. El complejo colorante-proteína presenta un máximo de absorbancia a 549nm a pH=3, siendo posible su cuantificación en un espectrofotómetro para proteínas en solución. La detección por teñido con azul de coomassie tiene un límite de detección de entre 0,1 – 0,5 µg de proteína [82]. Si bien este método no es tan sensible como la tinción con plata, involucra muchos menos pasos [82]. Su incorporación es proporcional a la cantidad de proteína, permitiendo detectar bandas de hasta 100 ng de proteína. Se trata de un derivado de trifenilmetano (ver Figura 3.6) capaz de unirse a través de interacciones electroestáticas con aminoácidos básicos protonados [83]. La tinción no se une al gel de poliacrilamida con alta afinidad, pero penetra a la matriz del gel y establece interacciones inespecíficas, haciendo necesaria una etapa de desteñido del gel, la cual procede durante 2 hs con solución de desteñido (ver anexo). Al desteñir el gel, solo aquellas moléculas a las cuales el colorante se encuentre unidas en forma específica conservarán la coloración.

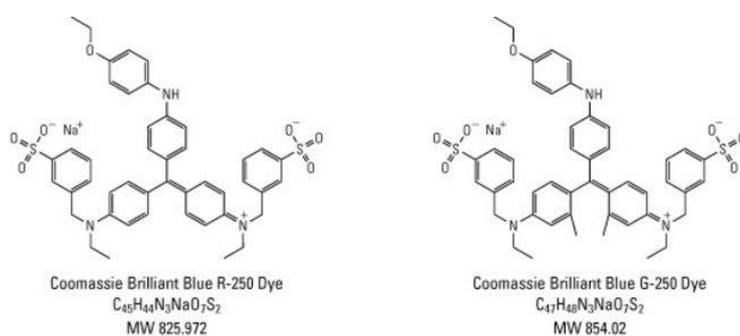


Figura 3.6. Estructura química de los colorantes azul de Coomassie R250 y G250.

El peso molecular de las proteínas se estima en comparación con un estándar de peso molecular PageRuler™ Unstained Protein Ladder 26614 (ThermoFisher) (ver Figura 3.7), sembrándose 10 µl del mismo en cada gel.

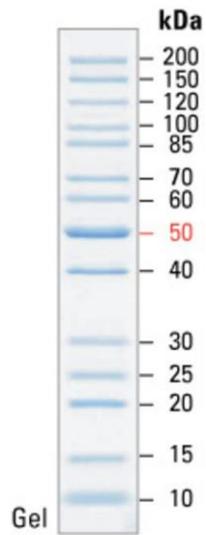


Figura 3.7. PageRuler™ Unstained Protein Ladder 26614 (Thermo Fisher). Patrón de bandas obtenido al realizar una electroforesis SDS-PAGE, 30mA durante 45 minutos. Los pesos moleculares de las diferentes bandas se indican en KDa.

4. RESULTADOS

4.1. Construcción del plásmido pBSK-inserto

Los insertos *gh1* y *cdt* y el pBSK se digieren con *XmaI* y *KpnI* y purificadon a partir de un gel de agarosa (datos no mostrados). Se realiza la ligación con la enzima T4 ADN Ligasa, generándose de esta manera el plásmido pBSK conteniendo el inserto. Se transforman células *E. coli* DH5 α y se seleccionan aquellas colonias blancas como posibles clones con los insertos de interés.

4.2. Screening de los clones positivos por PCR de colonia

Las colonias blancas se someten a PCR de colonia con el fin de comprobar la presencia del gen de interés por aparición de un producto de amplificación (Figura 4.1).

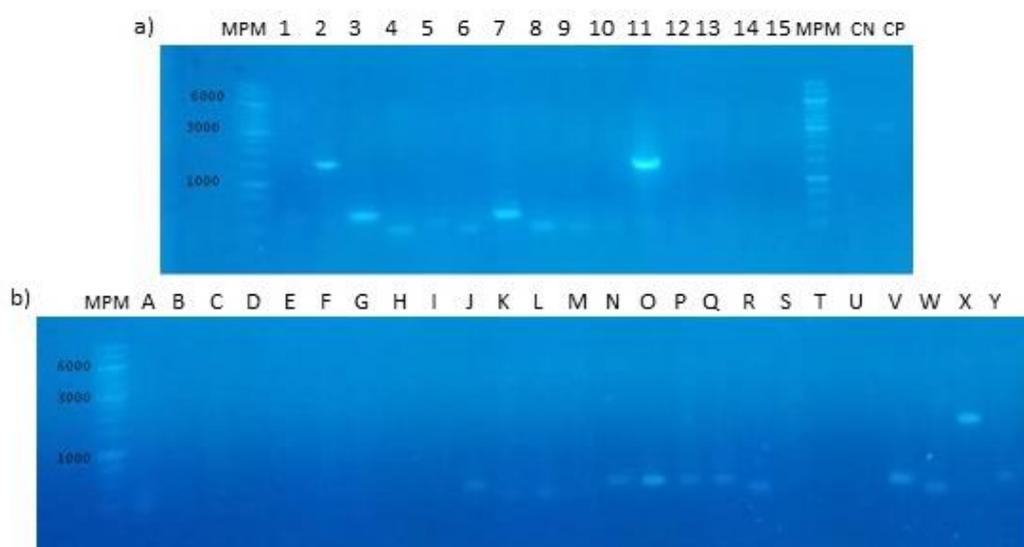


Figura 4.1. Electroforesis en gel de agarosa de PCR de colonia; (a) pBSK-*gh1* carriles 1-15; (b) pBSK-*cdt*, carriles A-Y. MPM: Marcador de peso molecular GeneRuler 1Kb; CP: Control Positivo pBSK-*cbhII*; CN: Control Negativo (agua).

Como se observa en la figura 4.1, solo una de las colonias blancas (colonia X) tiene el tamaño esperado para el gen *cdt*, y solo las colonias 2 y 11 presentan el tamaño esperado para el gen *gh1*. Los tamaños esperados son 1495 pb para *gh1* y 1804 pb en el caso de *cdt*. El control positivo, realizado con el pBSK-*cbh II*, se observa sumamente tenue, pero del tamaño esperado (1513pb). Las bandas presentes a la altura del frente de corrida se trata de los dNTPs, los cuales son agregados en exceso a la reacción de PCR. En el carril correspondiente al control negativo no se observa amplificación. Este control se genera al no agregar ADN molde a la reacción de PCR.

4.3. Ensayo de restricción para comprobar la presencia de clones positivos

Con el fin de confirmar la presencia del inserto, se siembran primarios de las colonias que hayan dado lugar a un producto de amplificación del tamaño esperado en la reacción de PCR. Al día siguiente se realiza una extracción del ADN plasmídico y este se somete a un análisis de restricción. Inicialmente este análisis se lleva a cabo con las enzimas *BamHI* y *XhoI*, las cuales presentan sitios de corte en los extremos de los genes a clonar. Por tanto, se esperaba obtener una banda de 2939 pb correspondiente al pBSK digerido previo a la ligación y otra de 1461 o 1770 pb correspondiente a *gh1* o *cdt* respectivamente. Tras la incubación, el producto de restricción se corre en un gel de agarosa para comprobar dichos patrones de banda (Figura 4.2).

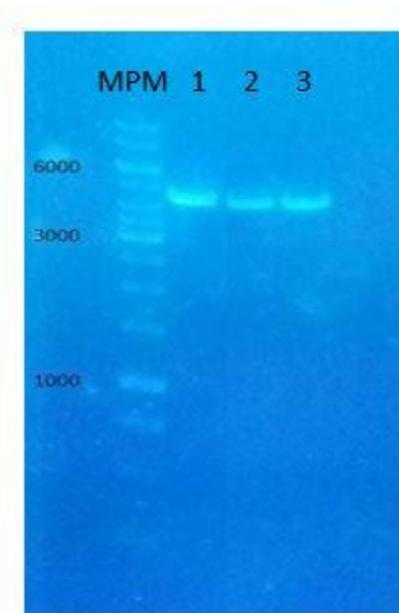


Figura 4.2. Electroforesis en gel de agarosa del ensayo de doble digestión con *BamHI* y *XhoI* del producto de clonado en pBSK. Carriles 1 y 2: vector pBSK-*gh1*; Carril 3: vector pBSK-*cdt*; MPM: Marcador de peso molecular GeneRuler 1Kb DNA Ladder.

El producto de restricción de los plásmidos obtenidos a partir de las colonias presuntamente positivas, que habían amplificado en la PCR de colonia, presentan un tamaño de banda de 4500 pb aproximadamente. Esto sugiere en un primer momento que el pBSK contiene el inserto y resulta linealizado al someterlo a las enzimas de restricción. Sin embargo, el tratamiento con las enzimas de restricción debería liberar el inserto dando lugar a dos bandas, como se explica anteriormente. Estos resultados son tomados como negativos y se vuelve a comenzar el proceso de clonado.

Se repiten los pasos anteriormente mencionados, desde el clonado de los genes sintéticos *gh1* y *cdt* en el pBSK. Una vez alcanzada la etapa de análisis del producto de restricción, se vuelve a observar una banda de 4500pb para el ADN plasmídico de las colonias presuntamente positivas. Esto sucede en repetidas ocasiones, por lo que se decide realizar la digestión con las mismas enzimas responsables de la digestión del vector y de los insertos, previo a su ligación (*XmaI* y *KpnI*), paralelamente a su digestión con *BamHI/XhoI* (tanto de la línea Anza como de la línea Thermo Scientific) para analizar si las mismas se encuentran activas (ver Figura 4.3).

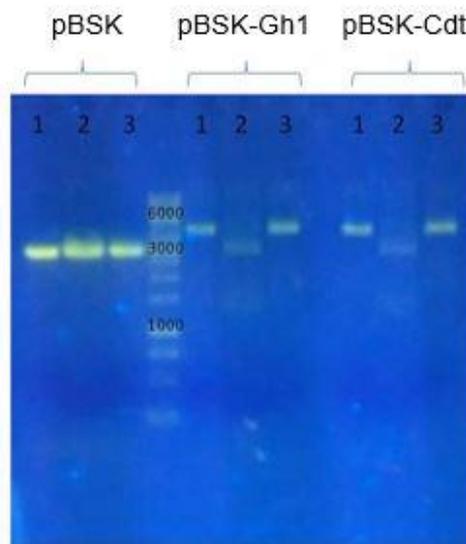


Figura 4.3. Electroforesis en gel de agarosa del ensayo de doble digestión con las enzimas *XhoI* y *BamHI* de la línea Anza, *XmaI* y *KpnI* (Thermo Scientific) y *XhoI* y *BamHI* (Thermo Scientific). Los carriles con número 1 muestran la digestión con *XhoI* y *BamHI* de la línea Anza. Los carriles con número 2 muestran la doble digestión con *XmaI* y *KpnI* (Thermo Scientific). Los carriles con número 3 muestran la doble digestión con *XhoI* y *BamHI* de la línea Thermo Scientific. En cada caso las dobles digestiones se realizaron para el pBSK como control, a pBSK-*gh1* y a pBSK-*cdt*. MPM: Marcador de peso molecular GeneRuler 1Kb DNA Ladder.

Dado que para la doble digestión con *XmaI/KpnI* se libera un fragmento, contrariamente a lo que sucede con el ensayo de restricción de *BamHI/XhoI*, se plantea la hipótesis de que una de estas enzimas no está digiriendo. Para poner a prueba dicha hipótesis, se llevan a cabo ensayos de restricción de las enzimas por separado, con la finalidad de linealizar el plásmido en cuestión (ver Figura 4.4).

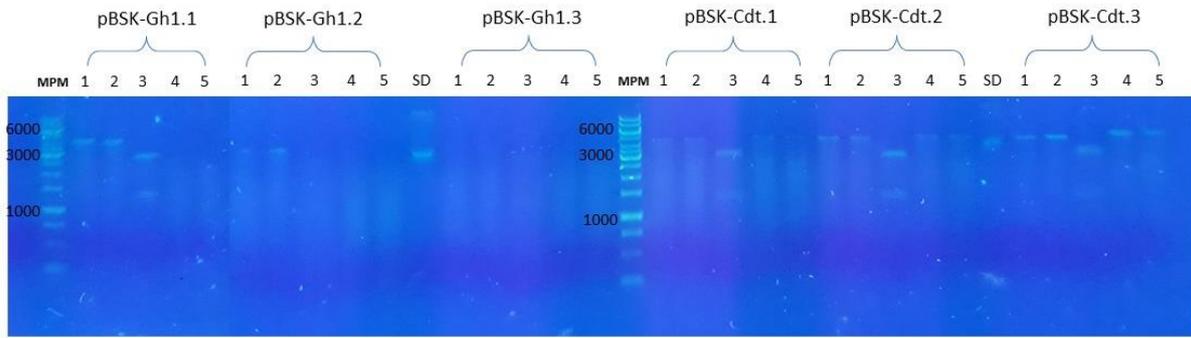


Figura 4.4. Electroforesis en gel de agarosa del ensayo restricción del pBSK-inserto obtenido por miniprep con las enzimas *XhoI*, *BamHI* de la línea Anza, *XmaI/KpnI*, *XhoI* y *BamHI* (Thermo Scientific). Los carriles número 1 muestran la digestión con *XhoI* de la línea Anza. Los carriles número 2 muestran la digestión con *BamHI* (línea Anza). Los carriles número 3 muestra la doble digestión con *XmaI* y *KpnI* (Thermo Scientific). Los carriles número 4 muestran la digestión con *XhoI* (Thermo Scientific). Los carriles número 5 muestran la digestión con *BamHI* (Thermo Scientific). MPM: Marcador de peso molecular GeneRuler 1Kb DNA Ladder; SD: pBSK-inserto sin digerir.

Se puede observar que las enzimas *BamHI* y *XhoI*, tanto de la línea Thermo Scientific como de la línea Anza, son enzimáticamente activas bajo las condiciones del ensayo (iguales condiciones que las utilizadas para la doble digestión).

4.4. Análisis bioinformático

Se decide enviar a secuenciar el inserto clonado en el pBSK de los clones presuntamente positivos. La comparación de las secuencias obtenidas con los genes de interés indica que no existe homología. Sin embargo, al alinear las secuencias clonadas purificadas mediante miniprep entre ellas mismas, se observa que las mismas presentan un Query cover del 98%, una homología del 99% y un e-value de cero (lo cual nos informa lo significativo del alineamiento). Las diferencias halladas entre dichas secuencias se encuentran al principio y al final de la misma, donde el cromatograma nos muestra que la secuenciación no es informativa (ver Figura 4.5).

Se plantea la hipótesis de que se trate de una contaminación proveniente de un transposón. Para poner a prueba dicha hipótesis, se analiza la secuencia obtenida en el software VecScreen. Este análisis muestra que la secuencia obtenida presenta un Query cover del 100% desde la posición 118 a la posición 1150, un porcentaje de homología del 99%, y un e-value de cero con la secuencia AF310136.1, que contiene una transposasa. Del nucleótido 19 al 113 existe homología con el vector de clonado pBSK.

Se analizó la secuencia en el software NEBCutter de BioLabs, donde se desprende la presencia de marco de lectura abierto (ORF) de 312 aminoácidos de la posición 126 a la posición 1064 de la secuencia (ver Figura 4.5).



Figura 4.5. Marco de lectura abierto (ORF) detectado por el software NEBCutter para la secuencia clonada en pBSK. El mismo se extiende desde el nucleótido 126 hasta el 1064.

4.4.1. Análisis aminoacídico de la secuencia

Se realiza un blastx, el cual nos permite traducir el ORF a proteína y posteriormente se lo compara con la base de datos de proteínas en busca de secuencias homólogas. Para la traducción de la secuencia nucleotídica a aminoácidos es posible especificar la especie de la cual proviene el ADN dado que existen excepciones al código genético, donde se selecciona “Bacteria and Archaea”. La secuencia nucleotídica en cuestión (de la posición 126 a la posición 1064) corresponde en un 100% a una transposasa de la familia IS4 presente en microorganismos como *E. coli*, *Klebsiella pneumoniae* y *Salmonella enterica* (ver Figura 4.6). A su vez, el software muestra la secuencia clonada traducida a proteína alineada con la secuencia aminoacídica de la transposasa de la familia IS4 (ver figura 4.7).

Alignments Download GenPept Graphics ⚙							
	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	IS4 family transposase [Klebsiella pneumoniae]	677	677	94%	0.0	100%	WP_095463671.1
<input type="checkbox"/>	IS4 family transposase [Salmonella enterica]	677	677	94%	0.0	100%	WP_079830651.1
<input type="checkbox"/>	IS4 family transposase [Escherichia coli]	677	677	94%	0.0	100%	WP_061856962.1

Figura 4.6. Alineamientos del blastx. Se observa una homología del 100% entre la secuencia en estudio y la transposasa de la familia IS4 de varias especies bacterianas.

IS4 family transposase, partial [Escherichia coli]
 Sequence ID: [WP_095465503.1](#) Length: 383 Number of Matches: 1
[▶ See 1 more title\(s\)](#)

Range 1: 72 to 383 [GenPept](#) [Graphics](#) ▼ Next Match ▲ Previous Match

	Score	Expect	Method	Identities	Positives	Gaps	Frame
	647 bits(1669)	0.0	Compositional matrix adjust.	312/312(100%)	312/312(100%)	0/312(0%)	-1
Query	1064		MPIVLVDWSDIREQKRLMVLRASVALHGRSVTLYEKAFPLSEQCKKAHQDLADLASIL		885		
Sbjct	72		MPIVLVDWSDIREQKRLMVLRASVALHGRSVTLYEKAFPLSEQCKKAHQDLADLASIL		131		
Query	884		PSNTTPLIVSDAGFKVPWYKVEKLGWYWLSRVRGKVQYADLGAENWKPISNLHDMSSSH		705		
Sbjct	132		PSNTTPLIVSDAGFKVPWYKVEKLGWYWLSRVRGKVQYADLGAENWKPISNLHDMSSSH		191		
Query	704		SKTLGYKRLTKSNPISQCILLYKRSRSGRKNQRSTRTHCHHPSPKIYSASAKEPWLATN		525		
Sbjct	192		SKTLGYKRLTKSNPISQCILLYKRSRSGRKNQRSTRTHCHHPSPKIYSASAKEPWLATN		251		
Query	524		LPVEIRTPKQLVNIYSKRMQIEETFRLDKSPAYGLGLRHSRTSSSERFDIMLLIALMLQL		345		
Sbjct	252		LPVEIRTPKQLVNIYSKRMQIEETFRLDKSPAYGLGLRHSRTSSSERFDIMLLIALMLQL		311		
Query	344		TCWLAGVHAQKQGWKDFQANTVRNRNVLSTVRLGMEVLRHSGYITREDLLVAATLLAQ		165		
Sbjct	312		TCWLAGVHAQKQGWKDFQANTVRNRNVLSTVRLGMEVLRHSGYITREDLLVAATLLAQ		371		
Query	164		NLFTHGYALGKL	129			
Sbjct	372		NLFTHGYALGKL	383			

Figura 4.7. Alineamiento de la secuencia aminoacídica clonada con la secuencia de la transposasa de la familia IS4 de *E. coli*. La secuencia de estudio presenta un 100% de homología con la secuencia aminoacídica de la transposasa de la familia IS4.

La secuencia nucleotídica en su totalidad se analiza mediante el software ORF Finder (Open Reading Frame Finder) del NCBI, especificando que se trata de un código genético de bacteria. El resultado arrojado por el sistema (ver Figura 4.8) muestra dos marcos de lectura abierto solapantes.

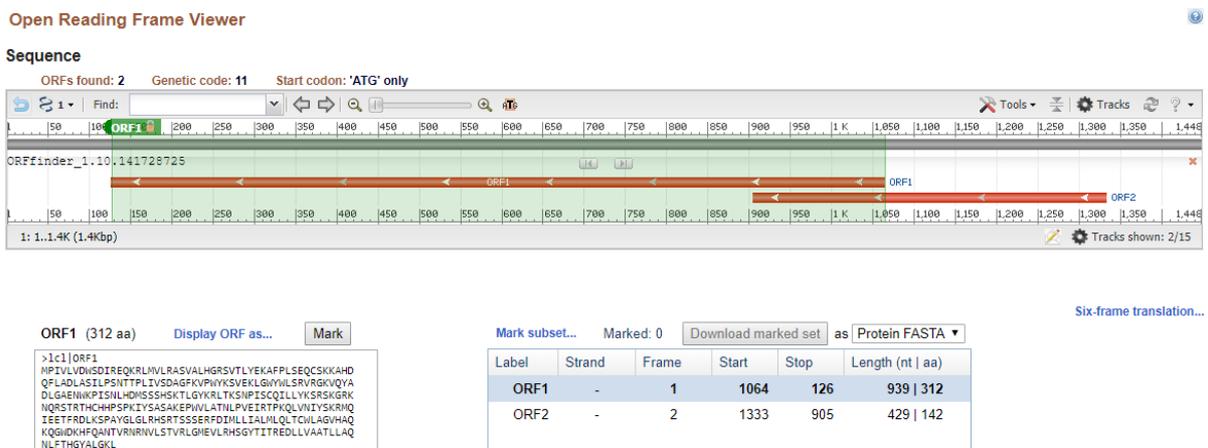


Figura 4.8. Visualización de los marcos de lectura identificados por el software ORF finder de NCBI, para la secuencia de trabajo. Nótese que se observan dos marcos de lectura, los cuales solapan.

4.4.2. Verificación de la identidad de la secuencia clonada

Se realiza un blast en la base de datos IS Finder [84], lo cual permite corroborar que la secuencia clonada en el pBSK se trata de un IS4, lo cual concordaría con lo indicado por los resultados obtenidos anteriormente por el blastx. Este análisis también muestra que la secuencia analizada se trata de una secuencia de inserción de la familia de IS4, subfamilia IS10 (ver Figura 4.8).

```
> IS10R Family: IS4
Length=1329

Score = 2573 bits (1298), Expect = 0.0
Identities = 1312/1314 (99%), Gaps = 2/1314 (0%)
Strand=Plus/Minus

Query 113 CTGAGAGATCCCCATAATTTCCCAAAGCGTAACCATGTGTGAATAAATTTGAGCTA 172
      |||
Sbjct 1329 CTGAGAGATCCCCATAATTTCCCAAAGCGTAACCATGTGTGAATAAATTTGAGCTA 1270

Query 173 GTAGGGTTGCAGCACGAGTAAGTCTTCCCTGTATTGTGTAGCCAGAATGCCGCAAAA 232
      |||
Sbjct 1269 GTAGGGTTGCAGCACGAGTAAGTCTTCCCTGTATTGTGTAGCCAGAATGCCGCAAAA 1210

Query 233 CTTCCATGCCTAAGCGAAGCTTTGAGAGTACGTTTCGATTTCTGACTGTGTTAGCCTGGA 292
      |||
Sbjct 1209 CTTCCATGCCTAAGCGAAGCTTTGAGAGTACGTTTCGATTTCTGACTGTGTTAGCCTGGA 1150

Query 293 AGTGCTTGTCCCAACCTTGTCTTCTGAGCATGAACGCCCGCAAGCCCAACATGTTAGTGAA 352
      |||
Sbjct 1149 AGTGCTTGTCCCAACCTTGTCTTCTGAGCATGAACGCCCGCAAGCCCAACATGTTAGTGAA 1090

Query 353 GCATCAGGGCGATTAGCAGCATGATCAAAAAGCTCTGAGCTGCTGTTCCGGCTATGGC 412
      |||
Sbjct 1089 GCATCAGGGCGATTAGCAGCATGATCAAAAAGCTCTGAGCTGCTGTTCCGGCTATGGC 1030

Query 413 GTAGGCTAGTCCGTGAGCAGGACTTTTCAAGTCTCGGAAGGTTTCTCAATCTGCATTC 472
      |||
Sbjct 1029 GTAGGCTAGTCCGTGAGCAGGACTTTTCAAGTCTCGGAAGGTTTCTCAATCTGCATTC 970

Query 473 GCTTCGAATAGATATTAAACAAGTTGTTGGGTGTTTCGAATTTCAACAGGTAAGTTAGTTG 532
      |||
Sbjct 969 GCTTCGAATAGATATTAAACAAGTTGTTGGGTGTTTCGAATTTCAACAGGTAAGTTAGTTG 910

Query 533 CTAGAACCCTATGGCTCTTTTCCGACGCTGAGTAGATTTTAGGTGACGGGTGGTACAAAT 592
      |||
Sbjct 909 CTAGAACCCTATGGCTCTTTTCCGACGCTGAGTAGATTTTAGGTGACGGGTGGTACAAAT 850

Query 593 GAGTCCGTGTCGAGCGCTGATTTTTTCGGCTTTAGAGCGAGATTTATACAATAGAATTT 652
      |||
Sbjct 849 GAGTCCGTGTCGAGCGCTGATTTTTTCGGCTTTAGAGCGAGATTTATACAATAGAATTT 790

Query 653 GGCATGAGATTGGATTGCTTTTAGTCAGCCTCTTATAGCCTAAAGCTTTGAGTGAATAG 712
      |||
Sbjct 789 GGCATGAGATTGGATTGCTTTTAGTCAGCCTCTTATAGCCTAAAGCTTTGAGTGAATAG 730

Query 713 ATGACATATCATGTAAGTTGCTGATAGGTTTCCAGTTTTCCGCTCCTAGGCTGCATATT 772
      |||
Sbjct 729 ATGACATATCATGTAAGTTGCTGATAGGTTTCCAGTTTTCCGCTCCTAGGCTGCATATT 670

Query 773 GTACTTTTCCCTTACTCGACTTAACAGTACCAACCCAGCTTCTCAACGGATTTATACC 832
      |||
Sbjct 669 GTACTTTTCCCTTACTCGACTTAACAGTACCAACCCAGCTTCTCAACGGATTTATACC 610

Query 833 ATGGCACTTTAAAGCCAGCATCACTGACAATGAGCGGTGTGGTGTACTCGGTAGAATGC 892
      |||
Sbjct 609 ATGGCACTTTAAAGCCAGCATCACTGACAATGAGCGGTGTGGTGTACTCGGTAGAATGC 550

Query 893 TCGCAAGGTCGGCTAGAAATGGTCAAGCTTTCTTTGAACATTGCTCGAAAGCGGGA 952
      |||
Sbjct 549 TCGCAAGGTCGGCTAGAAATGGTCAAGCTTTCTTTGAACATTGCTCGAAAGCGGGA 490

Query 953 ACGCTTCTCATAAAGAGTAACAGAACCAGCGTGTAGTGGCACTGAAGCTCGCAATACCA 1012
      |||
Sbjct 489 ACGCTTCTCATAAAGAGTAACAGAACCAGCGTGTAGTGGCACTGAAGCTCGCAATACCA 430

Query 1013 TAAGTCGTTTTTGTCTCAGCAATATCAGACCAGTCAACAAGTACAATGGCATCGTATTGC 1072
      |||
Sbjct 429 TAAGTCGTTTTTGTCTCAGCAATATCAGACCAGTCAACAAGTACAATGGCATCGTATTGC 370

Query 1073 CCGAACAGATAAAGCTAGCATGCCAACGGTAT -CAGCGAGTCGCTCTTTGTGGAGGTGAC 1131
      |||
Sbjct 369 CCGAACAGATAAAGCTAGCATGCCAACGGTATACAGCGAGTCGCTCTTTGTGGAGGTGAC 310

Query 1132 GATTACCTAACAATCGGTGATTTGATGTTATGTTTTGTTCTCGCTTTGGTTGGCA 1191
      |||
Sbjct 309 GATTACCTAACAATCGGTGATTTGATGTTATGTTTTGTTCTCGCTTTGGTTGGCA 250

Query 1192 GGTACGGCCAAAGTTCGGTAAGAGTGAGAGTTTTACAGTCAAGTAAATGCGTGGCAAGCCA 1251
      |||
Sbjct 249 GGTACGGCCAAAGTTCGGTAAGAGTGAGAGTTTTACAGTCAAGTAAATGCGTGGCAAGCCA 190

Query 1252 ACGTTAAGCTGTTGAGTGGTTTTAAGTGTAAATCGGGGCAAGATTTGGTAAAGAGTCTG 1311
      |||
Sbjct 189 ACGTTAAGCTGTTGAGTGGTTTTAAGTGTAAATCGGGGCAAGATTTGGTAAAGAGTCTG 130

Query 1312 GTAAAAATCGAGTTCGCACACTCTGTTGTCTGATTTGATTTTTTCGCGAAACCAATTTG 1371
      |||
Sbjct 129 GTAAAAATCGAGTTCGCACACTCTGTTGTCTGATTTGATTTTTTCGCGAAACCAATTTG 70

Query 1372 ATCATATGACAAGATGTGTATCCACCTTAACCTTAATGATTTT -ACCAAAATCAT 1424
      |||
Sbjct 69 ATCATATGACAAGATGTGTATCCACCTTAACCTTAATGATTTTACCAAAATCAT 16
```

Figura 4.9. Alineamiento de la secuencia clonada en el pBSK con la secuencia presente en la base de datos del IS Finder de un elemento IS4 de la subfamilia IS10. El mismo presenta un 99% de coincidencias desde la posición 108 de la secuencia clonada hasta la posición 1424.

De acuerdo con este software, la secuencia presenta un solo ORF desde la posición 108 hasta la posición 1316 (ver Figura 4.10).

Protein section		ORF number : 1			
ORF 1					
Length	Begin	End	Strand	Fusion ORF	
1209 bp 402 aa	108	1316	+	No	
ORF function : Transposase Chemistry : DDE ORF sequence : MCELDI LHD SL YQF CPELHLKRLNSL TLACHALLDC KTL TL TELGRNLP TKART KHNKRI DRLLGNRHLH KERLAVYRWHAS FIC SGNTMPI VLVDWSD IREQKR L MVLRASVALHGRSVTLYEKAFPLSEQC SKKAHQD FLADLASI LP SNTTFLIVSDAGFKV FWYKSVKLGWY WLSRVRGK VQYADLGAENWKPI SNLHDMSSSHS KTLGYKRLTKSNP IS CQI LLYKSRSKGRKNQRSTR THCHHPSPKI YSASAKE FWVLA TNLPVEIRTP KQLVN IYSKRMQI EETFRDLKS PAYGLGRHSRTSS SERFDIMLLIALMLQLT CWLAGVHAQKQGWDKHFQANTVRNRNVLSTVRLGMEVLRHSGYTI TREDLLVAAT LLAQNLFTHGVALG KL					

Figura 4.10. ORF reportado por el software IS Finder. En la imagen se indica la secuencia aminoacídica del ORF identificado, así como su posición y largo.

Dicho ORF coincide con los ORF1 y ORF2 reportados por el software ORF finder. Dado que el ORF reportado por el IS Finder coincide con los marcos de lectura abiertos reportados anteriormente, se plantea la hipótesis de que el software ORF finder haya detectado dicho gen como dos marcos de lectura abiertos. Para poner a prueba dicha hipótesis, se analiza el solapamiento de ambas secuencias en el programa informático MEGA7. Se observa que ambos ORF reportados por el ORF finder son secuencias solapantes.

4.4.3. Repetidos de la secuencia clonada

Se identifican los repetidos inversos característicos de las Secuencias de Inserción (ISs) realizando un blast con el software IS Finder (ver Figura 4.11). Los mismos se extienden desde la posición 114 a 163 y desde la posición 1394 a la posición 1436. Sin embargo, no se observan repetidos directos en la secuencia, los cuales se buscaron con el servidor FAIR [85].

IR Length : 17/22

IRL : CTGATGAATCCCCTAATGATTTTGGTAAAAATCATTAAAGTTAAGGTGGAT

IRR : CTGAGAGATCCCCTCATAATTTCCCCAAAGCGTAACCATGTGTGAATAAA

DNA sequence

CTGATGAATCCCCTAATGATTTTGGTAAAAATCATTAAAGTTAAGGTGGATACACATCTTGTGCATATGATCAAAATGGTTTCGCGAAAAATCAATAATCAGACAAACAAGATGTGCGAACTCGATATTTTACACGACTCTCTTTACCAATTTCTGCCCGAATTACACTTAAAACGACTCAACAGCTTAACGTTGGCTTGCCACGCATTACTTGACTGTAARACTCTCACTCTTACCGAACTTGGCCGTAACTGCCAACCRAAGCGAGAACAAAACATAACATCAAAACGAAATCGACCGATTGTAGGTAATCGTCACCTCCACAAAGAGCGACTCGCTGTATACCGTTGGCATGCTAGCTTTATCTGTTCGGGCAATACGATGCCCATTTGTACTTGTGACTGTCTGTATTTCTGAGCAAAAACGACTTATGGTATTGCGAGCTTCAGTCGCACCTACACGGTCTGTTACTCTTTATGAGAAAGCGTTCCCGCTTTCA GAGCAATGTTCAAAGAAAGCTCATGACCAATTTCTAGCCGACCTTGCAGACATTTACCCGAGTAACACCACCCGCTCATTTGTCAGTGTATGCTGGCTTTA AAGTGCCATGGTATAAATCCGTTGAGAAAGCTGGTTGGTACTGGTTAAGTCGAGTAAGAGGAAAAGTACAATATGCAGACCTAGGAGCGGAAAACCTGGAA ACCTATCAGCAACTTACATGATATGTCATCTAGTCACTCAAAGACTTTAGGCTATAAGAGGCTGACTAAAAGCAATCCAAATCTCATGCCAAATTTCTATTG TATAAATCTCGCTCTAAAGGCCGAAAAAATCAGCGCTCGACACGGACTCATTTGTCACCACCCGCTCACCTAAAATCTACTCAGCGTCGGCAAAAGGAGCCAT GGGTTCTAGCAACTAATCTACCTGTTGARATTCGAACACCCAAACAACTTGTTAATATCTATTGGAAGCGAATGCAGATTGAAGAAACCTTCCGAGACTT GAAAAGTCCTGCCTACGGAAGCTAGGCTAGCCATAGCCGAAACGAGAGCTCAGAGCGTTTGGATATCATGCTGCTAATCGCCCTGATGCTCAACTAACA TGTTGGCTTCCGGCGCTTCATGCTCAGAAACAAAGGTTGGGACAAAGCACTTCCAGGCTAACACAGTCAGAAATCGAAACGCTACTCTCAACAGTTTCGCTTAG GCATGGAAGTTTTCGGCGCTTTCGCTACACAAACAAAGGAAAGACTTACTCGTGGCTGCAACCCCTACTAGCTCAAAATTTTATTTCACACATGGTTACGC TTTGGGGAATTTATGAGGGGATCTCTCAG

Figura 4.11. Identificación de los repetidos inversos en la secuencia clonada. Se observa el repetido inverso izquierdo (IRL) en naranja y el repetido inverso derecho (IRR) en verde.

4.4.4. Determinación del sitio de inserción del elemento transponible

Se verifica también la presencia del sitio de restricción *XmaI* en el extremo 5' de la secuencia clonada utilizando el software NEBCutter. El sitio de restricción de *KpnI* en posición 3' no se observa en la secuencia. Paralelamente se determina la posición en la que ocurre la inserción del elemento transponible. Análisis en el software NEBCutter muestran que la inserción no se dio entre sitio de restricción *XmaI* y *KpnI* dado que se mantiene la secuencia del pBSK entre ambos sitios. En el producto de clonado se identifica la secuencia GGTA C. Sin embargo, el nucleótido terminal de esta secuencia de reconocimiento (C) no se encuentra presente en el producto de clonado. De haberse dado la inserción entre *XmaI* y *KpnI*, el sitio de clonado múltiple (MCS) de pBSK se vería interrumpido. Se realiza un blast entre la secuencia del IS4, subfamilia IS10 hallado en la base de datos IS Finder y la secuencia en estudio, permitiendo determinar que la secuencia de inserción se extiende desde el nucleótido 114 al 1436 (ver Figura 4.12).

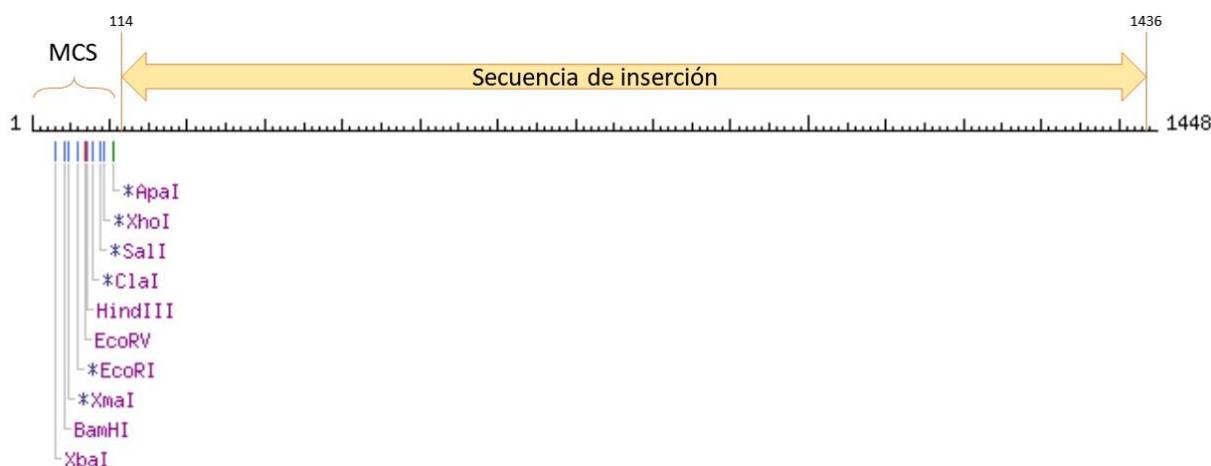


Figura 4.12. Diagrama de la secuencia clonada. Se observa el sitio de clonado múltiple (MCS) intacto y la secuencia de inserción en posición 3' respecto a este. En la imagen no se observa el sitio de restricción de *KpnI* pues al encontrarse en forma parcial no es reconocido por el software NEBCutter.

Analizando el sitio de restricción *KpnI* vemos que al mismo le falta un nucleótido, correspondiente a la base nitrogenada terminal del extremo 3', la cual es separado de la secuencia de reconocimiento al momento de la digestión (ver Figura 4.13).



Figura 4.13. Sitio de reconocimiento y mecanismo de corte de *KpnI*.

4.5. Expresión de las proteínas

Se realiza un gel de electroforesis desnaturalizando al 15% el cual no permite resolver completamente las proteínas de mayor peso molecular. Por tanto, se realiza un gel de poliacrilamida al 10% (ver Figura 4.14). A su vez se ensayan geles a distintas concentraciones de extracto celular ($Abs_{620} = 2$, $Abs_{620} = 5$ y $Abs_{620} = 10$), determinándose que la Abs_{620} que permite la óptima visualización de las proteínas en el extracto celular es $Abs_{620} = 10$.



Figura 4.14. Gel de electroforesis SDS-PAGE al 10% de proteínas totales presentes en el extracto celular total con $Abs_{620} = 10$. Wt: wild type; MPM PageRuler™ Unstained Protein Ladder; BglCA: control interno de cepa capaz de expresar las proteínas en estudio; BglCF: β – glucosidasa de *Thermobifida fusca* recodificada para su expresión en levaduras; XynB: xilanasa de *Aspergillus niger*; CbhI celobiohidrolasa I de *Rasamsonia emersonii*; CbhII: celobiohidrolasa II de *Myceliophthora thermophila*.

No se observa una expresión diferencial de ninguna de las proteínas transformadas. Esto podría deberse a que las mismas no se encuentran en el extracto celular, sino que son excretadas al medio, resultado que se esperaba dado que a los genes que codifican para las proteínas en cuestión se les había adicionado un péptido señal de excreción celular.

5. DISCUSIÓN

El objetivo de este trabajo consistía en otorgarle el fenotipo celulolítico a una cepa de *S. cerevisiae* a través de la incorporación de genes sintéticos que codifican para enzimas celulolíticas. Dicho objetivo no fue alcanzado dado que no fue posible la construcción de un vector que contuviera dichos genes. En su lugar, se clonó una IS4 de la subfamilia IS10.

5.1. Secuencias de Inserción de la familia IS4

Las secuencias de inserción (ISs) son un tipo particular de elementos transponibles, definidos por ser pequeñas unidades móviles de ADN con único gen en su secuencia el cual codifica para una transposasa, flanqueado por repetidos invertidos en los extremos [86], y en algunos casos también repetidos directos (ver Figura 5.1). Esta transposasa es la responsable de la movilidad de la secuencia de inserción dado que interactúa con los repetidos invertidos en los extremos de una manera sitio específica, cortando el ADN en ambos extremos y transfiriéndolo a otra región donde no es necesario que exista homología de secuencia [87].

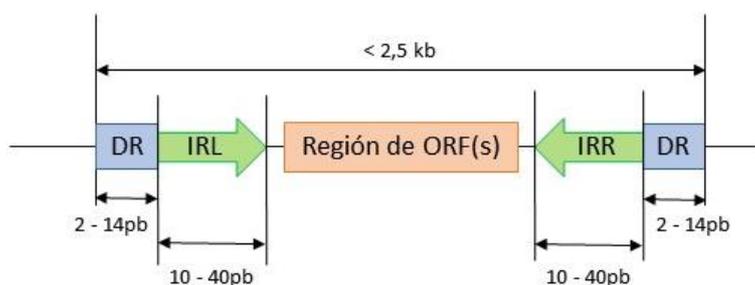


Figura 5.1. Estructura típica de un elemento IS. Los elementos IS contienen las repeticiones directas (DR) que son el sitio blanco de la duplicación, las repeticiones invertidas izquierda y derecha (IRL e IRR) y los marcos de lectura abiertos (ORFs) que codifican la transposasa.

Sin embargo, no es necesariamente cierto que los transposones se integran en forma aleatoria [88]. Se ha reportado que el transposón de levadura Ty1 presenta una fuerte preferencia por integrarse por fuera de los marcos de lectura abiertos [89]. Se ha reportado que el sitio de inserción en la trasposición puede verse influenciado por la topología del ADN en dicha región, la presencia de proteínas del huésped en el sitio de unión; por ejemplo, algunas transposasas evitan regiones transcritas [90].

La familia de las secuencias de inserción IS4 es un grupo bastante heterogéneo clasificado como tal por las similitudes observadas en sus secuencias IRs (Inverted Repeats) [91]. El IS10, puede verse involucrado en el transposón compuesto Tn10, donde comprende un par de secuencias de inserción que flanquean un gen de resistencia a la tetraciclina [92].

Las transposasas de la subfamilia IS10 presentan una tríada aminoacídica conservada en su sitio activo denominada motivo DDE (ver Figura 5.2). Este motivo se encuentra conservado en la mayor parte de las transposasas. El mismo se compone de dos aspartatos y un glutamato.

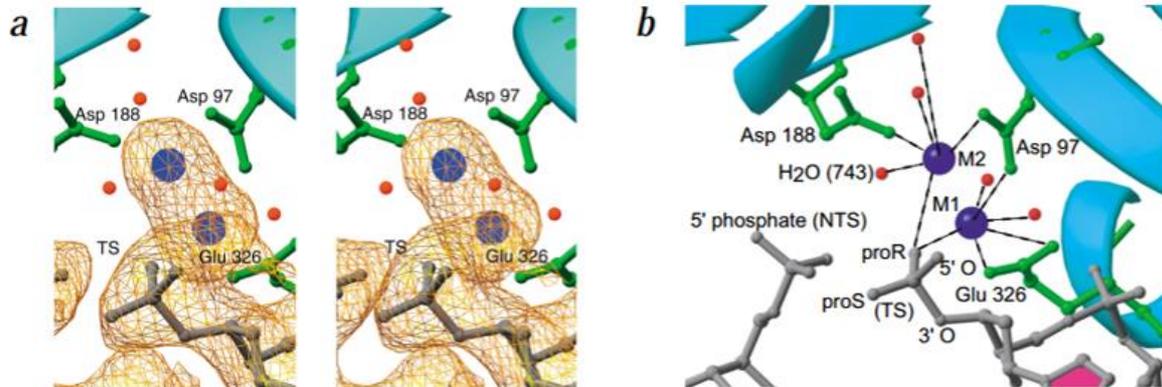


Figura 5.2. Vista del sitio activo con motivo DDE de la transposasa Tn5 de la familia IS4. Se observa en verde la tríada aminoacídica Asp/Asp/Glu coordinados a dos metales divalentes representados en azul (en este caso Mn^{2+}). (a) Mapa de densidad electrónica de los metales divalentes. (b) Coordinación de los metales divalentes en el sitio activo. Las líneas sólidas en negro representan interacciones covalentes y las moléculas de agua se representan como esferas rojas. Tomado de [93].

Los elementos IS se clasifican dentro de la Clase I de elementos transponibles y son constituyentes normales de muchos cromosomas y plásmidos bacterianos, habiéndose descrito tanto en bacterias Gram negativas como Gram positivas [94]. En los extremos de esta secuencia se observan repetidos de entre 2-20 pb, los cuales surgen durante el proceso de transposición, cuando la transposasa genera extremos cohesivos en la molécula de ADN, que posteriormente son reparados por el Sistema de reparación del ADN de la célula huésped (ver Figura 5.3) [95]. La longitud de la duplicación de destino es característica para cada elemento IS. Además se encuentran generalmente flanqueados por repeticiones invertidas, que se cree que son el sitio de reconocimiento y unión para la transposasa durante la transposición (ver Figura 5.1) [96]. Por otra parte, la clase II de elementos transponibles agrupa el conjunto de elementos transponibles ancestralmente relacionados que se conocen como la familia de transposones Tn3, mientras que la clase III cubre el bacteriófago Mu transponible y los fagos relacionados [94].

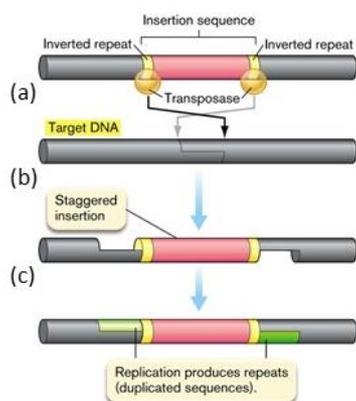


Figura 5.3. Mecanismo de inserción de una secuencia de inserción. (a) La transposasa se une al sitio blanco, donde realiza un corte en la secuencia generando extremos cohesivos. (b) La secuencia de inserción se adhiere a los extremos cohesivos del ADN blanco. (c) Los mecanismos de reparación del ADN de la célula reparan las mellas, generando repetidos en los extremos de la IS.

Algunas secuencias de inserción forman elementos transponibles complejos, encontrándose asociados a una región de ADN que codifica para genes de resistencia antibiótica o patogenicidad, siendo denominados en este caso, transposones [94]. Los transposones están involucrados en la rápida adaptación de las comunidades bacterianas, dada su fácil transferencia horizontal [97], denominándose así al proceso de adquisición de material genético a partir de una fuente distinta que del genoma de la progenie [88]. Además, se encuentran implicadas en el proceso de generación de diversidad genética de una población bacteriana al introducir mutaciones al momento de desplazarse de una región de genoma a otra [87].

La transposición es llevada a cabo mediante la unión de la transposasa al ADN, dando lugar a un corte en la secuencia. Hay incluso evidencias de que la asociación de la transposasa con el ADN también requiere la replicación activa del ADN del huésped [98]. Recientes estudios demuestran que la transposasa se encuentra activa en forma de oligómero; un ejemplo de esto es la transposasa de Mu, la cual en su forma activa es un tetrámero unido a ambos extremos del elemento transponible [99].

El evento de transposición puede verse inducido por condiciones de estrés. Actualmente se ha logrado la eliminación del genoma de ADN móvil y de genes de virulencia críptica de las cepas de *E. coli* K12 MG1655 y W3110 [100]. No se observaron diferencias en la tasa de crecimiento y expresión de proteínas recombinantes, pero éstas cepas permiten una transformación por el método de electroporación más eficiente y pueden propagar plásmidos que resultan inestables en otras cepas [101].

5.2. Repetidos de la secuencia de inserción clonada

Fue posible la identificación de los repetidos inversos en los extremos de la IS4. Los mismos presentan un solapamiento parcial con la secuencia codificante para la transposasa, fenómeno observado para otros IS [102] (ver Figura 5.4).

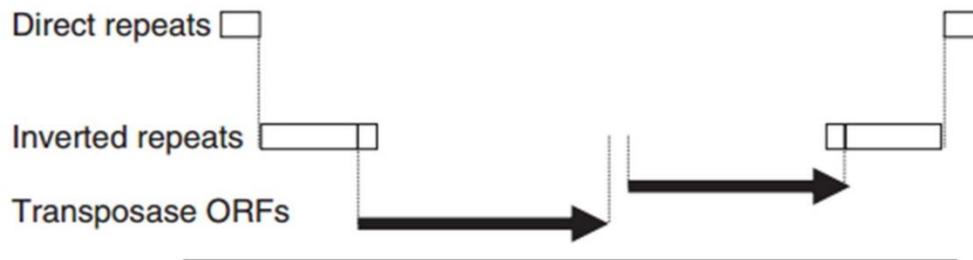


Figura 5.4. Esquema de un elemento IS donde es posible evidenciar el solapamiento parcial entre los repetidos inversos y los ORF. Modificado de [102].

Por otra parte, los repetidos directos no fueron hallados. De acuerdo con lo reportado en la bibliografía, los repetidos directos de los IS4 presentan un largo de entre 9-14pb [103]. Si bien generalmente se generan como producto de la transposición, su presencia no es requisito indispensable para la identificación de una secuencia de inserción [104].

5.3. Sitio de inserción del elemento IS

Dado que el sitio de reconocimiento de *KpnI* se encuentra incompleto, esto nos lleva a pensar que la inserción tuvo lugar en el sitio de restricción de *KpnI* una vez ocurrida la digestión. Además, la secuencia desde el sitio de corte de *KpnI* hacia 3' presenta una homología del 99% con una secuencia de inserción de la familia IS4, subfamilia IS10; mientras que al analizar la secuencia desde la posición 1-113 se observa una homología del 100% con el pBSK II SK (+) (ver Figura 5.5).



Figura 5.5. Sitio de inserción del transposón. Se observa en dirección 5' parte del sitio de clonado múltiple del pBSK con sus respectivos sitios de restricción y en dirección 3', posterior al sitio parcial de restricción de *KpnI* se puede observar parte de la IS4, subfamilia IS10.

Los distintos tipos de IS presentan diferentes niveles de selectividad por la secuencia blanco. Se ha reportado la especificidad de alguno de ellos por secuencias ricas en A-T; mientras que otros tienden a insertarse dentro de otras secuencias de inserción [103]. Se cree que en este

caso la inserción en la posición 114 se dio aprovechando la digestión parcial con la enzima *KpnI*. Dicha hipótesis se apoya en la conservación del sitio de reconocimiento de *XmaI* (el cual de haber sido digerido habría causado la escisión de un fragmento de 56pb en el vector).

5.4. Contaminación del material de trabajo con IS4

Las IS4 son elementos transponibles de uso frecuente en los laboratorios. Esto podría haber dado lugar a una inserción de la IS4 en el genoma de la célula huésped *E. coli* DH5 α wild type. El experimento se repitió en variedad de condiciones y cambiando los reactivos, pero estos no afectaron los resultados obtenidos lo cual apoya nuestra hipótesis. Otros investigadores trabajando con las mismas células obtuvieron la misma contaminación en sus clonados.

Otro factor a tener en cuenta son las condiciones bajo las cuales se tienen a las células competentes. Recordar que las condiciones de estrés aumentan la capacidad transformante de las células como mecanismo de generación de variantes que permitan la supervivencia de las mismas.

5.5. Ligación

Dado que el proceso de ligación con la enzima ADN Ligasa T4 es poco eficiente se podría haber utilizado otro mecanismo para la clonación de *gh1* y *cdt* en pBSK. Una alternativa en creciente expansión es la técnica Gibson Assembly. Se trata de una reacción isotérmica de un solo paso que ensambla rápidamente segmentos de ADN con terminaciones homólogas de entre 15-20pb. Para ello se hace uso de tres enzimas: una 5'-exonucleasa que expone la secuencia sobresaliente para el reconocimiento específico de los segmentos de ADN complementarios, una ADN polimerasa encargada de rellenar los *gaps* y una ADN ligasa que une los segmentos (ver Figura 5.5) [105]. Se trata de una técnica muy robusta que permite la unión de fragmentos de ADN de hasta 600kb [106]. Supone una gran ventaja en comparación con la ligación convencional, donde las incompatibilidades de enzimas de restricción, la necesidad de pasos de ligación intermedios o el tamaño limitado de los segmentos capaces de unirse resultan un obstáculo en la construcción de ensamblados [107].

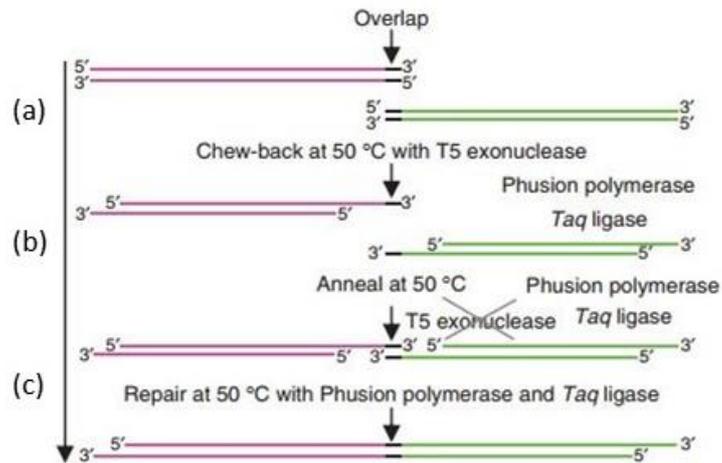


Figura 5.5. Esquema de la técnica Gibson Assembly. Dos fragmentos adyacentes de ADN (violeta y verde) que presentan secuencias terminales solapantes (negro), se ligan mediante una reacción isotérmica de un solo paso. (a) La enzima T5 exonucleasa expone la secuencia saliente para el reconocimiento de los segmentos complementarios de ADN. (b) La ADN polimerasa se encarga de rellenar los gaps. (c) La ADN ligasa sella las mellas. Tomado de [105].

5.6. Expresión de proteínas en el extracto celular

No se observa una expresión diferencial de ninguna de las proteínas transformadas. Esto podría deberse a que las mismas no se encuentran en el extracto celular, sino que son excretadas al medio, resultado que se esperaba dado que a los genes que codifican para las proteínas en cuestión se les había adicionado un péptido señal de expresión celular. Sin embargo, también podría ser consecuencia de una baja expresión. Para probar dicha hipótesis es necesario la realización de una técnica más sensible como es el caso del Western Blot, donde se da una unión específica de la proteína de interés a un anticuerpo marcado. Esto sería posible dado que los genes insertados fueron adicionados con secuencias FLAG.

6. CONCLUSIONES Y PERSPECTIVAS

6.1. Conclusiones

No fue posible la clonación de los genes *gh1* y *cdt* (codificantes de una β -glucosidasa y un transportador de celobiosa respectivamente) en el vector de expresión pBSK. En el sitio de clonado múltiple de pBSK (sobre el sitio de restricción del *KpnI*) se insertó una secuencia de inserción de la familia IS4, de la subfamilia IS10 lo cual se adjudica a la posible contaminación de las cepas de trabajo con dicho transposón.

Ensayos independientes de SDS-PAGE permitieron analizar los extractos celulares de *S. cerevisiae* transformadas con los genes de una β – glucosidasa de *Thermobifida fusca* (BgICF), recodificada para su expresión en levaduras; una xilanasa de *Aspergillus niger* (XynB); y dos celobiohidrolasas, observándose que las proteínas no se encuentran presentes en la fracción celular o la misma no se expresa en las cepas transformadas.

6.2. Perspectivas

Dado que no fue posible alcanzar los resultados esperados, a futuro nos proponemos:

1. La repetición del trabajo con un huésped en condiciones apropiadas (el cual no presente contaminaciones por elementos transponibles), permitiendo analizar el clonado del sistema β -glucosidasa/transportador de dextrinas.
2. La modificación del sistema de selección dado por el vector. Una alternativa es la implementación del sistema de selección URA3, lo cual evitaría falsos positivos al momento de analizar las placas.
3. Dado que la reacción de ligación es altamente ineficiente, esto podría resultar un factor limitante al momento de clonar los genes de interés en el pBSK. Por tanto podría ensayarse la unión del vector pBSK y el inserto mediante la técnica de Gibson Assembly.
4. Sería pertinente la realización de SDS-PAGE (así como técnicas con mayor límite de detección como es el caso del Western Blot) para analizar la expresión de las proteínas en el sobrenadante previamente concentrado y la realización de ensayos de actividad enzimática para analizar no solo si las enzimas se expresan, sino que las mismas se encuentran activas.

7. BIBLIOGRAFÍA

- [1] M. Vohra, J. Manwar, R. Manmode, S. Padgilwar, and S. Patil, "Bioethanol production: Feedstock and current technologies," *J. Environ. Chem. Eng.*, vol. 2, no. 1, pp. 573–584, Mar. 2014.
- [2] M. V. Rodionova *et al.*, "Biofuel production: Challenges and opportunities," *Int. J. Hydrogen Energy*, vol. 42, no. 12, pp. 8450–8461, Mar. 2017.
- [3] H. Zabed, J. N. Sahu, A. Suely, A. N. Boyce, and G. Faruq, "Bioethanol production from renewable sources: Current perspectives and technological progress," *Renew. Sustain. Energy Rev.*, vol. 71, pp. 475–501, May 2017.
- [4] K. Day-Rubenstein *et al.*, "Research Investments and Market Structure in the Food Processing, Agricultural Input, and Biofuel Industries Worldwide," *U.S. Dept. Agric. Econ. Res. Serv.*, 2011.
- [5] P. Azadi, R. Malina, S. R. H. Barrett, and M. Kraft, "The evolution of the biofuel science," *Renew. Sustain. Energy Rev.*, vol. 76, pp. 1479–1484, Sep. 2017.
- [6] V. Parisutham, T. H. Kim, and S. K. Lee, "Feasibilities of consolidated bioprocessing microbes: From pretreatment to biofuel production," *Bioresour. Technol.*, vol. 161, pp. 431–440, Jun. 2014.
- [7] A. Alaswad, M. Dassisti, T. Prescott, and A. G. Olabi, "Technologies and developments of third generation biofuel production," *Renew. Sustain. Energy Rev.*, vol. 51, pp. 1446–1460, Nov. 2015.
- [8] F. Saladini, N. Patrizi, F. M. Pulselli, N. Marchettini, and S. Bastianoni, "Guidelines for energy evaluation of first, second and third generation biofuels," *Renew. Sustain. Energy Rev.*, vol. 66, pp. 221–227, Dec. 2016.
- [9] F. Alam, Mobin S, and Chowdhury H, "Third generation biofuel from Algae," *Elsevier*, vol. 105, pp. 763–768, 2015.
- [10] E. M. Rubin, "Genomics of cellulosic biofuels," *Nature*, vol. 454, no. 7206, pp. 841–845, Aug. 2008.
- [11] A. Heredia, A. Jimenez, and R. Guillen, "Composition of plant cell walls," *Z. Lebensm. Unters. Forsch.*, vol. 200, no. 1, pp. 24–31, Jan. 1995.
- [12] V. Juturu and J. C. Wu, "Microbial cellulases: Engineering, production and applications," *Renew. Sustain. Energy Rev.*, vol. 33, pp. 188–203, May 2014.
- [13] V. Menon and M. Rao, "Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept," *Prog. Energy Combust. Sci.*, vol. 38, no. 4, pp. 522–550, Aug. 2012.
- [14] A. Gupta and S. R. Turner, "Cellulose," in *Encyclopedia of Applied Plant Sciences*, Elsevier, 2017, pp. 185–190.
- [15] H. V. Scheller and P. Ulvskov, "Hemicelluloses," *Annu. Rev. Plant Biol.*, vol. 61, no. 1, pp. 263–

- 289, Jun. 2010.
- [16] D. Shallom and Y. Shoham, "Microbial hemicellulases," *Curr. Opin. Microbiol.*, vol. 6, no. 3, pp. 219–228, Jun. 2003.
- [17] J. Pérez, A. J. Muñoz-Dorado, T. De La Rubia, and J. Martinez, "Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview."
- [18] B. C. Behera, B. K. Sethi, R. R. Mishra, S. K. Dutta, and H. N. Thatoi, "Microbial cellulases – Diversity & biotechnology with reference to mangrove environment: A review," *J. Genet. Eng. Biotechnol.*, vol. 15, no. 1, pp. 197–210, Jun. 2017.
- [19] A. V. Gusakov, "Alternatives to *Trichoderma reesei* in biofuel production," *Trends Biotechnol.*, vol. 29, no. 9, pp. 419–425, Sep. 2011.
- [20] S. Horn, G. Vaaje-Kolstad, B. Westereng, and V. G. Eijsink, "Novel enzymes for the degradation of cellulose," *Biotechnol. Biofuels*, vol. 5, no. 1, p. 45, 2012.
- [21] R. H. Doi and A. Kosugi, "Cellulosomes: plant-cell-wall-degrading enzyme complexes," *Nat. Rev. Microbiol.*, vol. 2, no. 7, pp. 541–551, Jul. 2004.
- [22] E. A. Bayer, H. Chanzy, R. Lamed, and Y. Shoham, "Cellulose, cellulases and cellulosomes," *Curr. Opin. Struct. Biol.*, vol. 8, no. 5, pp. 548–557, Oct. 1998.
- [23] Y. Shoham, R. Lamed, and E. A. Bayer, "The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides," *Trends Microbiol.*, vol. 7, no. 7, pp. 275–281, Jul. 1999.
- [24] M. Desvaux, "*Clostridium cellulolyticum* : model organism of mesophilic cellulolytic clostridia," *FEMS Microbiol. Rev.*, vol. 29, no. 4, pp. 741–764, Sep. 2005.
- [25] P. Biely, M. Vršanská, M. Tenkanen, and D. Kluepfel, "Endo- β -1,4-xylanase families: differences in catalytic properties," *J. Biotechnol.*, vol. 57, no. 1–3, pp. 151–166, Sep. 1997.
- [26] J. C. Hurlbert and J. F. Preston, "Functional Characterization of a Novel Xylanase from a Corn Strain of *Erwinia chrysanthemi*," *J. Bacteriol.*, vol. 183, no. 6, pp. 2093–2100, Mar. 2001.
- [27] H. Chen *et al.*, "A review on the pretreatment of lignocellulose for high-value chemicals," *Fuel Process. Technol.*, vol. 160, pp. 196–206, Jun. 2017.
- [28] S. Alfenore and C. Molina-Jouve, "Current status and future prospects of conversion of lignocellulosic resources to biofuels using yeasts and bacteria," *Process Biochem.*, vol. 51, no. 11, pp. 1747–1756, Nov. 2016.
- [29] M. Galbe and G. Zacchi, "Pretreatment: The key to efficient utilization of lignocellulosic materials," *Biomass and Bioenergy*, vol. 46, pp. 70–78, Nov. 2012.
- [30] N. Gaurav, S. Sivasankari, G. Kiran, A. Ninawe, and J. Selvin, "Utilization of bioresources for sustainable biofuels: A Review," *Renew. Sustain. Energy Rev.*, vol. 73, pp. 205–214, Jun. 2017.
- [31] M. J. Taherzadeh and K. Karimi, "Pretreatment of Lignocellulosic Wastes to Improve Ethanol and Biogas Production: A Review," *Int. J. Mol. Sci.*, vol. 9, no. 9, pp. 1621–1651, Sep. 2008.
- [32] G. Brodeur, E. Yau, K. Badal, J. Collier, K. B. Ramachandran, and S. Ramakrishnan, "Chemical and Physicochemical Pretreatment of Lignocellulosic Biomass: A Review," *Enzyme Res.*, vol. 2011, pp. 1–17, 2011.

- [33] L. J. Jönsson and C. Martín, "Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects," *Bioresour. Technol.*, vol. 199, pp. 103–112, Jan. 2016.
- [34] M. M. Ibrahim, W. K. El-Zawawy, Y. R. Abdel-Fattah, N. A. Soliman, and F. A. Agblevor, "Comparison of alkaline pulping with steam explosion for glucose production from rice straw," *Carbohydr. Polym.*, vol. 83, no. 2, pp. 720–726, Jan. 2011.
- [35] Y.-S. Cheng, Y. Zheng, C. W. Yu, T. M. Dooley, B. M. Jenkins, and J. S. VanderGheynst, "Evaluation of High Solids Alkaline Pretreatment of Rice Straw," *Appl. Biochem. Biotechnol.*, vol. 162, no. 6, pp. 1768–1784, Nov. 2010.
- [36] S. McIntosh and T. Vancov, "Enhanced enzyme saccharification of Sorghum bicolor straw using dilute alkali pretreatment," *Bioresour. Technol.*, vol. 101, no. 17, pp. 6718–6727, Sep. 2010.
- [37] Z. Qiu, G. M. Aita, and M. S. Walker, "Effect of ionic liquid pretreatment on the chemical composition, structure and enzymatic hydrolysis of energy cane bagasse," *Bioresour. Technol.*, vol. 117, pp. 251–256, Aug. 2012.
- [38] M. J. Earle and K. R. Seddon, "Ionic liquids. Green solvents for the future," *Pure Appl. Chem.*, vol. 72, no. 7, Jan. 2000.
- [39] I. Loaces, "Obtención de nuevas enzimas para la producción de biocombustibles," 2016.
- [40] H. Olivier-Bourbigou and L. Magna, "Ionic liquids: perspectives for organic and catalytic reactions," *J. Mol. Catal. A Chem.*, vol. 182–183, pp. 419–437, May 2002.
- [41] H. Zhao, C. L. Jones, G. A. Baker, S. Xia, O. Olubajo, and V. N. Person, "Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis," *J. Biotechnol.*, vol. 139, no. 1, pp. 47–54, Jan. 2009.
- [42] F. M. Gírio, C. Fonseca, F. Carvalheiro, L. C. Duarte, S. Marques, and R. Bogel-Lukasik, "Hemicelluloses for fuel ethanol: A review," *Bioresour. Technol.*, vol. 101, no. 13, pp. 4775–4800, Jul. 2010.
- [43] H. Chum, L. Douglas, D. Feinberg, and A. Schroeder, "Evaluation of pretreatment of biomass for enzymatic hydrolysis of cellulose," 1985.
- [44] L. E. Gollapalli, B. E. Dale, and D. M. Rivers, "Predicting Digestibility of Ammonia Fiber Explosion (AFEX)-Treated Rice Straw," *Appl. Biochem. Biotechnol.*, vol. 98–100, no. 1–9, pp. 23–36, 2002.
- [45] P. Kumar, D. M. Barrett, M. J. Delwiche, and P. Stroeve, "Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production," *Ind. Eng. Chem. Res.*, vol. 48, no. 8, pp. 3713–3729, Apr. 2009.
- [46] M. J. Taherzadeh and K. Karimi, "Pretreatment of Lignocellulosic Wastes to Improve Ethanol and Biogas Production: A Review," *Int. J. Mol. Sci.*, vol. 9, no. 9, pp. 1621–1651, Sep. 2008.
- [47] Y. Kim, N. S. Mosier, and M. R. Ladisch, "Enzymatic digestion of liquid hot water pretreated hybrid poplar," *Biotechnol. Prog.*, vol. 25, no. 2, pp. 340–348, Mar. 2009.
- [48] M. Gutierrez-Correa, L. Portal, P. Moreno, and R. P. Tengerdy, "Mixed culture solid substrate fermentation of *Trichoderma reesei* with *Aspergillus niger* on sugar cane bagasse," *Bioresour.*

- Technol.*, vol. 68, no. 2, pp. 173–178, May 1999.
- [49] D. Klein-Marcuschamer, P. Oleskowicz-Popiel, B. A. Simmons, and H. W. Blanch, “The challenge of enzyme cost in the production of lignocellulosic biofuels,” *Biotechnol. Bioeng.*, vol. 109, no. 4, pp. 1083–1087, Apr. 2012.
- [50] C. Vieille and G. J. Zeikus, “Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability,” *Microbiol. Mol. Biol. Rev.*, vol. 65, no. 1, pp. 1–43, Mar. 2001.
- [51] T. Kaper, S. J. J. Brouns, A. C. M. Geerling, W. M. DE Vos, and J. Van der Oost, “DNA family shuffling of hyperthermostable β -glycosidases,” *Biochem. J.*, vol. 368, no. 2, pp. 461–470, Dec. 2002.
- [52] F. H. Arnold, P. L. Wintrode, K. Miyazaki, and A. Gershenson, “How enzymes adapt: lessons from directed evolution,” *Trends Biochem. Sci.*, vol. 26, no. 2, pp. 100–106, Feb. 2001.
- [53] S. J. Benkovic, “A Perspective on Enzyme Catalysis,” *Science (80-.)*, vol. 301, no. 5637, pp. 1196–1202, Aug. 2003.
- [54] R. E. Cobb, N. Sun, and H. Zhao, “Directed evolution as a powerful synthetic biology tool,” *Methods*, vol. 60, no. 1, pp. 81–90, Mar. 2013.
- [55] H. Alper, C. Fischer, E. Nevoigt, and G. Stephanopoulos, “Tuning genetic control through promoter engineering,” *Proc. Natl. Acad. Sci.*, vol. 102, no. 36, pp. 12678–12683, Sep. 2005.
- [56] B. G. Schuster and M. S. Chinn, “Consolidated Bioprocessing of Lignocellulosic Feedstocks for Ethanol Fuel Production,” *Bioenergy Res.*, vol. 6, no. 2, pp. 416–435, 2013.
- [57] J. C. du Preez, M. Bosch, and B. A. Prior, “Temperature profiles of growth and ethanol tolerance of the xylose-fermenting yeasts *Candida shehatae* and *Pichia stipitis*,” *Appl. Microbiol. Biotechnol.*, vol. 25, no. 6, pp. 521–525, Mar. 1987.
- [58] J. C. du Preez, M. Bosch, and B. A. Prior, “Temperature profiles of growth and ethanol tolerance of the xylose-fermenting yeasts *Candida shehatae* and *Pichia stipitis*,” *Appl. Microbiol. Biotechnol.*, vol. 25, no. 6, pp. 521–525, Mar. 1987.
- [59] P. A. M. Claassen *et al.*, “Utilisation of biomass for the supply of energy carriers,” *Appl. Microbiol. Biotechnol.*, vol. 52, no. 6, pp. 741–755, Nov. 1999.
- [60] A. Limayem and S. C. Ricke, “Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects,” *Prog. Energy Combust. Sci.*, vol. 38, no. 4, pp. 449–467, Aug. 2012.
- [61] C. Martín, M. Galbe, C. F. Wahlbom, B. Hahn-Hägerdal, and L. J. Jönsson, “Ethanol production from enzymatic hydrolysates of sugarcane bagasse using recombinant xylose-utilising *Saccharomyces cerevisiae*,” *Enzyme Microb. Technol.*, vol. 31, no. 3, pp. 274–282, Aug. 2002.
- [62] T. Conway, Y. A. Osman, J. I. Konnan, E. M. Hoffmann, and L. O Ingram, “Promoter and Nucleotide Sequences of the *Zymomonas mobilis* Pyruvate Decarboxylase,” *J. Bacteriol.*, vol. 169, no. 3, pp. 949–954, 1987.
- [63] L. O Ingram, T. Conway, D. P. Clark, G. W. Sewell, and J. F. Preston, “Genetic Engineering of Ethanol Production in *Escherichia coli*,” vol. 53, no. 10.
- [64] I. Loaces, G. Bottini, G. Moyna, E. Fabiano, A. Mart??nez, and F. Noya, “EndoG: A novel

- multifunctional halotolerant glucanase and xylanase isolated from cow rumen," *J. Mol. Catal. B Enzym.*, vol. 126, pp. 1–9, 2016.
- [65] D. D. Songstad, P. Lakshmanan, J. Chen, W. Gibbons, S. Hughes, and R. Nelson, "Historical perspective of biofuels: learning from the past to rediscover the future," *Vitr. Cell. Dev. Biol. - Plant*, vol. 45, no. 3, pp. 189–192, Jun. 2009.
- [66] I. Loaces, S. Schein, and F. Noya, "Ethanol production by *Escherichia coli* from *Arundo donax* biomass under SSF, SHF or CBP process configurations and in situ production of a multifunctional glucanase and xylanase," *Bioresour. Technol.*, vol. 224, pp. 307–313, 2017.
- [67] E. I. Wiloso, R. Heijungs, and G. R. de Snoo, "LCA of second generation bioethanol: A review and some issues to be resolved for good LCA practice," *Renew. Sustain. Energy Rev.*, vol. 16, no. 7, pp. 5295–5308, Sep. 2012.
- [68] M. Wang, C. Saricks, and D. Santini, "Effects of Fuel Ethanol Use on Fuel-Cycle Energy and Greenhouse Gas Emissions," p. 39, 1999.
- [69] J. F. Mac Faddin, S. Rondinone, and O. Giovanniello, *Pruebas bioquímicas para la identificación de bacterias de importancia clínica*. Médica Panamericana, 2003.
- [70] S. Ehrt and D. Schnappinger, "Isolation of Plasmids from *E. coli* by Alkaline Lysis," in *E. coli Plasmid Vectors*, New Jersey: Humana Press, pp. 75–78.
- [71] J. M. Galazka, C. Tian, W. T. Beeson, B. Martinez, N. L. Glass, and J. H. D. Cate, "Cellodextrin Transport in Yeast for Improved Biofuel Production," *Science (80-.)*, vol. 330, no. 6000, pp. 84–86, Oct. 2010.
- [72] S.-J. Ha *et al.*, "Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 2, pp. 504–9, 2011.
- [73] A. C. Hayward, "Bacteriophage Sensitivity and Biochemical Group in *Xanthomonas malvacearum*," *J. gen. Microbiol.*, vol. 35, pp. 287–298, 1964.
- [74] A. A. Esener, J. A. Roels, and N. W. F. Kossen, "The influence of temperature on the maximum specific growth rate of *Klebsiella pneumoniae*," *Biotechnol. Bioeng.*, vol. 23, no. 6, pp. 1401–1405, Jun. 1981.
- [75] A. Tsuji, Y. Kaneko, K. Takahashi, M. Ogawa, and S. Goto, "The Effects of Temperature and pH on the Growth of Eight Enteric and Nine Glucose Non-Fermenting Species of Gram-Negative Rods," *Microbiol. Immunol.*, vol. 26, no. 1, pp. 15–24, Jan. 1982.
- [76] Quiagen, "QIAquick® Spin Handbook QIAGEN Sample and Assay Technologies," *Qiagen*, vol. Volume 20, no. March, pp. 9–10, 2008.
- [77] Thermo Fisher Scientific, "DNA Insert Ligation into Vector DNA." 2012.
- [78] A. Tomasz, "The Mechanism of the Irreversible Antimicrobial Effects of Penicillins: How the Beta-Lactam Antibiotics Kill and Lyse Bacteria," *Annu. Rev. Microbiol.*, vol. 33, no. 1, pp. 113–137, Oct. 1979.
- [79] J. Watson *et al.*, *Molecular biology of the gene*, 7th ed., vol. 7th editio. New York: Pearson, 2013.
- [80] H. A. Erlich, "Polymerase Chain Reaction," *J. Clin. Immunol.*, vol. 9, no. 6, 1989.

- [81] K. B. Mullis and F. A. Faloona, "Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction," *Methods Enzymol. Sci.*, vol. 155, no. 4839, pp. 335–350, 1987.
- [82] J. L. Brunelle and R. Green, "Coomassie Blue Staining," 2014, pp. 161–167.
- [83] T. H. Steinberg, "Chapter 31 Protein Gel Staining Methods," 2009, pp. 541–563.
- [84] P. Siguier, J. Perochon, L. Lestrade, J. Mahillon, and M. Chandler, "ISfinder: the reference centre for bacterial insertion sequences."
- [85] R. Senthilkumar *et al.*, "FAIR: A server for internal sequence repeats," *Bioinformatics*, vol. 4, no. 7, pp. 271–275, 2010.
- [86] T. Nagai, L.-S. Phan Tran, Y. Inatsu, and Y. Itoh, "A New IS4 Family Insertion Sequence, IS4Bsu1, Responsible for Genetic Instability of Poly-Glutamic Acid Production in *Bacillus subtilis*," *J. Bacteriol.*, vol. 182, no. 9, pp. 2387–2392, 2000.
- [87] B. Hallet and D. J. Sherratt, "Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements," *FEMS Microbiol. Rev.*, vol. 21, no. 2, pp. 157–178, Jan. 2006.
- [88] R. H. A. Plasterk, "Molecular mechanisms of transposition and its control," *Cell*, vol. 74, no. 5, pp. 781–786, Sep. 1993.
- [89] H. Ji *et al.*, "Hotspots for unselected Ty1 transposition events on yeast chromosome III are near tRNA genes and LTR sequences," *Cell*, vol. 73, no. 5, pp. 1007–1018, Jun. 1993.
- [90] N. L. Craig, "Target site selection in transposition," *Annu. Rev. Biochem.*, vol. 66, no. 1, pp. 437–474, Jun. 1997.
- [91] M. Chandler and P. Siguier, "Insertion Sequences," *Brenner's Encycl. Genet.*, vol. 62, no. 3, pp. 86–94, 2013.
- [92] N. Kleckner, R. M. Chalmers, D. Kwon, J. Sakai, and S. Bolland, "Tn10 and IS10 Transposition and Chromosome Rearrangements: Mechanism and Regulation In Vivo and In Vitro," 1996, pp. 49–82.
- [93] S. Lovell, I. Y. Goryshin, W. R. Reznikoff, and I. Rayment, "Two-metal active site binding of a Tn5 transposase synaptic complex," *Nat. Struct. Biol.*, vol. 9, no. 4, pp. 278–281, Apr. 2002.
- [94] H.-M. Tan, "Bacterial catabolic transposons," *Appl Microbiol Biotechnol*, vol. 51, pp. 1–12, 1999.
- [95] N. D. F. Grindley and R. R. Reed, "Transpositional Recombination in Prokaryotes," *Annu. Rev. Biochem.*, vol. 54, no. 1, pp. 863–896, Jun. 1985.
- [96] P. Gamas, M. G. Chandler, P. Prentki, and D. J. Galas, "Escherichia coli integration host factor binds specifically to the ends of the insertion sequence IS1 and to its major insertion hot-spot in pBR322," *J. Mol. Biol.*, vol. 195, no. 2, pp. 261–272, May 1987.
- [97] C. Dahlberg and M. Hermansson, "Abundance of Tn3, Tn21, and Tn501 Transposase (tnpA) Sequences in Bacterial Community DNA from Marine Environments," *Appl. Environ. Microbiol.*, vol. 61, no. 8, pp. 3051–3056, 1995.
- [98] B. M. Harshey and A. 1 Bukhari, "A mechanism of DNA transposition," *Genetics*, vol. 78, no. 2, pp. 1090–1094, 1981.

- [99] T. A. Baker and K. Mizuuchi, "DNA-promoted assembly of the active tetramer of the Mu transposase.," *Genes Dev.*, vol. 6, no. 11, pp. 2221–2232, Nov. 1992.
- [100] G. Posfai, "Emergent Properties of Reduced-Genome Escherichia coli," *Science (80-.)*, vol. 312, no. 5776, pp. 1044–1046, May 2006.
- [101] P. H. Oliveira, K. J. Prather, D. M. F. Prazeres, and G. A. Monteiro, "Structural instability of plasmid biopharmaceuticals: challenges and implications," *Trends Biotechnol.*, vol. 27, no. 9, pp. 503–511, Sep. 2009.
- [102] A. Wagner, C. Lewis, and M. Bichsel, "A survey of bacterial insertion sequences using IScan," *Nucleic Acids Res.*, vol. 35, no. 16, pp. 5284–5293, 2007.
- [103] J. Mahillon, C. Léonard, and M. Chandler, "IS elements as constituents of bacterial genomes," *Res. Microbiol.*, vol. 150, no. 9–10, pp. 675–687, Nov. 1999.
- [104] A. Biswas, D. T. Gauthier, D. Ranjan, and M. Zubair, "ISQuest: finding insertion sequences in prokaryotic sequence fragment data," *Bioinformatics*, vol. 31, no. 21, pp. 3406–3412, Nov. 2015.
- [105] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith, "Enzymatic assembly of DNA molecules up to several hundred kilobases," *Nat. Methods*, vol. 6, no. 5, pp. 343–345, May 2009.
- [106] D. G. Gibson, H. O. Smith, C. A. Hutchison, J. C. Venter, and C. Merryman, "Chemical synthesis of the mouse mitochondrial genome," *Nat. Methods*, vol. 7, no. 11, pp. 901–903, Nov. 2010.
- [107] M. R. McAllaster, A. N. Sinclair-Davis, N. A. Hilton, and C. L. de Graffenried, "A unified approach towards Trypanosoma brucei functional genomics using Gibson assembly," *Mol. Biochem. Parasitol.*, vol. 210, no. 1–2, pp. 13–21, Nov. 2016.

8. ANEXO

8.1. Medios de cultivo

➤ Luria Bertani Caldo (LB)

Triptona	10 g/L
Extracto de levadura	5 g/L
NaCl	10 g/L

➤ LB Agar

LB-Caldo	
Agar	16 g/L

➤ Yeast Extract-Peptone-Dextrose (YPD)

Extracto de levadura	10 g/L
Peptona bacteriológica	20 g/L
Glucosa	20 g/L

8.2. Soluciones para la extracción del ADN plasmídico

➤ Solución A

Tris.HCl 50 mM

EDTA 10 mM

pH = 8

➤ Solución B

NaOH 0,2 M

SDS 1% (m/v)

➤ Solución C

Acetato de Potasio 3M

Ácido acético 11,5% (V/V)

pH = 4,8

➤ **Buffer TE**

10 mM Tris-HCl

0.1 mM EDTA

pH = 8.0

8.3. Electroforesis en gel de agarosa

➤ **Buffer TAE 1X**

Tris-acético 0,04 M

EDTA 0,001 M

pH= 8,0

8.4. Purificación del ADN del gel de agarosa

➤ **Buffer EB**

Tris-Cl 10 mM

pH = 8.5

8.5. Electroforesis desnaturizante en gel de poliacrilamida SDS

➤ **Gel de resolución 12%**

Buffer de resolución 2,5 mL

Acrilamida 30% / Bisacrilamida 0,8% 4,04 mL

H₂O 3,56 mL

TEMED 10% 5 µL

Amonio Persulfato 10% 50 µL

➤ **Gel de resolución 15%**

Buffer de resolución 2,5 mL

Acrilamida 30% / Bisacrilamida 0,8% 5,0 mL

H₂O 2,55 mL

TEMED 10% 5 µL

Amonio Persulfato 10% 50 µL

➤ **Acrilamida 30% / Bisacrilamida 0,8%**

Acrilamida 58,4 g

Bisacrilamida 1,6 g

H₂O Csp 200 ml

➤ ***Buffer* de resolución**

Tris base 72,6 g

SDS 1,6 g

H₂O 300 mL

pH = 8,8

Llevar a 400 ml de H₂O

➤ **Gel apilamiento**

Buffer compactador 1,25 mL

Acrilamida 30% / Bisacrilamida 0,8% 0,8mL

H₂O 2,95 mL

TEMED 10% 5 µL

Amonio Persulfato 10% 30 µL.

➤ ***Buffer* compactador**

Tris base 24 g

SDS 1,6 g

H₂O 300 mL

pH = 6,8

Llevar a 400mL de H₂O

➤ **Buffer de unión (pH = 8,0)**

H₂O 49,9 mL

Imidazol stock 5M 100 µL

Buffer His-Tag 2X 50mL

➤ **Buffer His-Tag 2X**

2X His Tag Buffer

100 mM NaH₂PO₄

600 mM NaCl

➤ **Buffer de carga 2X**

0,5 M Tris pH 6,8 12mL

Glicerol (20%) 10mL

SDS 20% 10mL

0,1% azul bromo fenol 5mL

H₂O Csp 45mL

Antes de usar agregar β- mercaptoetanol 10%

➤ **Buffer de corrida Tris-Glicina**

Tris 302,75 g

Glicina 721 g

SDS 50 g

H₂O 5 L

Antes de usar diluir a partir de 10X

➤ **Solución de teñido rápido**

0,2%(m/V) Azul Brillante de Coomassie en:

Metanol 500 mL/L

Ácido acético glacial 100 mL/L

H₂O 400 mL/L.

➤ **Solución de desteñido rápido**

Metanol 500 mL/L

Ácido acético glacial 100 mL/L

H₂O 400 mL/L