



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**“Evolución dirigida de una quimera  
de shikimato deshidrogenasa”**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**B I Ó L O G A  
P R E S E N T A:**

**VALERIE YSELLE DE ANDA TORRES**



**DIRECTOR DE TESIS:  
DR. LORENZO PATRICK SEGOVIA FORCELLA  
(2012)**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



Esta tesis se realizó bajo la dirección del Dr. Lorenzo Patrick Segovia Forcella, en el Departamento de Ingeniería Celular y Biocatálisis del Instituto de Biotecnología de la Universidad Nacional Autónoma de México con el apoyo de una beca de alimentación y vivienda (mayo 2011-diciembre 2011).

El comité que asesoró y participó en el desarrollo y realización de este proyecto:

Dr. Lorenzo Patrick Segovia Forcella

Dra. Claudia Martínez Anaya

Dr. Joel Osuna Quintero

Dr. Humberto Flores Soto

El jurado del examen de Licenciatura, estuvo conformado por:

Presidente: Dr. Víctor Manuel Valdés López

Vocal: Dr. José Adolfo Escalante Lozada

Secretario: Dr. Lorenzo Patrick Segovia Forcella

Suplente: Dr. Joel Osuna Quintero

Suplente: Dr. Humberto Flores Soto

## **Agradecimientos académicos**

*En primer lugar agradezco a mi casa de estudios la UNAM, la Facultad de Ciencias y al Instituto de Biotecnología por haberme formado en la investigación científica.*

*“Al Dr. Lorenzo Segovia”, por brindarme la oportunidad de realizar mi tesis de Licenciatura en su laboratorio. Por guiarme y apoyarme no sólo en la realización del proyecto, si no también en lo personal. Gracias!*

*“A la Dra. Claudia Martínez”, por su ejemplo de disciplina, dedicación y rigor en el trabajo científico. Por todos sus consejos, comentarios y apoyo técnico en el laboratorio, así como el tiempo dedicado en la revisión del trabajo escrito.*

*“Al Dr. Adelfo Escalante y Dr. Víctor Valdés” mis sinceros agradecimientos por sus enriquecedores comentarios y sugerencias para la mejora de este escrito.*

*“Al Dr. Humberto Flores Soto” por su ayuda y apoyo técnico durante mi estancia en el IBT. Así como también por sus consejos, aportaciones y comentarios que me fueron de gran ayuda en la parte experimental y en el trabajo escrito.*

*“Al Dr. Joel Osuna” por su apoyo y valiosos comentarios para la realización de este trabajo.*

*“Al Dr. Ernesto Ortiz” por compartirme el secreto de las células súper competentes.*

*A mis compañeros de laboratorio: Katy, Dago, Mario, Banda, Fer y Miguel, estos tres últimos mejor conocidos como los “Labengers”. Gracias por su paciencia, consejos, apoyo y risas en el laboratorio.*

*“A Don Juan”, por su apoyo en la preparación de medios y soluciones, y por las risas compartidas.*

*A toda la comunidad del IBT por hacerme más cómoda mi estancia en Cuernavaca.*

## **Agradecimientos personales**

*“A mi madre” que es la esencia de todos mis triunfos (que también son tuyos). Jamás existirá una forma de retribuirle todo lo que has hecho por mí, sin escatimar esfuerzo alguno, has sacrificado gran parte de tu vida para formarme, educarme y amarme. Por ti he podido seguirme preparando en mi formación profesional, y sé que sin ti, simplemente no lo hubiera logrado.*

*“A Emi”, mi amor, que a pesar de la distancia, nunca te rendiste y dejaste todo para estar conmigo. Te agradezco todo lo que haces por mí, tu apoyo incondicional, tu paciencia infinita, tus consejos que han servido tanto y tu gran amor, que me hacen la mujer más feliz del mundo. Eres parte esencial de mi vida y siempre lo serás. Te amo!*

*“A Luchi” por ser el pilar fundamental de la familia y por ser un ejemplo a seguir. Gracias por hacerme una mejor persona a través de tus consejos, apoyo y cuidados. Gracias por estar conmigo siempre.*

*“A mi padre” por el cariño, la comprensión, y apoyo que me han servido para ser una mejor persona y me han ayudado para culminar mi carrera profesional.*

*“A mi familia” hermanos, tíos y primos, gracias por el apoyo y palabras de aliento. Esto también es para ustedes.*

*“A Toma y Emilio”, por dejarme ser parte de su familia, a pesar de la distancia, siempre los he tenido cerca en los momentos más críticos de mi vida. Les agradezco de corazón los consejos y el amor que he recibido y sobretodo, gracias por el apoyo incondicional.*

*“A Adri, Muegs, Duqueshis y amigos” por todo lo que hemos vivido juntos durante la carrera, por seguir este camino en paralelo. Gracias por todo el cariño, consejos, paciencia, risas y sobretodo por soportarme en momentos de estrés perpetuo!*

*A todos los profesores que me han acompañado durante mi trayectoria estudiantil, por brindándome siempre su orientación con profesionalismo ético que me han servido para consolidar mi formación como estudiante universitario.*

**“Gracias a todos por su apoyo, y por ser parte importante de mi vida”.**

*Dedicada a mi madre, esposo y familia que son el pilar fundamental de todo lo que soy. Sin su apoyo incondicional, cariño y comprensión, no hubiera llegado hasta donde estoy ahora. Gracias*

# TABLA DE CONTENIDO

<b>RESUMEN</b> .....	<b>1</b>
<b>I. INTRODUCCIÓN</b> .....	<b>2</b>
I. I Recombinación de dominios.....	2
I. II Enzima shikimato deshidrogenasa .....	4
I. III Estructura de AroE y YdiB .....	6
I. III. I Dominio catalítico .....	8
I. III. II Plegamiento Rossmann .....	9
<b>II. ANTECEDENTES</b> .....	<b>12</b>
II. I Recombinación de dominios de shikimato deshidrogenasas .....	12
II. II Ingeniería de proteínas .....	14
II. III Evolución dirigida .....	15
II. IV Mutagénesis al azar .....	17
II. V Frecuencia de mutagénesis.....	19
II. VI Método de selección.....	20
<b>III. PLANTEAMIENTO</b> .....	<b>21</b>
<b>IV. HIPÓTESIS</b> .....	<b>22</b>
<b>V. OBJETIVOS</b> .....	<b>22</b>
V. I Generales .....	22
V. II Particulares .....	22
<b>VI. MATERIALES Y MÉTODO</b> .....	<b>23</b>
VI. I Cepas .....	23
VI. II Plásmidos .....	23
VI. III Mutagénesis al azar .....	24
VI. IV Generación de los bancos de mutagénesis .....	25
<b>VII. ESTRATEGIA EXPERIMENTAL</b> .....	<b>30</b>
<b>VIII. RESULTADOS</b> .....	<b>31</b>
VIII. I Bancos de mutagénesis .....	31
VIII. II Tamaño de los bancos .....	33
VIII. III Variabilidad general de los bancos .....	33
VIII. IV Método de selección.....	35
VIII. V Retransformación.....	45
<b>IX. DISCUSIONES</b> .....	<b>48</b>
<b>X. CONCLUSIONES</b> .....	<b>52</b>
<b>XI. PERSPECTIVAS</b> .....	<b>53</b>
<b>XII. BIBLIOGRAFÍA</b> .....	<b>54</b>
<b>ANEXOS</b> .....	<b>59</b>

## ÍNDICE DE FIGURAS

<b>Figura 1</b>	Vía del shikimato en <i>E.coli</i>	5
<b>Figura 2</b>	Estructura de AroE de <i>E. coli</i>	6
<b>Figura 3</b>	Estructuras de las enzimas AroE y YdiB	7
<b>Figura 4</b>	Reacciones catalizadas por AroE y YdiB	8
<b>Figura 5</b>	Topología del plegamiento Rossmann	9
<b>Figura 6</b>	Secuencias con plegamiento Rossmann	10
<b>Figura 7</b>	Plásmidos utilizados	24
<b>Figura 8</b>	Amplificación de la construcción pT4-EYE-CAT con diferentes temperaturas de extensión	31
<b>Figura 9</b>	Extracción de plásmido de la construcción pT4-EYE-CAT	31
<b>Figura 10</b>	Digestión para la obtención de vector de clonación	32
<b>Figura 11</b>	Crecimiento de la cepa BW25113 AroE <sup>-</sup> Km <sup>r</sup> , transformada con las construcciones AroE, AroBS, YdiB, pT4 y EYE, en medio M9 con diferentes concentraciones de AAA.	36
<b>Figura 12</b>	Alineamiento de la variante 1 del banco C9 con el gen de la quimera EYE parental	39
<b>Figura 13</b>	Alineamiento de la variante 2 del banco C9 con el gen de la quimera EYE parental	40
<b>Figura 14</b>	Alineamiento de la variante 1 del banco C10 con el gen de la quimera EYE parental	41
<b>Figura 15</b>	Alineamiento de la variante 2 del banco C10 con el gen de la quimera EYE parental	42
<b>Figura 16</b>	Alineamiento de aminoácidos de la variante 1 del banco C9 con la quimera EYE parental	43
<b>Figura 17</b>	Alineamiento de aminoácidos de la variante 2 del banco C9 con la quimera EYE parental	43
<b>Figura 18</b>	Alineamiento de aminoácidos de la variante 1 del banco C10 con la quimera EYE parental	44
<b>Figura 19</b>	Alineamiento de aminoácidos de la variante 2 del banco C10 con la quimera EYE parental	44
<b>Figura 20</b>	Ensayo para determinar la concentración mínima de aminoácidos aromáticos	47

## ÍNDICE DE TABLAS

<b>Tabla 1</b>	Ejemplos de generación de quimeras por recombinación de dominios	3
<b>Tabla 2</b>	Los 9 plegamientos más ancestrales, y ampliamente distribuidos del árbol filogenómico que describe la diversificación del mundo de las proteínas	11
<b>Tabla 3</b>	Cepas utilizadas	23
<b>Tabla 4</b>	Soluciones de reacción en volúmenes de $\mu\text{L}$ . Condiciones de mutagénesis	25
<b>Tabla 5</b>	Células transformadas con 6 construcciones, e inoculadas en 3 medios diferentes	28
<b>Tabla 6</b>	Determinación del tamaño de cada banco (número de variantes)	33
<b>Tabla 7</b>	Análisis de la variabilidad de los bancos	34
<b>Tabla 8</b>	Número de variantes sencillas, dobles, triples. Cobertura de cada banco	34
<b>Tabla 9</b>	Número de colonias del banco C4 en 3 diferentes medios	37
<b>Tabla 10</b>	Número de colonias del banco C9 en 3 diferentes medios	37
<b>Tabla 11</b>	Número de colonias del banco C10 en 3 diferentes medios	37
<b>Tabla 12</b>	Análisis de las variantes obtenidas en medio mínimo	45
<b>Tabla 13</b>	Número de colonias en 3 diferentes medios	46

## ABREVIATURAS

<b>Abreviatura</b>	<b>Significado</b>
SDH	Shikimato deshidrogenasa
AAA	Aminoácidos aromáticos
Phe	Fenilalanina
Trip	Triptófano
Tyr	Tirosina
NADPH	Forma reducida nicotinamida adenina dinucleótido fosfato
NADP <sup>+</sup>	Forma oxidada nicotinamida adenina dinucleótido fosfato
FAD <sup>+</sup>	Forma oxidada dinucleótido de flavina-adenina
FADH <sub>2</sub>	Forma reducida de dinucleótido de flavina-adenina
PEP	Fosfoenol piruvato
Eri-4P	Eritrosa 4-fosfato
DAHP	3-deoxy-d-arabinoheptulosonato 7-fosfato
DHQ	dehidroquinato
DHS	dehidroxishikimato
5-EPS-3-P	5-enolpiruvilshikimato 3-fosfato
EPS sintasa	Enolpiruvilshikimato sintasa
<i>E.coli</i>	<i>Escherichia coli</i> ,
Gen de SDH	<i>aroE</i>
Gen de YdiB	<i>ydiB</i>
Proteína de SDH	AroE
Proteína de YdiB	YdiB
Asp, D	Aspartato
SCOP	Clasificación estructural de proteínas
Km	Kanamicina
CAT	Cloranfenicol acetil transferasa
CIM	Ensayo de concentración mínima inhibitoria
SDS page	Electroforesis en gel de poliacrilamida con dodecilsulfato sódico
$K_m$	Constante de Michaelis-Menten (medida de afinidad al sustrato)
$V_{max}$	Velocidad máxima de reacción
epPCR	PCR propensa a errores
<i>DNA shuffling</i>	Barajeado de DNA
dGTP	Deoxi-guanosin trifosfato
dNTP	Deoxi -nucleótido trifosfato
PCR	Reacción en cadena de la polimerasa
LB	Medio Luria-Bertani ( <b>Lysogeny Broth</b> )
SOC	Medio óptimo de crecimiento, con represión catabólica ( <b>Super Optimal broth with Catabolite repression</b> )
SOB	Medio de crecimiento súper optimo ( <b>Super Optimal Broth</b> )
pb	Pares de bases
rpm	Revoluciones por minuto
D.O	Densidad óptica
M9	Medio mínimo
C-T	Carboxilo terminal
N-T	Amino terminal

**RESUMEN**

Se analizó la posibilidad de obtener variantes activas a partir de una quimera no activa (EYE), producto de la recombinación entre dominios de las enzimas xenólogas shikimato deshidrogenasa (SDH) y quinato/shikimato deshidrogenasa (YdiB) de *E. coli*, mediante evolución dirigida usando mutagénesis al azar. Esta técnica es una herramienta muy poderosa para elucidar relaciones de estructura-función de proteínas, además de que permite la modificación de las mismas para mejorar o alterar sus características. Se utilizó la técnica de PCR mutagénico, modificando las condiciones de reacción para obtener distintas tasas de mutagénesis. Basándonos en estudios de evolución dirigida anteriores, utilizamos 3 tasas de mutagénesis. La primera corresponde aproximadamente a 2 mutaciones/1000 pb, la segunda genera alrededor de 3.5 mutaciones/1000 pb y la última genera 8.1 mutaciones/1000 pb. Los productos de PCR mutagénico se introdujeron en el vector de expresión pT4 y los bancos resultantes fueron caracterizados mediante ensayos de complementación en medio mínimo M9 de una cepa carente de la actividad SDH. En este trabajo se obtuvieron 4 bancos (C1, C4, C9, C10) cuya nomenclatura refleja la tasa de mutagénesis utilizada. Se obtuvieron un total de: 65000, 47600, 125000 y 585000 (C1, C4, C9, C10) variantes respectivamente, sin embargo en ningún caso se encontraron variantes activas.

## I. INTRODUCCIÓN

### II Recombinación de dominios

Durante el curso de la evolución, los procesos de duplicación génica, divergencia de secuencias y recombinación génica han sido una de las fuentes principales para la generación de proteínas con nuevas funciones. Los genes involucrados en estos procesos usualmente codifican dominios enteros [Basthon M y Chothia C, 2007], razón por la cual se consideran a los dominios como las unidades evolutiva del proteoma [Thornton *et al.*, 1999].

Existen estudios detallados enfocados en investigar el papel de la duplicación génica, la divergencia de secuencias y la recombinación de dominios para la creación de proteínas con nuevas funciones [Basthon M y Chothia C, 2007]. Este último proceso (recombinación de dominios) es un mecanismo recurrente en la naturaleza ya que posibilita incrementar la especificidad de enzimas, actuar como conexión entre dominios con diferencias funcionales y regular la actividad de proteína. Además de esto, la recombinación de dominios provee el contexto estructural para la búsqueda de nuevas funciones con aplicación en la industria y en la biotecnología, razón por la cual es un recurso ampliamente utilizado en la Ingeniería de proteínas [Basthon M y Chothia C, 2007].

En la literatura existe una gran cantidad de reportes donde se aborda la recombinación de dominios en búsqueda de novedades [Segatori L, *et al.*, 2004; Giesecke A, Fang R, y Joung J. 2006; Thulasiram H, Erickson H, y Poulter C. 2007; Wheeldon IR, Campbell 2009; Peisajovich SG, *et al.*, 2010]. La recombinación de dominios de diferentes proteínas, da lugar a proteínas quiméricas, las cuales se definen como proteínas híbridas codificadas por secuencias de nucleótidos empalmados de dos o más genes completos o parciales [Shanmugaratnam S, Eisenbeis S y Höcker B 2012].

La mayoría de los estudios donde llevan a cabo recombinación de dominios, se basan en recombinar secuencias de proteínas homólogas ya que esto permite utilizar el material preexistente en la naturaleza e imitar el fenómeno evolutivo para investigar los

determinantes de la función y el plegamiento de las proteínas [Espinosa A, 2010, Carbone MN, Arnold FH, 2007]. Sin embargo, la obtención de novedades funcionales no se limita a recombinar proteínas homólogas, también se pueden recombinar dos o más dominios que no se presentan juntos en la naturaleza (**Ver tabla 1**).

<b>AUTOR</b>	<b>RESULTADOS OBTENIDOS</b>
Segatori, <i>et al.</i> , (2004).	Realizaron el intercambio de dominios entre 2 tioredoxinas que participan en la formación de enlaces disulfuro en las proteínas de membrana en <i>E. coli</i> . La tioredoxina DsbA forma los nuevos enlaces, mientras que DsbC tiene un dominio que isomeriza los enlaces, y también presenta función de chaperona. La quimera resultante de la combinación del dominio de dimerización de DsbC con el de tioreductasa de DsbA generó una proteína con novedad funcional, siendo que además de formar enlaces también los isomeriza. El dominio de dimerización no contribuye a la catálisis, si no más bien contribuye a la interacción de DsbC con otra proteína en la vía que permite su reducción constante.
Giesecke A, Fang R, y Joung J (2006).	Los dedos de zinc Cys2His2 (C2H2 ZFs), median una gran cantidad de interacción de proteína-DNA y proteína-proteína. En este trabajo recombinaron los (C2H2 ZFs) del factor de transcripción de dimerización DZF, es decir de unión a DNA, dando lugar a una quimera con especificidad de interacción de proteína-proteína. La recombinación de dominios dio lugar a nuevos sitios de reconocimiento y cambio de especificidad.
Thulasiram H, Erickson H, y Poulter C (2007).	Generaron 11 quimeras a partir de la recombinación de enzimas que sintetizan isoprenoides por reacciones de elongación, ciclopropanación y ramificación. Las enzimas recombinadas fueron: crisantemil difosfato sintasa y farnesil difosfato sintasa con el 75 % de identidad. Las quimeras resultantes, sintetizaron isoprenoides por ciclobutanización, el cual es un mecanismo que no se presenta en las enzimas recombinadas parentales.
Wheeldon IR, Campbell E, y Banta S (2009).	Generaron un quimera, producto de la recombinación de tres dominios que no se presentan juntos en la naturaleza: una aldo-ceto reductasa termoestable, dos dominios zipper $\alpha$ -hélice de leucina y un dominio randomly-coiled. La quimera dio a la formación de hidrogel y actividad de reductasa generando un biomaterial activo, lo cual es un avance importante para la biotecnología, ingeniería de tejidos, la bioelectrocatalisis y la generación de biosensores.
Peisajovich <i>et al.</i> , (2010).	Construyeron 66 quimeras a partir de 11 proteínas involucradas en la vía de división de levaduras. Las recombinaciones resultaron en una mayor diversidad de respuestas en la vía. Los resultados obtenidos sugieren que la recombinación de dominios preexistentes puede ser un mecanismo importante en la evolución de redes de proteínas y en la generación de nuevos comportamientos fenotípicos.

**Tabla 1. Ejemplos de generación de quimeras por recombinación de dominios.**

Por otra parte, existe una gran cantidad de reportes donde se recombinan dominios de enzimas deshidrogenasas [Lebbink JH *et al.*, 1995; Tripura C, y Podile A 2007; Goihberg E, *et al.*, 2010; Sharkey y Engle 2009; Kataoka *et al.*, 1994] las cuales son enzimas clave en el

metabolismo energético y secundario de las células procariontes y eucariontes. Muestro a continuación dos ejemplos ilustrativos de recombinación de enzimas deshidrogenasas:

En Tripura C, y Podile A (2007) reportaron la generación de una quimera producto de la combinación de dominios de la enzima glucosa deshidrogenasa (GDH), dependiente de pirrol quinolina quinona (PQQ) de *E. coli*, (enzima que tiene un papel crucial en la solubilización de fosfatos minerales, y también importancia industrial ya que es utilizada como biosensor de glucosa), junto con la GDH (PQQ) de *Serratia marcescens*, enzima eficaz solubilizando fosfatos, pero no se expresa en *E.coli*. Los dominios recombinados N-terminal de *E. coli* y C-terminal de *S. marcescens*, generaron una quimera, que se logró expresar en *E. coli* con un mayor porcentaje de tolerancia a altas temperaturas (58%), comparado con la GDH parental de *E. coli*, la cual solo tiene un porcentaje de 27.5%

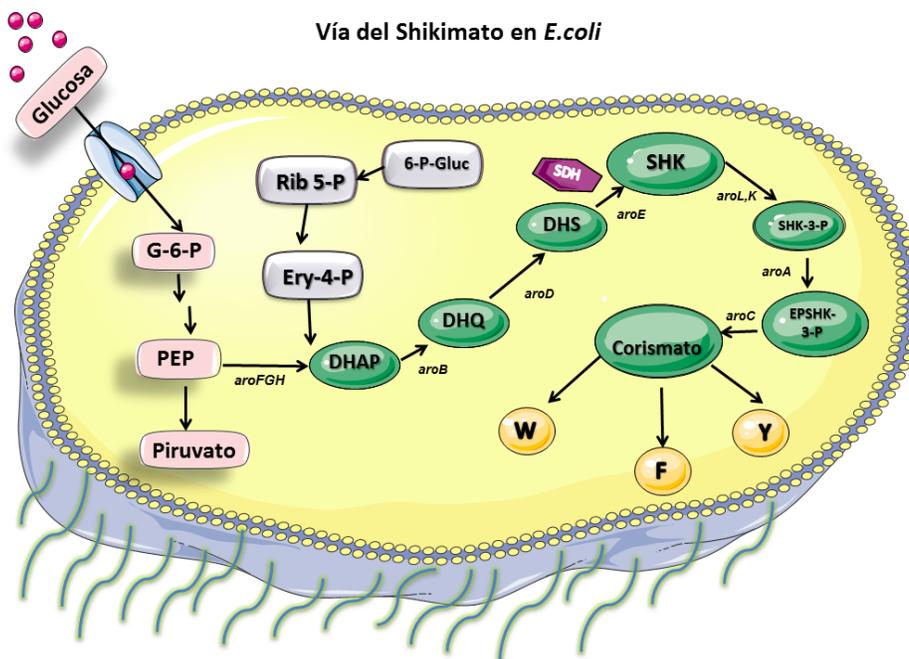
En Kataoka *et al.*, (1994) recombinaron el dominio N-terminal de la fenilalanina deshidrogenasa de *Thermoactinomyces intermedius* con el dominio C-terminal de la leucina deshidrogenasa de *Bacillus stearothermophilus*. La quimera presentó los mismos valores de  $K_m$  para la fenilalanina y la leucina que la enzima parental fenilalanina deshidrogenasa y una menor actividad específica ante fenilalanina como sustrato; sin embargo, presentó mayor actividad frente a isoleucina y valina, actividad que no se encuentra en la fenilalanina deshidrogenasa parental. Se concluyó que ambos dominios se pueden plegar independientemente y que la combinación genera un nuevo sitio activo, consistente con la nueva especificidad.

## **I. II Enzima shikimato deshidrogenasa**

En el presente trabajo se utilizó una shikimato deshidrogenasa perteneciente a la superfamilia de aminoácido deshidrogenasas dependientes de NAD (P). A continuación se presentan las características de la enzima.

La enzima shikimato deshidrogenasa (SHD) cataliza la reducción reversible del DHS a shikimato utilizando NADPH como cofactor oxidándolo a  $\text{NADP}^+$  (EC 1.1.1.25). Esta reacción es de particular importancia en el metabolismo central de la mayoría de los organismos ya que forma parte de la ruta de biosíntesis de compuestos aromáticos entre

los que se encuentran los aminoácidos aromáticos: (Phe), (Trp), y (Tyr). La ruta de biosíntesis es conocida como "Vía del Shikimato" debido a que históricamente el shikimato fue el primer intermediario encontrado, que servía como único requerimiento para el crecimiento de ciertas cepas auxotróficas para AAA [Sprenger GA, 2007]. La vía del shikimato se encuentra en la mayoría de las bacterias [Pittard J 1996], parásitos apicomplexos, [Roberts F *et al.*, 1998, Roberts CW *et al.*, 2002], hongos y plantas [Herrmann KM y Weaver LM 1999]. Por su parte, los organismos metazoarios carecen de la vía del shikimato, por lo que deben adquirir AAA del medio para su sobrevivencia [Sprenger GA, 2007]. La vía del shikimato comienza con la generación de DAHP a partir de la condensación de la Eri-4P y PEP (provenientes de la ruta de las pentosas fosfato y ruta glucolítica respectivamente) generando una serie de intermediarios que darán lugar a corismato, precursor universal de los AAA [Sprenger GA, 2007]. Las reacciones de la vía del shikimato se llevan a cabo en el citosol de las bacterias y hongos (Figura 1) y en plantas se lleva a cabo en organelos plásticos [Roberts CW, *et al.*, 2002].



**Figura 1.** Vía del shikimato en *E.coli*. Los precursores, (PEP) y (Eri4P) provienen de vías metabólicas centrales: glucolisis (Rosa) y vía de pentosas fosfato (gris). En verde la vía del shikimato. Se muestran los genes que codifican las enzimas que llevan a cabo cada reacción. La SDH se muestra en morado. Modificado de Sprenger GA, 2007.

### I. III Estructura de AroE y YdiB

La caracterización de la estructura y función de la enzima shikimato deshidrogenasa ha sido de principal interés para la generación de antimicrobiales e insecticidas dirigidos a esta enzima o a otras en la misma vía, así como para la producción de antivirales como el oseltamivir (tamiflú) [Espinosa A, 2010]. En *Escherichia coli* se han encontrado 2 enzimas de SDH codificadas por los genes *aroE* (NADPH-dependiente) [Anton JA, y Coggins JR 1988] y *ydiB* (NADH/NADPH-dependiente) [Michael G *et al.*, 2003]. Las dos enzimas (AroE y YdiB) han sido ampliamente estudiadas, y hasta la fecha se cuenta con una gran cantidad de información con respecto a su estructura tridimensional, determinada por cristalografía de rayos X [Michel G *et al.*, 2003; Benach J *et al.*, 2003; Padyana AK y Burley SK, 2003; Singh S, *et al.*, 2005, 2006; Ye S, *et al.*, 2003; Gan, J *et al.*, 2007, Han C, *et al.*, 2009]. Las estructuras de AroE y YdiB comparten el mismo plegamiento tridimensional el cual comprende dos dominios separados por una hendidura que permite alternar una conformación cerrada y abierta. Estos dominios tienen arquitecturas  $\alpha/\beta$  conectados por  $\alpha$ -hélices y un pequeño *linker*. El dominio N-terminal está interrumpido por el dominio C-terminal completado por una cola de 34 aminoácidos codificados en el extremo 3' de la secuencia del gen [Michel G *et al.*, 2003] (Figura 2).

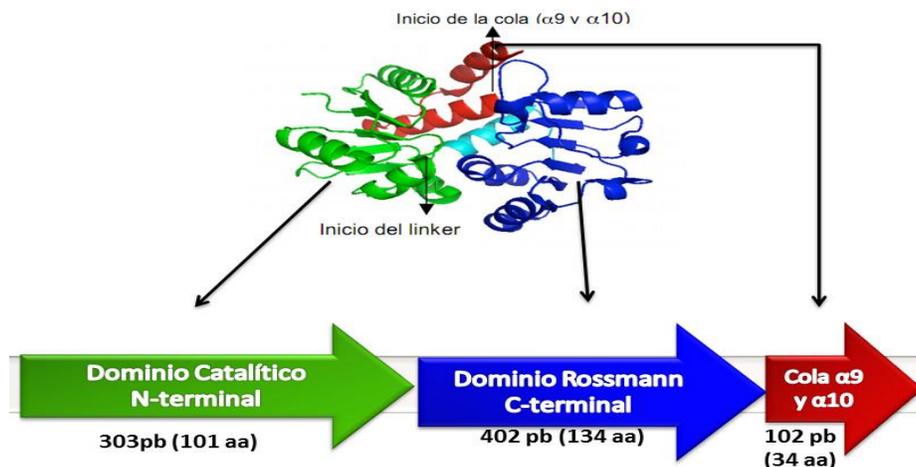
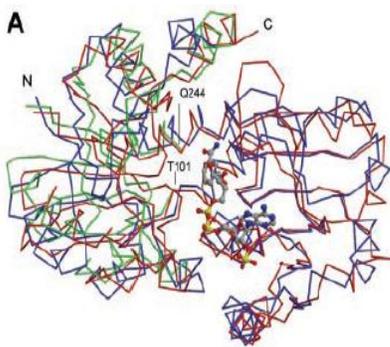


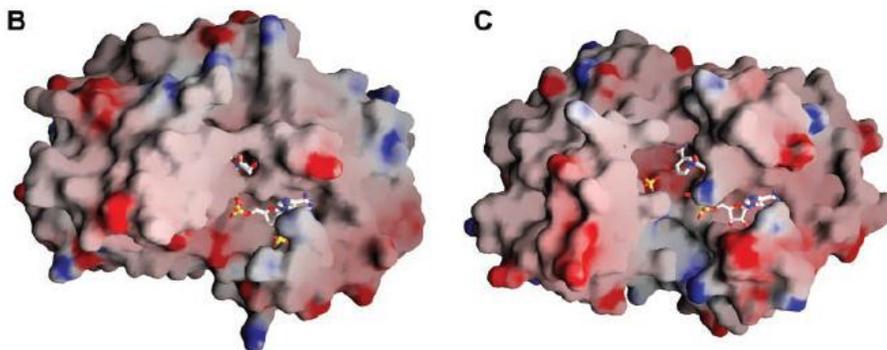
Figura 2. Estructura de AroE de *E.coli*. Las flechas de colores simbolizan la estructura del gen *aroE*. Las flechas verde y roja pertenecen al dominio N-terminal, se observa como el dominio Rossmann (azul) interrumpe la secuencia del primer dominio (Modificado de Espinosa A 2010).

El arreglo de estos dominios a lo largo de las hélices conectadas crea una hendidura en donde se localiza el cofactor  $\text{NADP}^+$  ó ( $\text{NAD}^+$ ). Tanto AroE como YdiB llevan a cabo un cambio conformacional al cerrar los dos dominios una vez unido el cofactor y sustrato. Los dos residuos de unión a sustrato se repositionan para poder llevar a cabo la catálisis (**Figura 3**) [Michel G *et al.*, 2003]. AroE y YdiB comparten un porcentaje de identidad del 52 %. Los genes *aroE* y *ydiB* que codifican a las isoenzimas, presentan un nivel de identidad del 26.7%.



**Figura 3. Estructuras de AroE y YdiB.**

**A)** Moléculas de AroE y YdiB superpuestas en el sitio de unión a nucleótidos con NADPH en representación *ball and stick*. En rojo se muestra la conformación cerrada de AroE, mientras que en azul se muestra YdiB en conformación abierta. En verde se representa la conformación intermedia de AroE. Los dos residuos que actúan como bisagras para el cambio conformacional se representan numerados en AroE. **(B)** Representaciones de superficie de AroE, en conformación cerrada **(C)** Representaciones de superficie de YdiB en conformación abierta. **B** y **C** se colorean de acuerdo a su potencial electrostático con  $\text{NADP}^+$  y  $\text{NAD}^-$ . (Tomada de Michael G *et al.*, 2003).



A continuación, se hará una breve descripción de la estructura y características específicas de los dominios que conforman a las enzimas AroE y YdiB.

### I. III. I Dominio catalítico

El dominio catalítico o de unión a sustrato posee un dominio sandwich  $\alpha\beta\alpha$  el cual también incluye una estructura de hairpin  $\alpha$ -hélice al final de su C-terminal. El dominio activo está definido por la posición del anillo de nicotinamida del cofactor y está delimitado mayoritariamente por los residuos del dominio N-Terminal. El sitio activo es un hueco formado por el extremo del dominio C-terminal y cadenas  $\beta$  del dominio C-terminal y el N-terminal, las hélices  $\alpha_1$ , una cara de la hélice  $\alpha_9$ , el loop extendido entre  $\beta_1$  y  $\alpha_1$ , y los primeros residuos de la hélice conectora  $\alpha_5$  [Michael G *et al.*, 2003].

En Michael *et al.*, (2003) se reporta que la reacción que se lleva a cabo en el sitio activo de AroE es la oxidación de shikimato utilizando NADPH como cofactor; a su vez YdiB lleva a cabo la misma reacción utilizando indistintamente  $\text{NAD}^+$  o  $\text{NADP}^+$  como cofactores, y también lleva a cabo la oxidación del DHQ para producir quinato. En cuanto al reconocimiento del sustrato, AroE es altamente específica para shikimato utilizando solamente NADPH como cofactor cuando YdiB tiene una eficiencia catalítica más baja ( $\sim 400$ ) para el shikimato comparada con AroE [Michael *et al.*, 2003]. Por otra parte, en Cortés-Tolalpa (2011) y Escalante 2012 (comunicación personal), se sugiere que la catálisis de YdiB, parece estar dirigida solamente en la síntesis de quinato, sin tener ninguna participación en la síntesis de shikimato a partir de DHS, y coinciden con Michael *et al.*, (2003) en cuanto la actividad de AroE (Figura 4).

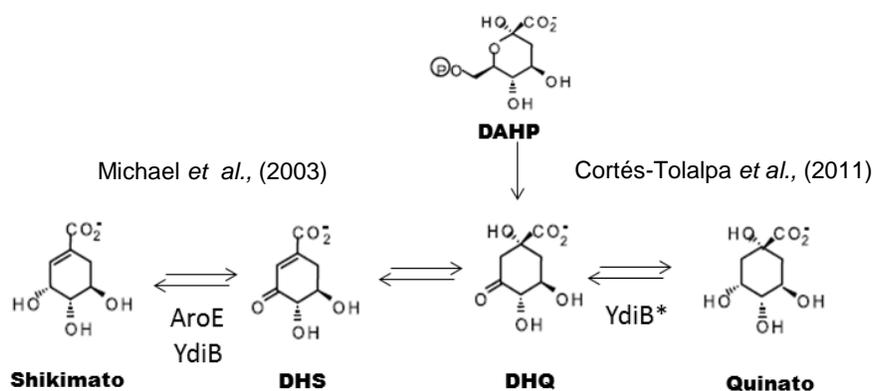


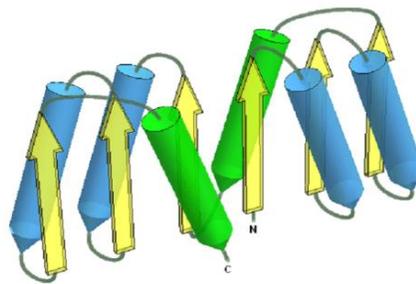
Figura 4. Reacciones catalizadas por AroE y YdiB. En Michael *et al.*, (2003), se reporta actividad de quinato/shikimato para YdiB, mientras que en Cortés-Tolalpa *et al.*, (2011), se observa que YdiB\* solo presenta actividad de quinato deshidrogenasa. Ambos trabajos coinciden con la actividad de shikimato deshidrogenasa para AroE.

### I. III. II Plegamiento Rossmann

El dominio C-terminal es un dominio de unión a dinucleótidos (NADP/NAD) que adopta la forma canónica del plegamiento Rossmann [Michel G *et al.*, 2003; Padyana AK y Burley SK 2003; Lindern HA, *et al.*, 2005]

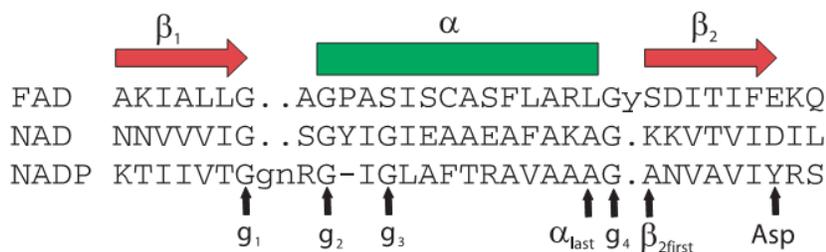
El plegamiento Rossmann es característico de enzimas deshidrogenasas, las cuales catalizan la oxidación o reducción de un sustrato empleando coenzimas que actúan como aceptores o donadores de electrones y protones:  $\text{NAD}^+/\text{NADH}$ ,  $\text{NADP}^+/\text{NADPH}$ ,  $\text{FAD}^+/\text{FADH}_2$  siendo respectivamente la forma oxidada y reducida de la coenzima.

La estructura del dominio Rossmann está compuesta de 6 hojas- $\beta$  paralelas, en orden 3-2-1-4-5-6, y  $\alpha$ -hélices a ambos lados (**Figura 5**) [Bottoms CA *et al.*, 2002, García-Guevara 2012].



**Figura 5. Topología del plegamiento Rossmann. Se encuentra estructurado con 6 hojas- $\beta$  paralelas (amarillo) [Tomado de García Guevara, 2012].**

En la primera parte del dominio Rossmann ( $\beta 1\alpha 1\beta 2$ ) se encuentran 3 residuos de glicina rodeados por residuos hidrofóbicos. Los residuos al final de la hoja  $\beta 2$  ayudan en la identificación de la coenzima (i.e. si la enzima se une a  $\text{FAD}^+$ ,  $\text{NAD}^+$  O  $\text{NADPH}$ ) [Kallberg Y, y Persson B 2006] (**Figura 6**). La presencia de residuos cargados negativamente indican que  $\text{FAD}^+$  o  $\text{NAD}^+$  es preferiblemente el cofactor, por el otro lado, si se encuentran residuos básicos en la posición que precede del segundo residuo de glicina indica preferencia de la enzima por  $\text{NADP}$  [Kallberg Y y Persson B 2006].



**Figura 6. Secuencias con plegamiento Rossmann,  $\alpha$  y  $\beta$  denotan elementos de la estructura secundaria. Las fechas indican la posición de importancia crítica para la predicción de la especificidad de la coenzima. [Tomado de Kallberg Y y Persson B, 2006].**

Existe una molécula de agua conservada en el dominio Rossmann, lo que genera puentes de hidrógeno con el dinucleótido y el asa rica en glicinas; además parece estabilizar el asa y mantener al cofactor en una conformación extendida, así como contribuir entálpicamente a la energía libre del cofactor [Bottoms CA *et al.*, 2002].

El dominio Rossmann se puede encontrar en combinación con otro tipo de plegamientos, en el N-terminal o en el C-terminal [Brenner SE *et al.*, 1996, Bottoms CA *et al.*, 2002]. En el caso de las SDH's el dominio Rossmann está insertado en el dominio catalítico, a nivel de secuencia génica como se observa en la **figura 2** [Bashton M y Chothia C 2002].

Además de ser característico de enzimas deshidrogenasas, el dominio Rossmann también se ha reportado en enzimas con amplias actividades catalíticas tales como hidrolasas, liasas e isomerasas, siendo así un dominio ampliamente distribuido y abundante en la naturaleza [Kallberg Y y Persson B, 2006]. Por lo anterior se ha propuesto que el plegamiento Rossmann es uno de los 9 dominios más antiguos. En el trabajo de Caetano-Anollés (2007) se describe la evolución de 776 plegamientos tomando como base la clasificación estructural de proteínas (SCOP). Mediante la reconstrucción de un árbol filogenético, se muestra que el dominio Rossmann es uno de los plegamientos que aparecieron más tempranamente en la evolución, además de estar ampliamente distribuido en los proteomas de todos los organismos que han sido completamente secuenciados (**Tabla 2**) [Caetano-Anollés *et al.*, 2007].

Clave SCOP	PLEGAMIENTO
c.37	Loop-P que contiene el plegamiento nucleosido trifosfato hidrolasa
a.4	Plegamiento de unión a DNA/RNA (haz triple helicoidal
c.1	TIM $\beta/\alpha$ -barrel
c.2	Rossmann. de unión a dinucleotidos NAD(P)
d.58	Ferredoxina-like
c.23	Flavodoxina-like
c.55	Motivo Ribonucleasa H-like
b.40	Plegamiento OB
c.66	S-adenosill-L-metionina-dependiente de metiltransferasas

**Tabla 2. Los 9 plegamientos más ancestrales, y ampliamente distribuidos del árbol filogenómico que describe la diversificación del mundo de las proteínas [Caetano-Anollés *et al.*, 2007].**

## II. ANTECEDENTES

### II. I Recombinación de dominios de shikimate deshidrogenasas

Los miembros de la familia de shikimate deshidrogenasa (SDH) son un modelo interesante ya que como se mencionó anteriormente, cuentan con dos dominios claramente definidos: un dominio versátil y antiguo - el plegamiento Rossmann- y un dominio catalítico. Las enzimas SDH representan a una familia ampliamente distribuida con un papel esencial en el metabolismo secundario, en el cual se sintetizan metabolitos que no están involucrados directamente en la reproducción o crecimiento del organismo, si no que tienen funciones vitales complementarias, como por ejemplo la síntesis de antibióticos, pigmentos, toxinas, entre otros (Demain AL 1998).

La familia SDH se divide en 4 grupos de acuerdo a diferencias bioquímicas y funcionales: AroE, YdiB, SdhL y RfiI [Michel G *et al.*, 2003; Ye S *et al.*, 2003; Guo J, y Frost JW., 2004].

En Singh *et al.*, (2008) se realizó un análisis de ~250 secuencias homólogas de la familia SDH, pertenecientes a genomas bacterianos que representan la mayoría de los grupos taxonómicos. Se determinaron las relaciones filogenéticas de los miembros de la familia SDH mencionados anteriormente. Cabe resaltar que uno de los resultados de este análisis filogenético fue la observación de que YdiB de *E. coli*, se encuentra relacionada a los homólogos de otras bacterias entéricas Gram Positivas (Firmicutes). Debido a que el análisis obtenido para esta proteína, no es congruente con el árbol filogenético de los taxones bacterianos obtenido por análisis del gen del 16S rDNA, se propone que la secuencia de YdiB fue adquirida por transferencia horizontal, es decir es una proteína xenóloga en *E. coli* [Singh S *et al.*, 2008].

Por lo anterior surgen diferentes interrogantes relacionadas con la diversificación de la misma familia SDH y los papeles biológicos de sus miembros: ¿Qué presiones selectivas han impulsado la diversificación a lo largo de esta familia? ¿Cuántas veces se ha producido la diversificación de la familia? ¿El conjunto de homólogos SDH de un linaje puede proporcionar información sobre la biología del organismo? Y finalmente ¿es posible

recombinar dominios entre enzimas shikimato deshidrogenasa de diferentes especies y generar quimeras funcionales? [Singh *et al.*, 2008].

Tomando en cuenta la última cuestión, en el grupo del Dr. Lorenzo Segovia se llevó a cabo la recombinación de genes representantes de la familia SDH: AroE de *E. coli* (**E**), YdiB de *E. coli* (**Y**) y AroE de *B. subtilis* (**B**). Se construyeron 6 quimeras dobles, al substituir exclusivamente el dominio Rossmann que se encuentra interrumpiendo la secuencia del dominio catalítico.

Las quimeras se nombraron de acuerdo al origen del N-terminal (primera letra), origen del C-terminal o dominio Rossmann (segunda letra) y cola de 34 aa (última letra) resultando en las siguientes quimeras: EBE, BEB, EYE, YEY, BYB, YBY. Las quimeras se clonaron en el vector de expresión pT4, y fueron caracterizadas mediante dos análisis: el primero fue el análisis de la capacidad de plegamiento de las quimeras por medio del ensayo de CMI (Concentración Mínima Inhibitoria a cloranfenicol) y geles de SDS PAGE los cuales sirvieron para determinar el nivel de expresión y solubilidad de las quimeras. El segundo análisis fue funcional, en el que se realizaron ensayos de complementación en medio mínimo M9 sin aminoácidos aromáticos, de una cepa carente de la actividad SDH [Espinosa A, 2010; Banda J 2011; Martínez-de la Escalera 2012].

En general se observó que las quimeras son inestables e incapaces de complementar una cepa de *E.coli* carente del gen *aroE* (AroE) salvo en el caso de la quimera EBE, la cual mostró niveles de complementación en medio mínimo M9, semejantes a la cepa transformada con el gen *aroE* de *E. coli*. En los trabajos de Espinosa (2010), Banda (2011) y Martínez de la Escalera (2012), a nivel estructural, no se sabe a donde ocurrieron los cambios generados durante la recombinación de dominios, que impiden un buen plegamiento y por lo tanto nula actividad en las quimeras construidas.

La única quimera activa EBE, tiene una semejanza con la quimera EYE en cuanto al origen del dominio N-terminal y la cola de 34 aminoácidos, los cuales provienen del gen *aroE* de *E. coli*. Por otra parte el dominio Rossmann proviene del gen *aroE* de *B. subtilis*. La diferencia en el origen del dominio Rossmann puede representar una limitación en cuanto

a semejanza estructural, sin embargo se sabe que el dominio Rossmann es un plegamiento estructuralmente muy conservado a lo largo de muchos linajes bacterianos. Por otra parte las quimeras EYE y EBE comparten un porcentaje de identidad a nivel de secuencia génica del 66.19%, y a nivel de secuencia de proteína del 54.55%. Por lo anterior se puede sugerir que la proteína quimérica EYE presenta un parecido estructural a EBE y por lo tanto podría ser una quimera funcional.

En este trabajo se sugiere que la quimera EYE necesita un mínimo de cambios para poder adquirir un plegamiento correcto, y por lo tanto una actividad que se refleje en los ensayos de complementación. Un acercamiento que puede posibilitar la obtención de variantes activas, es el enfoque de evolución dirigida de proteínas. Es importante tomar en cuenta que algunos autores consideran este enfoque como un apartado dentro de la ingeniería de proteínas [Bornscheuer U, Kazlauskas RJ 2011; Bloom JD *et al.*, 2005], sin embargo, en este trabajo se hará una clara distinción entre ingeniería de proteínas y evolución dirigida.

## **II. II Ingeniería de proteínas**

Uno de los primeros ejemplos donde se utilizó la ingeniería de proteínas, es el trabajo de [Estell DA *et al.*, 1985] en donde modificaron una proteasa tolerante al blanqueador detergente ampliamente utilizado en la industria. A partir de este trabajo, se ha incrementado de manera impresionantemente la lista de reportes en donde se modifican enzimas mediante ingeniería de proteínas, lo anterior ha tendido grandes contribuciones particularmente en el área del mejoramiento de actividades de enzimas, alteración de sustrato y especificidad del producto [Foo JL *et al.*, 2012]. Como ejemplo de ello se puede mencionar la modificación de enzimas para la síntesis de intermediarios farmacéuticos de sitagliptina, utilizado para el tratamiento de la diabetes [Savile CK *et al.*, 2010], modificación para la síntesis de poliésteres [McCarthy A, 2003] y manufactura de etanol celulósico [Gray KA *et al.*, 2006].

De esta forma, la ingeniería de proteínas ha sido un recurso ampliamente utilizado en la biotecnología. Sin embargo esta técnica cuenta con ciertas limitaciones que pueden ser cruciales para asegurar el éxito en la modificación de proteínas.

Con la ingeniería de proteínas se puede alterar la estructura de una proteína para mejorar o cambiar sus propiedades; introduciendo mutaciones en los sitios donde se “cree” que puedan tener un efecto sobre la función de la proteína y utiliza la información evolutiva para seleccionar el tipo de cambio que se desea. [Segovia L y Peimbert M 2010]. [Bornscheuer U y Kazlauskas RJ, 2011]. Esto se realiza con ayuda de la información de la variabilidad de cada residuo, la posición tridimensional y la interacción con otros residuos en el espacio [Anantharaman V *et al.*, 2003].

Por lo anterior, la ingeniería de proteínas depende del conocimiento *a priori* de una gran cantidad de información estructural y evolutiva que no siempre esta disponible. Aunado a esto, actualmente no se pueden hacer predicciones razonablemente precisas sobre los efectos que tendrán mutaciones específicas en las proteínas a nivel teórico [Segovia L y Peimbert M, 2010]. En ausencia de información estructural o evolutiva, es imposible diseñar cambios que permitan entender la relación estructura-función de una proteína. Sin embargo, existe otro enfoque que no se limita con estos parámetros para poder operar [Segovia L, Soberón X, 2008]. Este enfoque se denomina evolución dirigida, el cual se encarga de realizar cambios al azar en el gen de la proteína de estudio, y posteriormente seleccionar las clonas modificadas de acuerdo a las propiedades deseadas.

### **II. III Evolución dirigida**

Las proteínas son máquinas moleculares de notable complejidad, constituidas por cientos de aminoácidos lo que resulta en un impresionante sistema con varios miles de átomos, lo que conlleva indiscutiblemente a una ardua tarea para determinar experimentalmente la posición relativa de estos átomos en el espacio. Aunado a esto, se conocen solo algunas miles de estructuras de proteínas cuando se tienen registradas millones de secuencias proteicas [Segovia L, Soberón X, 2008]. Aunque existen técnicas cada vez mas poderosas que permiten suponer cual es la estructura tridimensional correcta a partir solo de la secuencia de aminoácidos, existen un sin fin de secuencias proteicas de las cuales no se tiene una idea clara de cual pueda ser su estructura, ya sea por limitaciones experimentales, falta de capacidad o tiempo para determinarlas [Segovia L, Peimbert M,

2010]. Aunque en teoría es posible llegar a conocer los sistemas proteínicos y modificarlos con base en predicciones específicas, actualmente existen muchas limitaciones para lograrlo de manera correcta [Segovia L y Soberón X, 2008].

En la actualidad existe un enfoque interesante que logra modificar proteínas en un periodo relativamente corto de tiempo. Este enfoque se denomina evolución dirigida, tomando en cuenta su símil con el proceso de la evolución natural que opera sobre tres componentes fundamentales: la variación genética, la reproducción y la selección de los individuos ganadores. La variabilidad genética es el motor de la evolución, con el cual la selección natural actúa para ejercer un efecto controlador sobre el número de variantes sobrevivientes y seleccionando las que confieren mayor capacidad de adaptación. En el caso de la evolución dirigida, la variación genética se obtiene a través de ciclos sucesivos de mutagénesis. La reproducción se logra por medio de la amplificación de las variantes por medio de ciclos de duplicación bacteriana y finalmente la selección se basa en identificar la propiedad deseada de las variantes, con un método de selección adecuado [Segovia L y Soberón X 2008; Segovia L y Peimbert M 2010].

Este principio de introducción de variabilidad acoplado a selección, es conceptualmente bastante sencillo desde un punto de vista práctico, sin embargo, existen limitaciones que han evitado que se despliegue de manera significativa su tremendo potencial. Estas limitaciones se presentan tanto en la generación de diversidad, como en los métodos de búsqueda y selección [Segovia L y Peimbert M 2010].

La evolución dirigida de proteínas es una importante rama de la biotecnología, además de ser un acercamiento poderoso para optimizar funciones y características de diversas proteínas o también para desarrollar proteínas con características totalmente diferentes o novedades funcionales. [Dougherty MJ y Arnold FH 2009; Böttcher D y Bornscheuer UT 2010; Kaur J y Sharma R 2006; Jackel C. *et al.*, 2008].

En la evolución dirigida, al igual que en la ingeniería de proteínas, se modifica el gen de la proteína de estudio, sin embargo no se hace de una manera racional, si no que se genera variación al azar dando como resultado una mayor diversidad de secuencias de DNA

alterado o mutado, que al traducirse, originan proteínas con residuos diferentes a la proteína parental, posteriormente pueden ser seleccionadas por la propiedad deseada [Bornscheuer U y Kazlauskas RJ 2011].

Las técnicas de evolución dirigida más utilizadas actualmente, se basan en métodos de mutagénesis al azar y recombinación de genes parentales (DNA *shuffling*) [Stemmer WP 1994a, 1994b; Ness JE *et al.*, 1999; Matsuura T y Yomo T 2006; Shen B 2002; Harayama S 1998]. Estas técnicas han demostrado ser una poderosa herramienta para la optimización y modificación de las proteínas: [Labrou NE 2010; Blagodatski A y Katanaev VL 2011].

Para el caso de este trabajo sería inútil la utilización de la técnica de DNA *shuffling* debido a que se crean mosaicos de quimeras que contienen fragmentos de diferentes proteínas, imposibilitando la utilización de la quimera EYE tal y como fue diseñada. Aunado a esto, las variantes generadas mediante esta técnica, contendrían demasiados cambios estructurales, por lo que se dificultaría la identificación de aquellos que contribuyeron a la mejora, razón por la cual se mencionarán solo los aspectos fundamentales del método de mutagénesis al azar, utilizado en este trabajo.

## II. IV Mutagénesis al azar

La mutagénesis al azar es una herramienta que permite elucidar las relaciones de estructura función de una proteína. Es ampliamente utilizada para la modificación de proteínas para mejorar o alterar sus características. La forma más común de introducir mutaciones sencillas, es a través de una reacción en cadena de la polimerasa, propensa a errores (epPCR) [Wilson DS y Keefe AD, 2000]. Al copiar el DNA en condiciones de reacción no óptimas, es decir, agregando grandes cantidades de  $Mg^{2+}$  o  $Mn^{2+}$  y concentraciones desbalanceadas de dNTP, se incrementa la probabilidad de incorporar nucleótidos incorrectos de 0.001% a 1% [Leung DW *et al.*, 1989; Cadwell RC y Joyce GF 1992]. Posteriormente los productos de la PCR con mutaciones, son introducidos en un vector de expresión y en el banco resultante, se pueden seleccionar variantes buscando los cambios deseados en la actividad de la proteína [Miyazaki K y Takenouchi M 2002; Miyazaki K 2003]. En métodos totalmente *in vivo*, se emplean cepas de *E. coli* carentes de algunas

enzimas reparadoras de DNA favoreciendo sustituciones de forma similar a la técnica de epPCR [Bornscheuer UT *et al.*, 1998, Blagodatski A y Katanaev VL, 2011].

La mutagénesis al azar permite identificar mutaciones benéficas en ausencia de información estructural, o cuando dichas mutaciones son difíciles de predecir a partir de la estructura de la proteína [You L y Arnold FH 1996; Wan L *et al.*, 1998; Bloom JD *et al.*, 2005]. Por anterior, la técnica de mutagénesis al azar es una herramienta muy poderosa y ampliamente utilizada en estudios de evolución dirigida siendo así que se ha empleado para mejorar la selectividad, estabilidad, y actividad de enzimas. Como muestra de esto se puede mencionar la creación de catálisis enantioselectiva, a partir variantes de lipasas de *B. subtilis* de estructura desconocida, en donde se utiliza una sola ronda de epPCR, para producir docenas de ciclohexano monooxigenasas con selectividad R o S [Reetz MT, *et al.*, 2004].

Otro ejemplo del uso de la técnica epPCR, para el mejoramiento de la selectividad fue el trabajo de Mahan SD, *et al.*, 2004, en el cual mejoraron drásticamente el uso de una citosina deaminasa diseñada para usarse en células tumorales.

Por otra parte, también existen reportes donde utilizan la técnica de mutagénesis al azar, donde además de mejorar la estabilidad, actividad, y catálisis de enzimas, se evidencian mutaciones seleccionadas en residuos inesperados difíciles de predecir mediante un enfoque de diseño racional. Como muestra de ello se pueden mencionar los trabajos de [Kumar S *et al.*, 2005; Kim D y Guengerich FP 2004] en donde utilizaron el citocromo P450 para someterlo a múltiples rondas de epPCR, aumentando la actividad para muchos componentes biológicos activos. Una gran cantidad de las mutaciones se encontraron distantes del sitio activo y no se pudieron haber predicho mediante un método de diseño racional.

A su vez, en Kim YW *et al.*, (2004) utilizaron 2 rondas de PCR para seleccionar y optimizar una glucosa sintasa de *Agrobacterium sp.*, aumentando su actividad catalítica y expandiendo el repertorio de sustratos. Esta glucosa sintasa había sido caracterizada previamente con diseño racional y al igual que en [Kumar S, *et al.*, 2005; Kim D, Guengerich

FP 2004], las mutaciones seleccionadas se encontraron distantes del sitio activo. De manera similar en [Lin H, *et al.*, 2004], se encontraron 4 sustituciones de aminoácidos distantes del sitio activo de mutantes glucosa sintasas optimizadas por epPCR.

Otro ejemplo de optimización de enzimas mediante mutagénesis al azar, ha sido la modificación enzimática de la TEM-1  $\beta$ -lactamasa, donde por mutagénesis y selección para la resistencia de un antibiótico  $\beta$ -lactámico de tercera generación (aztreonam), se crearon mutantes capaces de hidrolizar el antibiótico, aumentando la resistencia 150-veces [Camps M, *et al.*, 2003].

## II. V Frecuencia de mutagénesis

La mayoría de las estrategias de evolución dirigida involucran pocos cambios en la secuencia de la proteína, debido a que las enzimas ofrecen un rango de actividades débiles promiscuas, que son rápidamente alteradas con pocas mutaciones [Aharoni A *et al.*, 2005]. Para estudiar la función de una proteína, existen estudios que sugieren crear una sola sustitución de aminoácido, es decir una tasa de mutagénesis baja, lo cual corresponde aproximadamente a 1.5 mutaciones por cada 1000 pb [Vartanian JP *et al.*, 1996]. Esto permite caracterizar independientemente el efecto de cada sustitución de aminoácido, en la función de la proteína. Por otra parte, existen reportes donde describen que las tasas de mutagénesis que varían entre 2-7 mutaciones por cada 1000 pb, se consideran más eficientes para la creación de bancos o bibliotecas de proteínas mutantes con actividades mejoradas [Shafikhani S *et al.*, 1997; Wan L *et al.*, 1998; You L y Arnold FH, 1996]. También se ha reportado que las tasas de mutagénesis altas, las cuales superan 6 mutaciones por cada 1000 pb, generalmente resultan en una completa pérdida de la actividad de la proteína [Suzuki M *et al.*, 1996]. Sin embargo existen trabajos en donde describen que las proteínas pueden tolerar niveles extremos de mutagénesis [Vartanian JP *et al.*, 1996]. Siendo así, la frecuencia de mutación, un criterio a considerar para analizar la relación de estructura-función de proteínas, en estudios de evolución dirigida.

## II. VI Métodos de selección

Otro criterio de igual importancia para asegurar el éxito en un estudio de evolución dirigida, es el método de selección. Estudios demuestran que los mejores métodos de selección son los que se basan en la búsqueda directa de las variantes deseadas, ya sea por complementación de mutantes carentes de esta actividad, o por conferir resistencia a algún antibiótico [Arnold FH y Georgiou G 2003; Foo JL *et al.*, 2012; Jackel C *et al.*, 2008].

### III. PLANTEAMIENTO

La quimera EYE es una proteína que no muestra actividad en ensayos de complementación en medio mínimo M9. Aunado a esto, se desconoce el efecto de los cambios ocurridos durante la recombinación de dominios que pudieron afectar su función. Sin embargo se cuenta con información valiosa de las estructuras tridimensionales de las isoenzimas parentales (AroE y YdiB). Además existen estructuras homólogas las cuales se pueden modelar bien y finalmente se cuenta con la quimera activa EBE la cual es muy parecida estructuralmente a EYE.

Un acercamiento que permitirá obtener variantes activas de EYE, es la evolución dirigida que permite explorar la variación al azar. Esta variación se obtendrá al modificar las frecuencias de mutagénesis (baja, media y alta).

## IV. HIPÓTESIS

Se pueden obtener variantes activas de EYE por evolución dirigida.

## V. OBJETIVOS

### V. I Objetivo General

Someter a evolución dirigida mediante mutagénesis al azar, a la quimera EYE para la obtención de variantes activas.

### V. II Objetivos Particulares

- Construcción de 3 bancos o librerías de mutagénesis, usando como templado el gen de la quimera EYE.
- Analizar la variabilidad general de 10 variantes obtenidas en cada banco mediante análisis bioinformático.
- Identificar, aislar y caracterizar las clonas de cada banco que presenten actividad en el ensayo de complementación en medio mínimo.

## VI. MATERIALES Y MÉTODO

### VI. I Cepas

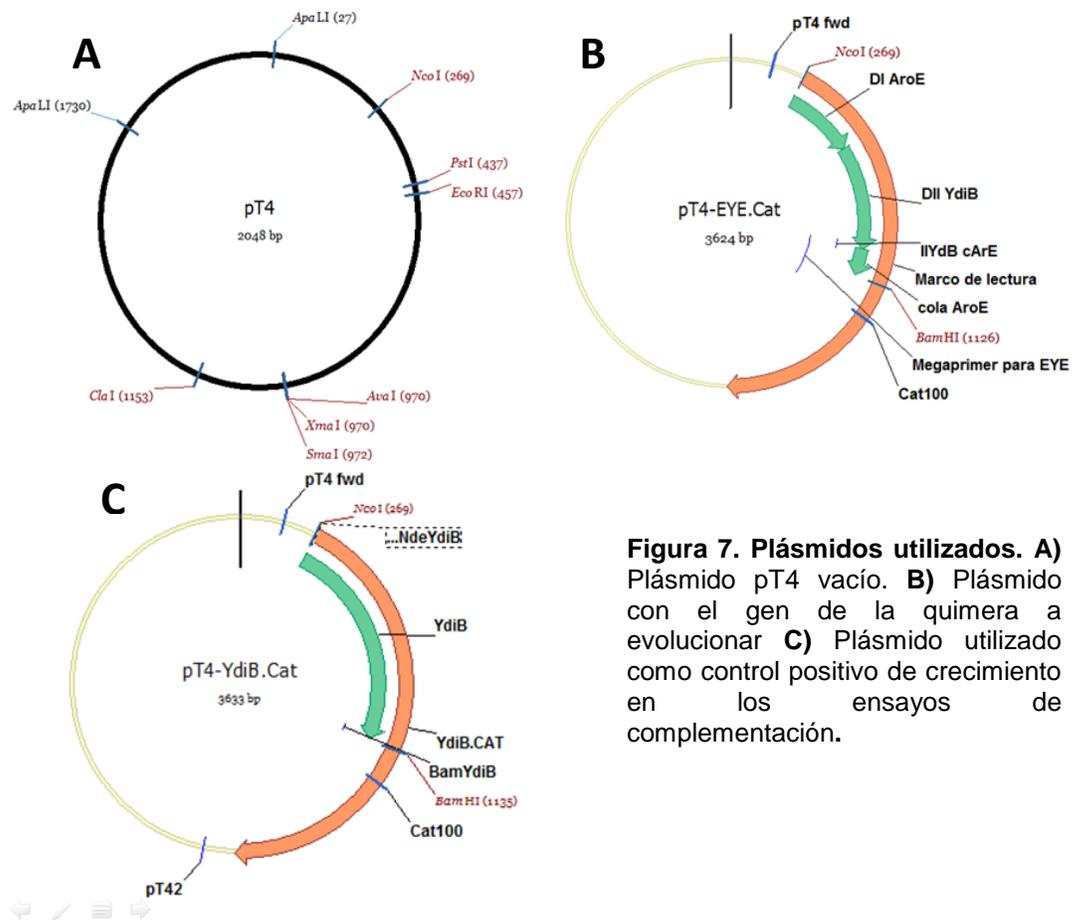
Las cepas utilizadas en este trabajo se muestran en la **tabla 3**.

<b>Cepa</b>	<b>Características</b>
MC1061TrpF	<p><b>Uso:</b> Se utilizaron para amplificar los bancos de mutagénesis (Ver anexo 1)</p> <p><b>Genotipo:</b> F- <math>\Delta(\text{ara-leu})7697</math> [araD139]B/r <math>\Delta(\text{codB-lacI})3</math> galK16 galE15 <math>\lambda</math>- e14- mcrA0 relA1 rpsL150(strR) spoT1 mcrB1 hsdR2(r-m+) TrpF</p>
BW25113:	<p><b>Uso:</b> Se utilizo para los ensayos de complementación en medio mínimo.</p> <p><b>Genotipo:</b> Gen <i>aroE</i> deletado (AroE-) de la colección de Keio de <i>E.coli</i> K-12 [Baba T <i>et al.</i>, 2006]. También se le eliminó el gen de resistencia a Km (AroE- Km-) [Espinosa A, 2010]. <math>\Delta\text{aroE}</math>: lacIq rrnB T14 <math>\Delta\text{lacZ}</math> WJ16 hsdR514 <math>\Delta\text{araBAD}</math> AH33 <math>\Delta\text{rhaBADLD78}</math> (Km<sup>r</sup>) (Keio collection).</p>

**Tabla 3. Cepas utilizadas.**

### VI. II Plásmidos

Las construcciones utilizadas en este trabajo fueron donados por [Banda J 2011] (**Figura 7**). Las secuencias YdiB y EYE, están clonadas en el vector pT4, el cual presenta un origen de replicación de alto número de copias (*ColE1*, presente en pUC), un promotor *pTrc* fuerte, un gen de resistencia a kanamicina (Km) y finalmente el gen de la enzima cloranfenicol acetil transferasa (CAT) como reportero de plegamiento. La actividad de resistencia cloranfenicol se correlaciona con un correcto plegamiento de CAT y de la proteína a la que se encuentra fusionada [Espinosa A, 2010].



**Figura 7. Plásmidos utilizados. A)** Plásmido pT4 vacío. **B)** Plásmido con el gen de la quimera a evolucionar **C)** Plásmido utilizado como control positivo de crecimiento en los ensayos de complementación.

### VI. III Mutagénesis al azar

Se utilizó el método de mutagénesis al azar, empleando el kit comercial de Diversify® *PCR Random Mutagenesis*, que permite controlar el nivel de mutación al azar, variando las concentraciones de manganeso y dGTP en la reacción de PCR. La tasa de mutagénesis aumenta, al incrementar la concentración de manganeso (cerca de 640  $\mu\text{M}$ ) y también si se aumenta la concentración de dGTP en la reacción. En este trabajo se utilizaron las tasas de mutagénesis baja, media y alta; las cuales crean 2, 3.5 y 8.1 mutaciones por cada 1000 pb respectivamente.

## VI. IV Generación de los bancos de mutagénesis

- 1. Extracción de DNA plásmidico para la obtención de vector:** Se puso a crecer un cultivo de 50mL LB/Km 25 µg/mL durante 18 h, 250 rpm, 30°C, con la construcción pT4-EYE-CAT, a partir del cual se hizo una extracción de plásmido utilizando Zyppy™ Plasmid Maxiprep. Se obtuvo una concentración de 100ng/µL, las alícuotas plásmido se analizaron en un gel de agarosa al 0.8%.
- 2. Reacciones de digestión:** En todas las reacciones de digestión se utilizó un volumen final de 60µL: 50 µL de plásmido pT4-EYE-CAT, 2 µL *Nco*I, 2µL *Bam*HI, 6 µL Buffer *Bam*HI (Fermentas). Las reacciones se incubaron durante 4 horas a 37°C.
- 3. Purificación del vector:** El vector correspondiente a 2770 pb, se purificó del gel de agarosa con el QIAquick Gel Extraction Kit, donde se obtuvo una concentración de 22.2 ng/µL, y también se purificó el inserto correspondiente 854 pb, con una concentración de 15ng/µL. Los fragmentos purificados se guardaron a -20°C.
- 4. PCR mutagénico:** La construcción EYE-CAT, fue amplificada con los oligos pT4 10pm/µL y CAT-100 10pm/µL (ver secuencia en tabla 4), bajo condiciones mutagénicas (1, 4 y 9) como se describe en Diversify® *PCR Random Mutagenesis*, Clontech (Tabla 4).

<i>Mutaciones por 1000pb</i>	<b>C1</b>	<b>C4</b>	<b>C9</b>
	<b>2</b>	<b>3.5</b>	<b>8.1</b>
Agua	39	36	31
10X TITANIUM <i>Taq</i> Buffer	5	5	5
MnSO <sub>4</sub> (8mM)	0	3	4
dGTP (2mM)	1	1	5
50X Diversify dNTP Mix	1	1	1
50X dNTP Mix	0	0	0
pT4 primer 10pm/µL. 5'-GACATATAAACGGTCTGGCA-3'	1	1	1
CAT-100 primer 10pm/µL. 5'-CCTATAACCAGACCGTTCAG-3'	1	1	1
pT4-EYE-CAT (templado)	1	1	1
TITANIUM <i>taq</i> Pol	1	1	1
Volumen total	50	50	50

Tabla 4. Soluciones de reacción en volúmenes de µL. Condiciones de mutagénesis C1, C4, C9.

- 5. Purificación de PCR:** Los productos de PCR de 1124 pb fueron purificados con el kit comercial GeneJET PCR Purification (Fermentas).

6. **Digestión de los productos de PCR:** Las digestiones de PCR se realizaron con los mismos volúmenes que se mencionaron anteriormente: 50  $\mu\text{L}$  de reacción de PCR mutagénica purificada. 2  $\mu\text{L}$  *Nco*I, 2  $\mu\text{L}$  *Bam*HI, 6  $\mu\text{L}$  Buffer *Bam*HI (Fermentas) 60  $\mu\text{L}$  volumen final.
7. **Purificación de PCR:** Los productos de las digestiones de PCR mutagénico (854 pb) fueron purificados por medio de gel de agarosa 1% usando GeneJET TM Gel Extraction Kit.
8. **Reacciones de ligación:** De la condición C1 y C9 se ligaron 141.9 ng de vector con 131.92 ng de fragmento correspondiente al gen mutagenizado (inserto). De la condición C4 se ligaron 100 ng de vector y 70.77 de inserto. Las reacciones de ligación se prepararon en una relación 1:3 a 16°C por 16 horas en un volumen total de 25  $\mu\text{L}$ . Los plásmidos ligados fueron precipitados con etanol absoluto, lavados con etanol al 70%, y resuspendidos en 15  $\mu\text{L}$  de agua tetradestilada estéril. De cada ligación se tomaron 2  $\mu\text{L}$  para electroporar las células competentes MC1061 $\Delta$ trp. Se recuperaron en medio SOC durante 1 h a 37 °C, 180 rpm. Se realizaron varias reacciones de electroporación para completar el volumen de ligación. Para la generación del último banco de mutagénesis, se pusieron bajo la condición de mutagénesis (9), 10 reacciones de mutagénesis (razón por la cual se nombro C10) como se describe en la **tabla 4**, reacción C9. Posteriormente se purificó de banda utilizando el QIAquick Gel Extraction Kit, donde se obtuvo una concentración de 17 ng/ $\mu\text{L}$ . De este banco C10, se pusieron 15 reacciones de ligación 100 ng de vector y 97.5 ng de inserto, relación 3:1, se dejaron 16 h, 16 °C. Se agregó a cada una de las 15 reacciones, 500  $\mu\text{L}$  de butanol, agitando cada una por vortex, posteriormente se centrifugaron a 13 000 rpm. A continuación se decanto el butanol, y el resto se evaporó en el Savant SpeedVac ® DNA 110 Concentrator por 20 minutos a temperatura media. Las 15 reacciones se resuspendieron en 5  $\mu\text{L}$  de agua, y se concentraron en una sola reacción.
9. **Transformación:** Todas las reacciones de ligación (C1, C4, C9, y C10), se transformaron por el método de electroporación en la cepa MC1061 $\Delta$ trpF. Lo

determinante en este paso fue la eficiencia de transformación de esta cepa, la cual se hizo electro competente por el método descrito en el **Anexo 1** (Ernesto Ortiz comunicación personal). Para los bancos C1, C4, C9 y C10, se obtuvieron 5, 17, 1 y 25 mL respectivamente de células recuperadas en medio SOC, las cuales se dejaron crecer durante 1 h, 37°C a 180 rpm.

- 10. Tamaño de los bancos:** Para estimar el tamaño de los bancos C1, C4, C9 y C10, se platearon 1, 5, 1 y 5 µl respectivamente de las células recuperadas MC1061Δtrp en medio LB/ KM 25 µg/mL, se incubaron 16 h a 30°C, hasta la aparición de colonias las cuales sirvieron para calcular el tamaño de los bancos.
- 11. Amplificación de los bancos:** Para amplificar los bancos, se partió del volumen restante de células transformadas, y se inocularon en medio LB/Km 25 µg/mL, para amplificar los bancos. Los cultivos incubaron 16 h, 30°C, 250 rpm. Una vez saturados los cultivos, se les hizo una extracción de plásmido utilizando Zyppy™ Plasmid Maxiprep Kit.
- 12. Variabilidad general de los banco:** Se extrajo DNA de plásmido de 10 colonias de cada banco, mediante el protocolo High pure Plasmid isolation Kit de Roche. Los productos se mandaron a secuenciar utilizando los oligos pT4 10pm/µL y CAT-100 10pm/µL. Las secuencias obtenidas se compararon con la secuencia de EYE-CAT parental, utilizando el programa CLUSTALW2 multiple sequence alignment.
- 13. Ensayo de complementación medio líquido:** Se transformaron las construcciones: pT4-AroE-CAT, pT4-AroBS-CAT, pT4-YdiB-CAT, pT4-EYE-CAT, y el plásmido vacío pT4, en la cepa BW25113 AroE<sup>-</sup>, Km<sup>-</sup>, competente por medio del método químico DMSO (**Ver Anexo 2**). Se agregaron a las células, ~60ng/µL de cada construcción y se dejaron reposar 10 min en hielo. Las células se recuperaron 1 hora a 37°C a 250 rpm. Posteriormente se inocularon 100ul de células transformadas, en 3 mL de LB/Km 25 µg/mL y los cultivos se dejaron crecer 16 h, 30°C a 250 rpm. Se centrifugó 1mL de los cultivos originales durante 2 minutos a 5000 rpm. Se decantó el sobrenadante, y los pellets se resuspendieron en 1mL de buffer de sales 1X (**Ver composición en anexo 3**). Se centrifugaron a 5000 rpm durante 2 min y volvió a

decanar el sobrenadante. Este proceso se repitió 5 veces y a continuación se resuspendieron los pellets de células en 1 mL de sales 1X. Se tomaron 20  $\mu\text{L}$  de las células anteriores y se inocularon en 2 mL de medio mínimo M9 (Ver anexo 3) con las siguientes concentraciones de aminoácidos aromáticos: 100  $\mu\text{g}/\text{mL}$ , 200  $\mu\text{g}/\text{mL}$ , 300  $\mu\text{g}/\text{mL}$ , 400  $\mu\text{g}/\text{mL}$  y 500  $\mu\text{g}/\text{mL}$ . Los cultivos se incubaron a 37°C, 250 rpm durante 30 horas monitoreándose la D.O a 600 nm en intervalos de 2 horas.

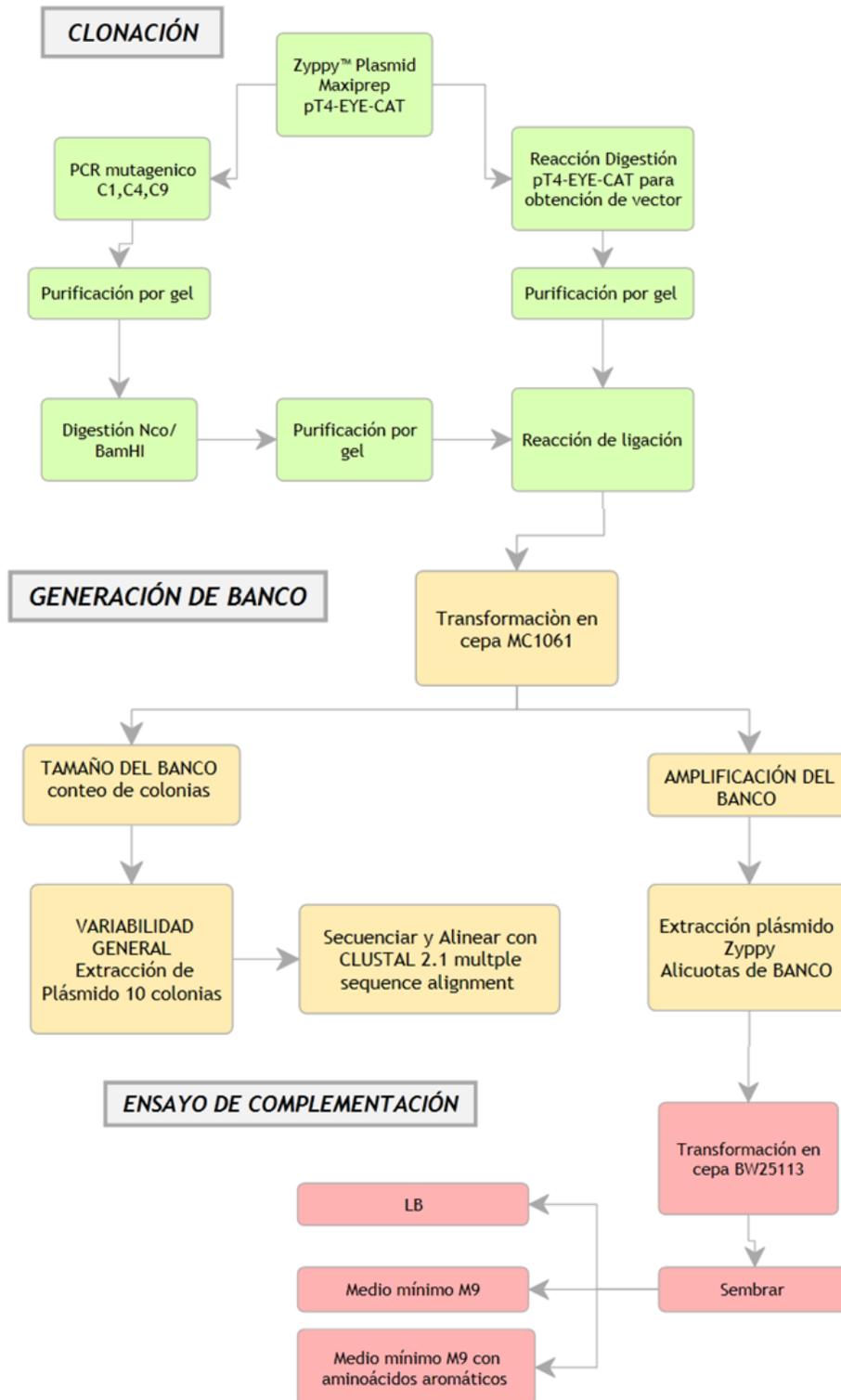
**14. Ensayo de complementación medio sólido:** En todos los casos se transformaron  $\sim 60\text{ng}/\mu\text{L}$  de los plásmidos: bancos (C4, C9, y C10), pT4-YdiB como control positivo y el plásmido vacío como control negativo, y la construcción pT4-EYE-CAT como control de referencia en la cepa BW25113 AroE<sup>-</sup>Km<sup>r</sup>. Se recuperaron en 1mL de medio SOC durante 1:30 horas, a 37°C, 180 rpm. Posteriormente el mL recuperado se centrifugó 2 min, a 5000rpm, y el pellet se resuspendió en 1 mL de sales 1X y se volvió a centrifugar bajo las mismas condiciones anteriores. Éste proceso se repitió 2 veces y finalmente, se resuspendió en 1mL de sales 1X. Una vez lavadas las células con sales, eliminando cualquier resto de medio rico, se inocularon en tres medios diferentes: 1) Medio LB/Km 25  $\mu\text{g}/\text{mL}$ , que sirve como control de viabilidad 2) Medio M9 con AAA, tiamina (Tm 10  $\mu\text{g}/\text{mL}$ ), y Km (25  $\mu\text{g}/\text{mL}$ ), que sirve como control positivo de crecimiento, y finalmente el medio de selección de variantes activas, 3) Medio M9/Km25  $\mu\text{g}/\text{mL}$  y Tm (10  $\mu\text{g}/\text{mL}$ ) sin AAA. Las células se inocularon en los 3 medios descritos anteriormente, de acuerdo a la **tabla 5**. Posteriormente se dejaron crecer a 37°C durante 96 horas hasta la formación de colonias.

<b>Construcción</b>	<b>LB/Km Dilución<math>1 \times 10^{-3}</math></b>	<b>M9/Km/Tm/ AAA Dilución<math>1 \times 10^{-3}</math></b>	<b>M9/Km/Tm</b>
pT4-YdiB	100 $\mu\text{L}$	100 $\mu\text{L}$	100 $\mu\text{L}$ Dilución $1 \times 10^{-3}$
pT4	100 $\mu\text{L}$	100 $\mu\text{L}$	100 $\mu\text{L}$ (Directos)
pT4EYE	100 $\mu\text{L}$	100 $\mu\text{L}$	100 $\mu\text{L}$ (Directos)
C4	100 $\mu\text{L}$	100 $\mu\text{L}$	800 $\mu\text{L}$ en 8 cajas
C9	100 $\mu\text{L}$	100 $\mu\text{L}$	800 $\mu\text{L}$ en 8 cajas
C10	100 $\mu\text{L}$	100 $\mu\text{L}$	800 $\mu\text{L}$ en 8 cajas

**Tabla 5. Células transformadas con 6 construcciones, e inoculadas en 3 medios diferentes 1) LB/Km 25 $\mu\text{g}/\text{mL}$ , 2) M9/Km 25 $\mu\text{g}/\text{mL}$ /Tm10 $\mu\text{g}/\text{mL}$ / AAA 500 mg/mL, 3) M9/Km 25 $\mu\text{g}/\text{mL}$ /Tm 10 $\mu\text{g}/\text{mL}$ .**

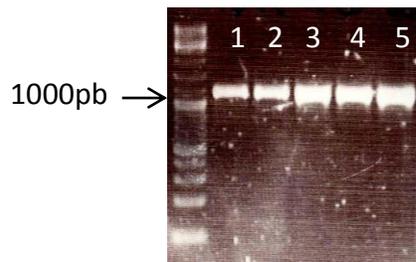
**15. Ensayo de concentración mínima de aminoácidos aromáticos:** Se transformaron por medio de electroporación células BW25113 AroE-Km- con 1  $\mu$ L de los plásmidos de las variantes 1 y 2 del banco C9, y C10, el plásmido vacío como control negativo, y la construcción pT4-EYE-CAT como control de referencia. Se recuperaron en 1 mL de SOC durante 1:30 horas, a 37°C a 180 rpm. Se inocularon en medio LB/Km 25  $\mu$ g/mL, se dejaron crecer durante 16 horas hasta la formación de colonias. Posteriormente, se tomó una colonia de cada transformación y se inoculó en 3 mL de medio LB/Km 25  $\mu$ g/mL, durante 13 horas. Se midió la D.O de los cultivos, y se igualaron a una D.O de 0.6 nm. Se tomó 1 mL de cada uno de los cultivos anteriores y se centrifugaron 2 min a 5000rpm. Se decantaron los sobrenadantes, y se resuspendieron en 1 mL de buffer de sales 1X, a continuación se volvió a centrifugar bajo las mismas condiciones anteriores. Éste proceso se repitió 2 veces y finalmente se resuspendió en 1 mL de sales 1X. Una vez lavadas las células con sales, se realizaron diluciones de  $10^0$  a  $10^{-7}$ . Se sembraron 5  $\mu$ l de las diluciones de cada construcción en 12 medios diferentes, i) LB Km 25  $\mu$ g/mL, ii) M9, Km/Tm 25  $\mu$ g/ml con concentraciones de AAA 1) 500, 2) 250, 3) 125, 4) 62.5, 5) 31.25, 6) 15.63, 7) 7.81, 8) 3.91, 9) 1.95, 10) 0.98, 11) 0  $\mu$ g/mL (medio mínimo M9). Las cajas se incubaron 96h a 37°C.

VII. ESTRATEGIA EXPERIMENTAL



## VIII. RESULTADOS

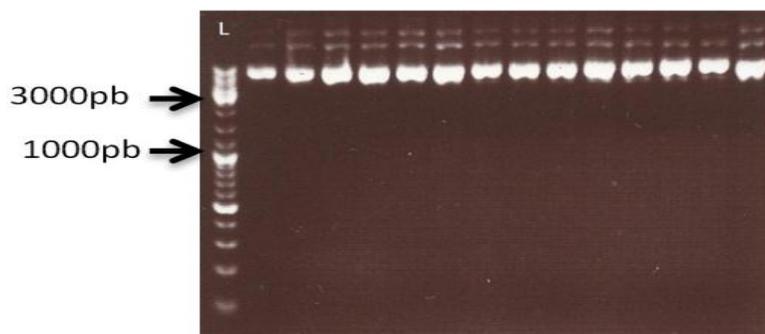
Antes de amplificar la construcción mediante PCR mutagénico, se decidió obtener la temperatura de extensión óptima de los oligos pT4 y CAT100, utilizando como DNA templado pT4-EYE-CAT, mediante PCR de gradiente con las siguientes temperaturas: 61, 60.6, 58.2, 54.7, 53 °C. Los ciclos fueron los siguientes: 94°C por 4 min, 94°C por 30 s, [T° Gradiente] 72 °C por 45 min, 72°C por 5 min, 25°C. Los productos obtenidos se analizaron en un gel de agarosa al 0.8%. Como se observa en la **figura 8**, la temperatura de extensión con la que se obtiene mayor cantidad de producto específico amplificado es 53°C. Por lo que para las reacciones de PCR mutagénicas se utilizó esta temperatura de extensión, con las condiciones anteriores.



**Figura 8. Amplificación de la construcción pT4-EYE.CAT con diferentes temperaturas de extensión: 1) 61°C, 2) 60.6°C, 3) 58.2, °C, 4) 54.7 °C, 5) 53 °C.**

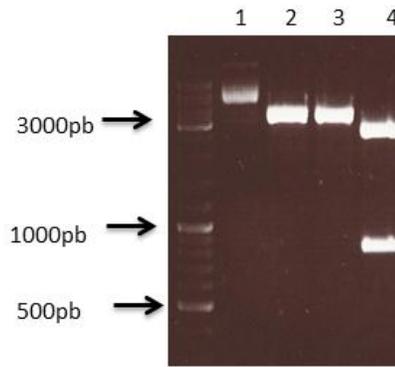
### VIII. I Bancos de mutagénesis

Debido a que la generación de bancos implica clonar los genes mutantes generados por PCR en vectores para su expresión, se procedió a obtener una gran cantidad de vector a partir de la construcción parental pT4-EYE-CAT (**Figura 9**).



**Figura 9. Extracción de plásmido pT4-EYE-CAT. L: marcador de peso molecular. Los carriles siguientes pertenecen al plásmido pT4-EYE-CAT (100ng/ µL).**

Una vez obtenido suficiente plásmido para digerir, se realizó una prueba para verificar el corte de cada una de las enzimas por separado (*Nco*I, *Bam*HI de Fermentas) en sus buffers correspondientes, además de que se hizo una doble digestión en buffer *Bam*HI (Figura 10).



**Figura 10. Digestión para la obtención de vector de clonación carril:**

- 1) Plásmido pT4-EYE-CAT
- 2) Digestión del plásmido con enzima *Bam*HI
- 3) Digestión del plásmido con enzima *Nco*I
- 4) Plásmido digerido con las dos enzimas *Bam*HI y *Nco*I.

Como se puede observar en la **figura 10**, el corte de las enzimas produce los fragmentos esperados, tanto con la enzima *Nco*I como con *Bam*HI, el plásmido se linealiza, y por otro lado, con la doble digestión se obtiene el fragmento correspondiente al gen de la quimera EYE (854 pb) y un fragmento de mayor tamaño, (2770 pb), el cual sirve como vector para clonación.

Se decidió probar 3 condiciones mutagénicas debido a 2 razones. La primera es por que no sabemos cuales son los cambios que necesita la quimera EYE para adquirir actividad y por lo tanto, al explorar diferentes frecuencias mutagénicas se aumenta la probabilidad de encontrar variantes activas. En segundo lugar se ha reportado en estudios de evolución dirigida, que las tasas de mutagénesis que varían entre 1 hasta 9 mutaciones por cada 1000 pb, son eficientes para la creación de bancos de mutagénesis.

### VIII. II Tamaño de los bancos

Para determinar el tamaño de los bancos (número de variantes), se partió del producto de ligación y el conteo de colonias de la transformación. En la **tabla 6** se resume la determinación del tamaño de los bancos.

BANCO	$\mu\text{L}$ de células recuperadas (1)	Colonias (2)	$\mu\text{L}$ Inoculados en LB/km (25 $\mu\text{g}/\text{mL}$ ) (3)	Tamaño Banco (Numero de Variantes) (1*2)/3
C1	5,000	13	1	65,000
C4	17,000	14	5	47,600
C9	1000	124	1	124,000
C10	25,000	117	5	585,000

**Tabla 6. Determinación del tamaño de cada banco (No de variantes).**

### VIII. III Variabilidad general de los bancos

Para determinar la variabilidad generada en los bancos se analizaron las secuencias de 10 colonias obtenidas en cada uno. En total se obtuvieron 40 secuencias crudas, las cuales se editaron manualmente, con ayuda del programa BioEdit y se analizaron mediante el programa CLUSTALW2 multiple sequence alignment (Ver materiales y métodos). Con los datos obtenidos se calculó el número total de cambios de nucleótidos, la tasa mutagénica, así como el tipo de cambio de purina a purina (transiciones) o purina por una pirimidina (transversiones). Por otra parte, las secuencias nucleotídicas se tradujeron a secuencias de proteínas con ayuda del programa ExPASy-Translate. A partir de estos datos, se contabilizó el número de cambios de aminoácidos, codones de paro, mutaciones sinónimas y no sinónimas (**Tabla 7**).

	Banco C-1	Banco C-4	Banco C-9	Banco C-10
No. Variantes	65,000	47,600	124,000	585,000
No. total de cambios de nucleótidos	22	29	84	76
Tasa mutagénica (Total de cambios de nucleótidos / 8540 pb * 100)	0.25% (2.5 cambios por gen)	0.33% (3.30 cambios por gen)	0.98% (9.8 cambios por gen)	0.88% (8.8 cambios por gen)
Transiciones	(100%)	(68.97%)	(88.10%)	(86.84%)
Transversiones	0	(31.03%)	(11.90%)	(13.16%)
Mutaciones sinónimas	100%	(48.28%)	(34.52%)	(47.37%)
Mutaciones no sinónimas	0	(48.28%)	(65.48%)	(52.63%)

**Tabla 7. Análisis de la variabilidad de los bancos.**

Además de los datos anteriores, también se calculó el número de combinaciones posibles, al mutar 1, 2 o 3 posiciones (mutaciones sencillas, dobles y triples respectivamente) en el gen de la quimera EYE (854 pb) (**Ver Anexo 5**). Tomando en cuenta los datos anteriores, también se calculó el tamaño de banco necesario para poder tener representadas a las variantes sencillas, dobles y triples, con una probabilidad de 0.95% para lo cual se utilizó la fórmula de cobertura (**Ver Anexo 6**). Los datos de los anexos 5 y 6 se resumen en la **tabla 8**.

MUTACIONES	VARIANTES	COBERTURA
Sencillas	2,562	7674
Dobles	3,278,079	9,820,246
Triplas	2,792,923,308	8,366,849,787

**Tabla 8. Número de variantes sencillas, dobles y triples posibles y datos de cobertura.**

Como se observa en la **tabla 8**, el número de combinaciones posibles que se pueden obtener al mutar el gen EYE aumenta casi mil veces al cambiar solo una posición, es decir, el número de variantes dobles, es más de mil veces mayor al número de variantes sencillas, así como el número de variantes triples, es casi mil veces mayor al número de variantes dobles. Esto dificulta la exploración de variantes en el laboratorio, ya que como se observa, puede ser extremadamente limitada en comparación con la realizada por la evolución natural, tomemos a consideración el siguiente ejemplo: una proteína de *E. coli* tiene en promedio 319 aminoácidos, para lo cual existen 6061 variantes de mutaciones sencillas, 18 millones de variantes dobles, 36 mil millones de variantes con tres cambios y  $19 \times 10^{319}$

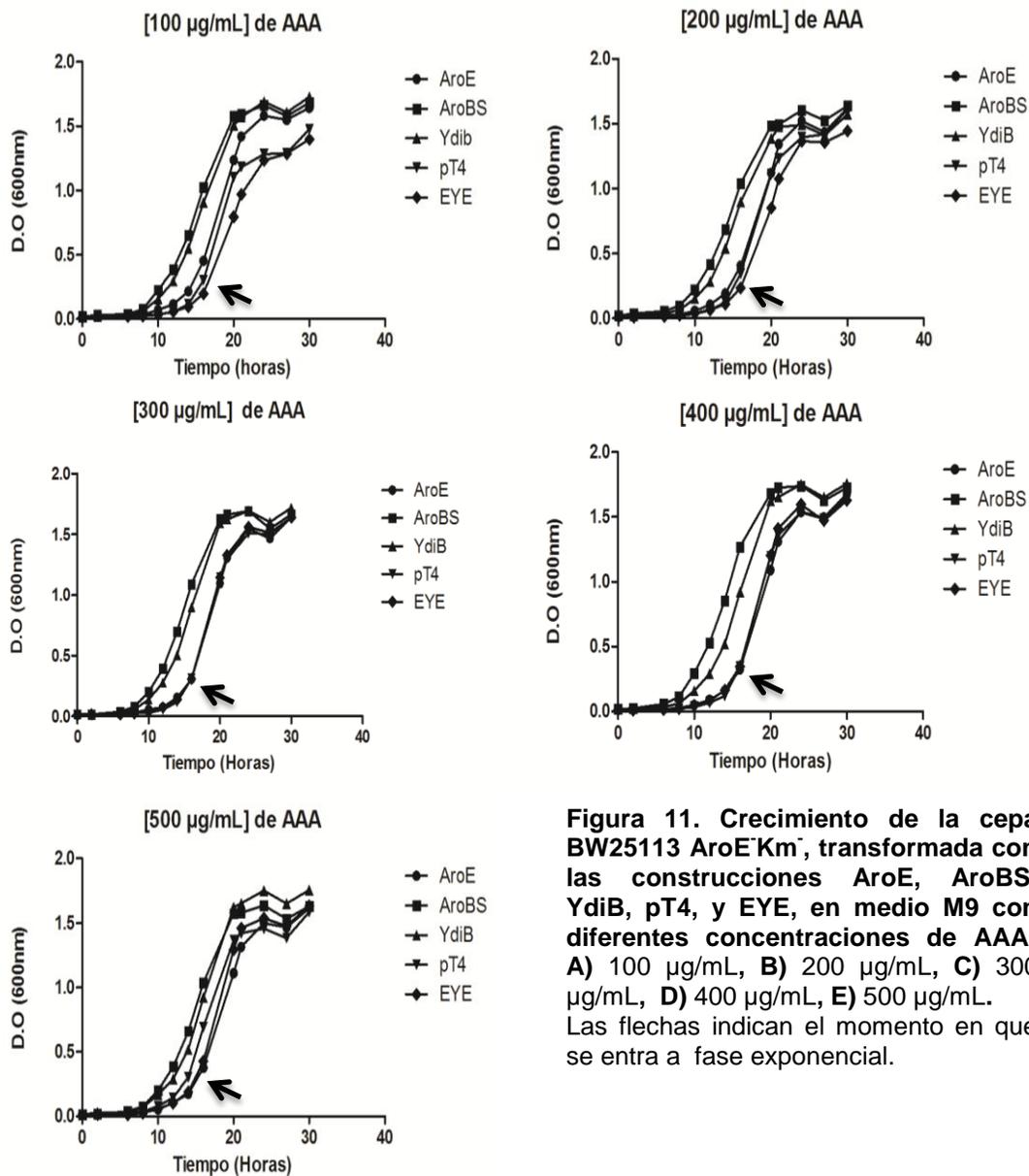
variantes con cambios en todas las posiciones. Este último número es mayor al número de partículas atómicas en el universo [Segovia L y Soberón X 2008]. Debido a lo anterior sólo es posible analizar un número extremadamente restringido de variantes en un experimento.

Un tamaño de banco con un gran número de variantes aumenta las probabilidades de encontrar por lo menos una mutación que sea benéfica para la característica buscada. En este trabajo se obtuvieron bancos de 65000, 47,600, 125,000 y 585, 000 (C1, C4, C9, C10) variantes respectivamente, el último banco C10, al tener mas de medio millón de variantes, aumentaba la posibilidad de encontrar una mutación que permitiera a EYE adquirir actividad de SDH por ser más grande. La actividad *in vivo* de las variantes se buscó a través de ensayos de complementación en medio sólido, el cual se describe en materiales y métodos.

#### VIII. IV Método de selección

Debido a la diversidad de reportes donde se describen diferentes concentraciones de aminoácidos aromáticos para preparar medio mínimo de selección M9 [Espinosa A, 2010, Banda 2011, Chengue 2010, y Flores Soto (comunicación personal)], se decidió diseñar un medio mínimo, con el objetivo de obtener un medio de control positivo de crecimiento en aminoácidos aromáticos (Ver ensayo de complementación en medio líquido en materiales y métodos) (**Figura 11**).

Para elegir el medio de crecimiento adecuado, se decidió utilizar aquel, en el cual la cepa BW25113  $\Delta$ aroE<sup>-</sup> Km<sup>r</sup>, transformada con cualquiera de las construcciones empleadas, tardara el menor tiempo posible en entrar a fase exponencial. Debido a lo anterior se eligió la concentración más alta de aminoácidos aromáticos 500µg/mL (**Figura 11E**). Sin embargo, no se realizaron replicas para obtener la desviación estándar y poder discernir entre diferencias significativas en el crecimiento, ya que el objetivo no era la estandarización del medio, si no un medio que sirviera como control positivo de crecimiento



**Figura 11. Crecimiento de la cepa BW25113 AroE<sup>Km</sup>, transformada con las construcciones AroE, AroBS, YdiB, pT4, y EYE, en medio M9 con diferentes concentraciones de AAA: A) 100 µg/mL, B) 200 µg/mL, C) 300 µg/mL, D) 400 µg/mL, E) 500 µg/mL. Las flechas indican el momento en que se entra a fase exponencial.**

Una vez obtenida la concentración de aminoácidos aromáticos necesaria para el óptimo crecimiento de la cepa BW25113 en medio líquido, se buscó la actividad *in vivo* de las variantes en medio sólido. Lo anterior debido a dos razones fundamentales. En primer lugar, en medio sólido se hace más fácil el aislamiento de las variantes. Y en segundo lugar, en medio líquido se facilita la contaminación de los cultivos de los bancos (datos no mostrados).

Debido a que en el banco C1, todas las variantes secuenciadas mostraron solo 1 o 2 cambios a nivel de gen, que resultaron en la misma secuencia de proteína EYE parental, se decidió no trabajar con este banco para la selección de variantes activas. En el caso de los bancos C4, C9 y C10, se realizó lo descrito en materiales y métodos para el ensayo de complementación en medio sólido. Las células se dejaron crecer a 37°C, durante 96 h hasta la formación de colonias. Con los controles positivos de viabilidad y crecimiento, se esperaba crecimiento de colonias. Por otro lado, en medio mínimo M9, se esperaba encontrar por lo menos una variante activa en cada caso. Las colonias se contabilizaron como se muestran en las **tablas 9, 10 y 11**.

Construcción	Colonias en LB/Km 25 µg/mL	Colonias en M9/ AAA 500 mg/mL	Colonias en M9
pT4-YdiB	600	473	500
pT4	289	251	0
pT4EYE	320	265	0
C4	267	163	0

**Tabla 9. Número de colonias del banco C4 en 3 diferentes medios.**

Construcción	Colonias en LB/Km 25 µg/mL	Colonias en M9/ AAA 500 mg/mL	Colonias en M9
pT4-YdiB	500	379	458
pT4	209	202	0
pT4EYE	200	210	0
C9	111	115	7

**Tabla 10. Número de colonias del banco C9 en 3 diferentes medios.**

Construcción	Colonias en LB/Km 25 µg/mL	Colonias en M9/ AAA 500 mg/mL	Colonias en M9
pT4-YdiB	280	160	132
pT4	241	93	0
pT4EYE	235	80	0
C10	287	100	2

**Tabla 11. Número de colonias del banco C10 en 3 diferentes medios.**

Como se puede observar en la **tabla 9**, no se obtuvo ninguna colonia en el medio de selección en el banco C4, por su parte, en el banco C9 (**tabla 10**) se obtuvieron un total de 7 colonias y finalmente en el banco C10, (**tabla 11**) se contabilizaron 2 colonias. Debido a que era posible que estas 9 colonias pudieran contener alguna quimera activa, se les realizaron los siguientes tratamientos: en primer lugar se restringieron en medio rico (LB/Km 25 µg/mL), posteriormente se pusieron en cultivos líquidos de 3mL LB/Km 25 µg/mL, para extracción de plásmido mediante el protocolo de High pure Plasmid isolation Kit (Roche).

Los productos se mandaron a secuenciar y se compararon con la secuencia de EYE-CAT parental, utilizando CLUSTALW2 multiple sequence alignment. Una vez obtenidas las secuencias de nucleótidos, se analizó el número de mutaciones, la tasa mutagénica y el número de cambios de aminoácidos que tenían en la secuencia proteica.

Mediante la comparación de secuencias de las 7 variantes obtenidas en el banco C9, se observó que 5, contenían la secuencia del gen *aroE*, siendo falsos positivos generados por un problema de contaminación. Sin embargo para términos de este trabajo, esta contaminación no representa un problema más grave, ya que las colonias que contenían la secuencia de *aroE*, se desecharon, y las otras colonias que contenían la secuencia EYE, se almacenaron como posibles colonias que pudieran contener variantes de EYE activas. El total se obtuvieron sólo cuatro posibles variantes activas: 2 secuencias del banco C9 y 2 secuencias del banco C10.

A continuación se muestra la comparación a nivel de secuencia de nucleótidos y aminoácidos, de las cuatro variantes con la secuencia EYE parental, utilizando el programa CLUSTALW2 multiple sequence alignment.

EYE	ATGGAACCTATGCTGTTTTGGTAATCCGATAGCCACAGCAAATCGCCATTTCAT	60
V1.C9	ATGGAACCTATGCTGTTTTGGTAATCCGATAGCCACAGCAAATCGCCATACATTCAT	60
	*****	
EYE	CAGCAATTTGCTCAGCAACTGAATATTGAACATCCCTATGGGCGCGTGTGGCACCATC	120
V1.C9	CAGCAATTTGCTCAGCAACTGAATATTGAACATCCCTATGGGCGCGTGTGGCACCATC	120
	*****	
EYE	AATGATTTTCATCAACACACTGAACGCTTCTTTAGTGTGGTAAAGGTGCGAATGTG	180
V1.C9	AATGATTTTCATCAACACACTGAACGCTTCTTTAGTGTGGTAAAGGTGCGAATGTG	180
	*****	
EYE	ACGGTGCCTTTTAAAGAAGAGGCTTTTGCCAGAGCGGATGAGCTTACTGAACGGGCAGCG	240
V1.C9	ACGGTGCCTTTTAAAGAAGAGGCTTTTGCCAGAGCGGATGAGCTTACTGAACGGGCAGCG	240
	*****	
EYE	TTGGCTGGTGTGTTAATACCCTCATGCGGTTAGAAGATGGACGCCTGCTGGGTGACAAT	300
V1.C9	TTGGCTGGTGTGTTAATACCCTCATGCGGTTAGAAGATGGACGCCTGCTGGGTGACAAT	300
	*****	
EYE	ACCGGCACGGGCATATTCGCGCCATTAAAGAGAGCGGTTTGGATATCAAAGGCAAACG	360
V1.C9	ACCGGCACGGGCATATTCGCGCCATTAAAGAGAGCGGTTTGGATATCAAAGGCAAACG	360
	*****	
EYE	ATGGTGTGTTAGGGGCCGTTGGTGCCTCAACGGCAATTGGCGCGCAGGGGCAATTGAA	420
V1.C9	ATGGTGTGTTAGGGGCCGTTGGTGCCTCAACGGCAATTGGCGCGCAGGGGCAATTGAA	420
	*****	
EYE	GGTTTAAAGAAATTAACCTCTTAAACCGTCGGGATGAGTTCTTCGATAAAGCCCTCGCC	480
V1.C9	GGTTTAAAGAAATTAACCTCTTAAACCGTCGGGATGAGTTCTTCGATAAAGCCCTCGCC	480
	*****	
EYE	TTCGCGCAGGGGTTAATGAAAACACCGATTGTGTCGTACGGTCACCGATCTCGCCGAT	540
V1.C9	TTCGCGCAGGGGTTAATGAAAACACCGATTGTGTCGTACGGTCACCGATCTCGCCGAT	540
	*****	
EYE	CAGCAAGCCTTTGCTGAAGCCCTGGCTTCCGCCGACATTTTAACCAATGGCACAAGTG	600
V1.C9	CAGCAAGCCTTTGCTGAAGCCCTGGCTTCCGCCGACATTTTAACCAATGGCACAAGTG	600
	*****	
EYE	GGTATGAAACCCCTTGAGAATGAATCATTGGTTAATGATATCAGTCTGTTACATCCGGA	660
V1.C9	GGTATGAAACCCCTTGAGAATGAATCATTGGTTAATGATATCAGTCTGTTACATCCGGA	660
	*****	
EYE	CTTCTGGTCACTGAATGCGTGTATAACCCGCATATGACGAAGTTATTGCAGCAGGCGCAA	720
V1.C9	CTTCTGGTCACTGAATGCGTGTATAACCCGCATATGACGAAGTTATTGCAGCAGGCGCAA	720
	*****	
EYE	CAAGCTGGTTGCAAAACGATTGATGGTTAGGAATGCTGGTGGCACAGGCGGCTCATGCC	780
V1.C9	CAAGCTGGTTGCAAAACGATTGATGGTTAGGAATGCTGGTGGCACAGGCGGCTCATGCC	780
	*****	
EYE	TTTCTTCTCTGGCACGGTGTCTGCCTGACGTAGAACCAGTTATAAAGCAATTGCAGGAG	840
V1.C9	TTTCTTCTCTGGCACGGTGTCTGCCTGACGTAGAACCAGTTATAAAGCAATTGCAGGAG	840
	*****	
EYE	GAATTGTCCGCGTG	854
V1.C9	GAATTGTCCGCGTG	854
	*****	

Figura 12. Alineamiento de la variante 1 del banco C9 con el gen de la quimera EYE parental. Las flechas indican las 5 mutaciones encontradas.

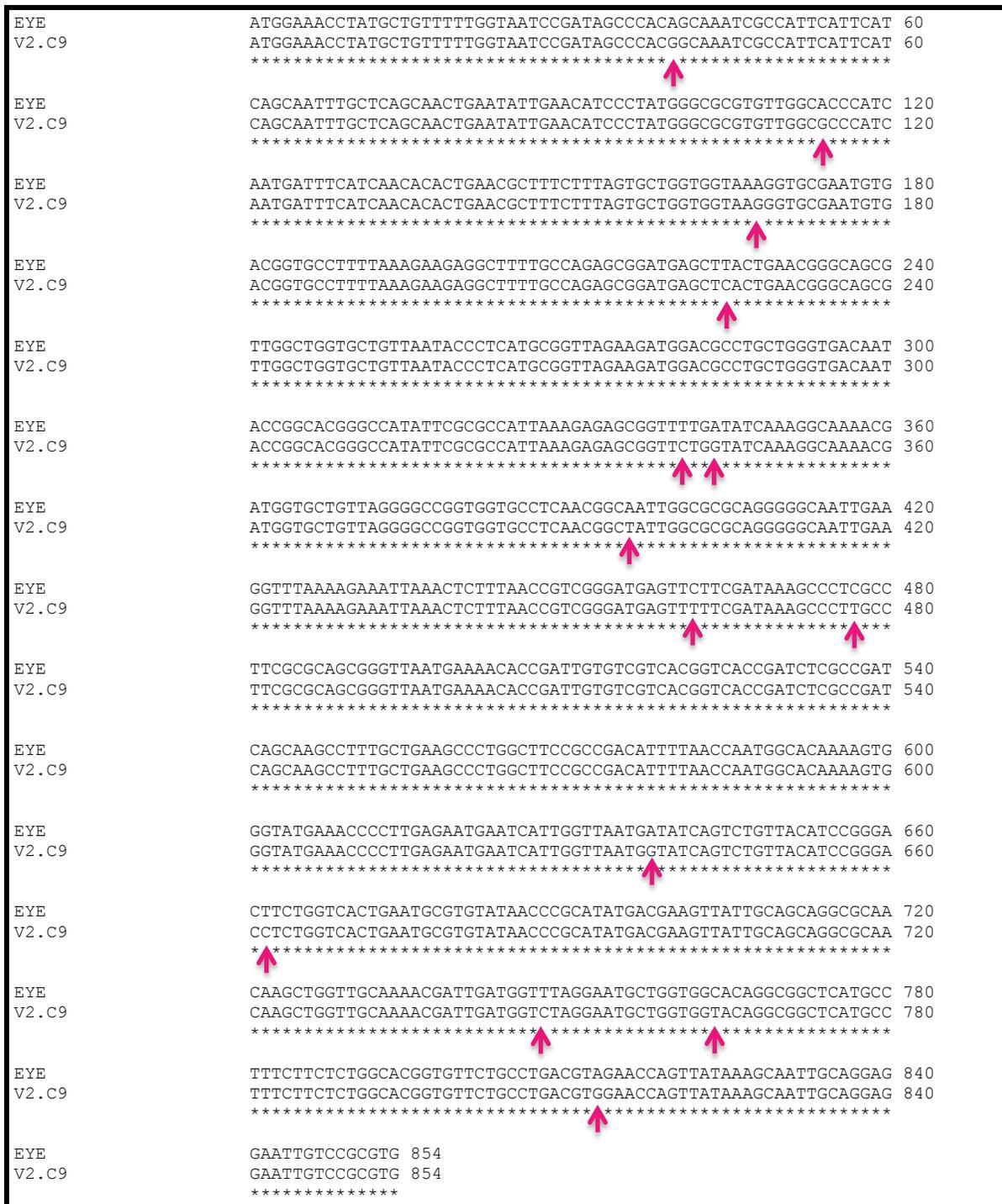


Figura 13. Alineamiento de la variante 2 del banco C9 con el gen de la quimera EYE parental. Las flechas indican las 14 mutaciones encontradas.

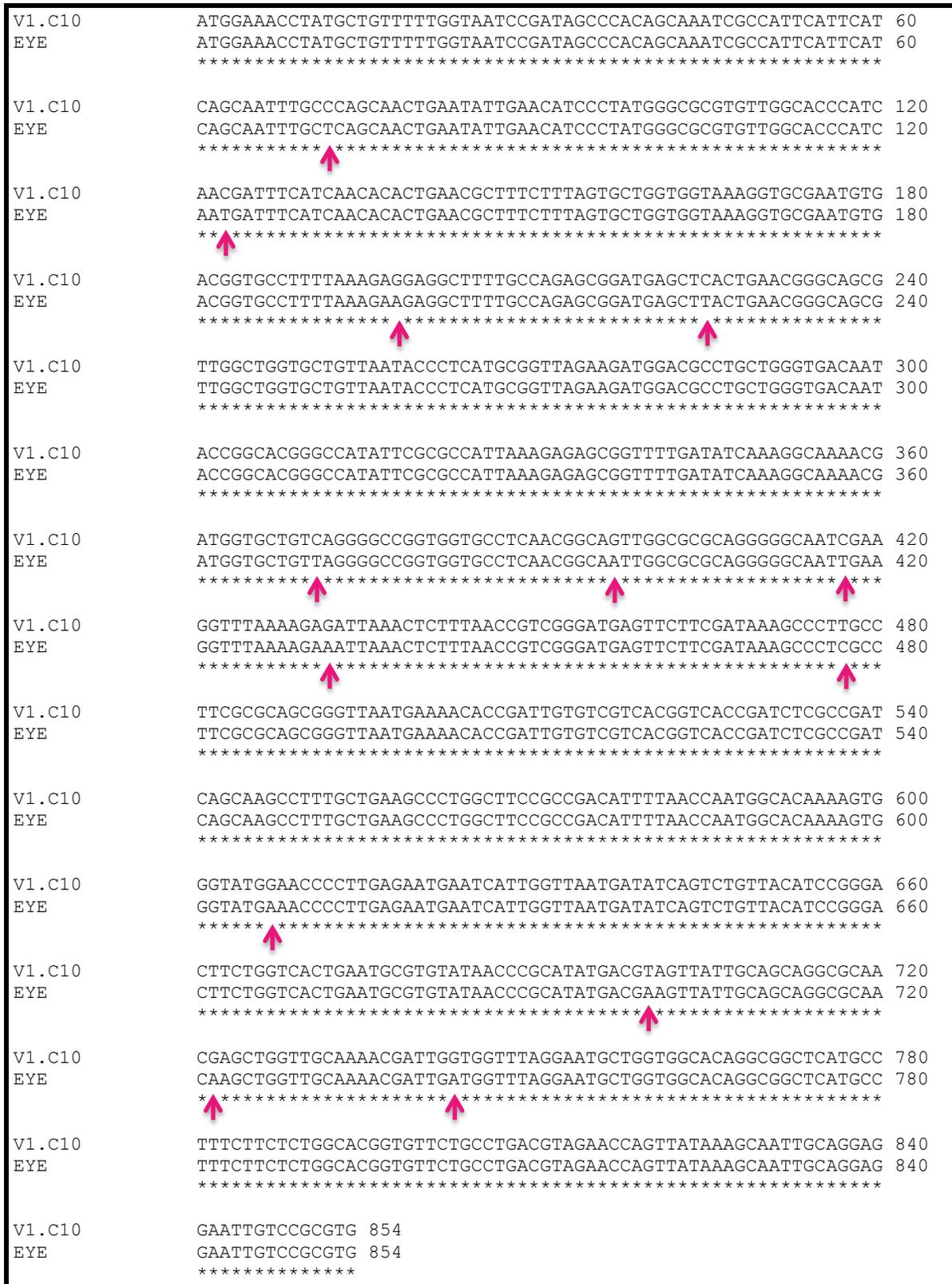
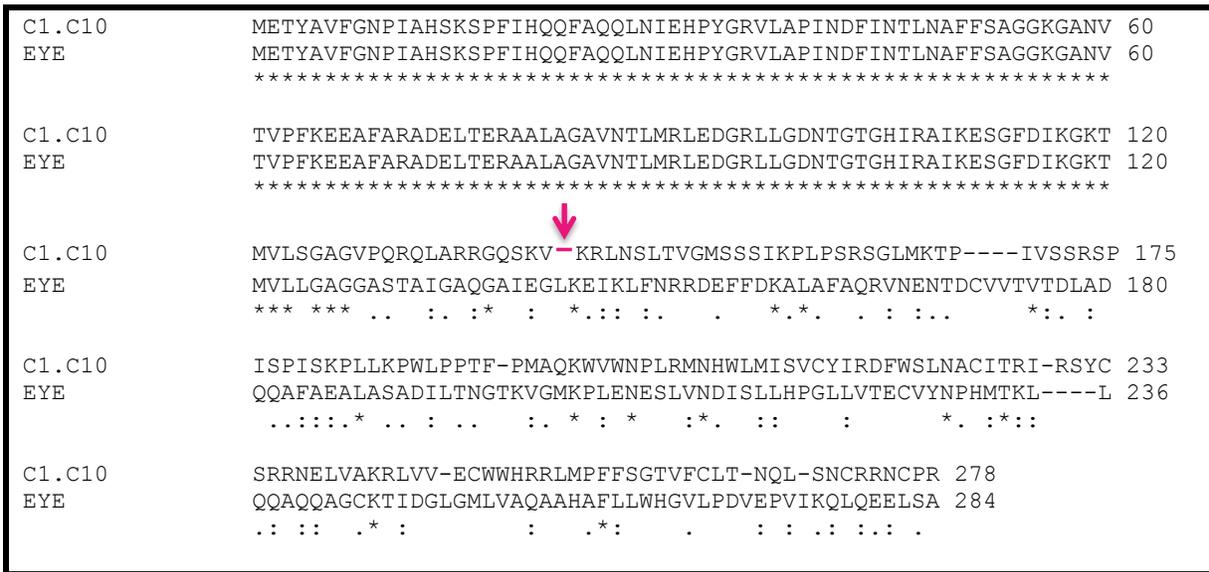


Figura 14. Alineamiento de la variante 1 del banco C10 con el gen de la quimera EYE parental. Las flechas indican las 13 mutaciones encontradas.

V2.C10 EYE	ATGGAACCTATGCTGTTTTTGGTAATCCGATAGCCACAGCAAATCGCCATTCATTCAT ATGGAACCTATGCTGTTTTTGGTAATCCGATAGCCACAGCAAATCGCCATTCATTCAT *****	60 60
V2.C10 EYE	CAGCAATTTGCTCAGCAACTGAATATTGAACATCCCTATGGGCGCGTGTGGCACCCATC CAGCAATTTGCTCAGCAACTGAATATTGAACATCCCTATGGGCGCGTGTGGCACCCATC *****	120 120
V2.C10 EYE	AATGATTTTCATCAACACACTGAACGCTTTCTTTAGTGCTGGTGGTAAAGGTGCGAATGTG AATGATTTTCATCAACACACTGAACGCTTTCTTTAGTGCTGGTGGTAAAGGTGCGAATGTG *****	180 180
V2.C10 EYE	ACGGTGCCTTTTAAAGAAGAGGCTTTTGCCAGAGCGGATGAGCTCACTGAACGGGCAGCG ACGGTGCCTTTTAAAGAAGAGGCTTTTGCCAGAGCGGATGAGCTTACTGAACGGGCAGCG *****	240 240
V2.C10 EYE	TTGGCTGGTGTGTTAATACCCTCATGCGGTTAGAAGATGGACGCCTGCTGGGTGACAAT TTGGCTGGTGTGTTAATACCCTCATGCGGTTAGAAGATGGACGCCTGCTGGGTGACAAT *****	300 300
V2.C10 EYE	ACCGGCACGGGCCATATTCGCGCCATTAAGAGAGCGGTTTTGATATCAAAGGCAAACG ACCGGCACGGGCCATATTCGCGCCATTAAGAGAGCGGTTTTGATATCAAAGGCAAACG *****	360 360
V2.C10 EYE	ATGGTGCTGTTAGGGGCCGGTGGTGCCTCAACGGCAATTGGCGCGCAGGGGGCAATTGAA ATGGTGCTGTTAGGGGCCGGTGGTGCCTCAACGGCAATTGGCGCGCAGGGGGCAATTGAA *****	420 420
V2.C10 EYE	GGTTTAAAAGAGATTAAACTCTTTAACCGTCGGGATGAGTCTTCGATAAGGCCCTTGCC GGTTTAAAAGAAATTAAACTCTTTAACCGTCGGGATGAGTCTTCGATAAGGCCCTTGCC *****	480 480
V2.C10 EYE	TTCGCGCAGCGGGTTAATGAAAACACCGATTGTGTGTCACGGTCACCGATCTCGCCGAT TTCGCGCAGCGGGTTAATGAAAACACCGATTGTGTGTCACGGTCACCGATCTCGCCGAT *****	540 540
V2.C10 EYE	CAGCAAGCCTTTGCTGAAGCCCTGGCTTCCGCCGACATTTTAACCAATGGCACAAAAGTG CAGCAAGCCTTTGCTGAAGCCCTGGCTTCCGCCGACATTTTAACCAATGGCACAAAAGTG *****	600 600
V2.C10 EYE	GGTATGAAACCCCTTGAGAAATGAGTCATTGGTTAATGATATCAGTCTGTTACATCCGGGA GGTATGAAACCCCTTGAGAAATGAGTCATTGGTTAATGATATCAGTCTGTTACATCCGGGA *****	660 660
V2.C10 EYE	CTTCTGGTCACTGAATGCGTGTATAACCCGCATATGACGAAGTTATTGCAGCAGGCGCAA CTTCTGGTCACTGAATGCGTGTATAACCCGCATATGACGAAGTTATTGCAGCAGGCGCAA *****	720 720
V2.C10 EYE	CAAGCTGGTTGCAAACGATTGGTGGTTTAGGAATGCTGGTGGCACAGGCGGCTCATGCC CAAGCTGGTTGCAAACGATTGGTGGTTTAGGAATGCTGGTGGCACAGGCGGCTCATGCC *****	780 780
V2.C10 EYE	TTTCTTCTCTGGCGCGGTGTTCTGCCTGACGTAGAACCAGTTATAAAGCAATTGCAGGAG TTTCTTCTCTGGCACGGTGTCTGCCTGACGTAGAACCAGTTATAAAGCAATTGCAGGAG *****	840 840
V2.C10 EYE	GAATTGTCCGCGTG GAATTGTCCGCGTG *****	854 854

Figura 15. Alineamiento de la variante 2 del banco C10 con el gen de la quimera EYE parental. Las flechas indican las 7 mutaciones encontradas.





**Figura 18. Alineamiento de aminoácidos de la variante 1 del banco C10 con la quimera EYE parental. La flecha indica el codón de paro que modifiqué la secuencia y generó una proteína truncada.**



**Figura 19. Alineamiento de aminoácidos de la variante 2 del banco C10 con la quimera EYE parental. Las flechas indican los 2 cambios de aminoácido.**

El análisis de las variantes 1 y 2 del banco C9 (**Figura 12 y 13**), reveló que éstas contenían 5 y 14 cambios respectivamente a nivel de la secuencia génica. Por otro lado, las variantes 1 y 2 del banco C10, contenían 13 y 7 cambios respectivamente (**Figura 14 y 15**). A nivel de secuencia de aminoácidos, se encontró que la variante 2 del banco C9 (**Figura 17**) y la variante 2 del banco C10 (**Figura 19**) contenían 6 y 2 cambios respectivamente a nivel de residuos. A su vez, la variante 1 del banco C9 (**Figura 16**) y la variante 1 del banco C10 (**Figura 18**), contenían cambios en la secuencia primaria del gen que dieron lugar a codones de paro en la proteína, dando lugar a proteínas truncadas. En la **tabla 12** se resume el análisis de los cambios que se generaron en las cuatro variantes de los bancos C9 y C10.

Banco	Variantes activas	No. Mutaciones	Tasa mutagénica	Cambio aa
C9	1	5	.58%	Proteína truncada
	2	14	1.5%	6
C10	1	14	1.6%	Proteína truncada
	2	7	0.81	2

**Tabla 12. Análisis de las variantes obtenidas en medio mínimo.**

### VIII. V Retransformación

Una vez analizadas las variantes, se decidió observar si repetían el fenotipo, complementando la cepa BW25113 AroE-Km-. Para esto, se utilizaron los plásmidos de las cuatro variantes y los plásmidos descritos en materiales y método. Se transformo mediante el mismo procedimiento que se describe en materiales y métodos para complementación en medio sólido. Las células se pusieron a crecer en los medios: 1) LB/Km 25 µg/mL, 2) M9/Km 25 µg/mL/Tm 10 µg/mL/AAA 500 mg/mL, 3) M9/Km 25 µg/mL/Tm 10 µg/mL, en este último se esperaba que la variante 2 del banco C9 y la variante 2 del banco C10 repitieran el fenotipo. Por otra parte, no se esperaba crecimiento de la variantes 1 del banco C9 y variante 1 del banco C10, ya que éstas daban lugar a proteínas truncadas. Sin embargo en ningún caso se observó crecimiento (**Tabla 13**).

Tabla 13. Número de colonias obtenidas en 3 diferentes medios.

Construcción	Colonias en LB/Km (25 µg/mL) Dilución $1 \times 10^{-3}$	Colonias en M9 /AAA (500 mg/mL) Dilución $1 \times 10^{-3}$	Colonias en M9 100 µL directos
pT4-YdiB	460	360	487
pT4	440	202	0
pT4EYE	365	341	0
V1.C9	460	432	0
V2.C9	433	418	0
V1.C10	455	395	0
V2.C10	426	397	0

Debido a que ninguna de las 4 variantes complementó en medio mínimo (**tabla 13**) se exploró la posibilidad de que alguna de ellas, pudiera tener una mejor actividad comparada con la quimera parental EYE. Se planteó que alguna de las variantes requeriría mínimas de cantidades de aminoácidos aromáticos para poder para poder crecer en un medio mínimo para lo cual se realizó un ensayo de concentración mínima de aminoácidos aromáticos en medio M9 que se describe en materiales y métodos. Las cajas se incubaron hasta la formación de colonias. Como se observa en la **figura 20**, en todos los casos donde el medio contenía aminoácidos aromáticos en concentraciones 1) 500, 2) 250, 3) 125, 4) 62.5, 5) 31.25, 6) 15.63, 7) 7.81, 8) 3.91, 9) 1.95, 10) 0.98 µg/mL, se pudo observar crecimiento. Por lo anterior podemos concluir que no hubo ninguna diferencia entre las variantes obtenidas, la quimera EYE parental y el plásmido vacío, lo cual nos ayudo a confirmar el fenotipo observado en el ensayo de complementación inicial. Esto es que las variantes no crecen en medio mínimo M9 y no existe ninguna diferencia de crecimiento con la quimera EYE parental.

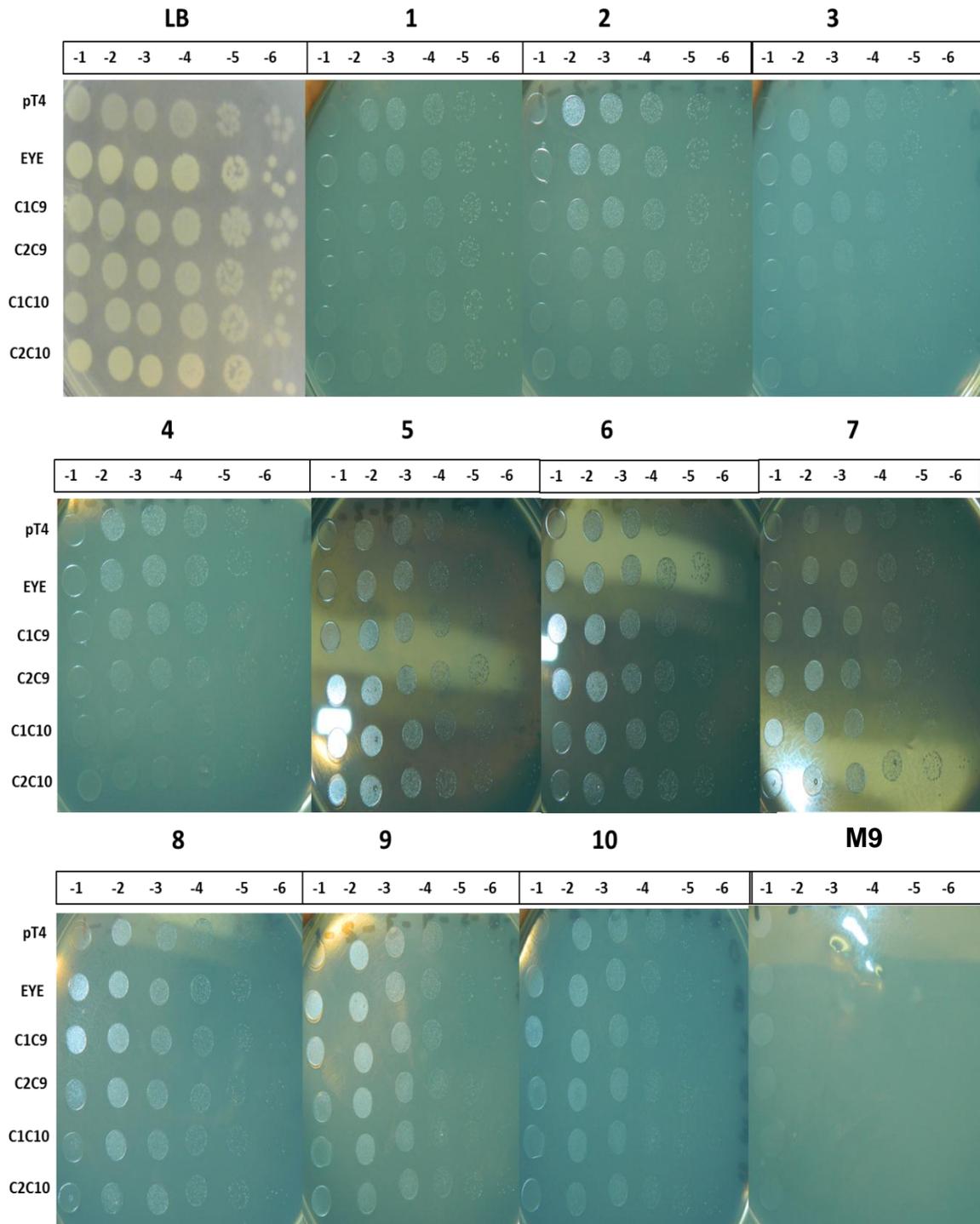


Figura 20. Ensayo de para determinar la concentración mínima de aminoácidos aromáticos.

## IX. DISCUSIONES

Las cuatro variantes obtenidas en los bancos C9 y C10 no tienen actividad de shikimato deshidrogenasa, y su comportamiento fue igual al de la quimera parental EYE, la cual no muestra actividad. Lo anterior nos lleva a suponer que el primer crecimiento de las variantes en medio mínimo M9 (Tabla 10 y 11) se pudo deber a 3 posibles fenómenos:

1. "Fenómeno de Cross-feeding": Ninguna de las mutantes inoculadas es capaz de crecer en medio mínimo, sin embargo se puede llevar a cabo un fenómeno conocido como "cross-feeding" o complementación intrabacteriana. En donde una mutante puede acumular intermediarios que pueden estimular el crecimiento de la mutante vecina, permitiendo su desarrollo en medio mínimo. De esta forma pudieron obtenerse las 4 variantes que en una primera instancia crecieron en medio mínimo, sin embargo cuando se extrajo el plásmido de cada una de ellas y se retransformó en la cepa BW25113 carente de la actividad SDH, se perdió el fenotipo observado.
2. El crecimiento de la cepa BW25113 "*per se*" permitió el crecimiento de las 4 variantes en medio mínimo y no fue debido al plásmido que contenía la cepa: Lo anterior puede explicarse al tomar en cuenta que el potencial global para la generación de diversidad es muy alto en cultivos celulares, pudiendo expandirse de miles a millones. En este caso bajo condiciones de oligotrofia, la cepa de *E.coli* utilizada BW25113 pudo acumular mutaciones en su cromosoma provocando un cambio en el fenotipo, dirigido a incrementar su adaptación a este medio mínimo. Al momento de extraer el plásmido de la bacteria, se perdieron dichas mutaciones el cromosoma y por lo tanto no se observó el mismo fenotipo.
3. Bajo el mismo esquema anterior, la cepa utilizada BW25113 pudo acumular mutaciones en su cromosoma permitiéndole una mayor expresión de YdiB, y debido a que esta enzima tiene actividad de shikimato deshidrogenasa, pudo completar la vía del shikimato, sintetizando los aminoácidos necesarios para su crecimiento en medio mínimo. De la igual manera, al momento de extraer el

plásmido de la bacteria BW23113, se perdió el genotipo adquirido, en este caso, la sobreexpresión de YdiB.

Ahora bien, la falta de actividad en las variantes obtenidas, nos lleva a suponer varios posibles escenarios que expliquen por que no se pudieron obtener variantes activas:

Primero: la técnica de epPCR, no es completamente aleatoria ya que en la incorporación de nucleótidos es más probable cambiar A por T (unidas por 2 puentes de hidrogeno), que G por C (unidas por 3 puentes de hidrogeno) y aún menos probable intercambiar A por C. El gran cambio estructural, al sustituir una purina de dos anillos, con una pirimidina de un anillo, hace este tipo de mutaciones (transversiones) menos probable que el cambio de una purina por una purina, o pirimidina por otra pirimidina (transiciones) [Wong TS *et al*, 2006], y esto se observa claramente en la **tabla 7**, donde el número de transiciones en todos los bancos es significativamente mayor, por ejemplo en el Banco C1, todas las mutaciones obtenidas fueron transiciones 100%, en el caso del banco C4 el 68.97% fueron transiciones, y solo el 31% transversiones, a su vez, se observa que en el banco C9, y C10, se obtiene un porcentaje similar en transiciones 88.10% y 86. 84% respectivamente, y 11.90% y 13% de transiciones obtenidas respectivamente.

Además, del sesgo que se observó en cuanto a la introducción de transiciones con respecto a las transversiones, también se puede decir que existe una limitación en el muestreo de 10 colonias por banco, ya que este último número no es estadísticamente significativo de las variantes obtenidas en cada banco, esto puede explicar las diferencias en la frecuencia de transversiones, transiciones y mutaciones que se muestran en la **tabla 7**.

Por otra parte, podemos afirmar que el número de variantes obtenidas en cada uno de los bancos (65000, 47,600, 125,000 y 585, 000; C1, C4, C9, C10 respectivamente), fue suficiente para poder representar a todas las variantes sencillas, debido a que sólo se necesitan 7674 variantes para representaras variantes con solo una mutación, sin embargo, el tamaño de los bancos obtenidos no fue suficiente para poder representar las variantes dobles y triples, ya que se necesitaría un banco de más de 9 millones de variantes para

poder representar a las mutaciones dobles, y mas de 8 mil millones para representar a todas las mutaciones triples, posibilitando analizar solo una fracción infinitesimal, del vasto número de las posibles combinaciones. En cuanto a este trabajo, se obtuvo solo un banco mayor a medio millón de variantes. Con lo anterior se puede sugerir que la quimera EYE, necesita más de dos o tres cambios para poder adquirir actividad de SDH, los cuales no se vieron reflejados en las variantes analizadas.

Para evitar las limitaciones que pudieran surgir con la técnica de mutagenesis al azar, podemos tomar en cuenta otros métodos que permitan obtener variantes con la característica de actividad deseada, por lo que se puede indicar que en la actualidad, existen muchos reportes en donde se utiliza una combinación de dos acercamientos: diseño racional para limitar los cambios a cierta regiones, e incorporando mutagénesis dirigida para buscar esos cambios que producen mutantes que exhiben las propiedades deseadas.

Para aumentar las probabilidades de obtener quimeras activas, se puede seguir el ejemplo de Sharkey MA y Engel PC (2009), en donde se optimizó el sistema de expresión, con vectores más regulables y con ayuda de chaperonas. En su trabajo, generaron 4 quimeras producto de la recombinación de dos glutamato deshidrogenasas provenientes de *Clostridium symbiosum* y *E. coli*, con el fin de determinar la autonomía de los dominios. Las deshidrogenasas presentan el 54% de identidad a nivel de secuencia de aminoácidos. Las quimeras fueron insolubles, solamente una quimera se pudo caracterizar gracias a la optimización del sistema de expresión con vectores más regulables, y junto con la expresión de una chaperona GroEl/Es.

También se podría utilizar un acercamiento bioinformático y estructural para optimización de quimeras. Como ejemplo de ello se puede mencionar los trabajo de Eisenbeis S *et al.*, (2012) y Bharat TA *et al.*, (2008), donde construyeron una quimera al combinar partes de la enzima de respuesta CheY, y la enzima imidazol glicerol fosfatasa sintasa, ambas de *Thermotoga maritima*. La proteína quimérica resultante, dio lugar a un monómero no estable. Mediante la estructura cristalográfica, y mediante acercamiento de diseño

computacional, se optimizaron las interfaces entre los fragmentos con mutaciones puntuales, dando lugar a una proteína estable. Como muestra de otro reporte donde se utiliza un enfoque bioinformático se encuentra el trabajo de Tripura C, Podile A (2007), en donde introdujeron mutaciones puntuales en regiones conservadas en alineamientos múltiples de unas quimeras de glucosa deshidrogenasa, mejorando así la afinidad de las quimeras a distintos sustratos.

Por lo anterior, sería útil incluir un análisis bioinformático, donde se pudiera conocer la secuencia del mayor número de ortólogos posibles de SDH para determinar la variabilidad de las secuencias, identificando las zonas de menor variabilidad, asociadas al sitio activo o a zonas reguladoras, construir un alineamiento múltiple identificando secuencias homólogas e identificar los sitios en donde han caído las mutaciones seleccionadas.

Independientemente del acercamiento bioinformático de secuencias, sería interesante dilucidar la estructura tridimensional de la quimera EYE con ayuda de la cristalografía de rayos X, para explorar la naturaleza conservativa de las mutaciones. Sin embargo para esto, se necesitaría primero que la quimera se produjera soluble y en gran cantidad.

## X. CONCLUSIONES

Se utilizó el enfoque de evolución dirigida mediante mutagénesis al azar, para obtener variantes activas a partir de la quimera EYE. Esta última es producto de la recombinación entre dominios de las enzimas xenólogas: shikimato deshidrogenasa AroE de *E. coli* y quinato/shikimato deshidrogena YdiB de *E.coli*.

En este trabajo, se generaron cuatro bancos utilizando tres diferentes tasas de mutagénesis: baja, media, y alta. Los bancos con tasa de mutagénesis alta generaron el mayor número de variantes.

Se extrajo DNA plásmidico de 10 variantes seleccionadas de cada banco, y las secuencias fueron compradas con la quimera EYE parental, a nivel de secuencia de nucleótidos y aminoácidos, utilizando el programa CLUSTALW2 multiple sequence alignment.

Para observar la variabilidad, se analizaron 10 secuencias de cada banco, lo que corresponde a 3, 416, 000 nucleótidos. Mediante formulas de combinatoria binomial, se determinó que los bancos obtenidos fueron lo suficientemente robustos para representar todas las combinaciones entre variantes sencillas, sin embargo no lo fueron para representar las dobles, triples etc.

En un primer ensayo de selección en medio mínimo, se obtuvieron cuatro variantes, a las cuales se les extrajo el plásmido y se compararon con la secuencia EYE parental. Se observó que dos de estas variantes tenían codones de paro en la proteína codificada lo que daba lugar a una proteína totalmente diferente a la quimera EYE parental. Las otras dos variantes tenían seis y dos cambios de residuos. Para comprobar si estas variantes eran activas y repetían el fenotipo se volvieron a crecer en medio mínimo M9, sin embargo en ningún caso se obtuvo crecimiento de variantes.

## XI. PERSPECTIVAS

Existen muchos criterios a considerar para asegurar el éxito en la funcionalidad y actividad de la quimera producida por recombinación de dominios. Esto conlleva a sugerir que se requiere una descripción más completa de como las propiedades de los dominios individuales se combinan para producir la función deseada en la proteína, siendo de suma importancia contar con información de la quimera resuelta (estructura 3D), los dominios recombinados por separado, secuencias y estructuras homólogas etc. Siendo así que el incorporar la información estructural y bioinformática, sin duda potenciara la posibilidad de adquirir actividad en la quimera EYE.

No en todos los trabajos de recombinación de dominios o evolución dirigida, se obtienen proteínas funcionales, por lo que sería interesante llevar a cabo un enfoque de diseño racional en conjunto con información estructural para mutar residuos que se consideren importantes, o que se encuentren conservados en la familia de shikimato deshidrogenasas.

Finalmente, se tiene que tomar en cuenta la limitación que implica el seleccionar colonias al azar de bancos de miles de variantes, por lo anterior puede optimizarse el sistema de selección en medio mínimo M9 para poder tener representadas a más variantes en los bancos obtenidos.

## XII. BIBLIOGRAFÍA

- Aharoni A, Gaidukov L, Khersonsky O, McQ Gould S, Roodveldt C y Tawfik DS (2005). *The 'evolvability' of promiscuous protein functions*. Nat Genet; 37(1):73-6.
- Anantharaman V, Aravind L y Koonin EV (2003). *Emergence of diverse biochemical activities in evolutionarily conserved structural scaffolds of proteins*. Curr Opin Chem Biol; (1):12-20.
- Anton JA y Coggins JR (1988). *Sequencing and overexpression of the Escherichia aroE gene encoding shikimate dehydrogenase*. Biochem J; 249:319-326.
- Arnold FH y Georgiou G., eds. (2003). *Directed enzyme evolution: screening and selection methods*. Methods in Molecular Biology Humana Press, Totowa, N.J
- Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL y Mori H (2006). *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection*. Mol Syst Biol; 2:2006.0008
- Bashton M y Chothia C (2007). *The generation of new protein functions by the combination of domains*. Structure; (1):85-99.
- Benach J, Lee I, Edstrom W, Kuzin A, Chiang Y, Acton T, Montelione G T, y Hunt, J F (2003). *The 2.3-Å crystal structure of the shikimate 5-dehydrogenase orthologue YdiB from Escherichia coli suggests a novel catalytic environment for an NAD-dependent dehydrogenase*. J. Biol. Chem; 278, 19176-19182.
- Banda J (2011) *Análisis de proteínas quiméricas entre AroE y YdiB de Escherichia coli*. Tesis de licenciatura. UNAM.
- Bharat TA, Eisenbeis S, Zeth K y Höcker B (2008). *A beta alpha-barrel built by the combination of fragments from different folds*. Proc Natl Acad Sci U S A; 105(29):9942-7.
- Blagodatski A y Katanaev VL (2011). *Technologies of directed protein evolution in vivo*. Cell Mol Life Sci; 68(7):1207-14.
- Bloom JD, Meyer MM, Meinhold P, Otey CR, mACmILLAN y Arnold FH (2005) *Evolving strategies for enzyme engineering*. Curr Opin Struct Biol; 15 (4): 447-52.
- Bornscheuer U y Kazlauskas RJ (2011). *Survey of protein engineering strategies*. Curr Protoc Protein Sci. 9; Cap 26: Un 26.7.
- Bornscheuer UT, Altenbuchner J y Meyer HH (1998). *Directed evolution of an esterase for the stereoselective resolution of a key intermediate in the synthesis of Epothilones*. Biotechnol. Bioeng; 58:554-559.
- Bottoms CA, Smith PE y Tanner JJ (2002). *A structurally conserved water molecule in Rossmann dinucleotide-binding domains*. Protein Sci; 11(9):2125-37.
- Böttcher D y Bornscheuer UT (2010). *Protein engineering of microbial enzymes*. Curr Opin Microbiol; (3):274-82.
- Brenner SE, Chothia C, Hubbard TJP y Murzin AG (1996). *Understanding protein structure: using scop fo fold interpretation*. Methods Enzymol; 266, 635-643.
- Cadwell RC y Joyce GF (1992). *Randomization of genes by PCR mutagenesis*. PCR Meth. App; 2:28-33.
- Caetano-Anollés G, Kim HS y Mittenenthal JE (2007). *The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture*. Proc Natl Acad Sci USA; 104(22):9358-63.
- Carbone MN y Arnold FH (2007). *Engineering by homologous recombination: exploring sequence and function within a conserved fold*. Curr Opin Struct Biol; 17(4):454-9.
- Camps M, Naukkarinen J, Johnson BP y Loeb LA (2003). *Targeted gene evolution in Escherichia coli using a highly error-prone DNA polymerase I*. Proc Natl Acad Sci USA; 100(17):9727-32.

- Cortes-Tolalpa L, Carmona SB, Cervantes-Salinas A, Garcia MA, Gosset G, Escalante A, Bolivar F (2011). *Ingeniería de vías metabólicas en Escherichia coli para la producción de Shikimate como precursor para la síntesis de compuestos antivirales contra influenza*. Biotecnología y Bioingeniería, 15, 30-47.
- Chenge, M (2010). *Generación de quimeras por intercambio de dos dominios entre dos proteínas parálogas*. Tesis de licenciatura, UNAM.
- Demain AL (1998). *Induction of microbial secondary metabolism*. Int Microbiol; 1(4):259-64.
- Dougherty MJ y Arnold FH (2009). *Directed evolution: new parts and optimized function*. Curr Opin Biotechnol; 20(4):486-491.
- Eisenbeis S, Proffitt W, Coles M, Truffault V, Shanmugaratnam S, Meiler J y Höcker B (2012). *Potential of fragment recombination for rational design of proteins*. J Am Chem Soc; 134(9):4019-22.
- Espinosa A (2010) *Obtención de una quimera activa por la recombinación de dominios entre dos Shikimate deshidrogenasas*. Tesis de Maestría en Ciencias Bioquímicas. UNAM.
- Estell DA, Graycar TP, y Wells JA (1985). *Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation*. J. Biol. Chem; 260:6518-6521.
- Foo JL, Ching CB, Chang MW y Leong SS (2012). *The imminent role of protein engineering in synthetic biology*. Biotechnol Adv; (3):541-9.
- García-Guevara JF (2012). *Sitios con acoplamiento estadístico en la familia CheY*. Tesis de Maestría en Ciencias Bioquímicas. UNAM.
- Gan, J, Wu, Y, Prabakaran P, Gu Y, Li Y, Andrykovitch M, Liu H, Gong Y, Yan H y Ji X (2007). *Structural and biochemical analyses of shikimate dehydrogenase AroE from Aquifex aeolicus: Implications for the catalytic mechanism*. Biochemistry; 46, 9513-9522.
- Giesecke A, Fang R y Joung J (2006). *Synthetic protein-protein interaction domains created by shuffling Cys2His2 zinc-fingers*. Mol Syst Biol; 2. 2:2006.2011.
- Goihberg E, Peretz M, Tel-Or S, Dym O, Shimon L, Frolow F y Burstein Y (2010). *Biochemical and structural properties of chimeras constructed by exchange of cofactor-binding domains in alcohol dehydrogenases from thermophilic and mesophilic microorganisms*. Biochemistry; 49(9):1943-53.
- Gray KA, Zhao L y Emptage M (2006). *Bioethanol*. Curr Opin Chem Biol; (2):141-6.
- Guo J y Frost JW (2004). *Synthesis of aminoshikimic acid*. Org Lett; 6:1585-1588.
- Han C, Hu T, Wu D, Qu S, Zhou J, Ding J, Shen X, Qu D, y Jiang H (2009) *X-ray crystallographic and enzymatic analyses of shikimate dehydrogenase from Staphylococcus epidermidis*. FEBS J; 276, 1125-1139.
- Harayama S (1998). *Artificial evolution by DNA shuffling*. Trends Biotechnol; 16(2):76-82.
- Herrmann KM y Weaver LM (1999). *The shikimate pathway*. Annu Rev Plant Physiol Plant Mol Biol; 50:473-503.
- Jackel C, Kast P y Hilvert D (2008). *Protein design by directed evolution*. Annu Rev Biophys; 37:153-173.
- Kallberg Y y Persson B (2006). *Prediction of coenzyme specificity in dehydrogenases/reductases*. A hidden Markov model-based method and its application on complete genomes. FEBS J; 273(6):1177-84.
- Kataoka K, Takada H, Tanizawa K, Yoshimura T, Esaki N, Ohshima T y Soda K (1994). *Construction and characterization of chimeric enzyme consisting of an amino-terminal domain of phenylalanine dehydrogenase and a carboxy-terminal domain of leucine dehydrogenase*. J Biochem; 116(4):931-6.
- Kaur J y Sharma R (2006). *Directed evolution: an approach to engineer enzymes*. Crit Rev Biotechnol; (3):165-99.

- Kim D y Guengerich FP (2004). *Selection of human cytochrome P450 1A2 mutants with enhanced catalytic activity for heterocyclic amine N-hydroxylation*. *Biochemistry*; 43(4):981-8.
- Kim YW, Lee SS, Warren RA y Withers SG (2004). *Directed evolution of a glycosynthase from Agrobacterium sp. increases its catalytic activity dramatically and expands its substrate repertoire*. *J Biol Chem*; 279(41):42787-93.
- Kumar S, Chen CS, Waxman DJ y Halpert JR (2005). *Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide*. *J Biol Chem*; 280(20):19569-75.
- Labrou NE (2010). *Random mutagenesis methods for in vitro directed enzyme evolution*. *Curr Protein Pept Sci*; 11(1):91-100.
- Lebbink JH, Eggen RI, Geerling AC, Consalvi V, Chiaraluce R, Scandurra R y de Vos WM (1995). *Exchange of domains of glutamate dehydrogenase from the hyperthermophilic archaeon Pyrococcus furiosus and the mesophilic bacterium Clostridium difficile: effects on catalysis, thermoactivity and stability*. *Protein Eng*; 8(12):1287-94.
- Leung DW, Chen E y Goeddel DV (1989). *A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction*. *Technique*; 1:11-15.
- Lin H, Tao H y Cornish VW. (2004) *Directed evolution of a glycosynthase via chemical complementation*. *J Am Chem Soc*; 126(46):15051-9.
- Lindner HA, Nadeau G, Matte A, Michel G, Ménard R y Cygler M (2005). *Site-directed mutagenesis of the active site region in the quinate/shikimate 5-dehydrogenase YdiB of Escherichia coli*. *J Biol Chem*; 280(8):7162-9
- Mahan SD, Ireton GC, Knoeber C, Stoddard BL y Black ME (2004). *Random mutagenesis and selection of Escherichia coli cytosine deaminase for cancer gene therapy*. *Protein Eng Des Sel*; 17(8):625-33
- Martínez de la Escalera- Fanjul, X (2012). *Creación y análisis de quimeras producto de intercambio de dominios entre dos proteínas homólogas*. Tesis de licenciatura. UNAM
- Matsuura T y Yomo T (2006). *In vitro evolution of proteins*. *J Biosci Bioeng*; 101(6):449-456.
- McCarthy, A (2003). *Metabolix, Inc. and Tepha, Inc. Bioplastics for industry and medical devices*. *Chem. Biol*; 10:893-894
- Michel G, Roszak AW, Sauve V, Maclean J, Matte A, Coggins JR, Cygler M y Laphorn AJ (2003). *Structures of shikimate dehydrogenase AroE and its paralog YdiB*. *J Biol Chem*; 278:19463-19472
- Miyazaki K (2003). *Creating random mutagenesis libraries by megaprimer PCR of whole plasmid (MEGAWHOP)*. *Methods Mol Biol*; 231:23-8.
- Miyazaki K y Takenouchi M (2002). *Creating random mutagenesis libraries using megaprimer PCR of whole plasmid*. *Biotechniques*; 33(5):1033-4, 1036-8.
- Ness JE, Welch M, Giver L, Bueno M, Cherry JR, Borchert TV, Stemmer WP y Minshull J (1999). *DNA shuffling of subgenomic sequences of subtilisin*. *Nat Biotechnol*; 17(9):893-896.
- Padyana, AK y Burley SK (2003). *Crystal structure of shikimate 5-dehydrogenase (SDH) bound to NADP: Insights into function and evolution*. *Structure*; (11), 1005-1013.
- Peisajovich SG, Garbarino JE, Wei P y Lim WA (2010). *Rapid diversification of cell signaling phenotypes by modular domain recombination*. *Science*; 328(5976):368-72.
- Pittard J (1996). *Biosynthesis of the aromatic amino acid*. In Neidhardt FC (ed) *Escherichia coli and Salmonella: cellular and molecular biology*, vol 1. ASM. Washington D.C. pp 458-484.
- Reetz MT, Soni P, Fernandez L, Gumulya Y, y Carballeira, JD (2010). *Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method*. *Chem. Commun*; 46:8657-8658.

- Reetz MT, Brunner B, Schneider T, Schulz F, Clouthier CM y Kayser MM (2004). *Directed evolution as a method to create enantioselective cyclohexanone monooxygenases for catalysis in Baeyer-Villiger reactions*. *Angew Chem Int Ed Engl*; 43(31):4075-8.
- Roberts F, Roberts CW, Johnson JJ, Kyle DE, Krell T, Coggins JR, Coombs GH, Milhous WK, Tzipori S, Ferguson DJ, Chakrabarti D y McLeod R (1998). *Evidence for the shikimate pathway in apicomplexan parasites*. *Nature*; 393:801-805.
- Roberts CW, Roberts F, Lyons RE, Kirisits MJ, Mui EJ, Finnerty J, Johnson JJ, Ferguson DJ, Coggins JR, Krell T, Coombs GH, Milhous WK, Kyle DE, Tzipori S, Barnwell J, Dame JB, Carlton J y McLeod R (2002). *The shikimate pathway and its branches in apicomplexan parasites*. *J Infect Dis*; 185 Suppl 1:S25-36.
- Savile CK, Janey JM, Mundorff EC, Moore JC, Tam S, Jarvis WR, Colbeck JC, Krebber A, Fleitz FJ, Brands J, Devine PN, Huisman GW y Hughes GJ (2010). *Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture*. *Science*; 329:305-309.
- Segatori L, Paukstelis PJ, Gilbert HF y Georgiou G (2004). *Engineered DsbC chimeras catalyze both protein oxidation and disulfide-bond isomerization in Escherichia coli: Reconciling two competing pathways*. *Proc Natl Acad Sci U S A*; 101(27):10018-23.
- Segovia L y Soberon X (2008). *Evolución experimental de proteínas en: Lopez-Munguía, A. Una ventana al quehacer científico*. Instituto de Biotecnología de la UNAM 25 aniversario, cap 26. México, D.F. UNAM. págs. 291-302.
- Segovia L y Peimbert M (2010). *Ingeniería de proteínas y evolución dirigida* Mensaje Bioquímico, 34, 135-141.
- Shafikhani S, Siegel RA, Ferrari E y Schellenberger V (1997). *Generation of large libraries of random mutants in Bacillus subtilis by PCR-based plasmid multimerization*. *Biotechniques*; 23(2):304-10.
- Sharkey MA y Engel PC (2009). *Modular coenzyme specificity: a domain-swapped chimera of glutamate dehydrogenase*. *Proteins*; 77(2):268-78.
- Shanmugaratnam S, Eisenbeis S y Höcker B. (2012). *A highly stable protein chimera built from fragments of different folds*. *Protein Eng Des Sel* [Epub ahead of print].
- Shen B (2002). *PCR approaches to DNA mutagenesis and recombination. An overview*. *Methods Mol Biol*; 192:167-174.
- Singh S, Korolev S, Koroleva O, Zarembinski T, Collart F, Joachimiak A y Christendat D (2005). *Crystal structure of a novel shikimate dehydrogenase from Haemophilus influenzae*. *J. Biol. Chem*; 280(17):17101-8.
- Singh SA, y Christendat D (2006). *Structure of Arabidopsis dehydroquinase dehydratase-shikimate dehydrogenase and implications for metabolic channeling in the shikimate pathway*. *Biochemistry*; 27; 45 (25):7787-96. Erratum in: *Biochemistry*. 2006; 45(34):10406.
- Singh SA, Stavriniades J, Christendat D y Guttman DS (2008). *A phylogenomic analysis of the shikimate dehydrogenases reveals broadscale functional diversification and identifies one functionally distinct subclass*. *Mol Biol Evol*; 25(10):2221-32.
- Sprenger GA (2007). *Aromatic Amino Acids*. *Microbiology monographs*; (5):93-127.
- Stemmer WP (1994a). *Rapid evolution of a protein in vitro by DNA shuffling*. *Nature*; 370(6488):389-391.
- Stemmer WP (1994b). *DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution*. *Proc Natl Acad Sci USA*; 91(22):10747-10751.
- Suzuki M, Christians FC, Kim B, Skandalis A, Black ME y Loeb LA (1996). *Tolerance of different proteins for amino acid diversity*. *Mol. Divers*; 2(1-2):111-118.
- Thornton JM, Orengo CA, Todd AE y Pearl FM (1999). *Protein folds, functions and evolution*. *J Mol Biol*. 293(2):333-42.

- Thulasiram H, Erickson H y Poulter C (2007). *Chimeras of two isoprenoid synthases catalyze all four coupling reactions in isoprenoid biosynthesis*. Science; 316(5821):73-6.
- Tripura C y Podile A (2007). *Properties of a chimeric glucose dehydrogenase improved by site directed mutagenesis*. J Biotechnol; 131(2):197-204.
- Vartanian JP, Henry M y Wain-Hobson S (1996). *Hypermutagenic PCR involving all four transitions and a sizeable proportion of transversions*. Nucleic Acids Res; 24(14):2627-31.
- Wan L, Twitchett MB, Eltis LD, Mauk AG y Smith M (1998). *In vitro evolution of horse heart myoglobin to increase peroxidase activity*. Proc Natl Acad Sci U S A; 95(22):12825-31.
- Wheeldon IR, Campbell E, y Banta S (2009). *A chimeric fusion protein engineered with disparate functionalities-enzymatic activity and self-assembly*. J Mol Biol; 392(1)129-42.
- Wilson DS y Keefe AD (2000). *Random mutagenesis by PCR*. Curr. Protoc. Mol. Biol; 51:8.3.1-8.3.9.
- Wong TS, Zhurina D y Schwaneberg U (2006). *The diversity challenge in directed protein evolution*. Comb Chem High Throughput Screen; 9(4):271-88.
- Ye S, Von Delft F, Brooun A, Knuth MW, Swanson RV y McRee DE (2003). *The crystal structure of shikimate de- hydrogenase (AroE) reveals a unique NADPH binding mode*. J Bacteriol; 185:4144-4151.
- You L y Arnold FH (1996). *Directed evolution of subtilisin E in Bacillus subtilis to enhance total activity in aqueous dimethylformamide*. Protein Eng; 9(1):77-83. Erratum in: Protein Eng 9(8):719.

## ANEXO 1. PROTOCOLO PARA PREPARAR CÉLULAS ELECTRO-COMPETENTES

### MATERIALES:

- 2 Matraces Erlenmeyer de 2.8 L
- 10 Celdas para medir en espectrofotómetro
- 4 Botes de centrifuga de 250 mL estériles (para rotor Beckman JA14)
- 2 Tubos para rotor JA20 estériles
- Hielo seco

### SOLUCIONES:

- 1 L de medio SB (500 mL para cada matraz) Esterilizar por autoclave.
- 100 mL de Glicerol al 10%. Esterilizar por autoclave.
- 2 L de medio HEPES 1mM (a partir del stock de 100 mM). Esterilizar por autoclave.

### PROCEDIMIENTO:

Todos los materiales y soluciones deben estar previamente en hielo o a 4°C antes de usarse.

1. Tomar 2 mL de medio SB estéril para utilizar como blanco en la curva de crecimiento.
2. Día 1: Estriar la cepa que se desea hacer competente. Dejar incubar 15 h con el antibiótico correspondiente.
3. Día 2: Tomar una colonia de la cepa estriada anteriormente y hacer un pre inoculo de 15 mL de medio con el antibiótico correspondiente, dejar incubar 15 h.
4. Día 3: Tomar del pre inoculo 5 mL para inocular cada matraz con medio SB, es decir, se toman 10 mL del pre inoculo. De cada matraz inoculado tomar 1 mL y medir su densidad óptica inicial. Dejar incubar a 37° C 250 rpm y medir cada media hora hasta que la densidad óptica se encuentre entre 0.5 y 0.7. Una vez alcanzada la densidad óptica colocar los matraces en hielo durante 15 min.
5. Todos los procedimientos siguientes se realizan en cuarto frio a 4 °C. Una vez enfriados los matraces colocar 250 mL de cultivo en cada bote de centrifuga, en total 4 botes.
7. Centrifugar a 3300 rpm por 15 minutos a 4°C, máx. Aceleración y deceleración (9 y 9)
8. Cuidadosamente descargar el sobrenadante y re suspender cuidadosamente el pellet en 150 mL de HEPES 100mM.
9. Distribuir la suspensión celular en 2 botes.
10. Centrifugar a 4500 rpm por 15 minutos a 4°C, máx. Aceleración y deceleración (9 y 9).
11. Cuidadosamente descargar el sobrenadante y re suspender cuidadosamente el pellet en 100 mL de HEPES 100 mM.
12. Distribuir la suspensión celular en 1 bote.
13. Centrifugar a 4500 rpm por 15 minutos a 4°C, máx. Aceleración y deceleración (9 y 9).
14. Cuidadosamente descargar el sobrenadante y re suspender cuidadosamente el pellet en 20 mL de glicerol al 10%.
15. Centrifugar a 4700 rpm por 15 minutos a 4°C, máx. Aceleración y deceleración (9 y 9).
16. Cuidadosamente descargar el sobrenadante y re suspender cuidadosamente el pellet en 1.2mL de glicerol al 10%.
17. Con la suspensión se hacen alícuotas de 50 µL y posteriormente se congelan en hielo seco o nitrógeno líquido y se guardan a -80°C.
18. Electroporar con 1pg de plásmido pUC19 para determinar la eficiencia de electroporación. Plaquear diluciones 1:100000.

**ANEXO 2. PROTOCOLO PARA PREPARAR CÉLULAS COMPETENTES DMSO****SOLUCIONES:***TSS:*

- 5 g de PEG 8000.
- 1.5 mL 1M MgCl<sub>2</sub> (o 0.30 g MgCl \*6H<sub>2</sub>O).
- Aforar a 47.5mL con LB.
- Ajustar pH a 6.5-6.7, esterilizar por membrana (0.22 µm).
- Agregar 2.5 mL de DMSO. Guardar a -20°C.

**PROCEDIMIENTO**

Tomar 2 mL de medio LB estéril para utilizar como blanco en la curva de crecimiento.

2. Día 1: Estriar la cepa que se desea hacer competente. Dejar incubar 15 h con el antibiótico correspondiente.
3. Día 2: Tomar una colonia de la cepa estriada anteriormente y hacer un pre inculo de 5 mL de medio con el antibiótico correspondiente, dejar incubar 15h.
4. Día 3: tomar 500 µL del cultivo anterior en 50 mL de LB fresco en un matraz de 200 mL. Dilución del cultivo diluir 1/100.
5. Crecer el cultivo diluido a una densidad óptica de 0.2-0.5 nm.
6. Centrifugar por 10 minutos a 3000 rpm (1000g) a 4°C.
7. Decantar el sobrenadante. El pellet de células debe de ser lo suficientemente compacto para realizar este paso. Pipetear cualquier resto de medio.
8. Resuspender cuidadosamente en 5 mL buffer TSS, no dejar restos de agregados celulares.
9. De la suspensión anterior, se hacen alícuotas de 100 µL y posteriormente se congelan en hielo seco o nitrógeno líquido y se guardan a -80°C.

**PARA TRANSFORMAR**

- 1) Descongelar las células en hielo durante 5 minutos.
- 2) Agregar 100 ng de DNA y dejar reposar en hielo durante 10 minutos, moviendo cuidadosamente las células cada 2 minutos.
- 3) Agregar 1 mL de LB, y recuperar las células a 37°C, a 180 rpm.
- 4) Platear 100 µL de células recuperadas en cajas de LB con el antibiótico correspondiente.

## ANEXO 3. SOLUCIONES PARA PREPARAR MEDIO MÍNIMO M9

Soluciones	□ STOCK	□ FINAL	Para 250mL
Sales	5X	1x	50
Glucosa	20%	0.40%	5
MgSO <sub>4</sub>	1M	2mM	0.5
CaCl <sub>2</sub>	200mM	0.1mM	0.125
Tiamina	10mg/mL	10ug/mL	0.25
Km	25mg/mL	25ug/mL	0.25
Triptófano	10mg/mL	500ug/mL	12.5
Fenilalanina	10mg/mL	500ug/mL	12.5
Tirosina	10mg/mL	500ug/mL	12.5
AGAR (autoclave)			6g

## SOLUCIONES

**Buffer de Sales M9 (5X)**

-Na<sub>2</sub>HPO<sub>4</sub> 16.89 g  
 -KH<sub>2</sub> PO<sub>4</sub> 7.5g  
 -NaCl 1.25g  
 -NH<sub>4</sub>Cl 2.5 g

Aforar con H<sub>2</sub>O a 500 mL esterilizar en autoclave (Almacenar a temperatura ambiente).

**Glucosa 20%**

Glucosa 4 g. Aforar con H<sub>2</sub>O a 200 mL y esterilizar en autoclave solo una vez. (Almacenar a temperatura ambiente).

**MgSO<sub>4</sub> 1M (PM= 246.48)**

MgSO<sub>4</sub> 6.1525 g. Aforar con H<sub>2</sub>O a 25 mL y esterilizar en autoclave (Almacenar a temperatura ambiente).

**CaCl<sub>2</sub> 1M**

CaCl<sub>2</sub> 14.7 g. Aforar con H<sub>2</sub>O a 100 mL y esterilizar en autoclave. (Almacenar a temperatura ambiente). Tomar 10mL y llevarlo a 50 mL para hacer solución 200 mM.

**Tiamina 10 mg/mL**

Tiamina 0.1 g en 10 mL.  
 Esterilizar por membrana y guardar a 4°C.

**Triptófano 10 mg/mL**

Pesar 0.5 g, agregar 40 mL de agua, agregar 40 µL de NaOH 10N y agitar por vortex hasta disolver, aforar a 50 mL (pH=8). Esterilizar por membrana y guardar a 4°C.

**Fenilalanina 10 mg/mL**

Pesar 0.5 g, agregar 40 mL de agua, agitar por vortex hasta disolver, aforar a 50 mL (pH=5 Esterilizar por membrana y mantener a 4°C. (Se disuelve en agua completamente).

**Tirosina 10mg/mL**

Pesar 0.5 g, agregar 30 mL de agua, agitar por vortex. Agregar HCl 5 M (1-2mL) aforar a 50 mL. (pH =1). Esterilizar por membrana y almacenar a 4°C.

**ANEXO 4. MEDIOS Y SOLUCIONES****SB (Super Broth)**

900 mL de H<sub>2</sub>O de fuerza iónica extra-baja  
35 g bacto- triptona  
20 g extracto de levadura  
5 g NaCl  
Mezclar todo hasta disolver  
Llevar a pH 7.5 con NaOH 5N  
H<sub>2</sub>O hasta 1 L  
Esterilizar por autoclave

**HEPES 100 mM**

2.3 g de HEPES  
90 mL de H<sub>2</sub>O de fuerza iónica extra- baja  
Llevar a pH 7 con NaOH 5N  
H<sub>2</sub>O hasta 100 mL  
Esterilizar por autoclave

**LB**

10g NaCl  
Extracto de Levadura 5g  
Tryptona 10g  
Agregar un litro de agua miliQ  
Esterilizar por autoclave

**SOC (Super Optimal broth with Catabolic repressor):**

2 g Bacto Triptona  
0.5 g de Extracto de Levadura  
0.05 g de NaCl  
Disolver en 75 mL de H<sub>2</sub>O  
Agregar 1 mL de KCl 250 mM  
Ajustar pH a 7 con NaOH 4M  
Aforar con H<sub>2</sub>O a 100 mL  
Esterilizar por autoclave

## ANEXO 5. ECUACIONES

**Coeficiente binomial**

$$\binom{n}{x} = \frac{n!}{x! (n-x)!} * c$$

Donde:

n= Tamaño del gen en pb (854)

x= Mutaciones sencillas (x=1), dobles (x=2) o triples (x=3)

C= Número de combinaciones restantes para cada base donde C=3 para sencillas, C=9 para dobles, C=27 para triples

**Variantes sencillas:**

$$v = \frac{854!}{1! (854-1)!} = \frac{854 * 853!}{1! (853)!} = 854 * 3 = 2562 \text{ variantes sencillas}$$

**Variantes dobles:**

$$v = \frac{854!}{2! (854-2)!} = \frac{854 * 853 * 852!}{2! (852)!} = 364231 * 9 = 3.27^6 \text{ variantes dobles}$$

**Variantes triples:**

$$v = \frac{854!}{3! (854-3)!} = \frac{854 * 853 * 852 * 851!}{3! (851)!} = 103441604 * 27 = 2.79^9 \text{ variantes triples}$$

## ANEXO 6. ECUACIONES

Cobertura

$$Nf = \frac{\ln(1 - f)}{\ln(1 - \frac{1}{v})}$$

Donde:

Nf= Número total de muestras a analizar, es decir la cobertura.

f=Probabilidad con que se desea completar la búsqueda en el banco (0. 95).

v= Número de variantes sencillas, dobles o triples.

Resolviendo las ecuaciones para cada variante, se obtiene lo siguiente:

**Cobertura para variantes sencillas:**

$$Nf = \frac{\ln(1 - 0.95)}{\ln(1 - \frac{1}{2562})} = 7674$$

**Cobertura para variantes dobles:**

$$Nf = \frac{\ln(1 - 0.95)}{\ln(1 - \frac{1}{3278079})} = 9,820,246$$

**Cobertura para variantes triples:**

$$Nf = \frac{\ln(1 - 0.95)}{\ln(1 - \frac{1}{2792923308})} = 8,366,849,787$$