

**Universidad Nacional Autónoma de México**

**Facultad de Ciencias**

**Generación de quimeras por intercambio de dominios entre  
dos proteínas parálogas**

*Marel Chenge Espinosa*

Tutor: *Dr. Lorenzo Patrick Segovia Forcella*

2009



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.

Este trabajo se realizó bajo la dirección del Dr. Lorenzo Segovia Forcella y la ayuda de la Dra. Mariana Peimbert Torres, en el Instituto de Biotecnología de la Universidad Nacional Autónoma de México.

## Índice

<b>Introducción</b>	<b>5</b>
Evolución de proteínas	
Modelo de estudio	
<i>Plegamiento Rossmann</i>	
<i>Shikimate deshidrogenasas: AroE y YdiB</i>	
<b>Antecedentes</b>	<b>12</b>
Intercambio de dominios	
La CAT como reportero de plegamiento	
<b>Hipótesis</b>	<b>16</b>
<b>Objetivos</b>	<b>16</b>
Objetivos particulares	
<b>Material y métodos</b>	<b>17</b>
Plásmidos	
Cepas	
Extracción de plásmido	
<i>Polymerase Chain Reaction (PCR)</i>	
Electroforesis de ADN en geles de agarosa	
Digestión del vector y de los productos de PCR	
Ligación del vector con el inserto	
Transformación de la ligación	
Concentración mínima inhibitoria (CMI)	
Extracción de proteínas y SDS PAGE	
<i>Western blot</i> Anti-CAT de la fracción soluble	
Ensayo de complementación	

<b>Resultados</b>	<b>25</b>
Obtención de las quimeras	
Concentración mínima inhibitoria	
Expresión de proteínas vista por SDS-PAGE	
<i>Western blot</i> Anti-CAT de la fracción soluble	
Ensayo de complementación	
<b>Discusión</b>	<b>31</b>
<b>Conclusiones</b>	<b>36</b>
<b>Perspectivas</b>	<b>37</b>
<b>Apéndice I: Soluciones y medios de cultivo</b>	<b>38</b>
<b>Apéndice II: Oligonucleótidos</b>	<b>41</b>
<b>Referencias</b>	<b>42</b>

## **Introducción**

Los seres vivos están formados por cuatro elementos principales: carbono, hidrógeno, oxígeno y nitrógeno, a partir de los cuales se forman las biomoléculas, éstas muestran un alto grado de complejidad. Sin embargo la mayoría de éstas pueden clasificarse dentro de cuatro grupos: carbohidratos, lípidos, ácidos nucleicos o proteínas.

Las proteínas, a diferencia del resto de las biomoléculas, participan en la gran mayoría, sino es que en todos, de los procesos que ocurren dentro de la célula. Muchas proteínas, llamadas enzimas, catalizan reacciones bioquímicas vitales para el metabolismo. También tienen importancia estructural ya que son componentes del citoesqueleto de las células, como es el caso de la actina. En la señalización celular, las proteínas juegan un papel importante como transductores de las señales, tanto intracelulares como extracelulares.

Su papel principal en la célula queda reflejado en el hecho de que la información genética se expresa en último término en forma de proteínas, es decir, existe para cada proteína un segmento de ADN que codifica la información que especifica su secuencia de aminoácidos.

### **Evolución de proteínas**

Al ser las proteínas agentes importantes para el desarrollo de la vida, el estudio sobre como evolucionan es de gran interés. Con el fin de entender los principios y las fuerzas que llevan al surgimiento de nuevos plegamientos y nuevas funciones.

La estructura primaria se refiere a la estructura química de la proteína, es decir, la secuencia de aminoácidos. La estructura secundaria se refiere a la estructura espacial local de segmentos de las cadenas polipeptídicas (como una  $\alpha$ -hélice, un giro, una hebra- $\beta$ ). La estructura terciaria define en si a la topología espacial global de estas estructuras (Patthy, 1999).

Las proteínas que se encuentran en la naturaleza han evolucionado, a través de la presión de selección, para realizar funciones específicas. Las propiedades funcionales de las proteínas dependen de sus estructuras tridimensionales. Estas estructuras surgen debido a que secuencias particulares de aminoácidos se pliegan para generar dominios compactos con estructuras tridimensionales específicas. Los dominios plegados pueden servir como módulos para construir estructuras más grandes tales como partículas virales, o bien

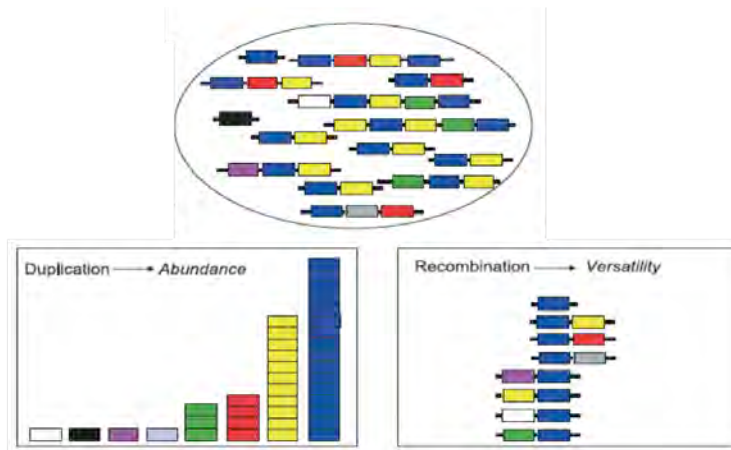
proveer de sitios catalíticos o de unión específicos como los que se encuentran en enzimas o proteínas que transportan oxígeno (Branden, Tooze, 1991).

La mayoría de las proteínas de gran tamaño están compuestas de varios dominios. Los dominios en estas proteínas muestran usualmente un alto grado de independencia estructural y tienen interacciones relativamente débiles entre ellos (Patthy, 1999).

Los dominios proteicos son descritos típicamente como las unidades primarias estructurales, funcionales y evolutivas de una proteína; como una unidad estructural, ya que se pliegan independientemente en formando su propio núcleo hidrofóbico; como una unidad evolutiva, ya que puede ser un segmento de secuencia de aminoácidos que se hereda y se conserva en distintos contextos; y por último, como una unidad funcional ya que es el mínimo de secuencia requerida para codificar una función en una proteína (Yeats, 2007).

Probablemente los primeros dominios proteicos surgieron como consecuencia de la recombinación y ensamblaje de un conjunto limitado de fragmentos estructurales capaces de plegarse independientemente. Estas estructuras se conocen como estructuras supersecundarias, las más comunes son los  $\beta\beta$  *hairpins*, los  $\alpha\alpha$  *hairpins* y los elementos  $\beta\alpha\beta$  (Söding, 2003).

En 2002 Vogel sugiere que las dos fuerzas evolutivas para la evolución de las proteínas multidominio, y en sí, del proteoma, son la duplicación y el barajeo por recombinación de dominios ya existentes (Figura 1). La duplicación conlleva a la expansión de un dominio, lo cual se refleja en la abundancia de éste y se asocia al incremento en la complejidad de un organismo. La recombinación de dominios brinda versatilidad al proteoma permitiendo así el surgimiento de nuevas funciones proteicas.



**Figura 1.** La duplicación y recombinación de dominios forman el repertorio proteico. La abundancia es el resultado de eventos de duplicación, mientras que la recombinación le confiere versatilidad al proteoma (Tomada de Vogel, 2004).

La combinación de dominios no sólo implica ventajas funcionales o evolutivas, al combinarse proteínas multidominio con otros dominios se pueden sumar las funciones individuales, o generar funciones nuevas por la cooperación entre los dominios, además, si partimos del hecho de la independencia estructural que tienen los dominios, las restricciones evolutivas sobre el plegamiento de la nueva proteína se reducen, pero surge la interacción entre dominios como una nueva restricción, teniendo que generar interfaces estables. Usualmente la función general del homólogo de la proteína de un dominio en la proteína multidominio ha sido conservada pero es modificada o más específica. En el caso de que la función no se conserve, se debe a que hay un rediseño de la función de la proteína por mutaciones progresivas en el dominio mismo para generar funciones diferentes (Bashton, 2007).

En el transcurso de la evolución, los dominios han pasado por varios eventos de recombinación, y las superfamilias de dominios aisladas que nunca recombinan son raras, en un análisis realizado por Han en el 2007, se observó que solo 86 de las 1439 superfamilias de dominios se pueden encontrar como proteínas de un solo dominio.

La combinación de dominios se rige por una distribución del tipo de la ley de potencias, es decir, unas cuantas superfamilias son muy abundantes y recombinan con muchos otros tipos de dominios, mientras que las superfamilias que son representadas por pocas proteínas solo se combinan con una cantidad pequeña de superfamilias de dominios (Han, 2007).



El orden secuencial de N-terminal a C-terminal de las combinaciones de dominios suele estar altamente conservado, es decir, si un dominio se combina con otro de manera A-B, siempre se va a encontrar en esa conformación, esto sucede tanto con proteínas de dos dominios como con proteínas multidominio (Vogel, 2004). La conservación de la posición de los dominios se puede deber a restricciones funcionales, es decir, para poder formar al sitio activo se debe de mantener cierto orden; aunque se puede ver desde una perspectiva evolutiva, la cual lo explicaría como que el orden que se presenta se debe a un solo evento de recombinación entre superfamilias (Bashton, 2002).

### **Modelo de estudio**

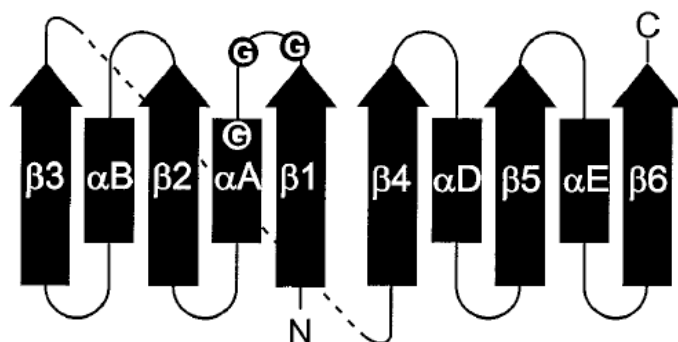
#### *Plegamiento Rossmann*

El plegamiento Rossmann es el más común entre los plegamientos  $\alpha/\beta$  y, en si, uno de los dominios más abundantes en la naturaleza, encontrándose en el 1% de las proteínas eucarióticas y en el 3% de las procarióticas. Los dominios con este plegamiento se encuentran en proteínas con una gran variedad de funciones, mostrando las seis clases de reacciones enzimáticas (Kallberg y Persson, 2006). Además se ha visto que estos dominios se combinan con dominios catalíticos de 7 superfamilias diferentes y se pueden definir cuatro tipos distintos de conexiones entre los dominios catalíticos y los dominio Rossmann, es decir, el dominio Rossmann puede estar unido al extremo N-terminal (tipo I) o al extremo C-terminal (tipo II) del dominio catalítico, también puede ocurrir que el dominio catalítico quede insertado dentro del Rossmann (tipo III) o viceversa (tipo IV).

Los dominios con plegamiento Rossmann tienen dos motivos estructurales  $\beta\alpha\beta\alpha\beta$  de unión a mononucleótidos unidos por sus extremos N-terminal, conformando un dominio de unión a dinucleótidos, que dan lugar a un plegamiento que consiste de seis hojas  $\beta$  rodeadas por  $\alpha$  hélices en un orden 321456 (Bottoms, 2002). Los dinucleótidos que las proteínas con plegamiento Rossmann unen más comúnmente son el dinucleótido flavin-adenina (FAD), el dinucleótido nicotinamida adenina (NAD) y el dinucleótido nicotinamida adenina fosfato (NADP). Estos dinucleótidos en presencia de la enzima y el sustrato transfieren un ion hidronio desde o hacia el átomo C4 del anillo nitrogenado (Carugo, 1997).

En la primera parte del plegamiento Rossmann ( $\beta_1\alpha_1\beta_2$ ) se encuentran tres glicinas rodeadas por aminoácidos hidrofóbicos, con la primera glicina al final de la hoja  $\beta_1$  y las otras dos al principio de la  $\alpha$  hélice (Figura 2). Las dos primeras glicinas están

involucradas en la unión al dinucleótido, mientras que la tercera está involucrada en el empaquetamiento cerrado de las hebras  $\beta$  y la  $\alpha$  hélice. Los residuos al final de la hebra  $\beta 2$  determinan si la enzima une FAD, NAD o NADP. Por lo general la presencia de un residuo cargado negativamente (Asp o Glu) determina el cofactor que se puede unir, NAD o FAD, en cambio, las enzimas que presentan un residuo básico (Arg generalmente) unen preferentemente NADP (Kallberg, 2006; Carugo, 1997).



**Figura 2.** Estructura secundaria y topología de un plegamiento Rossmann canónico (Tomada de Bottoms, 2002).

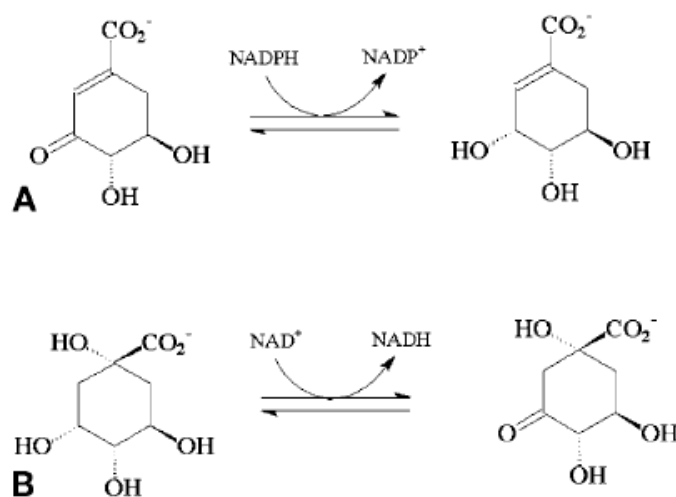
La teoría de que los dominios Rossmann de unión a dinucleótidos evolucionaron por duplicación de genes primordiales que codificaban para dominios más pequeños de unión a mononucleótidos ha sido ampliamente aceptada, sin embargo, ambas mitades del dominio de unión a dinucleótidos no están relacionadas evolutivamente, es decir, el dominio de unión a mononucleótido del N-terminal es menos variable que la mitad del C-terminal, probablemente porque el primero está involucrado con la unión de ADP del NADP, mientras que la segunda mitad está involucrada con la unión de la nicotina y se encuentra más cercana a la zona de unión del sustrato, la cual difiere de acuerdo a cada enzima (Carugo, 1997).

#### *Shikimato deshidrogenasas: AroE y YdiB*

La shikimato deshidrogenasa (SDH) es la cuarta enzima en la vía del shikimato para la síntesis de aminoácidos aromáticos. La SDH prototípica, AroE cataliza la reducción del dehidroshikimato a shikimato. Recientemente se han encontrados homólogos de la SDH AroE adicionales en algunas bacterias, estos homólogos se incluyen en cuatro grupos

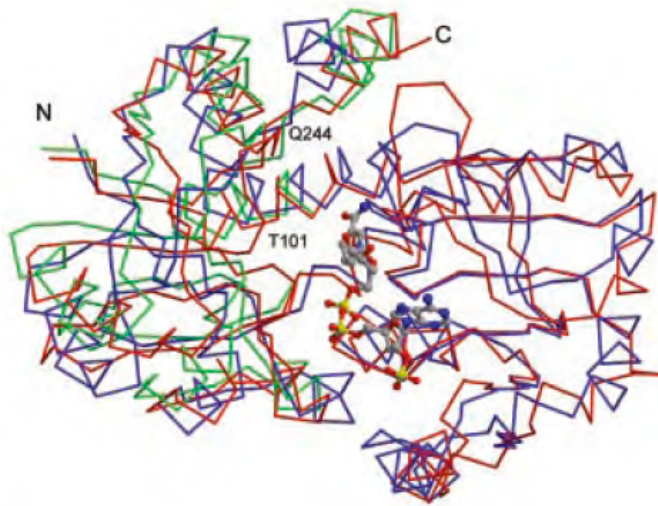
YdiB, RfiI, SdhL y AeiI, las cuales tienen actividades bioquímicas distintivas y funciones celulares particulares (Singh, 2008). De la familia de las SDH, sólo AroE y YdiB se encuentran en *Escherichia coli*. YdiB fue reportada previamente como paróloga de AroE, sin embargo, recientemente Singh (2008) realizó un análisis filogenético de las distintas SDH, llegando a la conclusión que la SDH YdiB codificada por bacterias entéricas Gram-negativas (como *Escherichia coli* o *Salmonella typhirium*) tiene su origen en las Firmicutes, las cuales son Gram-positivas, lo cual sugiere que YdiB fue adquirida por transferencia horizontal.

AroE oxida el ácido shikímico usando NADP como cofactor, mientras que YdiB puede hacerlo usando NADP o NAD, aunque en presencia de NAD parece ser más eficiente. Contrario a AroE, YdiB tiene también una actividad como quinato deshidrogenasa (Figura 3), sin embargo su actividad catalítica es baja para ambas reacciones, por lo que se cree que puede tener otro sustrato (Michel, 2003).



**Figura 3.** Reacciones catalizadas por AroE (A) y por YdiB (B). AroE oxida al ácido shikímico en presencia de NADP, mientras que YdiB también puede oxidar al quinato en presencia de NAD (Tomada de Michel, 2003).

A pesar de tener una identidad en secuencia relativamente baja (alrededor del 25%), ambas enzimas tienen una alta similitud estructural ya que adoptan el mismo plegamiento (Figura 4).

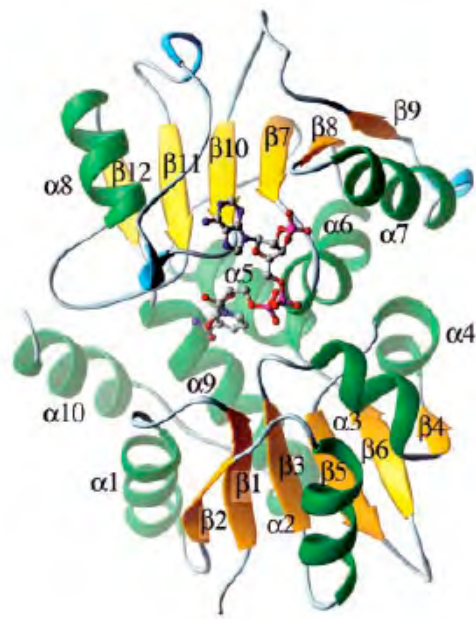


**Figura 4.** La estructura superpuesta de AroE y YdiB mostrando el esqueleto de  $\alpha$ . YdiB de azul, en una conformación abierta; de rojo, AroE en una conformación cerrada, y de verde AroE en una conformación intermedia (Tomada de Michel, 2003).

Ambas proteínas tienen dos dominios con arquitectura  $\alpha/\beta$  y están conectados por la hélice  $\alpha 5$  y un conector corto de 17 aminoácidos; el acomodo de estos dos dominios a lo largo de las hélices conectoras crea una hendidura profunda donde se localiza el cofactor (NADP en el caso de AroE y NAD en el de YdiB).

El dominio N-terminal consiste de seis hebras de hojas  $\beta$  y seis  $\alpha$  hélices; las hebras se acomodan en un orden 213564, siendo la  $\beta 5$  antiparalela al resto. Las tres primeras hebras  $\beta$  siguen una sucesión regular  $\beta/\alpha$ , con las hélices  $\alpha 1$  y  $\alpha 2$  paralelas a las  $\beta$ . La siguiente unidad  $\alpha/\beta/\alpha$  es irregular, con la hélice  $\alpha 3$  orientada a  $45^\circ$  en relación con la dirección de la hoja  $\beta$  y con la hélice  $\alpha 4$  de sólo un giro casi perpendicular a la hebra  $\beta 4$ .

El dominio C-terminal adopta la forma de un plegamiento Rossmann canónico, es decir, seis hebras paralelas de hojas  $\beta$  con un orden 321456 y hélices  $\alpha$  en ambos lados paralelas a las hebras  $\beta$ . La cuarta hélice  $\alpha$  presente en los dominios Rossmann está ausente en YdiB, mientras que en AroE la tercera y cuarta hélices son reemplazadas por asas irregulares, esto hace que los dominios Rossmann de ambas proteínas sean más pequeños de lo normal (Figura 5).



**Figura 5.** Representación de AroE de *E.coli*. El cofactor NADP se muestra con todos sus átomos y los enlaces entre estos (Tomada de Michel, 2003).

En cuanto a su estructura cuaternaria, AroE es una proteína monomérica en *E. coli* a diferencia de lo que sucede con otras deshidrogenasas. En el caso de YdiB, todo parece indicar que forma dímeros; los dos monómeros de esta proteína forman una interfase con los residuos de las hebras  $\beta 1$ ,  $\beta 2$  y la hélice  $\alpha 2$  de los dominios N-terminal, esto hace que se forme un dímero alargado. Los residuos que forman en YdiB esta interfase son hidrofóbicos y en su mayoría son reemplazados por residuos pequeños y polares en AroE, lo cual explica el por qué AroE se encuentra como monómero (Michel, 2003).

## Antecedentes

### **Intercambio de dominios**

Con el fin de entender la aportación de la presencia de ciertos dominios sobre la estabilidad y la función de ciertas proteínas, se han realizado varios trabajos donde se intercambian dominios entre distintas proteínas generando quimeras que en ciertos casos mantienen su estructura y su funcionalidad.

La fenilalanina deshidrogenasa de *Thermoactinomyces intermedius* actúa preferencialmente sobre la L-Fenilalanina y la L-Tirosina, en cambio la leucina deshidrogenasa de *Bacillus stearothermophilus* actúa exclusivamente sobre L-Leucina y sobre algunos L-aminoácidos ramificados. Las dos enzimas comparten 47% de identidad en secuencia. Con el objetivo de dilucidar el mecanismo para el reconocimiento del sustrato, Kataoka (1994) construyó una quimera que contenía el dominio N-terminal con la región de reconocimiento del sustrato de la fenilalanina deshidrogenasa, y el dominio C-terminal de unión al cofactor NAD de la leucina deshidrogenasa. Esta enzima quimérica actuaba sobre la fenilalanina y la leucina, mostrando valores de  $K_m$  similares a los de la fenilalanina deshidrogenasa, pero también mostró capacidad para actuar de manera más eficiente sobre la isoleucina y la valina que la proteína parental. Estos resultados mostraron que los dos dominios son capaces de plegarse independientemente, formando un sitio activo nuevo que explicaría la especificidad por los nuevos sustratos.

Buscando el papel que juegan los dominios en la actividad catalítica y la estabilidad de las glutamato deshidrogenasas (GDH), Lebbink y colaboradores (1995) intercambiaron dominios entre las GDH de la archaea hipertermófila *Pyrococcus furiosus* y la bacteria *Clostridium difficile*. Las GDH son proteínas de dos dominios, uno de unión al sustrato y otro de unión al cofactor, y forman hexámeros en solución. Al hacer el intercambio de los dominios, la quimera con el dominio de unión al glutamato de *C. difficile* y el dominio de unión al cofactor de *P. furiosus* no mostró ninguna actividad, sin embargo, la quimera contraria fue capaz de formar hexámeros y de tener actividad catalítica aunque mostró una afinidad reducida por el sustrato, pero una eficiente especificidad por el cofactor. Comparado con la GDH de *C. difficile*, la quimera archaea-bacteria fue ligeramente más termoactiva, menos termoestable, pero más estable ante la inactivación y desnaturalización con cloruro de guanidinio, mostrando que es posible intercambiar partes de proteínas homólogas de microorganismos que son filogenéticamente distantes.

En el 2003, Aliverti, Pandini y Zanetti, crearon proteínas quiméricas de dos isoformas de ferredoxin-NADP reductasas (FNR) de plantas, siendo una de hoja y otra de raíz. Ambas proteínas tienen un 48% de identidad en su secuencia de aminoácidos y tenían características estructurales y funcionales específicas. Las FNR tienen dos dominios, un dominio N-terminal que une FAD y uno C-terminal similar a un plegamiento Rossmann

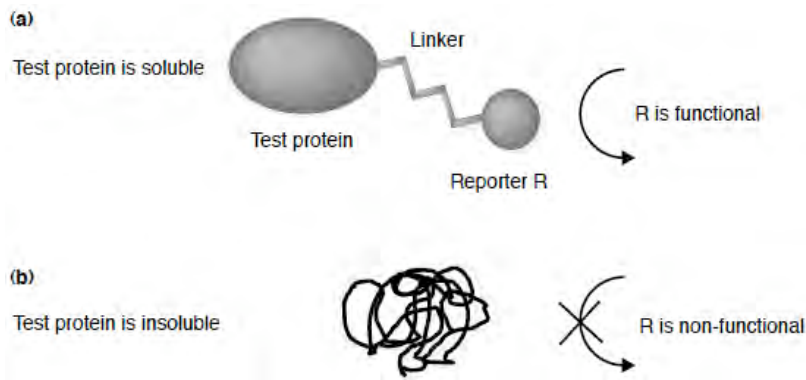
de unión a NADP. La quimera con el N-terminal de raíz y el C-terminal de la hoja fue inactiva y difícil de expresar, en cambio la quimera opuesta fue estable (aunque en menor grado que las proteínas parentales) y presentó actividad catalítica, inclusive tuvo una eficiencia catalítica mayor a la de las enzimas originales. Ellos explican que la disminución en la estabilidad se pudo deber a la pérdida de contactos entre aminoácidos de ambos dominios, en cambio en la quimera de hoja-raíz, el dominio de unión a FAD ayuda al dominio de unión a NADP a colocar adecuadamente al cofactor en el sitio activo, razón por la cual, la enzima quimérica es más eficiente.

En un trabajo más reciente, Li, Fast y Benkovic (2008) realizaron el intercambio de dominios con el objetivo de probar que nuevas funciones pueden surgir del ensamblaje de diferentes combinaciones de dominios funcionales. Para ello utilizaron dos enzimas distintas de la vía de síntesis de purinas, PurT y PurK. Estas proteínas tienen muy baja identidad en secuencia (27%) pero comparten similitud en su estructura terciaria, con 3 dominios, en la organización del sitio activo y en su mecanismo de reacción. Como resultado del intercambio de los dominios se crearon seis híbridos que fueron incapaces de catalizar las reacciones completas de las enzimas silvestres, pero cada una de las quimeras pudo catalizar reacciones parciales. Después de ser sometidas a una modificación en uno de las asas de un dominio, las quimeras fueron capaces de restaurar la actividad de PurK casi en su totalidad. Los autores sugieren que estos resultados podrían ilustrar aspectos del intercambio de dominios que ocurren durante la evolución de enzimas. A pesar de generar híbridos menos estables, probablemente porque la interfase entre dominios no está del todo optimizada, cada dominio es capaz de catalizar reacciones parciales, lo que sugiere que la unión funcional y los sitios activos fueron mantenidos parcialmente.

### **La CAT como reportero de plegamiento**

Una de las estrategias más utilizadas es la fusión de las proteínas de interés a reporteros de plegamiento, la cual tiene el objetivo de seleccionar a las proteínas que se encuentran en forma soluble dentro de un banco de proteínas, facilitando así el análisis funcional y estructural de las éstas. En esta estrategia el reportero de plegamiento posee una función o actividad biológica que pueda detectarse facilitando así la selección (Waldo, 2003).

En nuestro laboratorio trabajamos con la cloranfenicol acetil transferasa (CAT) como proteína reportera de plegamiento.



**Figura 6.** Principio de la metodología de fusión a un reportero de plegamiento. A) cuando la proteína se pliega, el reportero es funcional; B) cuando la proteína no se pliega, el reportero tampoco lo hace y pierde su función (Tomada de Waldo, 2003)

Maxwell (1999) demostró que las fusiones de CAT con proteínas insolubles conferían una menor resistencia al cloranfenicol que aquellas fusiones realizadas con proteínas solubles, por lo que era posible seleccionar a las proteínas solubles por su nivel de resistencia al antibiótico. El crecimiento de bacterias en medio que contiene niveles progresivamente más elevados de cloranfenicol confiere entonces una presión de selección para las proteínas. Usando esta técnica pudieron encontrar variables solubles de la integrasa del virus VIH a partir de un banco de integrasas insolubles.

Sieber (2001) utilizó este método de fusión a CAT para seleccionar proteínas solubles de un banco de variantes de citocromos P450, esto con el fin de poder separarlas de la membrana y hacerlas así más solubles. De esta manera sólo realizó los ensayos de actividad enzimática a las proteínas que fueron previamente seleccionadas por su solubilidad. Esta selección también les ayudó a eliminar todas las variantes que no eran completamente traducidas, es decir, que tenían el marco de lectura desfasado.



## **Hipótesis**

El intercambio de dominios entre las dos shikimato deshidrogenas, AroE y YdiB, generará quimeras capaces de plegarse.

## **Objetivos**

Obtener las proteínas quiméricas resultantes del intercambio de dominio entre AroE y YdiB para analizar los efectos estructurales y funcionales de las nuevas combinaciones.

### *Objetivos particulares*

- Generar las dos quimeras provenientes del intercambio de dominios entre las dos SDH de *E. coli* (AroE y YdiB).
- Análisis estructural: Observar la capacidad de plegamiento de cada quimera por medio de ensayos de CMI (concentración mínima inhibitoria) y geles de SDS-page.
- Análisis funcional: Observar si la función biológica de las quimeras se mantiene por complementación de la cepa de *E.coli aroE*

## Material y métodos

### Plásmidos

- pT4K12M4 : Resistente a Kanamicina, y con promotor trc. Este vector contiene la secuencia de nucleótidos correspondiente al extremo N-terminal de AroE (el dominio Rossman) fusionado con el gen *cat* (cloranfenicol acetil transferasa). Además tiene un *cassette* con la característica principal de poseer secuencias que al ser transcritas codifican para codones de paro en los seis marcos de lectura, con el fin de evitar falsos positivos, ya que si el vector no se encuentra completamente digerido, estos codones de paro no permitirán la transcripción de la construcción. (Figura 7)

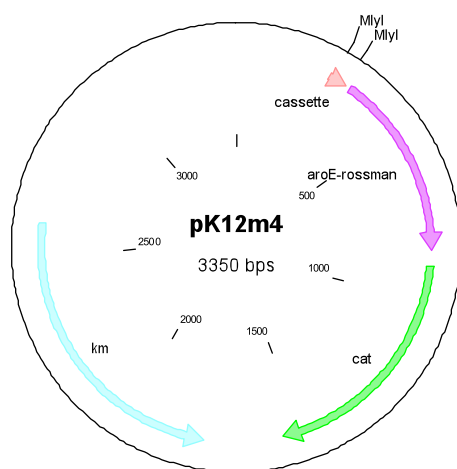


Figura 7. Vector pT4K12M4

- pTrc99a : Resistente a Ampicilina, promotor trc. Este vector es inducible con IPTG.

### Cepas de E. coli

- MC1061:  $F^-$  *araD139*  $\Delta$ (*ara-leu*)7696 *galE15 galK16*  $\Delta$ (*lac*)X74 *rpsL* (*Str<sup>r</sup>*) *hsdR2* ( $r_k^- m_k^+$ ) *mcrA mcrB1*
- JM101:  $F'$ *traD36 proAB lacI*  $\Delta$ (*lacZ*)M15/ $\Delta$ (*lac-proAB*) *glnV thi*

- K12 BW25113  $\Delta$ aroE: *lacI<sup>q</sup> rrnB<sub>T14</sub>  $\Delta$ lacZ<sub>WJ16</sub> hsdR514  $\Delta$ araBAD<sub>AH33</sub>  $\Delta$ rhaBAD<sub>LD78</sub> (*Km<sup>r</sup>*) (Keio collection)*

### Extracción de plásmido

El fundamento de esta metodología se basa en la lisis alcalina, la cual libera el plásmido de la bacteria y con ARNasa se elimina el ARN del lisado bacteriano. Para esta técnica se utilizó el *High Pure Plasmid Isolation Kit* de la marca *Roche*. El procedimiento que se realizó cada vez que se purificó un plásmido es como se indica en el protocolo del *kit*.

### PCR

El objetivo de la PCR (reacción en cadena de la polimerasa) es obtener un gran número de copias de un fragmento de ADN particular, partiendo de una cantidad mínima de ese fragmento original, o molde.

El protocolo para esta técnica fue el mismo para todos los productos que se obtuvieron en este trabajo, cambiando únicamente los oligonucleótidos o *primers* y el plásmido molde. Se utilizó la Taq polimerasa de la marca *Roche* con su respectivo *buffer*.

5  $\mu$ l *Buffer* para PCR 10X sin Mg<sup>+</sup>  
5  $\mu$ l dNTPs  
2  $\mu$ l oligo 1  
2  $\mu$ l oligo 2  
1  $\mu$ l plásmido (extraído por lisis alcalina con el *kit* de purificación)  
1  $\mu$ l Taq polimerasa  
Llevar a un volumen final de 50ul

El termociclador se programa de la siguiente manera:

1 ciclo a 92°C por 2 minutos  
30 ciclos de: 92°C por 40 segundos  
                  55°C por 40 segundos  
                  72°C por 1:20 minutos  
1 ciclo a 72°C por 5 minutos

Los productos de la PCR se purifican con el *High Pure PCR Product Purification Kit* de la marca *Roche*.

## **Electroforesis de ADN en geles de agarosa**

El fundamento de esta metodología se basa en la carga negativa del ADN, las muestras de ADN migran en un gel de agarosa, la cual es un polisacárido que permite separar moléculas según su tamaño. Al situar las moléculas de ADN en el gel y aplicar una diferencia de potencial eléctrico, éstas se mueven a diferentes velocidades según su tamaño, hacia el cátodo.

Se emplearon geles de agarosa grado Biología molecular (*Biorad*) al 1% y 1.5%, en solución amortiguadora TAE, los geles se corrieron en la misma solución. Para preparar la muestra se utilizó el *buffer* de carga 6x con glicerol al 30%.

## **Digestión del vector y de los productos de PCR**

Para eventualmente poder insertar el fragmento de ADN en el vector, es necesario primero que sean digeridos con enzimas de restricción, de preferencia con enzimas que dejen extremos cohesivos para que la inserción sea más exitosa.

Para la digestión:

- 35 µl del producto de la PCR que se desea cortar o del vector
- 1 µl de la enzima de restricción
- 5 µl del *buffer* 10X necesario para la enzima (el del proveedor)
- 5 µl de BSA (en el caso de que la enzima lo requiera)
- Ajustar con agua miliQ a un volumen final de 50 µl

Dejar la digestión por 16 horas (ON) a la temperatura que la enzima requiera. Detener la reacción calentado al tubo por 10 minutos 95°C o con columna de purificación. Todas las enzimas utilizadas fueron de la marca *Biolabs*.

## **Ligación del vector con el inserto.**

La reacción de ligación se realizó con una relación molar de vector: inserto de 1:4 respectivamente, para esto es necesario cuantificar la concentración tanto del vector como del inserto, con el fin de poner al final por lo menos 200 ng de vector. Se hizo un

control negativo de esta reacción con el fin de observar la eficiencia de la digestión del vector, el cual no debe recircularizarse en ausencia del inserto.

- X  $\mu$ l de vector digerido (200 ng)
- X  $\mu$ l de inserto digerido (para cumplir la relación 1:4)
- 2  $\mu$ l de *buffer* 10 x para ligasa
- 1  $\mu$ l de T4 ligasa *Biolabs*
- ajustar con agua miliQ a un volumen final de 20  $\mu$ l

Incubar la reacción por lo menos 12 horas a 16°C.

### **Transformación de la ligación**

Con el fin de transformar a las bacterias con el vector ligado, en condiciones estériles se transforma 1  $\mu$ l de la ligación en 50  $\mu$ l de células electrocompetentes MC1061 previamente descongeladas. Estas células con el plásmido se colocan en una celda de electroporación fría de 0.1cm y se aplica un pulso de 1.8 V en un electroporador ajustado a 25 uF y 200  $\Omega$ . Las células se recuperan en 1 ml de SOC y se incuban por una hora a 37°C y a 250 rpm. Esto mismo se hace con el control negativo. Se plaquean 300  $\mu$ l de las bacterias transformadas en cajas de LB sólido con el antibiótico para seleccionarlas y se incuban a 37°C.

### **Concentración mínima inhibitoria (CMI)**

La concentración mínima inhibitoria a un antibiótico, es aquella en la cual la mitad de las clonas ya no son resistentes, en este caso a cloranfenicol, el cual es el reportero de plegamiento.

- Se transforman las distintas construcciones fusionadas a la CAT. Se utiliza un vector de alguna proteína no fusionada a CAT como control negativo y a un vector con una proteína fusionada a CAT y que se sabe que si se pliega bien como control positivo.
- Se realizan diluciones de cultivos de cada cepa incubados en LB con 25  $\mu$ g/ml de kanamicina a 37°C por 16 horas (ON) a 250 rpm.
- Se hacen diluciones en agua estéril o en LB de  $10^{-1}$  a  $10^{-7}$  de los cultivos en una placa de 96 pozos.

- En cajas de LB sólido con concentraciones crecientes de cloranfenicol (20 µg/ml hasta 140 µg/ml), así como en una caja de LB sólido con kanamicina a 25 µg/ml se siembran gotas de 7 µl de cada una de las diluciones de cada cepa.
- Se incuban a 37°C

### **Extracción de proteínas y SDS PAGE**

- Inocular con una colonia de las células JM101 transformadas con el plásmido de expresión (pT4) en 3 ml de LB con kanamicina 25 µg/ml. Como control negativo se utilizar la cepa sin transformar. Incubar por 16 horas (ON) a 37°C a 250 rpm.
- Inocular con 1 ml de los cultivos anteriores 10 ml de LB con kanamicina. Incubar a 37°C a 250 rpm hasta alcanzar una densidad óptica (OD) a 600nm máxima de 1.
- Centrifugar 1.5 ml de cultivo en un tubo *ependorf* a 4000 rpm a 4°C por 5 minutos. Eliminar el sobrenadante y resuspender la pastilla en 1 ml de Tris 0.1M pH8.
- Sonicar por 30 segundo aproximadamente 3 veces, cuidando que la muestra no se caliente y que no se forme espuma.
- Centrifugar a 4000 rpm a 4°C por 20 minutos. Transferir el sobrenadante a un tubo nuevo, éste corresponde a la fase soluble de la muestra. La pastilla corresponde a la fase insoluble, y se debe de resuspender en 500 µl de Tris 0.1 M pH8 con SDS 1%.
- Tomar una fracción de cada muestra (50µl) y añadir B-mercaptoetanol y el colorante, las muestras se hierven durante 5 minutos.
- Correr las muestras en un gel desnaturalizante de acrilamida al 12.5% a 8 mA con Tris 0.1 M pH8 hasta que el colorante llegue hasta abajo.

Para hacer el gel de acrilamida al 12.5%:

- 2ml de LGS
- 3.35ml de acrilamida 30%
- 2.65ml de agua miliQ
- 40 µl de PSA 10%
- 20 µl de TEMED

Para hacer el gel concentrador:

- 1.25ml de UGS

- 0.75ml de acrilamida 30%
- 3ml de agua miliQ
- 15 µl de PSA 10%
- 5 µl de TEMED

Para teñir el gel con azul de Coomasie:

- Teñir por 1 hora con la solución I a temperatura ambiente en agitación
- Colocar por 30 minutos en la solución de desteñir I en agitación a temperatura ambiente.
- Usar la solución para desteñir II y dejar 8 horas a temperatura ambiente y en agitación. Posteriormente poner el gel en agua o secarlo

#### **Western blot Anti-CAT de la fracción soluble**

- Correr un gel de acrilamida al 12.5% de la fracción soluble de las muestras.
- Cortar de 3 papeles filtro gruesos y una membrana de nitrocelulosa del tamaño del gel.
- Remojar uno de los filtros en la solución cátodo, otro en la solución ánodo I y otro en la solución ánodo II.
- Colocar en orden el filtro con la solución, ánodo I, el filtro con la solución ánodo II, la membrana de nitrocelulosa, el gel y el filtro con la solución cátodo, cuidando no dejar burbujas de aire.
- Aplicar por una hora y media una corriente de 1.8 mA por cada cm<sup>2</sup> del gel.
- Bloquear la membrana de nitrocelulosa ON en una solución con 5% de leche en polvo y TBST
- Lavar la membrana tres veces durante diez minutos cada vez con TBST.
- Agregar el anticuerpo anti-CAT en una relación 1:5000 en solución de TBST con 0.1% de leche en polvo e incubar por una hora.
- Lavar la membrana tres veces con TBST.
- Posteriormente agregar el segundo anticuerpo Anti-DIG en una relación 1:5000, e igualmente incubar por 1 hora en TBST con 0.1% de leche.
- Lavar la membrana dos veces más con TBST.
- Agregar una solución con 1ml de BCIP y 1ml de NTB aforando a 10ml con agua

mQ, dejar en la solución hasta que se visualizaron las bandas y enjuagar con agua.

### **Ensayo de complementación**

Con el fin de observar la actividad *in vivo* de las proteínas quiméricas, se realizó un ensayo de complementación con la cepa K12 BW25113  $\Delta$ aroE de la colección Keio, la cual fue donada por el laboratorio del Dr. Guillermo Gosset del IBT.

La colección Keio está formada por mutantes, en donde cada una de ellas tiene un gen no esencial suprimido, incluyendo genes putativos. Este trabajo sistemático se realizó sustituyendo cada gen de interés por un *cassette* de kanamicina. Esto representó un problema, debido a que todas las construcciones con las que se estuvo trabajando estaban en el vector pT4K12M4, el cual es resistente a kanamicina, al igual que la cepa con la se harían los ensayos, razón por la cual fue necesario cambiar las construcciones de vector.

El vector que se eligió fue el pTrc99a, el cual también fue donado por el laboratorio del Dr. Gosset.

Este vector tiene un promotor *trc*, al igual que el vector PT4K12M4, y tiene resistencia a ampicilina.

- Los plásmidos se transforman en la cepa  $\Delta$ aroE, usando como control negativo al plásmido pTrc99a solo y como control positivo al pTrc99-aroE-cat.
- Se crecen en cajas de LB sólido con 100  $\mu$ g/ml de ampicilina para posteriormente picar una colonia de cada cepa y crecerla en 3 ml de medio LB líquido con ampicilina hasta alcanzar un OD de 0.6 (alrededor de cuatro horas).
- Centrifugar 1.5 ml de cada cultivo a 4000 rpm por 4 minutos, desechar el sobrenadante y el *pellet* se resuspende en 1ml de PBS, para nuevamente ser centrifugado. Este lavado se repite 2 veces más.
- Finalmente el *pellet* se resuspende en 500  $\mu$ l de PBS y se hacen diluciones de  $10^{-1}$  hasta  $10^{-7}$  en placas de 96 pozos.



- Sembrar gotas de 7  $\mu$ l de cada dilución en cajas de medio mínimo sólido, medio mínimo con vitaminas y aminoácidos aromáticos y LB. Las cajas se incuban a 37°C y a 30°C.

Este mismo protocolo se repite con el vector pT4.

## **Resultados**

### *Obtención de las quimeras*

La primera quimera que se construyó fue la que tendría el extremo N-terminal de YdiB y el extremo C-terminal de AroE.

Se diseñaron los oligonucleótidos para amplificar la secuencia correspondiente al dominio de unión al sustrato de YdiB, con los sitios de restricción para las enzimas NcoI y BstEII. Se utilizó el vector pT4K12M4, el cual ya contaba con la secuencia codificante para el dominio Rossmann de AroE. Tanto el vector como el producto de la PCR fueron digeridos por 16 horas a 37°C con la enzima NcoI en el *buffer* indicado por el proveedor (buffer 3) y BSA. Posteriormente se le añadió la enzima BstEII y se subió la temperatura a 60°C, a la cual la segunda enzima corta. El vector digerido se corrió en un gel de agarosa al 1%, con el fin de cortar la banda correspondiente a 3.3 Kb, después se extrajo el ADN del gel utilizando el *High Pure PCR Product Purification Kit* de Roche.

Para la ligación del vector con el inserto se utilizó la enzima T4-DNA ligasa. La reacción de ligación se llevó a cabo en una relación molar de vector:inserto de 1:4.

La segunda quimera a realizarse fue la que tendría el extremo C-terminal de AroE y el extremo N-terminal de YdiB.

Se diseñaron oligonucleótidos para amplificar el dominio de unión al sustrato de AroE, el oligonucleótido reverso tenía un fragmento del dominio Rossmann de YdiB. A la vez, uno de los oligonucleótidos (*forward*) para amplificar al dominio deseado de YdiB tenía un fragmento de la secuencia del dominio a amplificar de AroE.

Primero se realizaron las PCRs para obtener los dominios deseados de cada uno de los genes. Luego se realizó una tercera PCR con los productos de las PCRs anteriores con el fin de empalmarlos, ya que cada producto tenía una secuencia idéntica a la secuencia del otro producto. Por último se amplificó a la quimera con el oligonucleótido *forward* de *aroE* y el reverso de *ydiB*.

El vector pT4K12M4 fue digerido con las enzimas NcoI y Bam HI, con el fin de remover del vector la secuencia del dominio Rossmann de AroE y mantener la fusión a CAT de la nueva quimera.

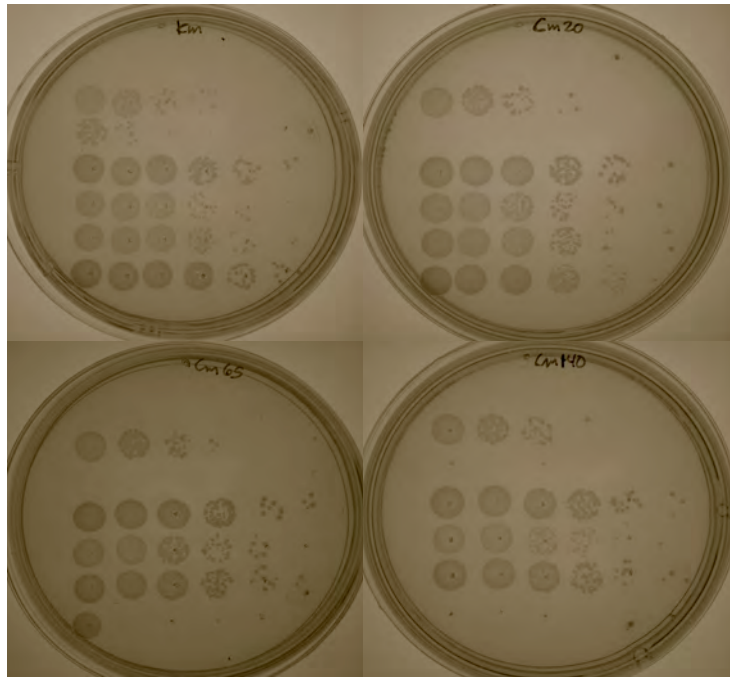
### *Concentración mínima inhibitoria*

Las quimeras fueron transformadas en la cepa MC1061, como control positivo se utilizó al vector pT4aroE-cat, como controles negativos se utilizó el pT4aroE sin CAT y el vector pT4K12M4, a la vez también se transformó con el vector pT4ydiB-CAT.

Se determinó la concentración mínima inhibitoria para las quimeras fusionadas a CAT y para AroE y YdiB. Las cepas que contenían el plásmido con *aroE* y con *aroE-cat* muestran un mayor grado de mortandad, ya que a pesar de que la densidad óptica de estos cultivos fue la misma que la de los demás (1 OD<sub>600nm</sub>), la viabilidad de las bacterias era menor. Sin embargo se puede apreciar que la fusión de AroE con CAT muestra un alto nivel de resistencia a cloranfenicol (140 ug/ml) y que la cepa con AroE no resiste ni la concentración más baja de cloranfenicol que se utilizó (20 ug/ml) (Fig.2).

En cuanto a YdiB, Quim1 (la quimera con el extremo N-terminal de YdiB y el extremo C-terminal de AroE) y Quim2 (la quimera con el extremo N-terminal de AroE y el extremo C-terminal de YdiB), las tres resistieron altas concentraciones de cloranfenicol al igual que AroE (140 ug/ml) sin notar una diferencia significativa entre ellas (Figura 8).

Por otro lado, el vector pT4K12M4, que contiene al dominio Rossmann de AroE fusionado a CAT y un *cassette* con codones de paro en los seis marcos de lectura, también mostró resistencia a cloranfenicol (alrededor de 50 µg/ml), esto indica que el dominio Rossmann es independiente para plegarse y no requiere del extremo C-terminal para su plegamiento.



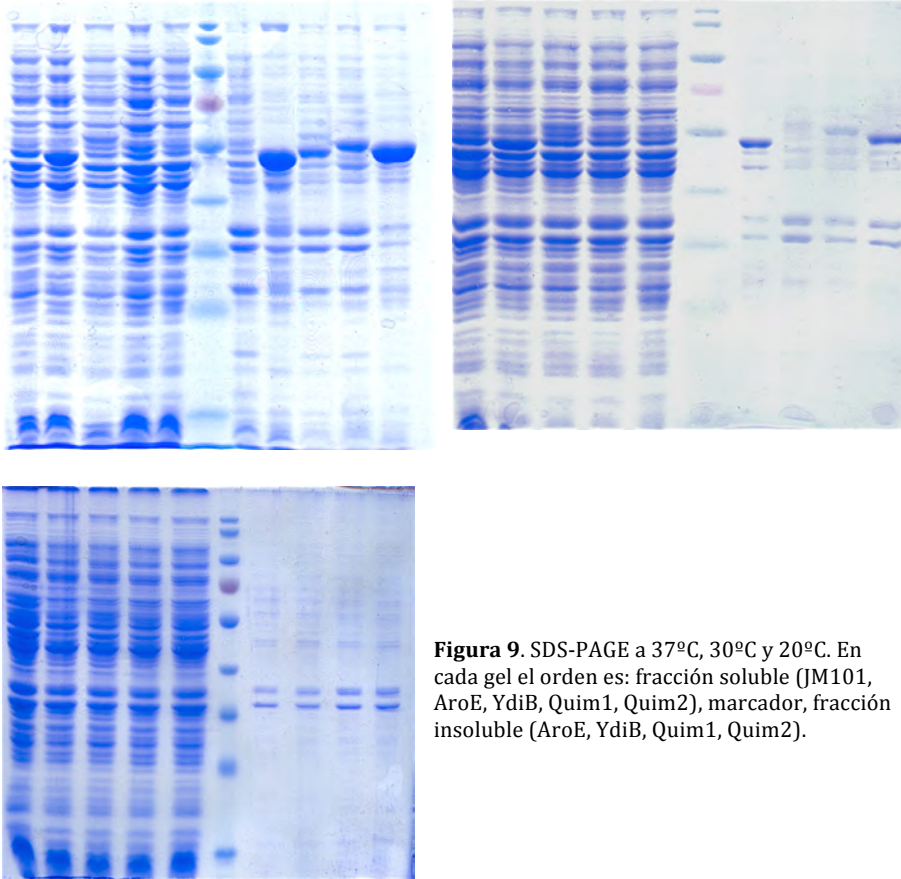
**Figura 8.** Concentración mínima inhibitoria. El orden en cada caja de arriba hacia abajo es: AroE-Cat, AroE, YdiB-Cat, Quim1-Cat, Quim2-Cat y pT4K12M4.

#### *Expresión de proteínas vista por SDS-PAGE*

Se transformaron los plásmidos de las quimeras Quim1, Quim2, de AroE y de YdiB en la cepa JM101.

Se crecieron las transformantes en tres temperaturas distintas (37°C, 30°C y 20°C) para observar como se afectaba la expresión de las proteínas en la fracción soluble.

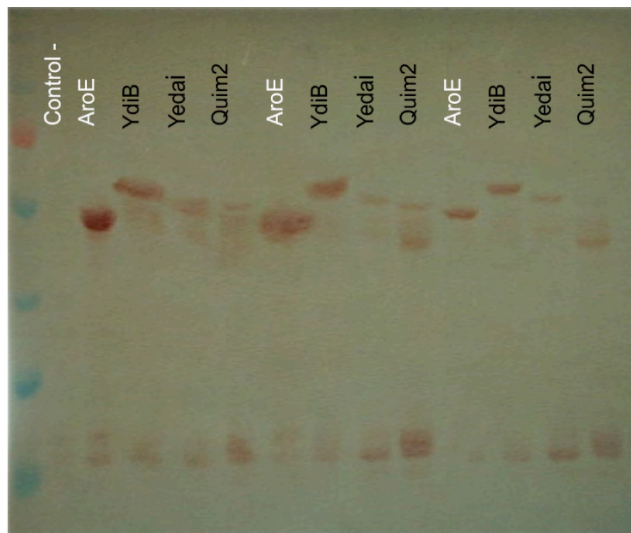
En los dos casos donde si pudieron observar las proteínas (37°C y 30°C) fue notorio que, tanto YdiB como Quim1, se expresaban en menor cantidad que AroE y la Quim2. (Figura 9)



**Figura 9.** SDS-PAGE a 37°C, 30°C y 20°C. En cada gel el orden es: fracción soluble (JM101, AroE, YdiB, Quim1, Quim2), marcador, fracción insoluble (AroE, YdiB, Quim1, Quim2).

#### *Western-blot Anti-Cat de la fracción soluble*

Debido a que no se veía más que a AroE en la fracción soluble a 37°C, se decidió hacer un *western blot* contra CAT, ya que todas las proteínas están fusionadas a esa proteína. Para el *western blot* se corrieron muestras de la fracción soluble de los cultivos crecidos a 37°C, 30°C y 20°C. Como control negativo se usó a AroE sin fusión a CAT, y como control positivo a AroE fusionada con CAT. En todas las muestras se observaron a las proteínas salvo en el control negativo, sin embargo también en todas, se puede apreciar a la CAT sola. Se observó que a 37°C hay mayor cantidad de proteínas que a 30°C, y que a su vez, se aprecia mayor cantidad de proteínas a 30°C que a 20°C. En el *western blot* se aprecia, a diferencia de los geles de acrilamida, que AroE y YdiB están en mayor proporción que ambas quimeras, probablemente porque éstas se pliegan mejor que las quimeras (Figura 10).



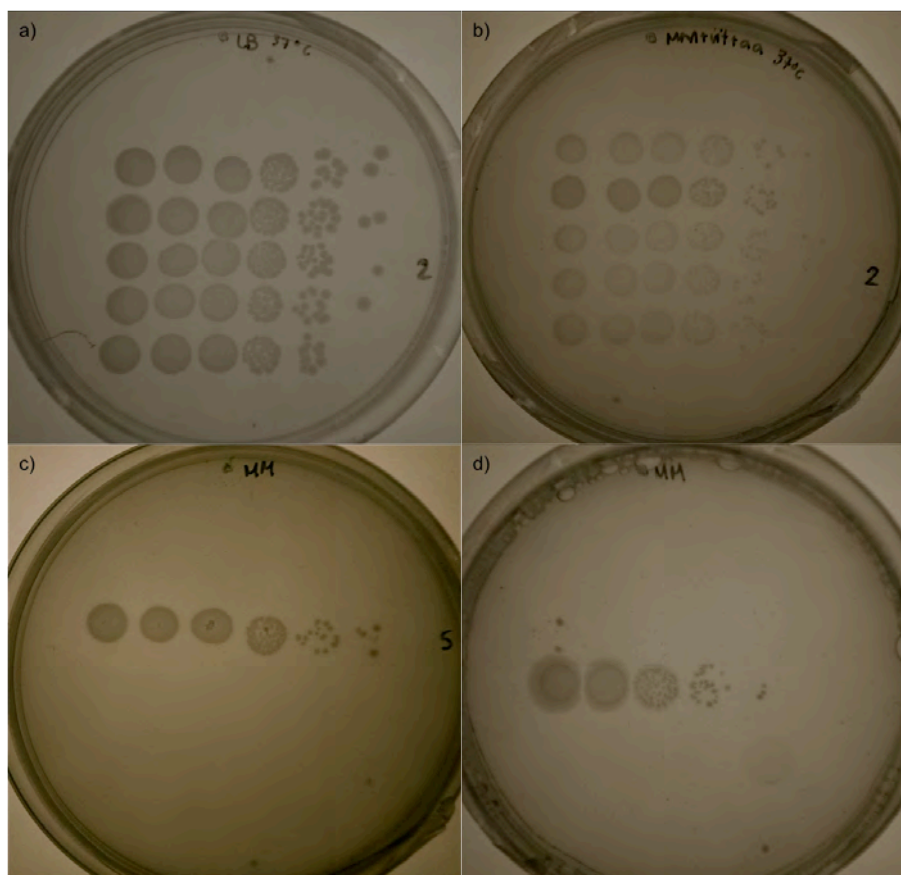
**Figura 10.** Western blot Anti-CAT de la fracción soluble. Después del control negativo siguen las muestras a 37°C, luego a 30°C y por último a 20°C. Se puede apreciar en la parte inferior de cada muestra una banda que corresponde a CAT libre.

### *Ensayo de complementación*

El ensayo de complementación fue realizado con la cepa K12 BW25113  $\Delta$ aroE, la cual es resistente a kanamicina, razón por la cual las construcciones se tuvieron que cambiar a un vector que tuviera resistencia a otro antibiótico. En este caso se eligió al vector pTrc99a, ya que tenía el mismo promotor que el vector con el que se estuvo trabajando previamente y además tenía resistencia a ampicilina, además este vector podía ser inducido con IPTG. Sin embargo, también se decidió hacer las complementaciones con el vector pT4, seleccionando con cloranfenicol en lugar de kanamicina, partiendo de los resultados de concentración mínima inhibitoria, donde tanto en el caso de AroE y YdiB, como en el de las quimeras, se vio que si eran capaces de resistir a cloranfenicol.

Los resultados con cada uno de los plásmido fueron distintos en cuanto a los controles. Con el vector pTrc99a sin ningún inductor el control de AroE fue positivo, es decir, si complementó a la cepa  $\Delta$ AroE, en cambio ni YdiB ni ninguna de las quimeras fueron capaces de complementar la función. Cuando se utilizó el vector pT4 los resultados mostraron una alta mortandad para la cepa complementada con AroE, y los pocos indicios de complementación se mostraban después de 1 semana de incubación; con este vector en

cambio, YdiB si complementó la función pero las quimera permanecieron sin poder hacerlo (Figura 11).



**Figura 11.** La cepa  $\Delta aroE$  complementada con AroE, YdiB y las quimera, Quim1 y Quim2, a) en LB; b) en medio mínimo adicionado con vitaminas y aminoácidos aromáticos, c) en medio mínimo en el vector pTrc99a, d) en medio mínimos en el vector pT4. En c) la única capaz de complementar fue AroE, mientras que en d) sólo lo hizo YdiB.

Para aclarar este hecho se realizó un tercer experimento donde se utilizó el vector pTrc99a pero induciéndolo con IPTG 0.1 mM. En este caso tanto AroE como YdiB complementaron la función permitiendo la supervivencia de la cepa  $\Delta aroE$ , sin embargo las quimeras siguieron sin mostrar función.

## **Discusión**

En este trabajo se construyeron dos quimeras a partir de dos proteínas con una identidad de secuencia de 25% y una homología estructural, AroE y YdiB, las cuales fueron analizadas estructuralmente y funcionalmente.

Se vio que las quimeras sí se pliegan en el ensayo de concentración mínima inhibitoria (CMI) ya que resistieron la misma concentración de cloranfenicol que las proteínas AroE y YdiB (140 mg/ml). También se vio que el dominio Rossmann fusionado a la CAT resistió un nivel medio de concentraciones del antibiótico (50 mg/ml), lo cual nos muestra que el dominio Rossmann es capaz de plegarse por sí mismo sin la necesidad del dominio de unión al sustrato.

Posteriormente se realizaron geles de SDS-PAGE para ver si las quimeras se encontraban en la fracción soluble de la célula, que era lo que se esperaba después de los resultados del CMI; sin embargo las quimeras nunca se pudieron apreciar en la fracción soluble, y de hecho, sólo a 37°C fue posible encontrar a AroE en esta fracción, mientras que todas las demás (incluyendo a la misma AroE) se encontraron en la fracción insoluble. A las demás temperaturas (30°C y 20°C) no se pudieron apreciar en la fracción soluble, sin embargo su cantidad disminuyó en la insoluble, esto indicaba que probablemente las quimeras y las otras proteínas estaban presentes en la fracción soluble, lo cual se comprobó al realizar un *western blot* de la fracción soluble de los experimentos a distintas temperaturas contra la CAT.

En el *western blot* fue posible ver las quimeras, a AroE y a YdiB en la fracción soluble, aunque también se apreció una banda que correspondía a la CAT libre. Probablemente la presencia de las proteínas que se acumularon en la fracción insoluble se debe a que el vector que se utilizó es de alto número de copias y tiene un promotor que se expresa constitutivamente, por lo que la cantidad de proteína que se produce es tanta que la célula sufre una carga metabólica tal que es incapaz de plegar bien todo el producto que se traduce y se forme cuerpos de inclusión con proteína mal plegada.

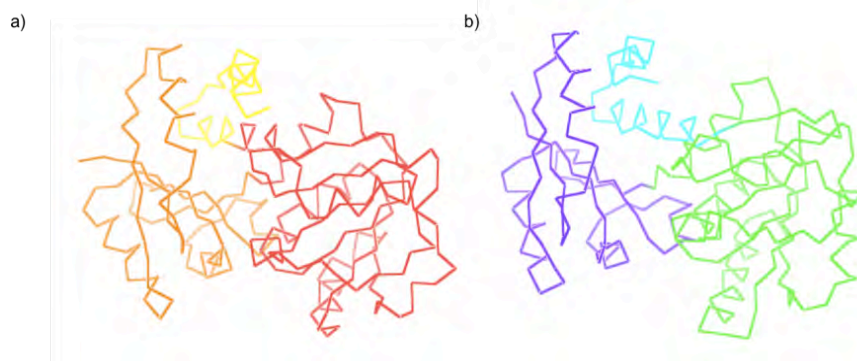
El hecho de encontrar a la CAT libre no descarta que las quimeras se plieguen, ya que en el *western blot* si se ven a las quimeras y a AroE y YdiB fusionadas con la CAT, este evento se puede tratar de explicar de dos maneras, una primera opción es que haya una secuencia alterna no canónica que el ribosoma reconozca como sitio de unión y traduzca a



partir de ahí la parte correspondiente a la CAT; otra posibilidad es que alguna proteasa corte al producto proteico dejando libre a la CAT.

Es importante recalcar que las quimeras se ven en la fracción soluble de la célula, es decir, que si son capaces de plegarse. Pero al parecer, las proteínas se enfrentan a ciertas limitaciones para que el plegamiento de todo el producto sea exitoso.

El intercambio de dominios puede generar modificaciones estructurales que impiden el plegamiento correcto de las proteínas. Analizando con más detalle la estructura de AroE y de YdiB se puede observar que el dominio N-terminal es completado por una estructura de horquilla  $\alpha$  helicoidal en el extremo C-terminal, que corresponde a las hélices  $\alpha 9$  y  $\alpha 10$  (Figura 5 y 12), esta estructura empaqueta a la hoja  $\beta$  al igual que la hélice  $\alpha 1$ . Al realizar las quimeras no se tomó en cuenta esto, lo cual posiblemente impidió el buen empaquetamiento del dominio de unión al sustrato en ambas quimeras, evitando a su vez que las enzimas pudieran tener actividad catalítica.



**Figura 12.** Estructura de la a) AroE, y b) YdiB, se puede apreciar en ambas proteínas que el extremo C-terminal (amarillo en AroE, azul en YdiB) complementa al dominio N-terminal (naranja en AroE, morado en YdiB).

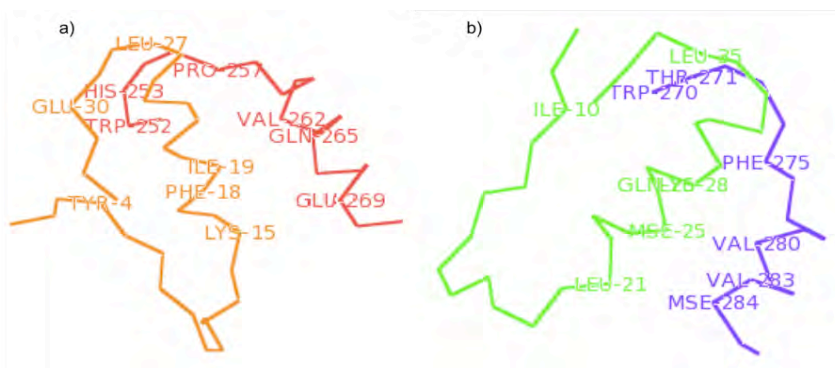
Se analizaron los aminoácidos que están en contacto entre el extremo C-terminal y el dominio de unión al sustrato N-terminal (Tabla 1 y 2, figura 13). Tanto en AroE como en YdiB la mayoría de los aminoácidos son hidrofóbicos, y los cambios de aminoácidos que hay entre ambas proteínas son similares y podrían no ser significativos. Sin embargo se podrían haber perdido contactos importantes para la estabilidad funcional de la proteína.

Residuo en AroE	Residuo en YdiB	Sustitución AroE : YdiB
E269	M284	Carga - : Hidrofóbico
Q265	V283	Sin carga : Hidrofóbico
V262	V280	Sin cambio (Hidrofóbicos)
P257	F275	Hidrofóbico : Hidrofóbico
H253	T271	Carga+ : Sin carga
W252	W270	Sin cambio (Hidrofóbicos)

**Tabla 1.** Correspondencia de aminoácidos de AroE y YdiB en el extremo C-terminal

Residuo en AroE	Residuo en YdiB	Sustitución AroE : YdiB
Y4	I10	Hidrofóbico : Hidrofóbico
K15	L21	Carga+ : Hidrofóbico
F18	M25	Hidrofóbico : Hidrofóbico
I19	K28	Hidrofóbico : Carga +
L27	Q26	Hidrofóbico : Sin carga
E30	L35	Carga - : Hidrofóbico

**Tabla 2.** Correspondencia de aminoácidos de AroE y YdiB en el extremo N-terminal



**Figura 13.** Aminoácidos que interaccionan entre el dominio N-terminal de las proteínas con la horquilla del extremo C-terminal. a) AroE, de naranja el extremo N-terminal, de rojo el extremo C-terminal; b) YdiB, de verde el extremo N-terminal y de morado el C-terminal.

En el análisis funcional ninguna de las quimeras fue activa, sin embargo YdiB si es capaz de complementar la función de AroE en la cepa  $\Delta aroE$ . Esto nos indicaba que ambas quimeras podrían ser capaces de complementar la función de shikimato deshidrogenasa,

es decir, que la quimera que contiene el dominio de unión a sustrato de YdiB podría complementar a la cepa deficiente al igual que la quimera que contenía el dominio de unión al sustrato de AroE, y que el hecho de que no complementarían la función puede deberse a una causa estructural.

Es posible que la interfase entre dominios también se haya visto alterada, afectando así la interacción del cofactor o del sustrato con los residuos catalíticos, impidiendo así la actividad.

Al igual que en el análisis estructural, el extremo C-terminal puede ser importante para el empaquetamiento correcto del extremo N-terminal, y por lo tanto para que la proteína sea funcional, ya que si el dominio Rossmann no está bien plegado es poco probable que sea capaz de realizar su función

El intercambio de los extremos C-terminal entre las quimera, sería entonces, una solución a los problemas estructurales y posiblemente también para los funcionales. Actualmente en el laboratorio esto se está realizando con quimeras construidas a partir de SDH ortólogas con resultados exitosos, ya que no sólo se están plegando mejor sino que al parecer también se mantiene la función, por lo tanto esta podría ser una manera de mejorar a las quimeras Quim1 y Quim2.

Sin embargo, por los resultados obtenidos se puede decir que la estabilidad estructural no garantiza la función de las proteínas quiméricas, y que no sólo se requiere que la proteína este plegada sino que los contactos entre los aminoácidos de la interfase sean los correctos para poder realizar una actividad completa. Es decir, no basta la independencia estructural y funcional aparente de los dominios, ya que finalmente la proteína se comportará como un todo en donde las partes tienen que estar bien complementadas entre sí.

Los dominios proteicos son más complejos que las definiciones estructurales o funcionales. El definir un dominio espacialmente requiere ver más allá de la secuencia, ya que, como se puede ver en este trabajo, la estructura de un dominio puede ser afectada por una serie de aminoácidos que se encuentran alejados en la secuencia pero que realmente son cercanos al plegarse, y que si estos no se tomaran en cuenta, la función entonces no podría realizarse.

Además tiene implicaciones en cuanto a la definición evolutiva de dominio, ya que no sólo evoluciona el dominio en sí, sino sus contactos con otros dominios deben de ser favorables para mantener la estructura y función de la proteína completa, es decir, la interacciones también deben de evolucionar.

Recientemente se publicó un artículo (Singh, 2008) donde se hizo un análisis filogenético de las shikimate deshidrogenasas. En ese trabajo se vio que YdiB es un xenólogo y no un parólogo de AroE y que filogenéticamente están alejadas, esto explica el bajo porcentaje de identidad de secuencia entre ambas (25%). Describen también a un nuevo grupo de SDH denominado Ael1 (AroE-like 1), este grupo es filogenéticamente más cercano a AroE. Sería interesante realizar quimeras entre AroE y alguna de las proteínas pertenecientes a ese grupo, ya que posiblemente se construyan quimeras más estables estructuralmente que inclusive puedan mantener la función.

## **Conclusiones**

De acuerdo a los antecedentes, es posible crear proteínas quiméricas a partir de dos proteínas homólogas parentales, en donde en la mayoría de los casos, una de las quimeras es funcional.

En este trabajo, se obtuvieron dos quimeras a partir de dos proteínas parálogas, AroE y YdiB, y a pesar de que si fueron capaces de plegarse no mantuvieron la función de shikimato deshidrogenasa. Esto nos indica que la estabilidad estructural no se traduce necesariamente en actividad catalítica.

A pesar de que los dominios se definen como estructuras independientes, tanto a nivel estructural como funcional, hay restricciones que impiden que la fusión entre dominios sea totalmente exitosa, se requiere que se forme una interfase estable, en donde los aminoácidos de ambas contrapartes sean compatibles. Además conservar la función representa un reto aún mayor, ya que se tienen que formar sitios nuevos que permitan la unión correcta de los sustratos, o que no se afecten por la fusión de los dominios los sitios ya existentes.

El entender más profundamente estas restricciones requerirá de un análisis más profundo, sin embargo, este es un acercamiento a las técnicas que se pueden utilizar para llegar a esa información.

## **Perspectivas**

El análisis de la estructura de las quimeras permitiría llegar a conclusiones más contundentes. Primero, sería importante ver si las quimeras siguen siendo capaces de plegarse aún en la ausencia de CAT, y de ser así volver a realizar los ensayos de complementación con las quimeras sin la fusión.

A pesar de que las quimeras no fueron capaces de complementar la función de AroE, es posible que si mantenga actividades parciales, por lo que sería apropiado hacer ensayos para probar si las quimeras son capaces de unir NAD o NADP, dependiendo el caso.

También se podría tratar de reparar a las quimeras, generando nuevas construcciones que contengan el extremo C-terminal que complementa al extremo N-terminal. Una vez obtenidas las quimeras con los extremos C-terminales intercambiados se deberían de repetir los experimentos con el reportero de plegamiento y la complementaciones de la cepa *ΔaroE*.

Para continuar con el análisis estructural, suponiendo que si son solubles las nuevas quimeras, se podrían realizar otros experimentos para ver la estabilidad de las proteínas quiméricas, como dicroísmo circular, ensayos calorimétricos y ensayos con agentes desnaturalizantes.

En caso de que estas quimeras no fueran activas se podría intentar hacer mutaciones puntuales al azar para ver si así recuperan su actividad.

## Apéndice I: Soluciones y medios de cultivo

### *Medio LB líquido*

NaCl 10g  
Extracto de levadura 5g  
Peptona 10g  
en un litro de agua miliQ

### *Medio LB sólido*

NaCl 2.5g  
Extracto de levadura 1.25g  
Peptona 2.5g  
Agar 3.75 g  
En un 250 ml de agua miliQ

### *Medio SOC*

Bacto triptona 2g  
Extracto de levadura 0.5g  
NaCl 0.05g  
en 75ml de agua miliQ  
Agregar:  
KCl 250mM 1 ml  
Ajustar a pH 7 con NaOH 4M  
Aforar a 100ml con agua miliQ

### *Sales M9 5X*

Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O 16g  
KH<sub>2</sub>PO<sub>4</sub> 3.75g  
NaCl 0.625g  
NH<sub>4</sub>Cl 1.25g  
Llevar a 250 ml con agua miliQ

### *Medio mínimo sólido*

Sales M9 5x 50ml  
Glucosa 20% 5ml  
MgSO<sub>4</sub> 1M 0.5ml  
CaCl<sub>2</sub> 1M 25µl  
Agar Noble 6g  
Llevar a 250ml con agua miliQ

### *PBS 10X*

NaH<sub>2</sub>PO<sub>4</sub> 2.28g  
Na<sub>2</sub>HPO<sub>4</sub> 11.5g  
NaCl 43.84g  
Agregar 450ml de agua miliQ  
Ajustar a pH 7.4  
Aforar a 500ml con agua miliQ

### *TAE 50X*

Tris Base 242g  
EDTA<sub>2</sub>NA 18.6g  
Ac. Acético glacial 60ml  
en un litro de agua miliQ  
Ajustar pH a 7

*Acrilamida 30%*

Acrilamida 29g  
Bis-acrilamida 1g  
Aforar a 100ml con agua miliQ  
Guardar a 4°C

*Upper Gel Solution (UGS)*

Tris-base 6.05g  
SDS 0.4g  
Ajustar a pH 6.8 con HCl  
Aforar a 100ml con agua miliQ

*Lower Gel Solution (LGS)*

Tris-base 18.1g  
SDS 0.4g  
Ajustar a pH 8.8 con HCl  
Llevar a 100 ml con agua miliQ

*Gel Loading buffer 4X*

Glicerol 2ml  
SDS 10% 4ml  
Tris 1M pH6.8 2.5ml  
Azul bromofenol 1% 0.5ml  
Antes de agregar a la muestra  
añadir 10% de βmercaptoetanol

*Teñir Coomasie*

Azul de Coomasie 0.2g  
Metanol 100ml  
Ácido acético glacial 20ml  
Llevar a 200ml con agua miliQ

*Desteñir I*

Metanol 500ml  
Ácido acético glacial 100ml  
Llevar a 1L con agua miliQ

*Desteñir II*

Metanol 100ml  
Ácido acético glacial 100ml  
Llevar a 1L con agua miliQ

*Cátodo*

Tris base 3g



Glicina 3g  
Metanol 200ml  
Aforar a 1L con agua miliQ

*Ánodo I*

Tris base 36.3g  
Metanol 200ml  
Aforar a 1L con agua miliQ

*Ánodo II*

Tris base 3g  
Metanol 200ml  
Aforar a 1L con agua miliQ

*TBS 20X*

Tris-HCl 15.7g  
NaCl 87.66g  
Llevar a 500ml con agua miliQ  
Guardar a 4°C

*TBST*

TBS 20X 25ml  
Tween 20 0.25ml  
Llevar a 500ml con agua miliQ  
Guardar a 4°C

## Apéndice II: Oligonucleótidos utilizados

- NcoAroE  
Secuencia (5'->3'): AGG TCC ATG GAA ACC TAT GCT G

- Bam AroE  
Secuencia (5'->3'): GCG GAT CCC GGC CAC GCG GAC AAT TCC

- BamydiB  
Secuencia (5'->3'): GTC AGC GGA TCC GGC ACC GAA CCC CAT GAC CTG TTT

- NdeYdiB  
Secuencia (5'->3'): GAG TTA CAT ATG GAT GTT ACC GCA AAA TAC GAA TTG

- NcoydiBsan  
Secuencia (5'->3'): AGGTCCATGGATGTTACCGCAAATAC

- BstEIIydiBsan  
Secuencia (5'->3'): GTATGGTCACCGTCGGTGTATAGCC

- YdiB-aro  
Secuencia (5'->3'): GAC AAT ACC GGC ACG GGC CAT ATT CG

- AroE-ydi  
Secuencia (5'->3'): GCC CGT GCC GGT ATT GTC ACC CAG CAG

## Referencias

- Aliverti, A. Pandini, V. Zanetti, G. 2004. Domain Exchange between isoforms of ferredoxin-NADP<sup>+</sup> reductase produces a functional enzyme. *Biochimica et Biophysica Acta* **1696**. 93-101.
- Bashton, M. Chothia, C. 2007. The Generation of New Protein Functions by the Combination of Domains. *Structure* **15**.
- Bashton, M. Chothia, C. 2002. The Geometry of Domain Combination in Proteins. *J. Mol. Biol.* **315**.
- Bottoms, A. Smith, E. Tanner, J. 2002. A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci* **11**. 2125-2137.
- Branden, C. Tooze, J. 1991. *Introduction to protein structure*. Garland Publishing. USA. 302pp.
- Carugo, O. Argos, P. 1997. NADP-dependent enzymes I: Conserved stereochemistry of cofactor binding. *Proteins* **28**. 10-28
- Carugo, O. Argos, P. 1997. NADP-dependent enzymes II: Evolution of the mono- and dinucleotide binding domains. *Proteins* **28**. 29-40
- Collinet, B. et al. 2000. Functionally accepted insertions of proteins within protein domains. *The Journal of Biological Chemistry* **275**. 17428-17433.
- Giesecke, A. Fang, R. Joung, J. 2006. Synthetic protein-protein interaction domains created by shuffling Cys2His2 zinc-fingers. *Mol Syst Biol.* **2**.
- Kallberg, Y. Persson, B. 2006. Prediction of coenzyme specificity in dehydrogenase/reductase. A hidden Markov model-based method and its application on complete genomes. *FEBS J* **273**. 1177-1184
- Kataoka, 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**. 931-936
- Lebbink, J.H.G. et al. 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 Engineering* **8**. 1287-1294
- Li, H. Fast, W. Benkovic, S.J. 2009. Structural and functional modularity of proteins in the de novo purine biosynthetic pathway. *Protein Science* **18**. 881-892.
- Lindner, H. et al. 2005. Site-directed Mutagenesis of the Active Site Region in the

Enrique-Rudino 5/11/09 04:16 PM

Comment: Paginas

- Quinate/ Shikimate 5-Dehydrogenase YdiB of *Escherichia coli*. *The Journal of Biological Chemistry* **280**. 7162-7169.
- Maxwell, K.L. et al. 1999. A simple in vivo assay for increased protein solubility. *Protein Sci*, **8**. 1908-1911.
  - Michel, G. et al. 2003. Structure of shikimate dehydrogenase AroE and its paralog YdiB. *The Journal of Biological Chemistry* **278**. 19463-19472.
  - Patthy, L. 1999. *Protein evolution*. Blackwell Science. UK. 228pp.
  - Sambrook, J. Fritsch, E. Maniatis, T. 1989. *Molecular Cloning. A laboratory manual*. 2a ed. Cold Spring harbor Laboratory Press. USA.
  - Sampaleanu, L.M. et al. 1999. Domain Exchange experiments in duck  $\delta$ -crystallins: Functional and evolutionary implications. *Protein Sci*. **8**. 529-537
  - Sieber, V. Martinez, C.A. Arnold F.H. 2001. Libraries of hybrid proteins from distantly related sequences. *Nat Biotechnol*, **19**. 456-60.
  - Singh, S. et al. 2008. A phylogenomic analysis of the shikimate dehydrogenases reveals broadscale functional diversification and identifies one functionally distinct subclass. *Mol. Biol. Evo* **25**. 2221-2232.
  - Thulasiram, H. Erickson, H. Poulter, C. 2007. Chimeras of two isoprenoid synthases catalyze all four coupling reactions in isoprenoid biosynthesis. *Science* **316**.
  - Waldo, G.S., 2003. Genetic screens and directed evolution for protein solubility. *Curr Opin Chem Biol*. **7**. 33-38.
  - Ye, S. et al. 2003. The crystal structure of shikimate dehydrogenase (AroE) reveals a unique NADPH binding mode. *Journal of Bacteriology* **185**. 4144-4151
  - Yeast, C.A. Orengo, C.A. 2007. Evolution of protein domains. *Encyclopedia of life science*.