



Universidad Nacional Autónoma de México

Facultad de Ciencias

Creación y análisis de quimeras producto del intercambio de dominios entre dos proteínas homólogas

TESIS

Que para obtener el título de: Bióloga
Presenta: Ximena Martínez de la Escalera Fanjul

Tutor: Dr. Lorenzo Patrick Segovia Forcella



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.

A mis padres

AGRADECIMIENTOS

A la UNAM por ser pública y gratuita, a la facultad de ciencias por la educación que tengo y los mejores años de mi vida y al IBT por permitirme realizar este trabajo en sus lujosas instalaciones.

A Lorenzo, mi director de tesis, por proponerme este proyecto tan padre y formativo.

A Claudia por enseñarme a trabajar en el laboratorio y a los miembros del laboratorio, en particular a Jesús por las charlas y la amena convivencia en el día a día
A mis sinodales Víctor, Lichi, Joel y Alfonso por sus correcciones, pero sobre todo a Lichi y a Vilchis por ser unos magníficos profesores y amigos.

A Alicia por la oportunidad que me dio en la licenciatura de trabajar en su laboratorio, por ser un modelo de científica apasionada y comprometida
A mi madre por inculcarme el espíritu científico y el no científico también, por presionarme y decirme que me echara un volado en los momentos correctos, por su apoyo y cariño incondicional, por la educación que me ha dado y por ser un gran ejemplo de mujer.

A mi padre por enseñarme en las caminatas tepoztecas todos los bichos y plantas, por apoyarme siempre y escuchar mis charlas biológicas e intentar entenderlas.

A mi abuelo por el apoyo a mi educación, por todos los ápidos y coleópteros de su gaveta y por ser un entomólogo frustrado.

A mi tía Nieves por ser mi segunda madre, por todo su cariño y consejos. A mi familia y en especial a Amaya por las muchas pláticas y por su afecto. Y a mis primos, todos distintos y todos divertidos.

A Santiago por la motivación que me inyectó sobre todo en momentos difíciles, por sus correcciones a este trabajo que estuvieron lejos de ser pocas, pero sobre todo, por el tiempo maravilloso que hemos pasado juntos.

A Uriel, Aura, Nando y Arturo, por hacer de este proceso algo mucho más enriquecedor, por sus consejos sobre mi proyecto y las pláticas ñoñas a la hora de comer.

A María, Nora, Piña y Daniela que en distintos momentos compartimos el salón de clase y lo hicimos muy bien y muy divertido.

A mis primas Julia y Natalia por ser mis hermanas. A Natalia por todas esas charlas sobre el quehacer científico y nuestro futuro en el área, por la ayuda cuando hay que tomar de decisiones. A Julia por ser mi compañera, no solo de facultad, sino de viajes y de vida, por ser mi más grande confidente, mi juez y mi abogada, por ser y estar.

A Alfredo por compartir nuestra vida en la facultad y a Natalia por ser una gran gran amiga. Ambos me han enseñado cosas que ya son parte de mi.

A las muchachas Julia, Fer, Sofía y Alexa por todas las risas, terapias y por convertir los momentos difíciles en una buena tarde en la terraza. Pero sobre todo a Julita por ser mi compañera de la carrera, mi equipo de trabajo, por las desveladas y por esas discusiones después de clase. A Fer por toda su ayuda en el diseño de este trabajo.

Y al resto de amigos, los del pulpo y los del prometeo, que compartieron esta época conmigo.

ÍNDICE

RESUMEN.....	1
INTRODUCCIÓN.....	2
ANTECEDENTES.....	8
MODELO DE ESTUDIO.....	15
Dominio Rossmann.....	15
Shikimato deshidrogenasas.....	17
HIPÓTESIS.....	26
OBJETIVOS.....	26
MATERIAL Y MÉTODO.....	27
Quimeras con un sitio de recombinación.....	27
Quimeras con dos sitios de recombinación.....	33
Concetración mínima inhibitoria.....	37
SDS PAGE.....	40
western blot.....	41
Complementación.....	42

RESULTADOS.....	44
Construcción de enzimas quiméricas.....	44
Concetración mínima inhibitoria.....	46
Complementación.....	48
SDS PAGE y Western blot.....	52
DISCUSIÓN Y CONCLUSIONES.....	55
PERSPECTIVAS.....	58
BIBLIOGRAFÍA.....	59
APÉNDICE 1.....	62
APÉNDICE 2.....	64

ÍNDICE DE FIGURAS

Figura 1	Representación de clasificación jerárquica de dominios de la base de datos SCOP.	5
Figura 2	Modelo de los posibles caminos de recombinación de un gen ancestral.	6
Figura 3	Esquema sobre método de las fusiones a proteínas reporteras del plegamiento y solubilidad.	12
Figura 4	Estructura secundaria y topología de un dominio Rossman clásico.	16
Figura 5	Vía del shikimato.	18
Figura 6	Shikimato deshidrogenasa YdiB de <i>E. coli</i> .	21
Figura 7	Estructura superimpuesta de AroE y YdiB	22
Figura 8	Mecanismo de reacción para la reducción del 3-dehidroshiquimato a shikimato.	23
Figura 9	Modelo de la estructura de AroE de <i>B. subtilis</i> .	25
Figura 10	Representación de la estructura de AroBS, YdiB y quimeras construidas.	27

Figura 11	Esquema representativo del primer PCR utilizado para generar la quimera BYB_CAT.	34
Figura 12	Esquema representativo del segundo PCR utilizado para generar la quimera BYB_CAT.	36
Figura 13	Representación de la estructura de AroBS, YdiB y quimeras construidas.	44
Figura 14	Cinéticas de complementación de la cepa $\Delta aroE km^-$ por la quimera BY_HIS.	49
Figura 15	Cinéticas de complementación de la cepa $\Delta aroE km^-$.	51
Figura 16	Expresión de proteínas.	52
Figura 17	Expresión de proteínas de quimeras en que se considera el dominio Rossman dentro del dominio catalítico.	54

ÍNDICE DE TABLAS

Tabla 1	Oligonucleótidos	37
Tabla 2	Cepas utilizadas	38
Tabla 3	Resultado CMI	47

ABREVIACIONES UTILIZADAS

ADH	Alcohol deshidrogenasa
AroBs	Shikimato deshidrogenasa de <i>Bacillus subtilis</i>
<i>aroBs</i>	Gen que codifica la shikimato deshidrogenasa de <i>Bacillus subtilis</i>
<i>aroD</i>	Gen que codifica la shikimato deshidrogenasa de <i>Bacillus subtilis</i>
AroE	Shikimato deshidrogenasa de <i>Escherichia coli</i>
<i>aroE</i>	Gen que codifica la shikimato deshidrogenasa de <i>Escherichia coli</i>
B	Proteína o dominio de shikimato deshidrogenasa de <i>Bacillus subtilis</i>
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
BY	Quimera con dominio N-terminal de AroBs y el C-terminal de YdiB
BYB	Quimera con dominio N-terminal y hélices $\alpha 9$ y $\alpha 10$ de AroBs y el C-terminal de YdiB
BYBmut	Quimera con dominio N-terminal y hélices $\alpha 9$ y $\alpha 10$ de AroBs y el C-terminal de YdiB con delección puntual
CAA	Medio mínimo (M9) con vitaminas y aminoácidos
CAT	Cloranfenicol acetil transferasa
CbADH	Alcohol deshidrogenasa de <i>Clostridium beijerinckii</i>
Cm	Cloranfenicol
CMI	Concentración mínima inhibitoria

CsGDH	Glutamato deshidrogenasa de <i>Clostridium symbiosum</i>
EcGDH	Glutamato deshidrogenasa de <i>Escherichia coli</i>
<i>E. coli</i>	<i>Escherichia coli</i>
EhADH1	Alcohol deshidrogenasa de <i>Entamoeba histolytica</i>
FAD ⁺	Flavín adenin dinuceótido
GDH	Glutamato deshidrogenasa
Km	Kanamicina
NAD(H)	Nicotamida adenina dinucleótido
NADP(H)	Nicotamida adenina dinucleótido fosfato
ON	<i>Over night</i> (16 h)
pIN-WT	Dominio catalítico de la integrasa HIV
PurK	Carboxilasa de la vía de síntesis <i>de novo</i> de purinas en <i>E. coli</i>
PurT	Formiltransferasa de la vía de síntesis <i>de novo</i> de purinas en <i>E. coli</i>
QSDH	Quinato shikimato deshidrogenasa
SAA	Medio mínimo (M9) con vitaminas
SCOP	<i>Structural Classification Of Proteins</i>
SDH	Shikimato deshidrogenasa
TbADH	Alcohol deshidrogenasa de <i>Thermoanaerobacter brockii</i>
wb	western blot

Y	Proteína o dominio de shikimato deshidrogenasa de <i>Escherichia coli</i>
YB	Quimera con dominio N-terminal de YdiB y el C-terminal de AroBs
YBY	Quimera con dominio N-terminal y hélices $\alpha 9$ y $\alpha 10$ de YdiB y el C-terminal de AroBs
YdiB	Quinato shikimato deshidrogenasa de <i>Escherichia coli</i>
ydiB	Gen que codifica la Quinato shikimato deshidrogenasa de <i>Escherichia coli</i>

RESUMEN

El estudio del intercambio de dominios entre proteínas permite entender la modularidad de éstas y los mecanismos evolutivos que hay detrás; es, también, uno de los pilares de la ingeniería de proteínas. En este trabajo se describe la construcción de cuatro quimeras, producto del intercambio entre dominios de la proteína shikimato deshidrogenasa (SDH) AroBs de *B. subtilis* (B) y la shikimato quinato deshidrogenasa (QSDH) YdiB de *E. coli* (Y). Las dos primeras quimeras se construyeron considerando el extremo C-terminal constituido por dominio el Rossman de la proteína nativa. Se denominaron BY y YB. El segundo par de quimeras se construyó considerando el dominio Rossman insertado dentro del dominio catalítico, por lo que aproximadamente los últimos 35 aa's del extremo C-terminal denominado cola (que son las hélices $\alpha 9$ y $\alpha 10$) pertenece a la secuencia de la proteína nativa de la que proviene el dominio catalítico (N-terminal). A partir de este intercambio se generaron las quimeras denominadas BYB y YBY. Se analizó el plegamiento y la solubilidad de las quimeras mediante la fusión a una proteína reportera, la enzima cloranfenicol acetil transferasa (CAT). La quimera que presentó mayor resistencia a cloranfenicol fue YB seguida de YBY. La única quimera que fue identificada en western blot fue YBY. Ninguna de las cuatro quimeras presentó actividad biológica en los ensayos de complementación de la cepa $\Delta aroE$. Finalmente se proponen algunas estrategias para entender y optimizar el intercambio de dominios entre YdiB y AroBs.

INTRODUCCIÓN

Los seres vivos se componen principalmente de cinco elementos: carbono, nitrógeno, hidrógeno, oxígeno y azufre. Estos átomos se enlazan y acomodan formando estructuras complejas como las biomoléculas, las cuales se clasifican en lípidos, carbohidratos, ácidos nucleicos y proteínas.

Dentro de las células, las proteínas desempeñan una gran diversidad de funciones. Usualmente se las asocia con el metabolismo; no obstante, participan en otros procesos como los de mantenimiento, corrección y decodificación del código genético, transporte y almacenamiento de pequeñas moléculas, comunicación celular y constituyen una parte fundamental en la estructura de las células, por lo que cada organismo cuenta con miles de diferentes proteínas.

Las enzimas son biomoléculas extremadamente complejas. Catalizan las reacciones que permiten la vida. La probabilidad de que las reacciones catalizadas por las enzimas ocurran de manera espontánea es tan pequeña que no ocurrirían ni en los años de vida que tiene el planeta Tierra. Las proteínas logran que estas reacciones sucedan en cuestión de minutos o segundos.

Las proteínas son biopolímeros cuyos monómeros son los L- α aminoácidos. Estos aminoácidos se unen entre sí generando una secuencia lineal denominada primaria o cadena polipeptídica. La secuencia primaria de una proteína es determinada por el gen que la codifica. Es sabido que para cada cadena polipeptídica existe un gen (Mathews, *et al.*, 2005).

Si bien un polipéptido se sintetiza de manera lineal, su forma funcional es plegada, el plegamiento comienza en cuanto los primeros aminoácidos salen del ribosoma, la fuerza que dirige este primer estadio del plegamiento son las interacciones hidrofóbicas (Brylinski, *et al.*, 2006). Los grupos hidrofóbicos en contacto con agua disminuyen la cantidad de puentes de hidrógeno que pueden formar las moléculas de agua, lo que coloca al agua en un estado energeticamente desfavorable, por lo que el agua promueve el agrupamiento de los residuos hidrofóbicos para disminuir la superficie al contacto con estos grupos, a este fenómeno se le denomina colapso hidrofóbico y la fuerza que lo dirige es la entropía (Williamson, 2012)

Aunque cada aminoácido dentro de la secuencia de la cadena polipeptídica es un componente independiente del sistema (la proteína), ésta debe de satisfacer restricciones globales (Steipe, 2004). Entre estas restricciones existen las de tipo energético, que deben cumplirse para que se dé un plegamiento correcto. La secuencia primaria tiene una energía libre elevada y en el proceso de plegamiento, esta energía libre va disminuyendo hasta alcanzar un estado termodinámico de mínima energía libre, por lo que el plegamiento resulta de la diferencia de energía libre entre el estado no plegado y el plegado, de manera que el plegamiento ocurre de manera espontánea ($\Delta G = \Delta H - T \Delta S < 0$)

La estabilidad neta de la proteína surge de unas cuantas interacciones del total que presenta (Arnold, *et al.*, 2001). A consecuencia de las interacciones (como son los puentes de hidrógeno) que se forman entre los residuos, se puede derivar un plegado regular local que se conoce como estructura secundaria. Las formas más conocidas son las hélices α , láminas β y *random coil*. Estas regiones a su vez se pliegan unas sobre otras formando una nueva estructura denominada terciaria. Este plegado da a la molécula su forma tridimensional total (Mathews, *et al.*, 2005). Finalmente en algunos casos, diferentes ca-

denas polipeptídicas se empaquetan juntas formando oligómeros lo que se conoce como estructura cuaternaria (Williamson, 2012).

Dentro de la forma global de la proteína se pueden reconocer dominios que se pueden definir de varias formas. Para este trabajo se manejarán como componentes independientes de carácter: estructural, funcional y evolutivo. Por lo general, las interacciones intradominio son mayores que las que se dan interdominios, por lo que se definen como estructuralmente independientes respecto al resto de la cadena polipeptídica. Con frecuencia, un dominio puede ser expresado como unidad conservando la capacidad de plegarse tomando su conformación tridimensional de manera correcta y manteniendo su función (Han, *et al.*, 2007). De esta forma, un dominio constituye el mínimo de secuencia requerida para codificar una función. Finalmente, los dominios se consideran una unidad evolutiva debido a que son bloques que contienen una secuencia de aminoácidos que se heredan y se conservan en distintos contextos (Yeats, 2007).

Existen bases de datos que clasifican y almacenan los dominios identificados. La base de datos *Structural Classification of Proteins* (SCOP) agrupa a las proteínas o a los dominios de proteínas en base a criterios de homología: a los dominios relacionados los agrupa en familias (>30% de similitud de secuencia) y a las familias relacionadas las agrupa en superfamilias (por origen evolutivo común basado en características estructurales y funcionales). Por criterio de analogía (estructurales), agrupa a las superfamilias en plegamientos y a los plegamientos en clases de estructura secundaria (Alva, *et al.*, 2010; Bashton y Chothia, 2002). Sin embargo, la definición de dominio considerada en esta base de datos es principalmente evolutiva y no estructural, lo que implica que si una proteína es dividida en dominios, se observarían éstos en otros contextos proteicos, es decir, unidos a otros dominios (Bashton y Chothia, 2007) Figura 1.

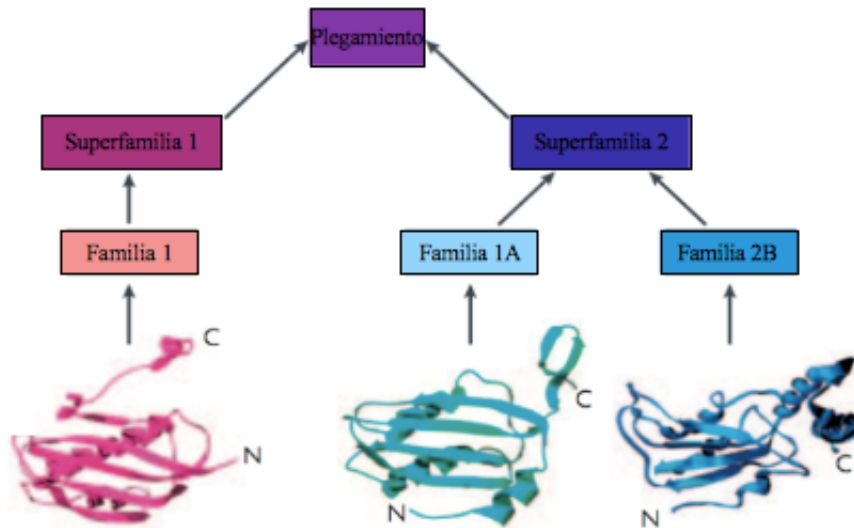


Figura 1 . Representación esquemática de la clasificación jerárquica de dominios de la base de datos.

SCOP (Modificada de Han, *et al.*,2007)

Hay proteínas que sólo constan de un dominio, pero la gran mayoría de las proteínas se componen de más de uno, a las cuales se les conoce como proteínas multidominio. Este último tipo de proteínas por lo general desempeñan funciones más específicas o complejas. Usualmente, los dominios que se encuentran en un contexto multidominio mantienen la función de su homólogo monodominio, sin embargo, el contexto en el que se encuentran permite que el dominio realice una función más específica, la altere o la expanda. No obstante, como siempre ocurre en la biología hay excepciones, pues hay dominios en un contexto multidominio que no realizan la misma función que su homólogo monodominio. Esto se explica por la acumulación de mutaciones y la acción de la selección natural (Bashton y Chothia, 2007). La divergencia puede deberse a procesos diferentes a la acumulación progresiva de mutaciones como recombinación no homóloga, permutaciones y barajeo. Una de las explicaciones que existen para la aparición de las

proteínas multidominio y su diversidad es la duplicación y posterior recombinación de dominios, seguida inevitablemente por la divergencia, que puede resultar en el cambio de función de un dominio (Vogel y Morea, 2006) Fig. 2.

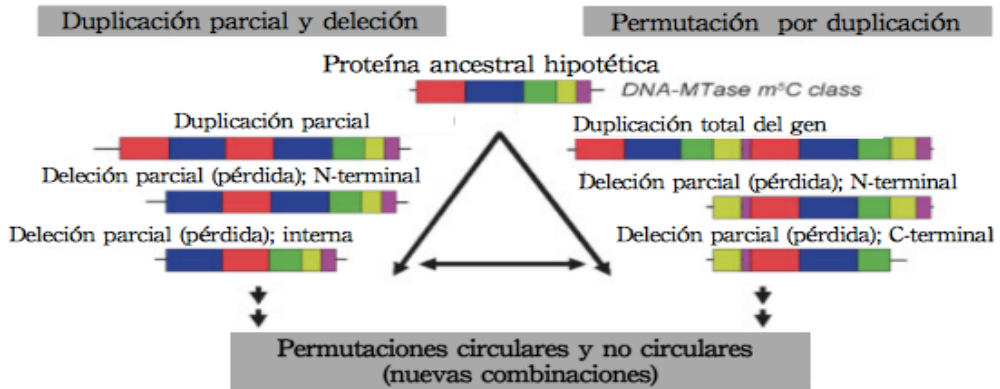


Figura 2. Esquema que ilustra los posibles caminos de recombinación de un gen ancestral. Los bloques de cada color representan un dominio. Izquierda duplicación parcial y deleción. Derecha permutación por duplicación y deleción de N o C terminal (modificada de Vogel y Morea, 2006).

Considerando un largo de secuencia promedio de 300 residuos, el número de posibles secuencias (20^{300}) excede al número estimado de partículas en el universo, por lo que resulta imposible que la vida haya explorado la totalidad de este espacio o siquiera la mitad. Las proteínas que conocemos hoy reflejan la descendencia de ancestros (éstos con un número limitado de formas), por lo que la duplicación, recombinación y divergencia de dominios explica la diversidad protéica que observamos actualmente (Alva, *et al.*, 2010).

El espacio de dominios no está completamente explorado. Sólo se conoce un número limitado de combinaciones de dominios. Unas pocas superfamilias de dominios son abundantes y se encuentran combinadas con muchas y muy diferentes superfamilias de dominios, mientras que la mayoría de las superfamilias se encuentran presentes en pocas

proteínas y sólo ocurren en un limitado número de combinaciones (Han, *et al.*, 2007). Dos dominios se pueden encontrar siempre unidos en la misma geometría, es decir, por el N-terminal o por el C-terminal. Este tipo de evidencia sugiere que la duplicación pueda darse también por bloques de dominios. Estas observaciones se vuelven robustas al encontrar que estos bloques de dominios se encuentran más conservados, lo cual sugiere que en vez de tratarse de restricciones evolutivas se trate de una geometría que responde a la descendencia de proteínas ancestrales (Figura 2). Es decir, más a la historia que a la función (Vogel, *et al.*, 2004), lo que no quiere decir que no puedan ocurrir nuevas combinaciones por procesos que se mencionaron antes.

La ingeniería en proteínas es el diseño y construcción de proteínas nuevas a través de la manipulación de sus genes. Es una rama reciente producto de los avances en el conocimiento de la secuencia, función, estructura y evolución de las proteínas; en donde a partir de la información *a priori* se realizan modificaciones en la secuencia de la proteína para generar una nueva función. La ingeniería en proteínas ha contribuido a entender la relación que existe entre la secuencia-estructura y estructura-función sin embargo aun existe mucho camino que recorrer en esta área del conocimiento (Fersht y Winter, 1992).

La primera proteína que fue trabajada mediante ingeniería de proteínas fue en 1982, las modificaciones que se realizaban en ese momento eran de uno o dos aminoácidos, pero pronto siguieron cambios de azas y hasta de dominios enteros, desde entonces la ingeniería en proteínas es un campo que ha crecido rápidamente gracias al uso cotidiano de oligonucleótidos. Hoy en día la ingeniería de proteínas ha encontrado aplicaciones en la industria, en la bioremediación y en el tratamiento de enfermedades (Brannigan, J. A., y Wilkinson, A. J., 2002).

ANTECEDENTES

En 1994 Kataoka, *et al.*, con el fin de comprobar el mecanismo de reconocimiento específico de sustratos, generaron un par de quimeras a partir de dos proteínas con 47% de identidad. Las enzimas utilizadas fueron la fenilalanina deshidrogenasa de *Thermoactinomyces intermedius*, cuyos sustratos preferencialmente son L-fenilalanina y L-tirosina, y la leucina deshidrogenasa de *Bacillus stearothermophilus*, que utiliza como sustrato L-leucina y algunos L-aminoácidos ramificados. Estas proteínas presentan en el extremo N-terminal la región encargada del reconocimiento del sustrato, mientras que el C-terminal se ocupa de la unión al cofactor NAD⁺. La quimera en la que el N-terminal correspondía a la leucina deshidrogenasa y el C-terminal a la fenilalanina deshidrogenasa, no pudo expresarse, mientras que la quimera contraria con N-terminal de fenilalanina deshidrogenasa y C-terminal de leucina deshidrogenasa no sólo se logró expresar, sino que presentó una Km igual para los sustratos de la fenilalanina deshidrogenasa, una menor actividad específica para la fenilalanina. En la deaminación oxidativa la especificidad por el sustrato fue igual que la de la fenilalanina deshidrogenasa, pero en la deaminación reductiva la especificidad por el sustrato fue la mezcla aditiva de la que presentan las enzimas parentales. Este trabajo demostró que los dominios presentan una capacidad de plegamiento autónomo, pero que mantienen la comunicación entre ellos y son capaces de interactuar para formar un sitio activo nuevo que resulta en la aparición de nuevas funciones.

En 2009 Li, *et al.* generaron seis proteínas quiméricas a partir de las combinaciones posibles del intercambio del primer o tercer dominio entre los tres dominios de las proteínas PurT y PurK de *E. coli*. con la finalidad de probar la intercambiabilidad de dominios, así como la generación de nuevas funciones a partir del intercambio de dominios. Estas

enzimas participan en la vía de síntesis *de novo* de purinas en donde PurT es una formiltransferasa y PurK una carboxilasa. A pesar de que difieren en su función, y con sólo 27% de identidad, sus estructuras tridimensionales son superponibles, unen sus respectivos sustratos de manera similar y comparten ciertas características en los mecanismos catalíticos. Ninguna quimera fue capaz de complementar a las respectivas cepas auxótrofas. La primera explicación para el resultado anterior fue que se debía a la inestabilidad de las quimeras, por lo que se procedió a generar una fusión traduccional de cada una de éstas a la proteína de unión a maltosa con la finalidad de minimizar la acumulación de las quimeras en cuerpos de inclusión. Sin embargo, tampoco de esta manera se logró que alguna quimera complementara. Lo que sí se observó fue la capacidad de las quimeras de catalizar reacciones parciales, sugiriendo con esto una incapacidad para mantener a los intermediarios de las reacciones. Sustituyendo la secuencia del loop-B cercano a la interface central de los tres dominios, Li, *et al.*, lograron obtener una quimera con la capacidad de realizar la reacción en su totalidad y por lo tanto de complementar a la cepa auxótrofa PurK⁻. Este trabajo demuestra que a pesar de que los dominios sean considerados tanto estructuralmente como funcionalmente independientes, es necesario que exista una compatibilidad entre ellos para que se pueda reconstituir la función en su totalidad (Li, *et al.*, 2009).

Sharkey y Engel en 2009 abordaron el tema de la intercambiabilidad entre dominios para poner a prueba la autonomía y las propiedades de cada dominio dentro de la estructura total de una proteína. Utilizaron como modelo de estudio dos glutamato deshidrogenasas, una proveniente de *E. coli* (EcGDH) y la otra de *Clostridium symbiosum* (CsGDH). Por simplicidad, utilizaron solo la primera letra de cada proteína E ó C respectivamente, para denotar al organismo al que pertenece cada dominio. Estas proteínas con 54% de identidad, en donde la arquitectura del gen es igual, cuentan con dos dominios. El N-

terminal corresponde al dominio de unión al sustrato. Éste no es continuo en la cadena de aminoácidos ya que se encuentra interrumpido por el C-terminal que corresponde al dominio de unión al cofactor, de tal forma que el dominio de unión a sustrato flanquea a ambos lados al dominio de unión al cofactor. Ambas enzimas realizan la misma función, sin embargo, difieren en varios aspectos como son: C utiliza como cofactor NAD^+ y presenta una gran dependencia a pH, mientras que E utiliza NADP^+ y es independiente del pH. Posteriormente, construyeron dos quimeras, CEC y ECE respetando la arquitectura del gen. El primer problema con que se encontró este grupo fue la expresión y solubilidad de las quimeras, sin embargo, lo resolvieron con un medio autoinducible y la expresión conjunta de chaperonas. En este momento obtuvieron a la quimera CEC de manera soluble. Este alentador resultado los llevó a construir una tercera quimera, CE, en la que no se respetó la arquitectura del gen, no obstante, esta quimera fue completamente insoluble. Lo anterior fue explicado por la falta de contactos entre la hélice correspondiente al dominio de unión al sustrato ubicada en la porción C-terminal, lo que aparentemente provocó un plegamiento deficiente. La quimera CEC formó homohexameros como la quimera CsGDH parental y fue capaz de complementar a las cepas auxótrofas ΔgdhA (mutante nula en EcGDH, sin actividad detectable de GDH dependiente de NADP^+). Esta quimera utilizó como cofactor al NADP(H) , pues su dominio de unión al cofactor pertenecía a E, lo que corrobora que los residuos que constituyen a este dominio son suficientes para unirse selectivamente a NADP(H) . CEC mostró una cooperatividad positiva con el L- glutamato al igual que CsGDH, pero difirió respecto a ésta en poder funcionar en un amplio rango de pH. La K_m para todos sus sustratos y la del cofactor fue menor, y la V_{max} mayor para estos. Este trabajo concluye que la interacción entre dominios es importante tanto para la afinidad de los sustratos como para el cofactor (Sharkey y Engle, 2009).

Goihberg, *et al.* en 2010, con la finalidad de comprender los determinantes de la termoestabilidad, llevaron a cabo el intercambio de un dominio entre tres alcohol deshidrogenasas (ADH) provenientes de distintos organismos: de una bacteria termófila *Thermoanaerobacter brockii* (TbADH), otra de la bacteria mesófila *Clostridium beijerinckii* (CbADH) y del protozoario *Entamoeba histolytica* (EhADH1). Estas enzimas cuentan con dos dominios bien definidos: el dominio de unión a cofactor (residuos 153-295), y el dominio catalítico (residuos 1-152 y 296-351/2). Estas proteínas forman homotetrámeros. Las quimeras que construyeron fueron denominadas TCT, CTC y ETE. Como se puede apreciar, utilizaron la primera letra del nombre de la enzima para denotar a qué organismo pertenece cada dominio. Tanto en TCT como en CTC se obtuvieron quimeras que formaban homotetrámeros y mantuvieron su actividad enzimática como las enzimas parentales. No obstante, en TCT la termoestabilidad se vio bastante disminuída respecto a la enzima parental TbADH. En la quimera CTC, la estabilidad térmica no se incrementó respecto a la de la enzima parental mesófila. La diferente termoestabilidad entre estas dos quimeras se atribuyó al ajuste imperfecto entre el dominio huésped y el anfitrión, pues a pesar que las diferencias entre los esqueletos de carbonos α sean no significativas, Goihberg *et al.*, argumentan que el intercambio puede resultar en la interferencia de elementos estabilizadores, implicados en la asociación entre dominios. En el caso de ETE se obtuvo una quimera monomérica cuya estructura secundaria se mantuvo respecto a las parentales. Esta quimera no presentó actividad y su estabilidad fue a temperaturas muy bajas. Estos resultados se atribuyeron a la pérdida de residuos clave, los cuales supuestamente impedían la formación del tetrámero. Comprobaron que la formación del tetrámero se ve involucrada en el aumento de la estabilidad a altas temperaturas, por lo que con una mutación puntual en la zona de asociación de monómeros lograron la formación de tetrámeros (Goihberg *et al.*, 2010)

Existen distintos métodos de tamizaje de solubilidad de proteínas entre los que se encuentran las fusiones a proteínas reporteras del plegamiento, en donde una proteína de interés es unida a una proteína reportera mediante una fusión traduccional. La proteína reportera debe presentar una actividad biológica fácilmente cuantificable, de tal manera, que tanto el plegamiento, como la solubilidad de la proteína que se quiere probar se traducirá en una actividad seleccionable, la realizada por la proteína reportera (Waldo, 2003).

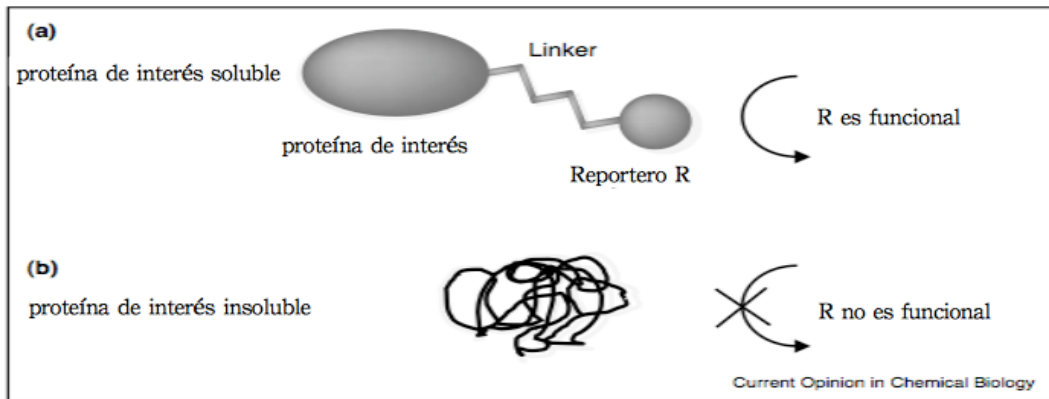


Figura 3. Principio en que se basa el método de las fusiones a proteínas reporteras del plegamiento y solubilidad. Un reportero con una actividad biológica fácilmente cuantificable se fusiona a la proteína de interés; cuando: a) la proteína de interés se pliega correctamente y es soluble el reportero también lo será y presenta actividad, la cual puede ser medible. b) cuando la proteína de interés es insoluble o se encuentra mal plegada, la función del reportero no es llevada a cabo, pues se ve impedido su plegamiento y por lo tanto su actividad (modificada de Waldo, 2003)

La enzima cloranfenicol acetil transferasa (CAT) acetila al cloranfenicol (Cm) impidiendo que se una al ribosoma bacteriano (bactericida), evitando de esta forma su efecto inhibitorio sobre la síntesis protéica. Con el fin de generar un nuevo método para poder identificar mutantes solubles generadas a partir de proteínas insolubles, en 1999 Maxwell, *et al.*, utilizaron la CAT como reportero de plegamiento basándose en que si la CAT

se encuentra fusionada a un dominio insoluble, la fusión precipitará dentro de la célula, lo que resultará en una pequeña cantidad de CAT intracelular y una baja resistencia a cloranfenicol. Por el contrario, si se encuentra unida a una mutante soluble se observará una mayor resistencia a Cm puesto que habrá mayor CAT intracelular (Figura 3). Para probar lo anterior, generaron dos fusiones traduccionales: La primera, entre el dominio catalítico de la integrasa VIH, que es insoluble en condiciones *in vitro* e *in vivo* cuando es expresado en *E. coli*, a la cual denominaron pIN-WT; La segunda, en la que substituyeron la Phe 185 por Lys, lo que hace soluble a este dominio tanto *in vivo* como *in vitro* llamada pIN-F185K. Debido a la diferencia evidente en la resistencia a cloranfenicol que presentaron las dos distintas fusiones (pIN-F185K 500 µg/ml y pIN-WT 250 µg/ml), se realizaron diluciones diferenciales de ambas construcciones (1:10, 000 respectivamente) para probar la capacidad de selección de mutantes solubles con este método, y el resultado fue que el 89% de las colonias que crecieron en LB con Cm (500 µg/ml) obtenidas fueron de pIN-F185K.

Más adelante, en 2001 Sieber, V. *et al.*, generaron una librería de quimeras resultado de la fusión del citocromo humano P₄₅₀ (1A2), que es membranal y el citocromo P₄₅₀ BM de *Bacillus megaterium*, que es soluble. Este grupo fusionó a la CAT las variantes de la librería para identificar aquellas que respetaran el marco de lectura y que fueran solubles. Aquellas variantes que fueran membranales impedirían la trimerización de la CAT, al igual no seleccionaría las variantes que no se plegaran. Un 80% de las candidatas seleccionadas por medio de las CAT mostró actividad.

En el laboratorio del Dr. Lorenzo Segovia, se construyeron quimeras producto del intercambio de los dominios entre tres diferentes Shikimato deshidrogenasas. Estas proteínas constan de dos dominios, el N-terminal o dominio I, que corresponde al dominio cat-

alítico o de unión a sustrato y el C-terminal o dominio II, que es el dominio de unión al cofactor. La arquitectura de estas proteínas es similar a la de las GDH, en donde el dominio catalítico es discontinuo en la secuencia de aminoácidos pues flanquea al dominio de unión al cofactor. Estas enzimas participan en la vía del Shikimato (detallada mas adelante). Dos de ellas provienen de *E. coli* AroE y YdiB, mientras que AroBs proviene de *B. subtilis*. Se denominaron a las quimeras utilizando las letras E y B para la identificación del organismo al cual pertenece cada dominio, y Y por el nombre de la proteína YdiB. Las primeras quimeras que se construyeron fueron EB, BE, EY y YE, en las que no se consideraba la discontinuidad del dominio N-terminal. Todas las quimeras se realizaron con una fusión traduccional a la enzima cloranfenicol acetil transferasa. Esta enzima funciona como un reportero de plegamiento y ayuda a la solubilización de las proteínas. Ninguna de estas proteínas tuvo la capacidad de complementar a la cepa Δ aroE de *E. coli*, a pesar de presentar altas resistencias a cloranfenicol. Puesto que la enzima CAT puede intervenir en la función de éstas proteínas, se decidió eliminarla, y al hacer esta modificación el gen de la quimera quedó fusionado con seis histidinas en fase. Sin embargo, tampoco así se logró que complementaran a esta cepa (Δ aroE). Lo siguiente que se realizó fue un cambio de cola (α 9 y α 10) del extremo C-terminal correspondiente al dominio N-terminal, de tal forma que se consideró la discontinuidad del dominio de unión a sustrato. Las quimeras que se generaron fueron: EBE, BEB, EYE y YEY. De estas sólo EBE logró complementar a la cepa Δ aroE, que resultó ser más soluble que una de las proteínas parentales AroBs, y mostró ser más eficiente que ambas proteínas parentales basándose en el coeficiente k_{cat}/K_m . Las otras quimeras fueron solubles cuando se encontraron unidas a la CAT, sin embargo, por sí solas (con la etiqueta de 6 His) no lo fueron, ya que se encontraron en cuerpos de inclusión o en algunos casos no se las encontró mediante western blot (wb) anti His, lo que puede explicarse por un plegamiento deficiente que puede resultar en una posterior y pronta proteólisis.

MODELO DE ESTUDIO

Dominio Rossman

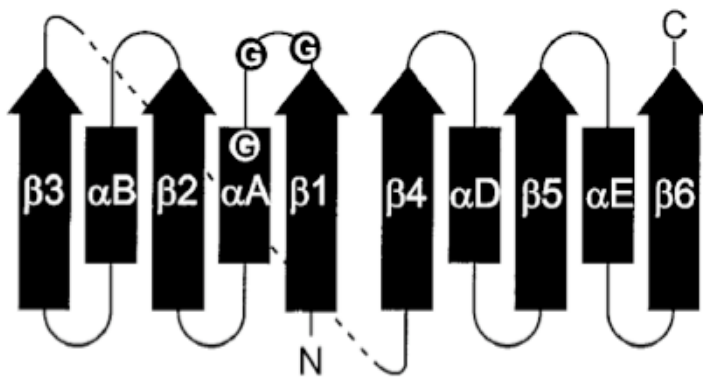
El dominio Rossman, también conocido como dominio de unión a dinucleótidos, se encuentra distribuido en todos los reinos. Es uno de los dominios más abundantes. Se encuentra en el 3% de las proteínas de procariontes y en el 1% de las de eucariontes (Kallberg y Persson, 2006).

Los dominios Rossman se encuentran conectados a una gran variedad de dominios. Estos últimos se conocen como dominios catalíticos y pertenecen a 17 familias de dominios que corresponden a siete superfamilias, por lo que la superfamilia Rossman es una de las pocas que se encuentra combinada con muchas otras superfamilias.

El dominio Rossman comúnmente une a los dinucleótidos: nicotinamida adenina fosfato (NADP), nicotinamida adenina (NAD⁺), o flavin adenina (FAD⁺). Cuando se forma la holoenzima, el cofactor (NADP⁺, NAD⁺ o FAD⁺) transfiere un ion hidronio desde o hacia el átomo C4 del anillo nitrogenado. (Carugo y Argos, 1997). Este participa en reacciones tanto de oxidación como de reducción.

El dominio Rossman se compone de dos motivos de unión a mononucleótidos cuya estructura secundaria es $\beta\alpha\beta\alpha\beta$. Éstos están unidos por la rotación de un *linker*, lo que resulta en seis hojas β paralelas flanqueadas por hélices α . Las hojas β presentan un orden relativo 321456 como se observa en la figura 4 (Bottoms, Smith y Tanner, 2002). Estos dominios muestran ciertas características conservadas:

- I) Tres glicinas, la primera al final de la hoja $\beta 1$ y las otras dos al inicio de la hélice αA . Las primeras dos glicinas están relacionadas en la unión del dinucleótido, mientras que la tercera se ve involucrada con el empaquetamiento de las hojas β y las hélices α (Figura 4). Estas glicinas presentan un patrón general $Gx(x)Gx(x)G$, lo que significa que el espacio entre glicinas puede ser de uno o dos residuos, sin embargo, hay excepciones a la regla, por ejemplo en la que
- II) Un residuo al final de la hoja $\beta 2$ relacionado con la identificación del cofactor; por lo general la presencia de un residuo cargado negativamente en esta posición (Asp o Glu) determina la unión del cofactor NAD^+ o FAD^+ , en cambio, las enzimas que presentan un residuo básico (Arg) suelen unir a $NADP^+$ (Kallberg y Persson, 2006; Carugo y Argos, 1997).
- III) La presencia de una molécula de agua en la apoenzima unida a la hélice αA presenta un patrón de puentes de hidrógeno que se dirigen a las tres glicinas conservadas y al pirofosfato del cofactor; se ha propuesto que ésta se ve involucrada en el reconocimiento del cofactor (Carugo y Argos, 1997; Bottoms, Smith y Tanner, 2002).



Existen cu

ectado

Figura 4. Estructura secundaria y topología de un dominio Rossmán clásico. Las flechas representan las hojas β y los rectángulos las hélices α . Los círculos representan las glicinas conservadas (Tomada de Bottoms, 2002).

al dominio catalítico: I) unido al extremo N-terminal del dominio catalítico, II) unido al extremo C-terminal, III) el dominio catalítico interrumpiendo el Rossman, IV) el dominio Rossman interrumpiendo el catalítico. Dentro de cada superfamilia de dominios catalíticos se mantiene la geometría de conexión (entre el dominio Rossman y el dominio catalítico). El *linker* que existe entre los dominios dentro de cada superfamilia sigue el mismo camino, y a excepción de una superfamilia, la gran longitud del *linker* permite posiciones similares en los cuatro tipos de conexiones que se presentan. Lo que sugiere que esta constancia en la geometría de unión entre los dominios, dentro de cada superfamilia, se debe a un solo evento de recombinación, es decir, que cada superfamilia surgió de una duplicación y la diversidad fue adquirida por divergencia más que por restricciones funcionales (Bashton y Chothia, 2002).

shikimato 5-deshidrogenasas

Este trabajo se realizó con dos enzimas que pertenecen a la superfamilia de las shikimato deshidrogenasas (SDH). Los miembros de esta superfamilia están compuestos por un dominio Rossman y un dominio catalítico. La superfamilia de las SDH ha sido dividida en cinco grupos funcionalmente diferentes (AroE, YdiB, Rf1, SdhL y Ael1) que se detallarán más adelante.

La enzima AroE es considerada como la prototipo de la superfamilia de las SDH. Las enzimas que pertenecen al grupo de AroE catalizan el cuarto paso de los siete que tiene la vía del shikimato (Singh *et al.*, 2008), vía que da como producto final al corismato. Este compuesto es el precursor de los aminoácidos aromáticos (Phe, Trp y Tyr) y metabolitos secundarios como se muestra en la Figura 5 (Pittard y Wallace, 1966). Se ha propuesto que los homólogos a la AroE, pero de otros grupos, tienen otras funciones. Algunos de

estos homólogos son proteínas esenciales, mientras que Sinhg, *et al.*, en 2008, proponen que otras pueden desarrollar roles condicionales o de nicho específico desviando la vía del shikimato para generar diversos productos.

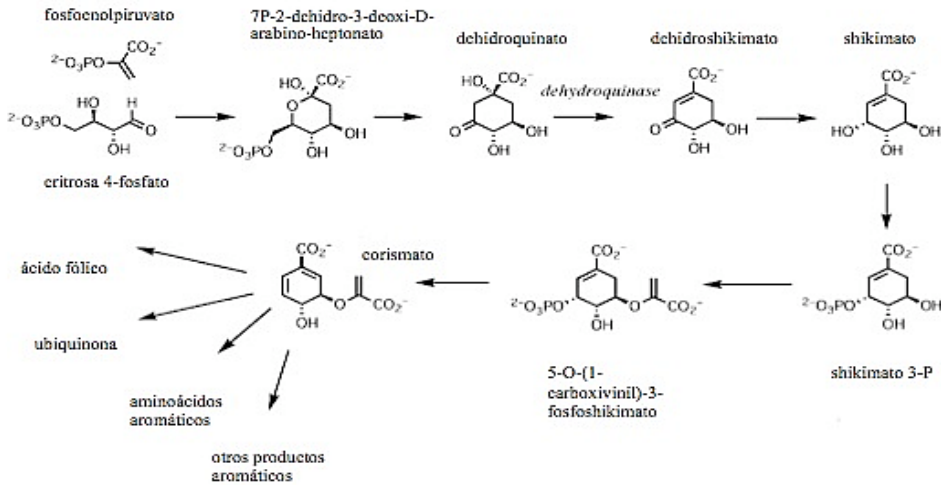


Figura 5. Representación de la vía del shikimato. Se observan los sustratos de la vía y el producto final (corimato) así como los compuestos que derivan de éste (Tomada de Espinosa, 2010.)

Varios organismos presentan más de una SDH (Sinhg, *et al.*, 2008). La vía del shikimato ha resultado de gran interés puesto que es esencial: en bacterias, arqueas, hongos, algas, plantas, parásitos del *Phylum* apicomplexa; pero no se encuentra en mamíferos (Roberts, *et al.*, 1998), por lo que puede ser blanco de insecticidas y fármacos no tóxicos para mamíferos (Michel, *et al.*, 2003 y Roberts, *et al.*, 1998).

La AroE de *E. coli* es la que da el nombre a este primer grupo. Este grupo de SDH pre-

presentan tres motivos característicos: SXS, NT y NTD. Estas enzimas se encuentran en las protobacterias. Estos motivos también se encuentran en homólogos de la AroE en los clados de las cianobacterias, actinobacterias y firmicutes lo cual es congruente con las relaciones establecidas por el 16SrRNA, sugiriendo que este tipo de enzimas ha evolucionado de manera vertical. Estas enzimas catalizan la oxidación del shikimato pero no del quinato, y unen preferentemente NADP(H). Una vez más, es importante recalcar que AroBs pertenece a este grupo (Singh, *et al.*, 2008).

El grupo de los homólogos de SdhL presentan los motivos NT y NTD, pero en vez del motivo SXS poseen uno (N/S/T) XG. La mayoría de los organismos que presentan una enzima de este grupo presentan más de una SDH (Singh, *et al.*, 2008).

El grupo de Rfi presenta los tres motivos pero de manera variable en cada subclado. Probablemente esto se deba a una diversificación funcional. Al igual que en el caso de las SdhL, todos los taxa que presentan una de las homólogas de Rfi también presentan más de una SDH que no pertenece al grupo de las Rfi (Singh, *et al.*, 2008).

Ael1 presenta los motivos (S/T)XS, N(A/C) y NFD, que difieren sustancialmente de los prototipos establecidos. Este grupo presenta el más amplio rango de especificidad por el sustrato. Las enzimas homólogas a Ael1 sólo se presentan en el linaje de las β -protobacterias (Singh, *et al.*, 2008).

YdiB presenta homólogos en varias bacterias entéricas además de E.coli. Este clado se agrupa dentro de un linaje de AroE Gram positivas (Firmicutes). La relación no es congruente con el árbol filogenético establecido a partir del 16S rRNA, por lo que se propone que el origen de estas enzimas fue en Firmicutes y mediante transferencia horizon-

tal fue adquirida por un ancestro de las bacterias entéricas que también la presentan. Estas enzimas contienen el motivo NT y NTD, pero el SXS puede ser SX (T/S). Todas las especies que presentan homólogos de YdiB también presentan más de una SDH no homóloga a YdiB.

Las enzimas homólogas a YdiB presentan un aspartato conservado en el motivo de unión al cofactor, indicando una preferencia sobre el NAD^+ , mientras que las AroE presentan una arginina que se relaciona con la especificidad sobre el NADP^+ (Singh *et al.*, 2008). YdiB de *E. coli* como la homóloga en *P. putida* cataliza la oxidación tanto de shikimato y quinato utilizando NAD (Michel *et al.*, 2003 y Singh, *et al.*, 2008).

Este trabajo se realizó con dos SDH con un 33% de identidad, YdiB de *E. coli* y AroBs de *B. subtilis*. YdiB es una proteína conformada por dos dominios con arquitectura α/β . El dominio N-terminal es discontinuo con un primer segmento (Met7- Thr 106) interrumpido por el dominio C-terminal (Gly124-Asp254) y completado por una cola correspondiente al dominio N-terminal (Gly 255- Phe 286). Los dominios están conectados por la hélice α_5 y un *linker* de 17 aminoácidos (Asp107- Lys123). El acomodo de estos dos dominios a lo largo de las hélices conectoras crea una hendidura profunda en donde se une el cofactor NADP^+ o NAD^+ gracias a residuos localizados en el dominio C-terminal, la reacción es más eficiente en presencia de NAD^+ (Figura 6) (Michel, *et al.*, 2003).

El dominio N-terminal pertenece a la superfamilia de las deshidrogenasas de aminoácidos. Consiste en seis hojas β y seis hélices α ; las hojas β acomodadas en un orden 213564 en donde la hoja β_5 es antiparalela al resto. Las tres primeras hojas β siguen una sucesión regular β / α con las hélices α_1 y α_2 paralelas a la hojas β . La siguiente unidad $\alpha / \beta / \alpha$

es irregular, en donde la hélice $\alpha 3$ está orientada a 45° en relación con la dirección de la hoja β y con la hélice $\alpha 4$ de sólo un giro prácticamente perpendicular a la hoja $\beta 4$. Este dominio es completado por las hélices $\alpha 9$ y $\alpha 10$, que se acomodan contra la hoja β del mismo lado que la $\alpha 1$ (Michel, *et al.*, 2003). Este dominio es capaz de unir tanto al shikimato como al quinato, por lo que esta proteína es considerada una shikimato quinato deshidrogenasa.

El dominio C-terminal adopta la forma de un dominio Rossmann canónico, es decir, seis hojas β paralelas con un orden 321456 y hélices α en ambos lados paralelas a las hojas β . La cuarta hélice α presente en los dominios Rossmann está ausente en YdiB, este dominio puede unir tanto NAD^+ como NADP^+ (Michel, *et al.*, 2003, Padyana y Burley, 2003 y Ye, *et al.*, 2003).

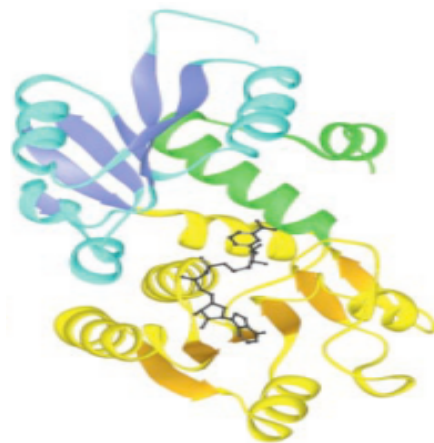


Figura 6. Representación de la SDH YdiB de *E. coli* en donde el dominio N-terminal (catalítico) es azul/cian, el dominio, C-terminal (Rossmann) amarillo/naranja, y la cola (N-terminal) verde. La unión de NAD^+ se muestra en negro en una representación de bolas y palos (tomada de Benach, *et al.*, 2003).

Al unirse el sustrato y el cofactor se promueve el paso de una conformación abierta a una cerrada, permitiendo la formación de un sitio activo. El movimiento es una rotación de aproximadamente 25° alrededor de un eje por el cual pasan los $C\alpha$ de la Lys32, y la Asp107, en donde la punta del dominio N-terminal, a la cual antes se denominó como la cola, recorre una distancia de 14\AA aproximadamente (Figura 7). Este cambio de conformación hace que cambien los puentes de hidrógeno e interacciones de Van der Walls entre los dominios. Ya que se formó el sitio activo (conformación cerrada), el C3 del 3-dihidrosikimato/3-dehidroquinato se posiciona muy cerca del C4 del anillo de nicotinamida para recibir el hidrógeno (Michel, *et al.*, 2003).

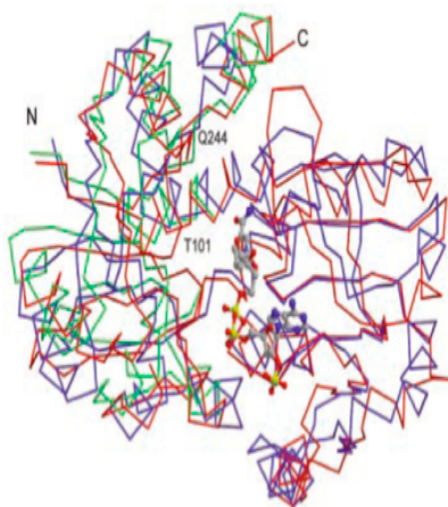


Figura 7. La estructura sobrepuesta de AroE y YdiB mostrando el esqueleto de $C\alpha$. YdiB de azul, en una conformación abierta; de rojo, AroE en una conformación cerrada, y de verde AroE en una conformación intermedia azul (Tomada de Michel, 2003).

Benach, *et al.*, en 2003 propusieron un mecanismo de reacción para YdiB en donde

plantean que, gracias al puente de hidrógeno entre la Asp107 y el grupo hidroxilo en el C5 del shikimato, la cadena ramificada actúa como el aceptor de un protón de este grupo hidroxilo cuando el hidruro es transferido hacia el NAD⁺ (de derecha a izquierda). Cuando la reacción va en sentido reverso (de izquierda a derecha) se requiere que el aspartato esté protonado antes de unirse al 3-dehidroshikimato, lo que será promovido por la carga negativa en el NADH adyacente (Benach, *et al.*, 2003), puede verse en la figura 8.

Investigaciones más recientes como la de Gan, *et al.* en 2007, proponen que el mecanismo de reacción utilizado por la SDH de *Aquifex aeolicus* incluye la transferencia de hidruros mediada por una molécula de agua unida mediante puentes de hidrógeno a la Asp107 y la Lys71, en donde una cadena de transferencia de hidruros resulta en la desprotonación del grupo hidroxilo del C3 del 3-dehidroshikimato, y el protón transferido de la Lys71 a la molécula de agua es facilitado por la asp107.

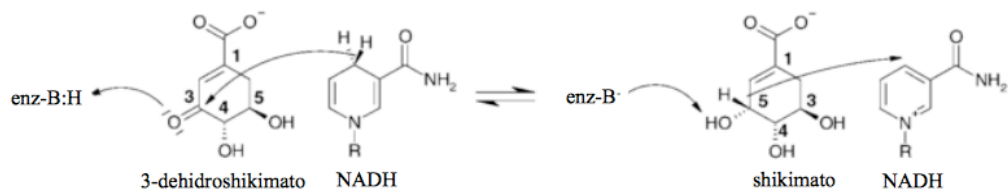


Figura 8. Mecanismo de reacción para la reducción del 3-dehidroshikimato a shikimato. La shikimato 5-dehidrogenasa cataliza la reacción de reducción del 3-dehidroshikimato a shikimato utilizando como cofactor NADH (tomada de Benach, *et al.*, 2003).

A pesar de tratarse de una proteína diferente, los residuos que se encuentran conservados en las SDH y que se plantea que son los involucrados en el mecanismo de reacción, también se encuentran en las proteínas con que se realizó este trabajo (es importante denotar que la Lys71, así como las Asp107 son los residuos de YdiB)

La segunda enzima que se utilizó en este trabajo es la SDH de *B. subtilis*. Esta enzima está codificada por el gen *aroD*, sin embargo, a lo largo de este trabajo se la denomina como *aroBs*. La enzima cuenta con 280 residuos de aminoácidos y a pesar de no ser conocida su estructura cristalográfica, en su secuencia se pueden reconocer los dos dominios que la componen: el dominio catalítico y el Rossman (Padyana y Burley, 2003). AroBs presenta los residuos involucrados en la unión al sustrato y al cofactor que se encuentran conservados en otras SDH (Padyana y Burley, 2003 y Ye *et al.*, 2003). Mediante un modelo realizado por Espinosa en 2010, se estableció que el dominio N-terminal (de unión a sustrato) es interrumpido por el dominio C-terminal al igual que YdiB. El dominio de unión a sustrato presenta seis hojas- β paralelas con un orden 213564 y seis hélices α , la β 5 antiparalela con respecto a las demás (Figura 9). Las tres primeras hojas β siguen una sucesión $\beta\alpha$ regular, con α 1 y α 2 paralelas a las hojas β . La siguiente subunidad $\alpha\beta\alpha$ es irregular debido a que la hélice α 3 está orientada 45° relativa a la orientación de las hojas β . Le sigue la hoja β 4, que es muy pequeña. El dominio es completado por un par de α hélices (α 9 y α 10) denominadas como la cola formando una horquilla que empaqueta a las hojas β del dominio N-terminal del mismo lado que α 1. El dominio C-terminal es un Rossman dependiente de NADP(H) canónico con seis hojas β paralelas en orden 321456. Las hélices α de la β 3 (β 9) y β 4 (β 10) están sustituidas por *loops*. Ambos dominios están unidos por un *linker* compuesto de la α 5 que cubre las hojas β del dominio C-terminal. El sitio activo, como es característico de las enzimas del grupo AroE, se espera que se encuentre en la interface entre los dominios (Figura 9).

Al comparar la secuencia de AroBs con tres proteínas con alto porcentaje de identidad y que se sabe que forman dímeros como unidad funcional: 2EGG A, AroE de *Geobacillus kaustophilus* (45% identidad); 1NVT A, AroE de *Methanococcus jannaschii* (40% identidad);

y YdiB de *E. coli*. (33% de identidad), se observó que no comparte la mayoría de los residuos involucrados en la formación de los dímeros de las tres estructuras.

Los dímeros de las tres estructuras se forman por residuos en el dominio N-terminal en los elementos: $\beta 1$, $\beta \alpha 2$ de 2EGG, en la $\beta \alpha 1$ y $\beta \alpha 2$ de 1NVT y en $\beta 1$ y $\beta \alpha 2$ en YdiB. A pesar de lo anterior en AroBs se puede observar un parche hidrofóbico similar al que estas enzimas presentan.

Por otra parte, cinco de los seis residuos que Michel, *et al.*, 2003 proponen como los determinantes para la dimerización de YdiB se observan en AroBs. Este residuo que no se conserva es la M40, y está sustituida por H35. La H35 la comparte con 2EGG, en donde se propone que forma puentes de hidrógeno en la interfaz del dímero, por lo que Espinosa (2010) propone que AroBS se encuentre como dímero en su estado funcional.

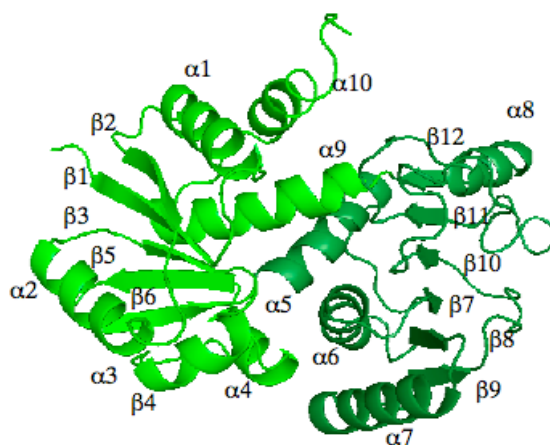


Figura 9. Modelo de la estructura de AroE de *B. subtilis*. El verde claro se presenta el dominio N-terminal o de unión a sustrato ($\beta 1$ - $\beta 6$; $\alpha 1$ - $\alpha 4$ y $\alpha 9$ - $\alpha 10$). En verde oscuro el conector ($\alpha 5$) y el dominio C-terminal o Rossman ($\beta 7$ - $\beta 12$; $\alpha 6$ - $\alpha 8$). Imágenes realizadas en PyMOL (Tomado de Espinosa, 2010).

HIPÓTESIS

El intercambio de dominios entre la enzimas shikimato deshidrogenas AroE de *B. subtilis* y la quinato shikimato deshidrogenasa YdiB de *E. coli* resultará en al menos una quimera activa.

OBJETIVOS

Estudiar la intercambiabilidad entre los dominios de las proteínas homólogas shikimato deshidrogenasa AroE de *B. subtilis* y quinato shikimato deshidrogenasa YdiB de *E. coli* desde el punto de vista estructural y funcional.

OBJETIVOS PARTICULARES

- Generación de cuatro quimeras provenientes del intercambio de dominios entre AroE de *B. subtilis* y YdiB de *E. coli*, tomando en cuenta diferentes consideraciones estructurales.
- Observar la expresión de las quimeras mediante SDS PAGE y Western blot.
- Observar la capacidad de plegamiento de cada quimera fusionada a la CAT por medio de ensayos de concentración mínima inhibitoria (CMI),
- Analizar la actividad biológica de cada quimera mediante ensayos de complementación de la cepa de *E. coli* Δ aroE Km⁻

MATERIALES Y MÉTODOS

CONSTRUCCIÓN DE QUIMERAS CON UN SITIO DE RECOMBINACIÓN

Las quimeras fueron generadas por el intercambio de dominios entre los genes de la shikimato deshidrogenasa AroBs de *B. subtilis* y quinato shikimato deshidrogenasa YdiB de *E. coli*. En total, se construyeron cuatro quimeras. Las dos primeras se realizaron considerando el extremo C-terminal como parte del dominio Rossman de la proteína nativa. Las segundas dos se realizaron considerando al dominio Rossman insertado dentro del dominio catalítico, por lo que el extremo C-terminal ($\alpha 9$ y $\alpha 10$) pertenece a la secuencia de la proteína parental de la que proviene el dominio catalítico (Figura 10).

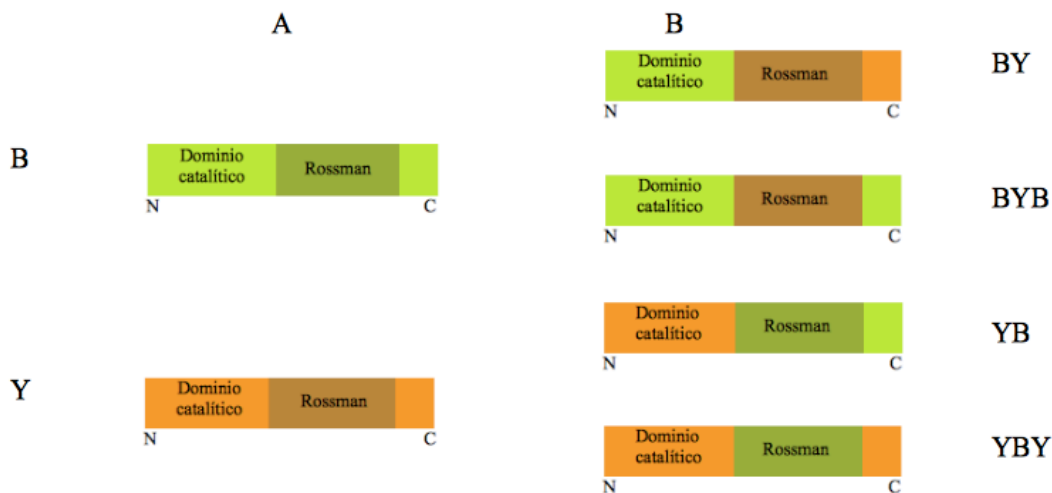


Figura 10. A) Representación de la estructuras, superior *aroBs* (B) e inferior *ydiB* (Y). B) Representación de secuencias de las cuatro quimeras construidas y analizadas en este trabajo.

Vector utilizado

El vector utilizado para todas las construcciones fue el pT4K12M4 previamente modificado por Escamilla (datos no publicados). Este vector presenta un origen de replicación de alto número de copias, un promotor pTrc fuerte, y un cassette de resistencia a kanamicina (Km). Tiene la cualidad de ser inducible por IPTG, sin embargo, el nivel de transcripción basal es suficiente para la expresión de las proteínas bajo este promotor.

pT4_BY_CAT

La quimera pT4_BY_CAT fue construida por Banda (2011)

pT4_YB_CAT

La construcción de la que se partió para generar la primera quimera pT4_YB_CAT fue la pT4_YE_CAT (Chenge, 2010), la cual se utilizó para amplificar el dominio N-terminal de YdiB sin incluir la cola ($\alpha 9$ y $\alpha 10$). Secuencia que se clonó en el vector de pT4_EB_CAT construido por Espinoza en 2010. Tanto el inserto, como el vector, fueron digeridos por las enzimas *NcoI* y *BstEII*.

Amplificación por PCR

La reacción de PCR para la amplificación del dominio N-terminal correspondiente a YdiB se realizó utilizando como DNA molde la construcción YE_CAT obteniendo un volumen final de 50 μ l con los reactivos de la siguiente manera:

-36 μ l agua tetradestilada,

-5 μ l *buffer* Pfu polimerasa (10X) de la marca *Roche*

-5 μ l dNTP's

-1 μ l oligo pT4

-1 μ l oligo Y1-BstEIIIR

El termociclador fue programado de la siguiente manera:

1 ciclo a: 92°C por 3 minutos

25 ciclos de: 95°C por 30 segundos

59°C por 30 segundos

72°C por 1 minuto

1 ciclo a : 72°C por 5 minutos

Purificación de fragmentos de PCR

La purificación de fragmentos de PCR se realizó utilizando el *High Pure Plasmid Isolation Kit (Roche)* siguiendo el instructivo, el cual consta de los siguiente pasos: ajustar el volumen total de la reacción de PCR a 100 μ l. Agregar 500 μ l de *Bindig Buffer* a cada 100 μ l de reacción de PCR y agitar bien. Insertar un filtro en un tubo recolector y transferir la mezcla al filtro y centrifugar 60 segundos a 14000 rpm. Desconectar el filtro y desechar el líquido del tubo recolector. Volver a colocar el filtro dentro del tubo recolector. Agregar al filtro 500 μ l de *Washing Buffer* y centrifugar 60 segundos a 14000 rpm. Desconectar el filtro y desechar el líquido del tubo recolector. Volver a colocar el filtro dentro del tubo recolector y agregar 200 μ l de *Washing Buffer* al filtro y centrifugar 60 segundos a 14000 rpm. Desconectar el filtro y desechar el líquido del tubo recolector. Conectar el filtro

ahora en un tubo para microcentrífuga de 1.5 ml. Agregar al filtro 50 µl de agua (en lugar de *Elution Buffer*) y centrifugar 60 segundos.

Digestión

Para generar extremos cohesivos del fragmento N-terminal correspondiente a YdiB obtenido mediante PCR, se realizó la reacción en un volumen final de 50 µl con los siguientes reactivos y en las siguientes cantidades, los cuales fueron de la marca *New England Biolabs(NEB)*:

- 30 µl DNA Molde (dominio N-terminal de YdiB)
- 5 µl *Buffer 3* (10X)
- 1 µl *NcoI*
- 2 µl *BsteII*
- 0.5 µl BSA (100X)
- 11.5 µl de agua tetradestilada

La reacción se incubó a 37°C durante 2 horas.

Ligación

La reacción de ligación se realizó con una relación molar de vector : inserto de 1: 3 respectivamente. Para esto fue necesario cuantificar la concentración tanto del vector, como del inserto, en este caso obtenida por *nanodrop* con el fin de agregar 150 ng de vector. Las reacciones se realizaron en un volumen final de 20 a 25 µl y fueron incubadas durante 1 h a temperatura ambiente. La mezcla contenía:

- x µl vector (150ng)
- x µl inserto (cumpliendo relación 1:3)

- x μ l agua tetradestilada
- 1 μ l de DNA ligasa T4 (*Fermentas*)
- 1x *Buffer* de DNA ligasa T4

Células DMSO competentes

Se dejó creciendo un cultivo de células durante 16 h (*over night* ON) a 30°C y 250 rpm en medio LB. Posteriormente, se diluyó el cultivo en 50 ml (dilución 1/100) de LB fresco en un matraz de 200 ml en donde se dejó crecer el cultivo diluido a 30°C y 250 rpm hasta una densidad óptica de 0.2 a 0.5 a 600 nm. El cultivo fue colocado en un tubo Falcon y fue centrifugado durante 10 minutos a 3000 rpm y a 4°C. Se decantó el sobrenadante y con una pipeta se extrajo el medio remanente. El Pellet fue resuspendido con 5 ml de *buffer* TSS (5g PEG 8000, 1.5 ml 1M MgCl₂, 2.5 ml DMSO y LB hasta 50 ml) en frío. Las alícuotas fueron de 100 μ l en tubos de microcentrífuga de 0.5 ml.

Transformación

Se transformaron 5 μ l producto de la reacción de ligación o 1 μ l de plásmido superenrollado en 100 μ l de células DMSO competentes de las cepas Δ *aroE Km^r* (de la colección keio de E.coli K-12 donada por el laboratorio del Dr. Gosset, IBT, y posteriormente modificada por la Mtra. Espinosa para la eliminación del cassette de kanamicina), JM101 ó MC6110. Las células transformadas fueron recuperadas en 1 ml de medio LB durante una hora a 30°C. Posteriormente, se centrifugaron a 9000 rpm durante 1h para sólo conservar el pellet, el cual fue resuspendido y finalmente las células se sembraron en cajas de LB con 25 μ g/ml de Kanamicina y en cajas de LB con 25 μ g/ml de cloranfenicol, que fueron incubadas a 37°C ON.

PCR de colonia y PCR de confirmación

En el caso de la PCR de confirmación se utilizó plásmido purificado. En el caso de la PCR de colonia se tomó con un palillo un poco de cada colonia (5 colonias) de la caja de LB sólido en donde se sembraron las transformaciones. Este fragmento de colonia se colocó en las paredes de un tubo de 200 μl y se resuspendió en 10 μl de agua tetrades-tilada. Estas células suspendidas en agua fueron hervidas a 95°C. Posteriormente, se agregaron los siguientes componentes dando como resultado un volumen total de 50 μl (10 μl para cada muestra):

- 5 μl *Buffer* taq
- 5 μl MgCl^2
- 5 μl μl *buffer* taq polimerasa (10X) de la marca *Roche*
- 5 μl dNTP's
- 3 μl oligo pT4
- 3 μl oligo reverso (cat100 ó pt42)
- 24 μl agua tetrades-tilada

El termociclador fue programado de la siguiente manera:

- 1 ciclo a: 95°C por 4 minutos

- 25 ciclos de: 95°C por 30 segundos
59°C por 30 segundos
72°C por 30 segundos

- 1 ciclo a : 72°C por 5 minutos

Secuenciación

La secuenciación fue realizada por la unidad de secuenciación del Instituto de Biotecnología de la Universidad Nacional Autónoma de México (UNAM). Para las construcciones con fusión traduccional a CAT se utilizaron los oligos pT4 (*fwd*) y cat100 (*rev*). En el caso de las construcciones con cola de histidinas el oligo reverso utilizado fue pT42.

CONSTRUCCIÓN DE QUIMERAS DE DOS SITIOS DE RECOMBINACIÓN

PCR para generar pT4_YBY_CAT y pT4_BYB_CAT

Se realizó una primera PCR utilizando como molde la construcción pT4_BY_CAT, un oligo *forward* (DIIYdB_cBs) con una secuencia de 23pb en el extremo 3' homóloga al inicio de la secuencia de la cola del DNA molde correspondiente a YdiB. El extremo 5' no era complementario al templado, sino a los últimos 21pb del dominio Rossman de Bs presente en la otra quimera (pT4_YB_CAT). El oligo reverso (cat100) es complementario a un sitio 100 pb río abajo del final de la secuencia quimérica, de tal manera que el fragmento amplificado consistió en el final del dominio Rossman correspondiente a Bs, toda la secuencia correspondiente a la cola de YdiB, y un fragmento de la CAT, ver Figura 11.



Figura 11. Esquema representativo de la primera PCR utilizado para generar la quimera BYB_CAT, la metodología es similar para la quimera YBY_CAT (otro molde y oligos).

Para construir pT4_BYB_CAT se utilizó la misma técnica pero naturalmente con otro oligo *forward* (DIIYdB_cBs). En este caso, el extremo 5' correspondía al dominio Rossman de YdiB y el 3' a la cola de Bs.

El producto de PCR fue un fragmento de 300pb aproximadamente, en donde el extremo 5' correspondía a los últimos 21 pb del dominio Rossman de la otra quimera y la cola homóloga al molde utilizado. Esta primera PCR se realizó en un volumen total de 50 μ l que contenía:

- 33 μ l agua tetradestilada,
- 5 μ l *buffer* Pfu polimerasa (10X) de la marca Roche
- 5 μ l dNTP's
- 1 μ l oligo *fwd* (DIIBs_colaY)
- 1 μ l oligo *rev* (CAT 100)

- 2 μ l DNA molde diluido 1:10 (pT4_YB_CAT)
- 1 μ l Pfu polimerasa de la marca *Roche*

El termociclador fue programado de la siguiente manera:

1 ciclo a: 92°C por 3 minutos

25 ciclos de: 95°C por 30 segundos
59°C por 30 segundos
72°C por 1 minuto

1 ciclo a : 72°C por 5 minutos

El producto de estas PCR se utilizó como oligo reverso (*megaprimer*), y ahora el molde, a diferencia de la primera PCR, fue la quimera contraria, de tal forma que la secuencia que hibridó del megaprimer fue la correspondiente al dominio Rossman. Como *forward* se utilizó el oligo pT4 que amplificó río arriba del promotor bajo el cual está el inicio del gen quimérico. Esta reacción dio como resultado el cambio de cola de las quimeras de un sitio de recombinación, ver Figura 12. La reacción se realizó en un volumen final de 50 μ l adicionando lo siguiente:

- x μ l agua tetradestilada,
- 5 μ l *buffer* Pfu polimerasa (10X) de la marca *Roche*
- 5 μ l dNTP's
- 1 μ l oligo *fwd* (pT4)
- x μ l oligo rev (megaprimer) concentración final 300ng/ μ l
- 1 μ l DNA molde (pT4_YB_CAT)

La Pfu polimerasa se adicionó en la primera repetición de la PCR en el paso del *annealig* y el oligo pT4 en el mismo paso pero de la sexta repetición en las siguientes cantidades:

- 0.5 μ l Pfu polimerasa de la marca *Roche*
- 2 μ l oligo *fixed* (pT4)

El termociclador fue programado de la siguiente manera:

- 1 ciclo a: 95°C por 3 min 35 ciclos de: 95°C por 50 segundos
 58°C por 2 minutos
 72°C por 1.5 minutos
- 1 ciclo a: 72°C por 10 minutos

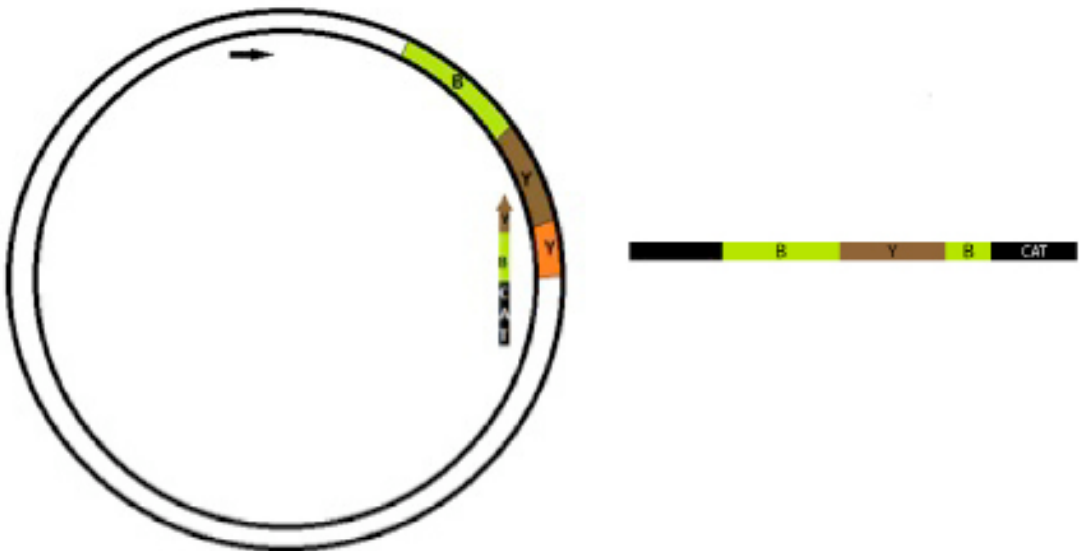


Figura 12. Esquema representativo de la segunda PCR utilizado para generar la quimera BYB_CAT, la metododología es similar para la quimera YBY_CAT (otro molde y oligos).

El producto de estas PCR se digirió con las enzimas *NcoI* y *BamHI* (de la marca *NEB*) y se clonó el vector pT4_YBY_CAT digerido del mismo modo que se explicó con anterioridad y utilizando la ligasa T4 de la marca *Fermentas*. Para la verificación se realizó una PCR de confirmación y posteriormente se mandaron a secuenciar.

Tabla 1.Oligonucleótidos

Nombre	Número/año	Tm	Secuencia (5'→3')
YI_BstEIIIR	232/2010	78	AAA GGT AAC CAC GCA GAT AGC CAT CAT
pT4	237/2004	60	GACATATAAACGGTTCTGGCA
pT42	238/2004	60	GCCAGTGAATTGTAATACGAC
DIIBs_cYdB	839/2010	126	CAAAAAGGCTTGAAAACACTCGATGGATACGGCATGTTGTTGTG
DIYdB_cBs	838/2010	126	AAGCTGGTTGCAAAACGATTGATGGTGTGGGCATGTTTGTGAA
Cat100	6497/2009	60	CTGAACGGTCTGGGTATAGG
aroE-47	5215/2009	60	AATCCGCGATGCCCTGAC
aroE+29R	5216/2009	62	TTCCCTGTCCGGAAACTGGAT

CONCENTRACIÓN MÍNIMA INHIBITORIA (CMI)

Este experimento permite calcular la concentración de un antibiótico (en este caso cloranfenicol) a la cual, la mitad de las clonas son resistentes. Las construcciones que se realizaron (como se especificó en los últimos dos apartados) tienen una fusión traduccional a CAT. La CAT funciona como un reportero de plegamiento bajo el principio de: que si

la proteína que está río arriba de la CAT se pliega, permitirá que la CAT se pliegue. De manera contraria, si no logra plegarse, impedirá que la CAT se pliegue. De tal manera que entre mejor sea el plegamiento de la CAT, mayor cantidad de proteína soluble se encontrará, al igual que la quimera que se encuentra unida a ésta. Así que, entre más resistencia a cloranfenicol tenga una clona, mayor proteína soluble tendrá.

Las cepas que se utilizaron para este ensayo fueron las JM101 y MC1061. El control negativo utilizado fue el plásmido pT4 (vacío). Los controles positivos fueron los vectores en donde estaban clonados los genes silvestres pT4_B_CAT y pT4_Y_CAT.

Se realizaron diluciones de los cultivos de cada cepa incubados en LB con 25 µg/ml de Km a 37°C por 16 h ON a 250 rpm. Se realizó una dilución 1/20 en agua de todos los cultivos. Se midió la concentración de cada uno a 600nm. Se hicieron diluciones seriales en LB de 10^{-1} a 10^{-7} de los cultivos en una placa de 96 pozos. En el primer pozo se inoculó la misma cantidad de células a partir del cual se realizaron las diluciones seriales. Se sembraron 5 µl de las diluciones correspondientes de 10^{-7} hasta la 10^{-3} en cajas de LB con concentraciones crecientes de Cm (100, 150, 200, 250, 300 µg/ml), y las cajas fueron incubadas ON a 37°C.

Tabla 2. Cepas utilizadas

JM101	<i>F traD36 proAB lacI Δ(lacZ)M15/ Δ(lac-proAB) glnV thi</i>
MC1061	<i>F⁻ araD139 Δ(ara-leu)7696 galE15 galK16 Δ(lac)X74 rpsL (Strr) hsdR2(rk - mk+) mcrA mcrB1</i>

BW25113 Δ aroE: *lacl_{grn}B_{T14} ΔlacZ_{W316} hsdR514 ΔaraBAD_{AH33} ΔrhaBAD_{LD78}* (*keiocollection*)

Eliminación de fusión traduccional a CAT

En el caso de pT4_BY_CAT y pT4_YBY_CAT para eliminar la fusión a CAT, se hizo una digestión con las enzimas *BamHI* y *BglII* en una reacción de 30 ml incubada durante una hora a 37°C con los siguientes componentes y en las siguientes cantidades:

- 20 µl DNA
- 1 µl *BamHI*
- 1 µl *BglII*
- 3 µl Buff 3 (*NEB*)
- 0.5 µl BSA

Los productos de estas digestiones fueron purificados con el kit de purificación de PCR (*Roche*) y posteriormente ligados con la ligasa T4 (*Fermentas*) durante una hora a temperatura ambiente.

En el caso de pT4_YB_CAT y pT4_BYB_CAT no se pudo realizar el procedimiento anterior puesto que existe un sitio *BglII* natural en la cola correspondiente a Bs, por lo que se digirió el inserto con las enzimas *NcoI* y *BamHI* adicionando 0.5 µl de BSA. Mientras que el vector fue digerido con *NcoI* y *BglII*, el sitio *BamHI* y *BglII*, al ser digeridos por sus respectivas enzimas, dejaron extremos complementarios entre sí. Ambas reacciones se realizaron en un volumen total de 30 µl con los siguientes componentes:

- 20 µl DNA

-1 μ l *NcoI*

-1 μ l BanHI o *BglII*

-3 μ l *Buffer 3 (NEB)*

Una vez que tanto el inserto como el vector fueron digeridos, se purificaron ambos a partir de una banda de gel de agarosa utilizando el kit de purificación de PCR (*Roche*), y posteriormente fueron ligados con las ligasa T4 (*Fermentas*) durante una hora a temperatura ambiente. Todas las ligaciones se transformaron en células DMSO competentes de la cepa Δ aroE.

SDS-PAGE

Una colonia de las células MC6110, transformada con una de las cuatro respectivas construcciones (pT4_BY_CAT, pT4_YB_CAT, pT4_BYB_CAT y pT4_YBY_CAT), fue inoculada en 3 ml de LB con Km 25 μ g/ml. Posteriormente, se incubaron por 16 horas ON a 37°C a 250 rpm. Una vez que se llegó a la saturación se centrifugaron los 3ml de cultivo en un tubo *Eppendorf* a 9,000 rpm a temperatura ambiente por un minuto con la finalidad de eliminar el sobrenadante y resuspender la pastilla en 500 μ l de PBS 1X y PMSF 100mM. Las muestras se sonicaron por un minuto cinco veces. Posteriormente, se centrifugaron a 4,000 rpm a 4°C por 20 minutos. El sobrenadante fue transferido a un tubo nuevo, pues éste corresponde a la fase soluble de la muestra, mientras que la pastilla corresponde a la fase insoluble. La fase insoluble fue resuspendida en 100 μ l de *buffer* con SDS 1%. A ambas fracciones se les adicionó *buffer* de carga y se hirvieron durante 10 minutos a 95°C. Las muestras se corrieron en un gel desnaturizante de acrilamida al 12.5% compuesto de un gel concentrador (1.25ml de UGS, 0.75ml de acrilamida 30%, 3ml de agua tetradesilada, 15 μ l de APS 10% y 5 μ l de TEMED) y un gel separador (2ml de LGS, 3.35ml de acrilamida 30%, 2.65ml de agua tetradesilada, 40 μ l de APS 10%

y 20 μ l de TEMED). Las muestras se corrieron a 4 mA durante 16 h. Para observar las muestras se realizaron tres lavados de 10 minutos en agua y se incubó el gel con azul de Coomassie (*BIO RAD*) durante 1 h a 56 rpm. Después fue desteñido con agua y papel absorbente en agitación a 56 rpm hasta que se hicieran visibles las bandas.

WESTERN BLOT

Se corrió un gel de acrilamida al 12.5% con las muestras igualmente preparadas que para el gel de SDS-PAGE. Se cortaron tres rectángulos de papel filtro gruesos y una membrana de nitrocelulosa del tamaño del gel. Estos fueron remojados en distintas soluciones y acomodados en orden en la cámara de transferencia semiseca. El orden de abajo hacia arriba fue: primer filtro con la solución ánodo I, segundo filtro con la solución ánodo II, la membrana de nitrocelulosa remojada en solución cátodo, el gel, y el tercer filtro con la solución cátodo, para transferirlo en una cámara de transferencia semiseca.

Una vez cerrada la cámara, se colocó un peso sobre ella y se le aplicó por una hora y media una corriente de 1.8 mA por cada cm² del gel. Posteriormente, la membrana de nitrocelulosa se bloqueó durante 1 h en una solución con TBST 1X y 5% de leche en polvo. Transcurrido este tiempo, se realizaron tres lavados de 10 minutos con TBST y finalmente se agregó el anticuerpo anti-CAT o anti-His, según el caso en una relación 1:3,000 en solución de TBST con 0.1% de leche en polvo, y se incubó durante una hora. La membrana se sometió a tres lavados de 10 ml con TBST.

En el caso de las construcciones fusionadas a CAT fue necesario agregar el segundo anticuerpo Anti-DIG en una dilución 1:5,000, y se incubó por 1 hora en TBST con 0.1% de leche y los tres posteriores lavados con TBST. Para revelar el último, se agregó una

solución con 1ml de BCIP y 1ml de NTB (ambos marca *Zimed*) aforando a 10 ml con agua mQ. Posteriormente, se dejó la solución hasta que se visualizaron las bandas, y el gel se enjuagó con agua antes de que se sobrerivelara. En el caso de las construcciones con cola de histidinas, el revelado se llevó a cabo aplicando 2 ml de Tetrametil Benzidina (TMB) (marca *Invitrogen*), esperando a que se visualizaran las bandas y enjuagaran con agua mQ.

COMPLEMENTACIÓN

La cepa $\Delta\text{aroE Km}^-$ fue transformada con los vectores pT4_B_HIS, pT4_Y_HIS, pT4_BY_HIS, pT4_YB_HIS, pT4_BYB_HIS y pT4_YBY_HIS; y con las construcciones, pero con la fusión traduccional a CAT. También la cepa se transformó con el vector pT4 vacío para ser utilizado como control negativo. Se tomó una colonia de cada construcción y se las creció en 3 ml de LB con ON a 30°C y 250 rpm. De los cultivos saturados se centrifugaron 400 μl durante 2 h a 5,000 rpm y se lavaron 3 veces con sales M9 1X centrifugando igual que la primera vez. Después de la última centrifugación se resuspendió el pellet en 400 μl de las sales M9 1X. Se inocularon 20 μl de las células lavadas en dos tipos de medios. Ambos contenían medio mínimo adicionado con 100 μl de kanamicina y vitaminas (16 μl de ácido p-aminobenzoico 62mg/ml, 200 μl de ácido 4-hidroxibenzoico 48 mg/ml 29 μl de ácido 2-3 dihidroxibenzoico 35mg/ml y 10 μl de tiamina 10 mg/ml), por lo que se prepararon juntos. Antes de aforar se separó la mitad de la mezcla y a una de las mitades se le adicionaron aminoácidos aromáticos (120 μl de Trp 4mg/ml, 200 μl de Phe 8mg/ml y 260 μl de Tyr 6.15 mg/ml) para que funcionaran como control positivo. Ambas mezclas se aforaron con agua tetradestilada estéril hasta 50 ml. Estos cultivos se mantuvieron a 30 °C y su crecimiento se siguió por densidad óptica a 600nm.

PCR de residuo de cassette de Km de cepa $\Delta aroE$ Km⁻

Para confirmar que las cepas fueran $\Delta aroE$ Km⁻ se realizó una PCR:

- 5 μ l MgCl²
- 5 μ l μ l *buffer* taq polimerasa (10X) de la marca *Roche*
- 5 μ l dNTP's
- 3 μ l oligo *fwd* (aroE-47)
- 3 μ l oligo *rev* (aroE +29R)
- 24 μ l agua tetradestilada

El termociclador fue programado de la siguiente manera:

1 ciclo a: 95°C por 4 minutos

25 ciclos de: 95°C por 30 segundos

59°C por 30 segundos

72°C por 30 segundos

1 ciclo a : 72°C por 5 minutos

RESULTADOS

CONSTRUCCION DE ENZIMAS QUIMÉRICAS

Se construyeron cuatro quimeras (Figura 13). Las dos primeras considerando el extremo N-terminal de una de las proteínas parentales y el C-terminal de la otra. A estas primeras se les denominó BY_CAT (Banda 2011) y YB_CAT, respectivamente. La segunda se construyó a partir de dos quimeras ya existentes en el laboratorio de la siguiente manera:

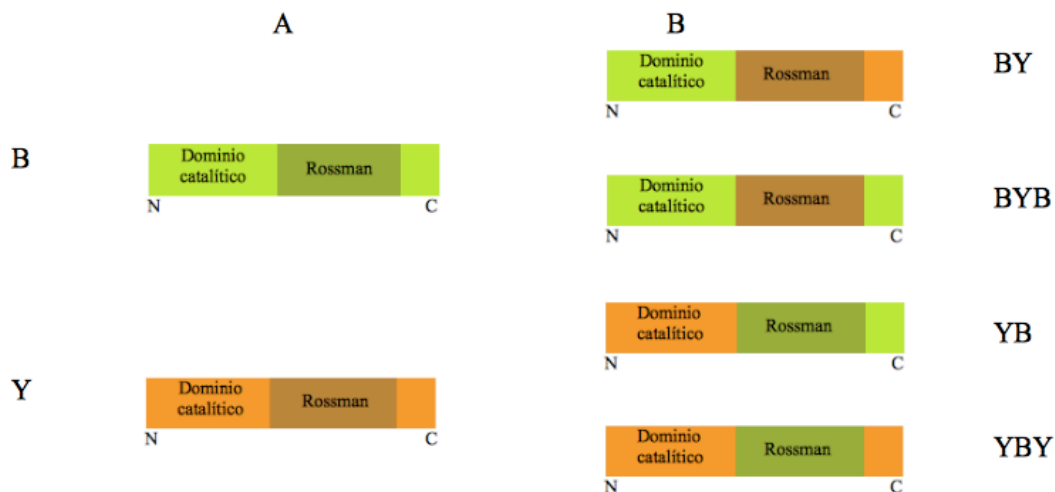


Figura 13. A) Representación de la estructuras, superior *aroBs* (B) e inferior *ydiB* (Y). B) Representación de secuencias de las cuatro quimeras construidas y analizadas en este trabajo.

Se amplificó el dominio N-terminal de YdiB utilizando como molde la construcción YE_CAT (Chenge, 2010). La reacción de PCR se especifica en la sección de materiales y métodos. Se aprovechó la quimera previamente construida (EB_CAT) por Espinosa en 2010, ya que contaba con la secuencia correspondiente al dominio C-terminal de la proteína AroBs.

Sobre esta quimera se realizó una digestión para extraer el fragmento correspondiente al dominio N-terminal (secuencia correspondiente a *aroE*). De esta forma, se pudo sustituir por el fragmento N-terminal de YdiB previamente amplificado. Tanto el vector como el inserto (producto de PCR) fueron digeridos con las mismas enzimas (Materiales y métodos).

Las quimeras BYB_CAT y YBY_CAT se distinguen por considerar al dominio Rossman dentro del dominio catalítico, es decir, que el extremo C-terminal de éstas corresponde al dominio catalítico. En este caso se debe realizar un cambio de las hélices $\alpha 9$ y $\alpha 10$ (de 40 y 36 aminoácidos), respectivamente, del extremo C-terminal de las quimeras BY y YB para que este fragmento ($\alpha 9$ y $\alpha 10$) pertenezca a la misma proteína a la que pertenece el dominio N-terminal.

Esto se logró mediante dos PCRs. La primera, amplificó la secuencia correspondiente a los aminoácidos del extremo C-terminal de BY y YB, por lo que las quimeras se utilizaron como molde. La particularidad de esta primera PCR reside en los respectivos oligos *forward* utilizados, pues fueron de carácter híbrido, es decir que, el extremo 5' corresponde al dominio Rossman de la quimera a la que se le quiere hacer el cambio de cola (secuencia que no corresponde al Rossman del molde utilizado), y el extremo 3' corresponde al inicio de la secuencia de las hélices $\alpha 9$ y $\alpha 10$ que se quiere sustituir (secuencia correspondiente al molde utilizado).

En la segunda PCR se utilizó como oligo *forward* el producto de la primera PCR y los moldes se intercambiaron. El producto de la amplificación a partir de YB en la PCR1 se utilizó sobre BY y viceversa, lo cual permitió la sustitución de los aproximadamente 300 pb del extremo 3' del molde utilizado (el procedimiento se describe en la sección de materiales y métodos).

El oligo *forward* que se utilizó para la primera PCR para generar BYB tenía una deleción puntual, sin embargo, no fue hasta mucho después que fue conocida, por lo que en los experimentos que siguieron debe de tomarse en cuenta que BYB en realidad no fue BYB, ya que la deleción produjo un cambio en el marco de lectura. A esta “quimera” se le denominó BYBmut. Esta construcción se volvió a hacer de manera correcta pero no ha sido analizada.

CONCENTRACIÓN MÍNIMA INHIBITORIA

Las construcciones clonadas fueron diseñadas para no terminar en un codón de paro después del último codón codificante con el fin de generar una fusión traduccional con la proteína cloranfenicol acetil transferasa. Esta enzima forma un homotrímero que es capaz de acetilar al Cm, lo que produce resistencia a este antibiótico.

La CAT se utilizó como reportero de plegamiento debido a que, si la proteína que está río arriba de la CAT se pliega, permitirá que la CAT se pliegue. De manera contraria, si no logra plegarse, impedirá que la CAT se pliegue. Por lo tanto, entre mejor sea el plegamiento de la CAT, mayor cantidad de la fusión de proteínas de manera soluble se encontrará.

Todas las construcciones y los controles fueron clonados en primera instancia en la cepa JM101 y posteriormente en la cepa MC1061. El vector vacío se uso como control negativo y se nombró pT4. Como controles positivos se utilizaron las proteínas silvestres parentales (AroE, AroBs y Ydib) fusionadas a la CAT. Se utilizaron concentraciones de Cm de 0 a 300 µg/ml.

La quimera con mayor resistencia fue YB (resistencia mayor a 200 $\mu\text{g/ml}$) seguida de BY y YBY (150 $\mu\text{g/ml}$). Por último, la que menor resistencia presentó fue BYBmut (cuya resistencia es menor que 100 $\mu\text{g/ml}$ de Cm) debido a que en realidad no se trata de BYB por la delección que presentaba.

Se muestran los resultados de dos ensayos en la tabla 3 en donde cada columna corresponde a ensayos diferentes. El primero, la columna de la izquierda, fue realizado con la cepa JM101 en donde la concentración máxima de Cm fue de 200 $\mu\text{g/ml}$. La columna de la derecha fue realizada con la cepa MC1061 con una concentración máxima de 300 $\mu\text{g/ml}$ de Cm. En el segundo ensayo no se incluyeron las quimeras BY y YB debido a los antecedentes, ya que las quimeras que se esperaba que tuvieran mayor resistencia fueran las que consideran la inserción del dominio Rossman dentro del dominio catalítico pero en ambos ensayos fueron más resistentes las que no lo hicieron.

Tabla 3. Resultados de ensayo de concentración mínima inhibitoria de cloranfenicol ($\mu\text{g/ml}$) ordenados de mayor resistencia a menor

pT4_E_CAT	> 200	> 300
pT4_Y_CAT	> 200	250
pT4_B_CAT	> 200	< 250
pT4_YB_CAT	>200	
pT4_YBY_CAT	150	150
pT4_BY_CAT	150	
pT4_BYB_CAT	50	<100
pT4	< 25	<100

ENSAYO DE COMPLEMENTACIÓN

El ensayo de complementación evidencia la actividad *in vivo* de las quimeras. Para realizar este experimento fue necesario utilizar la cepa $\Delta\text{aroE Km}^-$ que es incapaz de terminar la ruta de shikimato y, por lo tanto, de sintetizar los tres aminoácidos aromáticos (Phe, Tyr y Trp) y los productos secundarios de la vía como son las vitaminas aromáticas. La cepa fue transformada con las construcciones sin la fusión traduccional a la CAT, ahora en fase con seis histidinas.

Se logró dejar en fase a las His gracias a que después de la CAT se encontraban éstas, así que se realizaron digestiones para eliminar la CAT como se describe en la sección de materiales y métodos.

Se utilizó como control negativo el plásmido pT4 vacío y como positivos las proteínas silvestres AroBs, YdiB y en una ocasión AroE. Las cepas fueron crecidas en dos medios distintos. El primero, fue medio mínimo (M9) adicionado con vitaminas cuya síntesis es dependiente de la ruta del shikimato a este medio se le denominó Sin Aminoácidos (SAA). Y el segundo, adenina de lo que se agregó anteriormente, fue adicionado con los tres aminoácidos aromáticos, producto de la vía implicada denominado Con Aminoácidos (CAA), el cual funcionó como control de los reactivos, de los inóculos y de la viabilidad de las cepas.

En los primeros ensayos se observó crecimiento de las quimeras en el medio SAA después de un lag de 70 h. Se realizaron ensayos con un par de días de diferencia en donde se inocularon tubos con ON provenientes de la misma colonia y las quimeras (de la misma colonia) no siempre crecieron. Un ejemplo de esto se muestra en la Figura 14.

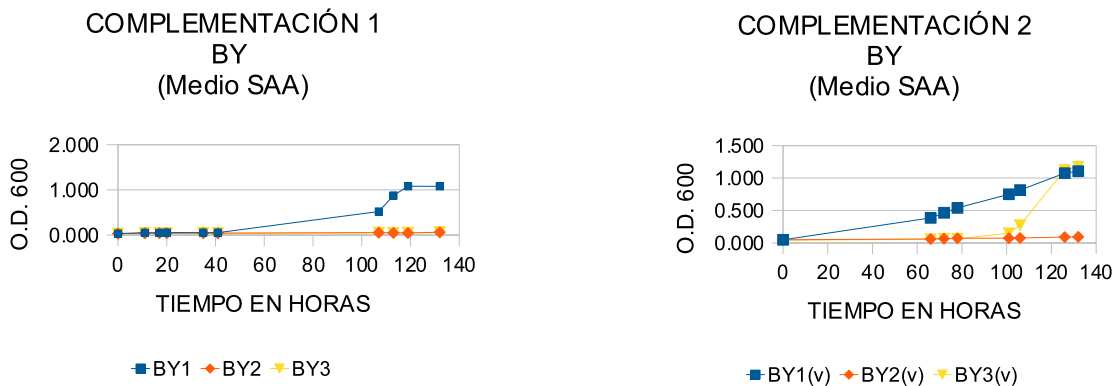


Figura 14. Cinéticas de complementación de la cepa $\Delta aroE Km^-$ por la quimera BY_HIS, en medio SAA. Ambas gráficas muestran el crecimiento de tres colonias diferentes, cada color representa una colonia, los ensayos de complementación se realizaron a partir de las mismas

En función de estos resultados surgieron dos hipótesis: como hubo un largo periodo de tiempo en el que no existió crecimiento, algunos tubos pudieron contaminarse. La segunda surgió de la observación de que los controles negativos (cepas transformadas con el plásmido vacío) nunca crecieron, fue que debido a la presión de selección que impone el medio SAA, se pudo ver favorecida una mutación espontánea que después de tantas horas comenzó a ser observable el crecimiento de la mutante.

Para probar la segunda hipótesis se realizó una complementación inoculando directamente con las cepas que ya habían complementado en el ensayo anterior, es decir, de los tubos que ya habían llegado a fase estacionaria. Ahora el crecimiento se observó sin el tiempo de lag, que era lo esperado de haberse tratado de una mutación espontánea.

Para probar que no se trató de una contaminación, lo primero que se realizó fue una PCR sobre la cicatriz del gen *aroE* (la colección keio, de la cual proviene esta cepa, está formada por mutantes, en donde cada una de ellas tiene un gen esencial, incluyendo genes putativos sustituido por un cassette de kanamicina. Pero el cassette de resistencia a Km fue eliminado (Espinosa 2010), lo que dejó una cicatriz de 170 pb. Este trabajo se realizó para posibilitar el uso del plásmido pT4 que presenta también un cassette de resistencia a Km con los oligos *aroE*-47F y *aroE*+29R.) Se utilizó como control la cepa JM101. En ningún caso se observó que alguna de las cepas contara con el gen *aroE* silvestre. El siguiente paso fue extraer plásmido de las cepas que complementaron sin el lag y mandar a secuenciar. Esto demostraría si se trataba de una contaminación o de una mutación en el plásmido. Las secuencias demostraron que se trataba de una contaminación de la cepa $\Delta\textit{aroE Km}^-$ transformada con el plásmido pT4_Y_HIS. Al saber que se trataba de una contaminación, se prepararon nuevamente todos los reactivos de ambos medios y se trabajó por separado a los controles positivos, es decir que se les manipuló hasta el final y posterior a su manipulación se lavaron las pipetas con alcohol.

Una vez que se eliminó la posibilidad de contaminación, tomando las medidas pertinentes, no se observó ninguna quimera con actividad *in vivo* (Figura 15A). En vista que las quimeras no complementaron, se realizó una nueva complementación con BY, YB, YBY y los controles con la fusión traduccional a CAT, ya que, como se sabe, ésta puede conferir estabilidad a la proteína a la que se fusiona (Figura 15B), pero tampoco se observó crecimiento en ninguna de las quimeras (no fue incluida BYBmut debido a los resultados del ensayo de CMI, las complementaciones anteriores, y su ausencia en los ensayos para la identificación de proteínas).

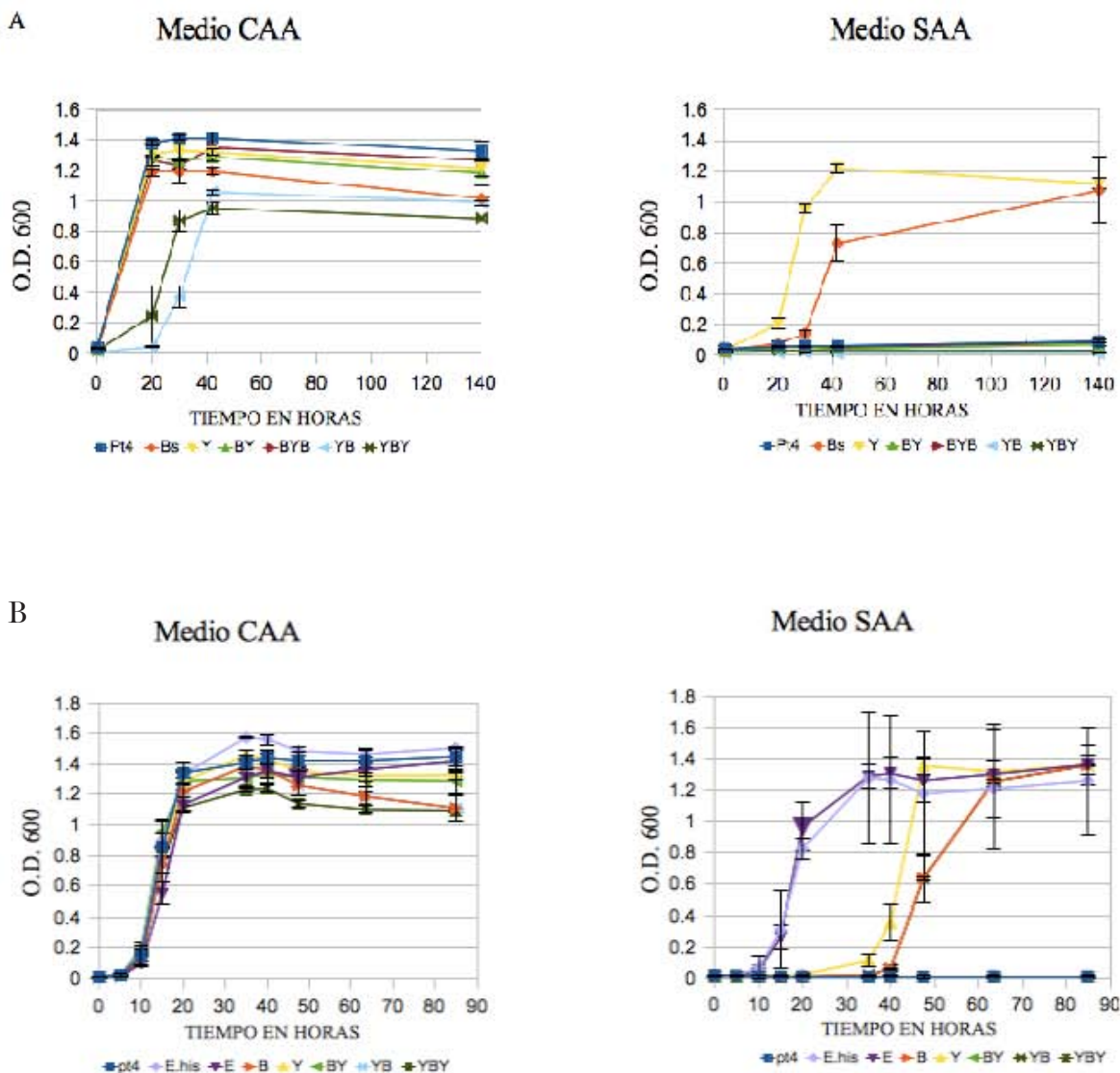


Figura 15. Cinéticas de complementación de la cepa $\Delta aroE$ Km. A) izquierda, Control de crecimiento (O.D. 600 nm) en medio SAA, cepas: pT4 (vector vacío), B_HIS, Y_HIS, pT4_BY_HIS, BYB_HIS, YB_HIS, pT4_YBY_HIS. Derecha crecimiento de las cepas anteriores en medio CAA. B) izquierda, Control de crecimiento (O.D. 600 nm) en medio SAA, cepas: pT4 (vector vacío), B_CAT, Y_CAT, BY_CAT, BYB_CAT, YB_CAT, YBY_CAT. Derecha crecimiento de las cepas anteriores en medio CAA.

SDS PAGE Y WESTERN BLOT

Se realizó tanto un gel de electroforesis SDS PAGE como un Western blot (wb) para confirmar la presencia de proteínas en tres distintos casos:

1) wb a partir de extractos totales de la células de la complementación en el medio CAA de las construcciones con una cola de 6 His (Figura 16) no se observó ninguna de las 4 quimeras en el Western blot anti HIS. Como control negativo se uso el plásmido pT4 vacío y como positivos las dos proteínas silvestres B y Y que ambas fueron visibles.

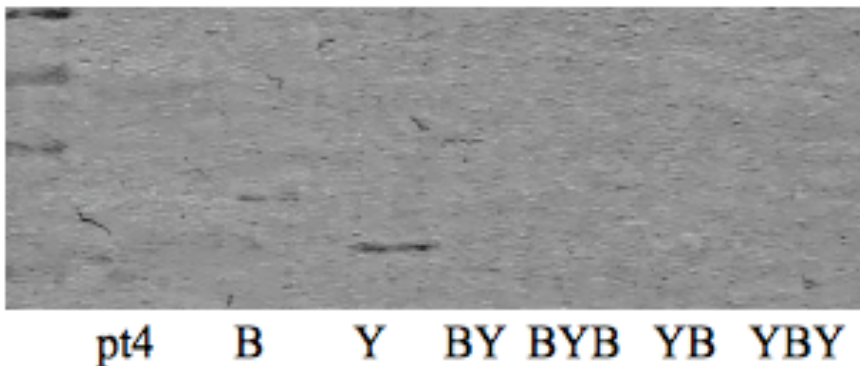


Figura 16. Expresión de cuatro quimeras y controles en vector pT4_HIS (de izquierda a derecha pT4, B, Y, BY, BYB, YB, YBY) wb anti HIS de extractos totales a partir de complementación en medio CAA en la cepa Δ aroE Km.

2) Se trabajó con los controles: pT4 como negativo y E, B y Y como positivos. Y con las quimeras fusionadas a CAT que consideraban el dominio Rossman insertado dentro del dominio catalítico (BYB y YBY) debido a los antecedentes. Las células se obtuvieron a partir de cultivos de LB incubados durante 16 h. a 37°C y 250 rpm. (Figura 17A). En el Western blot anti CAT se observó la fusión de la quimera YBY con la CAT y la CAT

libre en la fracción insoluble, mientras que en la soluble sólo se encontró la CAT libre. En ninguna de las fracciones se observó la fusión de BYBmut_CAT, ni la CAT libre. E se distinguió en ambas fracciones, B no se pudo observar, sólo apareció una banda tenue correspondiente a la CAT en la fracción insoluble, y la fusión Y_CAT sólo se observó en la fracción insoluble, mientras que en la soluble sólo se encontró la CAT libre (Figura 17B). En ese momento nos percatamos de que algo no estaba bien en la quimera BYBmut puesto que no se pudo visualizar la CAT libre, la quimera se mandó a secuenciar y se pudo observar que presentaba una deleción de un nucleótido (proveniente del oligo DIIYdB_cBs) al final del dominio Rossman lo que cambiaba el marco de lectura

3) A partir de las células provenientes de la complementación en el medio CAA en que se probaron las construcciones con fusión traduccional a CAT, se separó la fracción soluble de la insoluble. En la figura 17C se observa tanto la fusión YBY_CAT como la CAT libre en la fracción insoluble y soluble. No se incluyó a BYBmut. Los controles E y Y se observaron, en las dos fracciones, con menor CAT libre en la fracción insoluble. Solo se logró observar a B en la fracción soluble, sin embargo, sí se observó la CAT libre en ambas fracciones. Se incluyeron dos quimeras YEY y EYE como controles positivos (Figura 17C).

DISCUSIÓN Y CONCLUSIONES

Se construyeron cuatro quimeras a partir de dos proteínas homólogas con 33% de identidad, AroBs y YdiB. Dos de las quimeras se construyeron considerando el dominio Rossman unido por el extremo C-terminal al dominio catalítico, las cuales se nombraron BY y YB. Las otras dos quimeras se construyeron considerando al dominio Rossman dentro del dominio catalítico, las cuales fueron nombradas como BYB y YBY. La diferencia entre estos dos tipos de quimeras recae en la identidad de la $\alpha 9$ y $\alpha 10$. Debido a los antecedentes, se piensa que la identidad de esta cola es de suma importancia para la obtención de quimeras que se encuentren plegadas de manera correcta y que por lo tanto sean funcionales.

La delección que se encontró en la quimera BYBmut explica los resultados obtenidos en los diferentes ensayos que se practicaron. Debido a que la delección se encuentra en el dominio Rossman, el marco de lectura cambió, lo que resulta en que la CAT no fue sintetizada como se esperaba, explicando la baja resistencia a Cm. Esta delección impide que se pueda concluir algo acerca de esta quimera.

BY, YB y YBY, presentaron una alta resistencia a Cm, lo que indica que el plegamiento de las quimeras es suficientemente bueno para que se pliegue el reportero de plegamiento (CAT). Sin embargo, en ningún caso la resistencia a Cm llegó a ser tan alta como en las enzimas nativas. Como se observó en los western blots anti CAT, se muestra que

la CAT se escinde tanto en la fracción soluble como en la insoluble, por lo que los resultados del ensayo de CMI son falsos positivos. La resistencia observada probablemente se deba a la generada por la CAT libre. Para corroborar esta hipótesis se debería utilizar un control más, ya sea una proteína o un dominio que se sepa incapaz de plegarse fusionada a la CAT.

En los ensayos SDS PAGE y , en los que sólo se contaba con cola de histidinas, no se pudo observar ninguna de las quimeras. Cuando se utilizó a YBY_CAT se pudo observar la fusión con el reportero, lo que sugiere que la CAT se ve involucrada en la estabilidad de las quimeras, lo cual es congruente con los antecedentes. En el caso de las quimeras producto del intercambio de dominios de AroE y de AroBs, así como de AroE y YdiB, todas las quimeras se logran observar en wb anti CAT y sólo algunas (BE y EBE) cuando se cuenta con cola de His.

Los ensayos de complementación en los que se contaba con cola de histidinas mostraron que las quimeras no son capaces de complementar. Al no detectar las quimeras en el wb, a partir del medio CAA, se puede concluir que ninguna complementa debido a la ausencia de éstas proteínas. La ausencia no se debe a un problema de expresión puesto que en los casos en que se fusionaron a la CAT se observa la CAT tanto en el ensayo de CMI (por la resistencia a Cm) como en los SDS PAGE y wb. La presencia de la CAT no es un impedimento para la síntesis de la quimera, ya que la última se encuentra río abajo de la quimera, y como se mencionó, este reportero de plegamiento promueve la estabilidad de las quimeras.

La aparición de la CAT libre y la ausencia de las quimeras con cola de His sugiere que hay un ataque proteolítico, el cual puede ser el motivo de la ausencia de las quimeras cuando presentan la cola de histidinas. Sin embargo, los ensayos de complementación

con las quimeras fusionadas a la CAT tampoco complementan. Probablemente lo que sucede en este caso es que las quimeras que son solubilizadas por la CAT, y que no son sujeto de ataque proteolítico, no se encuentran lo suficientemente bien plegadas como para lograr tener actividad, o que el plegamiento independiente de los dominios no es suficiente para restablecer la función.

PERSPECTIVAS

La caracterización de la proteína nativa AroBs mediante un ensayo de dicroísmo circular, permitirá conocer la estructura secundaria de ésta puesto que sólo se tiene información al respecto por su secuencia y aproximaciones bioinformáticas. La técnica de permeación en gel proporcionará información acerca de su estructura cuaternaria. Obtener sus parámetros cinéticos tanto para el sustrato como para el cofactor. Resolver su estructura cristalográfica. En conjunto, esta información servirá para estudiar más a fondo las restricciones que tiene el intercambio entre dominios.

El generar modelos estructurales de manera bioinformática sobre las quimeras y análisis de las interfases, permitirá acercarse en mayor medida a las restricciones que presenta el intercambio de los dominios en este trabajo. La complementariedad entre las interfases surge de la identidad de los residuos que la conforman y pueden ser identificados *in silico*. Un análisis de las interfases servirá tanto para entender los resultados obtenidos, así como para dirigir la optimización de las quimeras.

Tomando en cuenta el trabajo de Sharkey y Engel (2009), por la similitud con el presente trabajo, la expresión y solubilidad de las quimeras puede resolverse utilizando un medio autoinducible y la expresión conjunta de chaperonas.

La optimización de quimeras por mutagénesis al azar (intentando generar sólo una mutación por variante) y seleccionar las variantes por la complementación de la cepa $\Delta aroE$ en medio SAA, permitirá conocer si existe un residuo que propicie la optimización de cada quimera.

BIBLIOGRAFÍA

- Alva, V., Remmert, M., Biegert, A., Lupas, A. N., y Söding, J. (2010). A galaxy of folds. *Protein science : a publication of the Protein Society*, 19(1), 124-30.
- Banda, J. (2011). Análisis de Proteínas Quiméricas entre AroE y YdiB de Escherichia coli. Reporte de Licenciatura, UNAM.
- Bashton, M., y Chothia, C. (2002). The geometry of domain combination in proteins. *Journal of molecular biology*, 315(4), 927-39.
- Bashton, M., y Chothia, C. (2007). The generation of new protein functions by the combination of domains. *Structure*, 15(1):85-99.
- Benach, J., Lee, I., Edstrom, W., Kuzin, A. P., Chiang, Y., Acton, T. B., et al. (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. *The Journal of biological chemistry*, 278(21), 19176-82.
- Bottoms, C. A., Smith, P. E., y Tanner, J. J. (2002). A structurally conserved water molecule in Rossmann dinucleotide-binding domains, *Protein Science*, 11(9): 2125-2137.
- Brannigan, J. A., y Wilkinson, A. J. (2002). Protein engineering 20 years on. *Nature Reviews Molecular Cell Biology*, 3(12): 964-970.
- Carugo, O., & Argos, P. (1997). NADP-dependent enzymes. I: Conserved stereochemistry of cofactor binding. *Proteins*, 28(1), 10-28.
- Chenge, M. (2010). Generación de quimeras por intercambio de dominios entre dos proteínas parálogas. (Tesis de Licenciatura – Universidad Nacional Autónoma de México), [En línea]. Disponible en: http://132.248.9.195:8080/tesdig/Procesados_tesis_2010/septiembre/0661980/Index.html[Consulta: 21 junio 2011].
- Espinosa, A. (2010). Obtención de una quimera activa por la recombinación de dominios entre dos shikimate deshidrogenasas. (Tesis de Maestría – Universidad Nacional Autónoma de México), [Consulta: 21 junio 2011]
- Fersht, A. y Winter, G. (1992). Protein engineering. *Trends in Biochemical Sciences* 17(8): 292-294.
- Goihberg, E., Peretz, M., Tel-Or, S., Dym, O., Shimon, L., Frolow, F., et al. (2010). Bio

- chemical 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.
- Han J, Batey S, Nickson A, Teichmann S. y Clarke J. (2007). The folding and evolution of multidomain proteins. *Nature Reviews; Molecular Cell Biology*. 8(4):319-30
- Kallberg, Y., & Persson, B. (2006). Prediction of coenzyme specificity in dehydrogenases/reductases. A hidden Markov model-based method and its application on complete genomes. *The FEFBS journal*, 273(6), 1177-84.
- Kataoka K, Takada H, Yoshimura T, Esaki N, Oshimura 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):931-936
- Li, H., Fast, W., y Benkovic, S. J. (2009). Structural and functional modularity of proteins in the de novo purine biosynthetic pathway. *Protein science*, 18(5), 881-92.
- Mathews CK., van Holde KE., y Ahern KG., (2005), *Bioquímica* (3a edición), Pearson educación, Madrid, 1368pp
- Maxwell, K. L., Mittermaier, a K., Forman-Kay, J. D., y Davidson, a R. (1999). A simple in vivo assay for increased protein solubility. *Protein science*. 8(9), 1908-11.
- Michel, G., Roszak, A. W., Sauvé, V., Maclean, J., Matte, A., Coggins, J. R., et al. (2003). Structures of shikimate dehydrogenase AroE and its Paralog YdiB. A common structural framework for different activities. *The Journal of biological chemistry*, 278(21), 19463-72.
- Padyana, A. y Burley, S. (2003). Crystal Structure of Shikimate 5-Dehydrogenase (SDH) Bound to NADP: Insights into Function and Evolution. *Structure*, 11(8), 1005-1013.
- Pittard, J., & Wallace, B. J. (1966). Gene controlling the uptake of shikimic acid by *Escherichia coli*. *Journal of bacteriology*, 92(4), 1070-5.

- Roberts, F., Roberts, C. W., Johnson, J. J., Kyle, D. E., Krell, T., Coggins, J. R., et al. (1998). Evidence for the shikimate pathway in apicomplexan parasites. *Nature*, 393(6687), 801-5.
- Sharkey MA y Engel PC. (2009). Modular coenzyme specificity: a domain-swopped chimera of glutamate dehydrogenase. *Proteins*, 77(2), 268-78.
- Sieber, V., Martinez, C. a, y Arnold, F. H. (2001). Libraries of hybrid proteins from distantly related sequences. *Nature biotechnology*, 19(5), 456-60.
- Singh, S., Stavrinides, J., Chistendat, D. y Guttman, D. S. (2008). A phylogenomic analysis of the shikimate dehydrogenases reveals broadscale functional diversification and identifies one functionally distinct subclass. *Molecular biology and evolution*, 25(10), 2221-32.
- Steipe, B. (2004). Consensus-based engineering of protein stability: from intrabodies to thermostable enzymes. *Methods in enzymology*, 388(1994), 176-86.
- Vogel, C., y Morea, V. (2006). Duplication, divergence and formation of novel protein topologies. *BioEssays*. 28(10): 973-8.
- Vogel, C., Bashton, M., Kerrison, N. D., Chothia, C., y Teichmann, S. a. (2004). Structure, function and evolution of multidomain proteins. *Current opinion in structural biology*, 14(2), 208-16.
- Waldo, G. S. (2003). Genetic screens and directed evolution for protein solubility. *Current Opinion in Chemical Biology*, 7(1), 33-38.
- Williamson, M. (2012), How Proteins work, (1a edición), Garland Science, New York, 468pp
- Ye, S., Delft, F. V., Brooun, A., Knuth, M. W., Swanson, R. V. y Mcree, D. E. (2003). The Crystal Structure of Shikimate Dehydrogenase (AroE) Reveals a Unique NADPH Binding Mode. *Journal of bacteriology*, 185(14), 4144-4151.
- Yeast, C.A. Orengo, C.A. (2007). Evolution of protein domains. Encyclopedia of life science.

Apéndice 1: Soluciones y medios de cultivo

Medio LB líquido		Aforar a 100ml con agua tetradesstilada
NaCl	10g	Guardar a 4°C
Extracto de levadura	5g	Upper Gel Solution (UGS)
Peptona	10g	Tris-base
En un litro de agua tetradesstilada		6.05g
		SDS
		0.4g
		Ajustar a pH 6.8 con HCl
Medio LB sólido		Aforar a 100 ml con agua tetradesstilada
NaCl	2.5g	
Extracto de levadura	1.25g	Lower Gel Solution (LGS)
Peptona	2.5g	Tris-base
Agar	3.75g	18.1g
En un 250 ml de agua tetradesstilada		SDS
		0.4g
		Ajustar a pH8.8 con HCl
PBS 10X		Llevar a 100 ml con agua tetradesstilada
NaH ₂ PO ₄	2.28g	Gel Loading <i>buffer</i> 4X
Na ₂ HPO ₄	11.5g	Glicerol
NaCl	43.84g	2ml
Agregar 450ml de agua tetradesstilada		SDS 10%
Ajustar a pH 7.4		4ml
Aforar a 500ml con agua tetradesstilada		Tris 1M pH6.8
		2.5ml
		Azul bromofenol 1% 0.5ml
		Antes de agregar a la muestra
		añadir 10% de mercaptoetanol
TAE 50X		Cátodo
Tris Base	242g	Tris base
EDTA2NA	18.6g	3g
Ac. Acético glacial		Glicina
En un litro de agua tetradesstilada		3g
Ajustar pH a 7		Metanol
		200ml
		Aforar a 1L con agua tetradesstilada
Acrilamida 30%		Ánodo I
Acrilamida	29g	Tris base
Bis-acrilamida	1g	36.3g
		Metanol
		200ml
		Aforar a 1L con agua tetradesstilada

Ánodo II

Tris base	3g
Metanol	200ml

Aforar a 1L con agua tetradestilada

TBS 20X

Tris-HCl	15.7g
NaCl	87.66g

Llevar a 500ml con agua tetradestilada
Guardar a 4°C

TBST

TBS 20X	25ml
Tween 20	0.25ml

Llevar a 500ml con agua tetradestilada
Guardar a 4°C

Apendice 2: Secuencias derivadas de las proteínas templado y quiméricas confirmadas por secuenciación de DNA, patrón de colores en función de figura 13.

B

M K K L Y G V I G N P I G H S M S P D I H N A S L K D L G
 L D G H Y H A F K V E E N D L E D A V K G I R A L G V Q G
 I N V T V P H K V S I M D Y L D H I D E S A K V L G A V N
 T V R R E G D K L V G Y H T D G E G F V K S L M K V L D K
 P I S E L S F L M I G A G G A A R A I F T T F V R N T P K
 K F D I C N R T L E K A K R L T E A T P S F H N K E V L S
 I K E A E E R L E Q Y D V I I H T T S V G M Y P N V D D V
 P L S L Q R A A S S A V V C D I V Y N P I Q T A L L K E A
 S Q K G L K T L D G V G M F V E Q A A L S F Q L W T G Q E
 P N I E

Y

M D V T A K Y E L I G L M A Y P I R H S L S P E M Q N K A
 L E K A G L P F T Y M A F E V D N D S F P G A I E G L K A
 L K M R G T G V S M P N K Q L A C E Y V D E L T P A A K L
 V G A I N T I V N D D G Y L R G Y H T D G T G H I R A I K
 E S G F D I K G K T M V L L G A G G A S T A I G A Q G A I
 E G L K E I K L F N R R D E F F D K A L A F A Q R V N E N
 T D C V V T V T D L A D Q Q A F A E A L A S A D I L T N G
 T K V G M K P L E N E S L V N D I S L L H P G L L V T E C
 V Y N P H M T K L L Q Q A Q Q A G C K T I D G Y G M L L W
 Q G A E Q F T L W T G K D F P L E Y V K Q V M G F

BY

M K K L Y G V I G N P I G H S M S P D I H N A S L K D L G
L D G H Y H A F K V E E N D L E D A V K G I R A L G V Q G
I N V T V P H K V S I M D Y L D H I D E S A K V L G A V N
T V R R E G D K L V G Y H T D G T G H I R A I K E S G F D
I K G K T M V L L G A G G A S T A I G A Q G A I E G L K E
I K L F N R R D E F F D K A L A F A Q R V N E N T D C V V
T V T D L A D Q Q A F A E A L A S A D I L T N G T K V G M
K P L E N E S L V N D I S L L H P G L L V T E C V Y N P H
M T K L L Q Q A Q Q A G C K T I D G Y G M L L W Q G A E Q
F T L W T G K D F P L E Y V K Q V M G F

BYB

M K K L Y G V I G N P I G H S M S P D I H N A S L K D L G
L D G H Y H A F K V E E N D L E D A V K G I R A L G V Q G
I N V T V P H K V S I M D Y L D H I D E S A K V L G A V N
T V R R E G D K L V G Y H T D G T G H I R A I K E S G F D
I K G K T M V L L G A G G A S T A I G A Q G A I E G L K E
I K L F N R R D E F F D K A L A F A Q R V N E N T D C V V
T V T D L A D Q Q A F A E A L A S A D I L T N G T K V G M
K P L E N E S L V N D I S L L H P G L L V T E C V Y N P H
M T K L L Q Q A Q Q A G C K T I D G V G M F V E Q A A L S
F Q L W T G Q E P N I E

YB

M D V T A K Y E L I G L M A Y P I R H S L S P E M Q N K A
 L E K A G L P F T Y M A F E V D N D S F P G A I E G L K A
 L K M R G T G V S M P N K Q L A C E Y V D E L T P A A K L
 V G A I N T I V N D D G Y L R G Y H T D G E G F V K S L M
 K V L D K P I S E L S F L M I G A G G A A R A I F T T F V
 R N T P K K F D I C N R T L E K A K R L T E A T P S F H N
 K E V L S I K E A E E R L E Q Y D V I I H T T S V G M Y P
 N V D D V P L S L Q R A A S S A V V C D I V Y N P I Q T A
 L L K E A S Q K G L K T L D G V G M F V E Q A A L S F Q L
 W T G Q E P N I E

YBY

M D V T A K Y E L I G L M A Y P I R H S L S P E M Q N K A
 L E K A G L P F T Y M A F E V D N D S F P G A I E G L K A
 L K M R G T G V S M P N K Q L A C E Y V D E L T P A A K L
 V G A I N T I V N D D G Y L R G Y H T D G E G F V K S L M
 K V L D K P I S E L S F L M I G A G G A A R A I F T T F V
 R N T P K K F D I C N R T L E K A K R L T E A T P S F H N
 K E V L S I K E A E E R L E Q Y D V I I H T T S V G M Y P
 N V D D V P L S L Q R A A S S A V V C D I V Y N P I Q T A
 L L K E A S Q K G L K T L D G Y G M L L W Q G A E Q F T L
 W T G K D F P L E Y V K Q V M G F

