



**UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO**

---

**PROGRAMA DE MAESTRIA Y DOCTORADO  
EN CIENCIAS BIOQUIMICAS  
INSTITUTO DE BIOTECNOLOGIA  
DEPARTAMENTO DE INGENIERÍA CELULAR Y BIOCÁTALISIS**

*“Diseño de proteínas a través de intercambio de  
asas catalíticas en barriles TIM”*

**TESIS**  
QUE PARA OBTENER EL GRADO ACADEMICO DE  
**MAESTRO EN CIENCIAS**

**PRESENTA:**

***Biol. Adrián Ochoa Leyva***

**DIRECTOR DE TESIS: DRA. GLORIA SAAB RINCÓN**



**CUERNAVACA, MORELOS, ABRIL 2008**



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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 quien decidió ser mi compañera por toda la eternidad, soportar mis malos ratos, y disfrutar lo bueno y lo malo de la vida juntos. Y sobre todo apoyarme en todo lo que implica el camino de la ciencia, así como disfrutar conmigo ayudándome a contar colonias aún sin convencerla de la importancia biológica para este trabajo.

**(A ti Liz Carol)**

A mis padres, hermana y abuela que siempre estuvieron allí para cualquier cosa que ocupaba. Gracias por recibirme cada vacaciones dispuestos a pasar unos buenos días en familia y decidieron apoyarme cuando decidí marchar hacia el camino de la ciencia y sobre todo alejarme bastante de mi tierra natal.

**(A ustedes: Blas Adrián, Josefina, Karina y Nana)**

A quien recientemente me uní a su familia y me han apoyado en los nuevos tiempos.

**(A ustedes: Familia García López)**

A la tía lulú que siempre quería que le recetara medicina porque yo estaba estudiando para "Dr." y la tía Bety que siempre aportaba para mi fundación en la ciencia. Y a todos los demás tíos que en las buenas y en las malas me daban la mano.

## **AGRADECIMIENTOS**

A la UNAM por ser mi segunda casa académica a la cuál admiro cada día más y por haberme apoyado para el desarrollo del presente trabajo. Por consecuencia al IBT por ser mi hogar día y noche.

A la Unidad de Docencia que siempre me apoyó para la asistencia a Congresos a presentar mi trabajo.

A el CONACYT por haber confiado en mí al otorgarme una beca para hacer posible mi manutención durante la realización de mi tesis.

A la Dra. Saab de quién aprendí muchísimo en el camino del quehacer científico y me enseñó a utilizar las herramientas del laboratorio, así como por haber sido mi asesora en este proyecto de investigación.

A el Dr. Soberón por permitirme formar parte de su grupo de investigación.

A mi comité tutorial por dedicar dos años de su tiempo a darle seguimiento a mi proyecto de investigación y a De La Vega por orientarme cuando estaba en desorden.

A mi comité de jurado, quienes se dieron la oportunidad de leer y enriquecer mi trabajo con sus atinados comentarios y discusiones científicas.

A el Biol. Filiberto Sánchez López, al Dr. Joel Osuna, al Dr. Manuel Heriberto Rivera y al Dr. Humberto Flores por toda la ayuda técnica para la realización de este trabajo y con quiénes me acaloraba con enriquecedoras discusiones científicas.

A Francisco Barona (Paco) por los comentarios y aportes valiosos para fundamentar este trabajo.

A José (Utrilla) quien siempre estaba dispuesto a darme la mano con cualquier cosa cuando se me atoraba la carreta y a darme un pretexto para un buen "beer brake".

A los "brothers" sinaloenses: el batillo y el Dago por esas cervezas que siempre nos echábamos discutiendo sobre nuestro futuro en la ciencia y hacerme sentir como si estuviera en Sinaloa.

A la bola social: Pol, Emilia, Cuahutemoc, Daniela, Eugenio, Sergio, Ericka, Joelito, David, Arcadio.

A los barona (Lianet, Pablo, Aldo, Víctor y genómicos) por permitirme interactuar con ustedes desde que se fundaron en nuestro laboratorio.

A quién siempre estaban dispuestos a contestar mis preguntas: Dr. Ricardo y Lety.

A los compañeros del labo: Mary, Gabriel, Alfredo, Etienne, Anelore, Nelly, Adriana, Bivi, Karina, Azucena, Gabriela y Luis Ledezma.

A los de la unidad de síntesis de oligos y secuenciación que siempre ponían cruces cuando llegaba por mis oligos y con mis secuencias. Jorge, Paúl, Santiago, Sol y Eugenio.

A la señora Juanita y a Pancho que siempre apoyaron en la preparación de reactivos y limpieza de material.

A todos los que por cuestión de tiempo y espacio no fueron incluidos en esta sección pero que de antemano saben que les agradezco por su apoyo.

El presente trabajo se realizó bajo la dirección de la Dra. Gloria Saab Rincón, en el Instituto de Biotecnología de la Universidad Nacional Autónoma de México. Durante la realización del mismo se recibió una beca del Consejo Nacional de Ciencia y Tecnología (CONACYT).

## COMITÉ TUTORAL:

Dra. Gloria Saab Rincón (IBT-UNAM)

Dr. Ricardo Caneck Rodríguez de la Vega Cuellar (IBT-UNAM)

Dr. Abel Moreno Cárcamo (IQ-UNAM)

## JURADO DE EXAMEN:

Dr. Lorenzo Segovia Forcella (IBT-UNAM)

Dr. Enrique Rudiño Piñera (IBT-UNAM)

Dr. David Romero (CCG-UNAM)

Dr. Luis G. Brieba de Castro (CINVESTAV-IRA)

Dra. Gloria Saab Rincón (IBT-UNAM)

# ÍNDICE

<b>RESUMEN</b> .....	1
<b>1. INTRODUCCIÓN</b>	
1.1 El mundo proteico .....	2
1.2 Diseño de proteínas y Evolución dirigida de proteínas .....	3
1.3 Generación de diversidad .....	5
1.4 Migración catalítica: el reto.....	7
1.5 El plegamiento de Barril TIM ( $\beta/\alpha$ ) <sub>8</sub> .....	8
1.6 Nuestra enzima modelo: fosforribosil antranilato isomerasa (PRAI).....	10
1.7 Reporteros de plegamiento en evolución dirigida .....	12
<b>2. ANTECEDENTES</b>	
2.1 Ingeniería de asas en barriles .....	15
2.2 Ingeniería de asas en otro plegamientos .....	16
2.3 Enzimas en estudio .....	18
<b>3. HIPÓTESIS Y OBJETIVOS</b> .....	21
<b>4. METODOLOGIA</b>	
4.1 Diseño de oligonucleótidos para la inserción de asas .....	23
4.2 Inserción de asas en la secuencia del gen de PRAI .....	25
4.3 Secuenciación de los bancos .....	27
4.4 Fusión de las librerías al gen del reportero de plegamiento (CAT) .....	28
4.5 Construcción de controles para el sistema de selección de plegamiento .....	29
4.6 Análisis del porcentaje de plegamiento de los bancos .....	30
4.7 Secuenciación bajo presión de selección .....	31
4.8 Análisis de “Western blot” .....	31
4.9 Sobreexpresión en la fracción soluble de las variantes sin fusión a CAT.....	32
4.9.1 Purificación de proteínas .....	33
4.9.2 Dicroísmo circular .....	33
<b>5. RESULTADOS</b>	
5.1 Inserción de las asas en la secuencia del gen de PRAI .....	34
5.2 Secuencias de los bancos sin selección .....	37
5.3 Fusión a CAT y análisis de los controles para el sistema de plegamiento .....	39
5.4 Secuencias obtenidas bajo presión de selección .....	43

5.5 Análisis in vitro del plegamiento (“Western-Blot”) .....	50
5.6 Sobreexpresión de variantes sin fusión a CAT .....	53
5.7 Purificación de proteínas para análisis de dicroísmo circular .....	53
5.8 Análisis de dicroísmo circular .....	56
<b>6. DISCUSIÓN</b> .....	<b>59</b>
<b>7. CONCLUSIÓN</b> .....	<b>61</b>
<b>8. ANEXOS</b>	
Anexo 1. Secuencias de oligonucleótidos .....	63
Anexo 2. Protocolos de biología molecular .....	65
Anexo 3. Soluciones .....	73
Anexo 4. Figuras del análisis de “western blot” .....	75
<b>8. BIBLIOGRAFÍA</b> .....	<b>76</b>



## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Estructura tridimensional de la enzima triosa fosfato isomerasa (TIM) .....	9
<b>Figura 2.</b> Vista lateral de la enzima fosforribosil antranilato isomerasa de <i>E. coli</i> . .....	10
<b>Figura 3.</b> Modelo de la enzima PRAI wt Cre-lox .....	12
<b>Figura 4.</b> Principio de selección de proteínas solubles por fusión a la proteína cloranfenicol acetil transferasa (CAT) como reportero de plegamiento .....	13
<b>Figura 5.</b> Diseño de oligonucleótidos que codifican para cada una de las asas .....	24
<b>Figura 6.</b> Diseño de oligos correspondientes a los conectores, para insertar las diferentes asas 4 al molde de PRAI .....	24
<b>Figura 7.</b> Proceso de ligación de cada asa con su conector para obtener los “megaprimers” ...	25
<b>Figura 8.</b> Reacciones de PCR para amplificar las mitades del gen .....	26
<b>Figura 9.</b> Reacción de PCR para la construcción de todo el gen de PRAI con la nueva asa insertada .....	27
<b>Figura 10.</b> Diseño de la fusión del gen de PRAI con el gen de CAT en el vector pDAN5-PRAI-CAT .....	28
<b>Figura 11.</b> Electroforesis en gel de acrilamida de las hibridaciones y ligaciones de oligos .....	34
<b>Figura 12.</b> Electroforesis en gel de agarosa de los fragmentos amino y carboxilo del gen de PRAI con la inserción del asa de la ureasa .....	35
<b>Figura 13.</b> Electroforesis en gel de agarosa de la amplificación por PCR para armar los diferentes bancos a partir de sus correspondientes mitades .....	35
<b>Figura 14.</b> Estructura tridimensional de PRAI señalando los sitios mutados para las asas 2, 4 y 6 .....	36
<b>Figura 15.</b> Esquema general del proceso de selección de plegamiento mediante crecimiento en presencia de cloranfenicol .....	41
<b>Figura 16.</b> Análisis de “western blot” para diferentes bancos .....	51
<b>Figura 17.</b> Fotografía de un gel de SDS con las fracciones solubles de proteínas .....	53
<b>Figura 18.</b> Cromatograma de purificación de PRAI Cre-lox WT .....	54
<b>Figura 19.</b> Cromatograma de purificación de PRAI y variantes en la columna Sephacryl S200 .....	55
<b>Figura 20.</b> Gel de SDS, donde se muestran nuestras proteínas al final de todos los pasos de purificación .....	56
<b>Figura 21.</b> Análisis por dicroísmo circular (CD) en el UV-lejano.....	57
<b>Figura 22.</b> Análisis de “western blot” para el banco del asa 4 de ureasa .....	75
<b>Figura 23.</b> Análisis de “western blot” para el banco del asa 2 de mandelato racemasa .....	75
<b>Figura 24.</b> Análisis de “western blot” para el banco del asa 6 de porfobilinógeno sintasa .....	75

## ÍNDICE DE TABLAS, ESQUEMAS Y GRÁFICAS

<b>Tabla 1.</b> Técnicas más utilizadas en evolución dirigida de proteínas .....	6
<b>Tabla 2.</b> Nombre, estructura y EC de las enzimas utilizadas en el presente estudio.....	19
<b>Tabla 3.</b> Tamaño de los bancos construidos del asa 4 .....	36
<b>Tabla 4.</b> Datos de las asas insertadas en PRAI Cre-lox WT.....	37
<b>Tabla 5.</b> Secuencias obtenidas sin presión de selección para los residuos Q79 y H81 de los diferentes bancos del asa 4.....	38
<b>Tabla 6.</b> Secuencias obtenidas sin presión de selección para el residuo I29 de los diferentes bancos del asa 2.....	38
<b>Tabla 7.</b> Secuencias obtenidas sin presión de selección para los residuos V124 y D126 de los diferentes bancos del asa 6.....	39
<b>Tabla 8.</b> Crecimiento de colonias de diferentes proteínas fusionadas a CAT.....	40
<b>Tabla 9.</b> Secuencias obtenidas con presión de selección para los residuos Q79 y H81 de los diferentes bancos del asa 4.....	43
<b>Tabla 10.</b> Secuencias obtenidas con presión de selección para el residuo I29 de los diferentes bancos del asa 2.....	43
<b>Tabla 11.</b> Secuencias obtenidas con presión de selección para los residuos V124 y D126 de los diferentes bancos del asa 6.....	44
<b>Tabla 12.</b> Secuencias obtenidas por duplicado para las posiciones Q79 y H81 en diferentes bancos del asa 4 de PRAI .....	44
<b>Tabla 13.</b> Secuencias obtenidas por duplicado para las posiciones V124 y D126 en diferentes bancos del asa 6.....	44
<b>Tabla 14.</b> Selección de residuos para las posiciones Q79 y H81 del asa 4.....	49
<b>Tabla 15.</b> Selección de residuos para las posiciones V124 y D126 del asa 6 .....	50
<b>Tabla 16.</b> Selección de residuos para la posición I29 del asa 2 .....	50
<b>Tabla 17.</b> Secuencias de las diferentes posiciones mutadas que presentaron fusión a CAT de manera soluble .....	52
<b>Tabla 18.</b> Secuencias de las diferentes posiciones mutadas que no presentaron fusión a CAT .....	52
<b>Tabla 19.</b> Predicción de estructura secundaria para cada una de las variantes de PRAI .....	57
<b>Tabla 20.</b> Tabla de oligonucleótidos que codifican las asas y los conectores .....	63
<b>Tabla 21.</b> Tabla de oligonucleótidos utilizados en el proyecto .....	64

<b>Esquema 1.</b> Rearreglo de amadori catalizado por PRAI .....	11
<b>Esquema 2.</b> Posible mecanismo de reacción catalizado por PRAI .....	11
<b>Gráfica 1.</b> Porcentajes de plegamiento para los 8 bancos del asa 4.....	41
<b>Gráfica 2.</b> Porcentajes de plegamiento para los 3 bancos del asa 2.....	42
<b>Gráfica 3.</b> Porcentajes de plegamiento para los 2 bancos del asa 6 .....	42
<b>Gráfico 4.</b> Diferencias entre frecuencias de aparición para el residuo Q79 (NNS) .....	46
<b>Gráfico 5.</b> Diferencias entre frecuencias de aparición para el residuo V124 (NNS) .....	47
<b>Gráfico 6.</b> Diferencias entre frecuencias de aparición para el residuo I29 (NNS) .....	47
<b>Gráfico 7.</b> Diferencias entre frecuencias de aparición para el residuo H81 (NNT) .....	48
<b>Gráfico 8.</b> Diferencias entre frecuencias de aparición para el residuo D126 (NNT) .....	49

## RESUMEN

Durante los primeros pasos de la evolución biológica, las proteínas adquirieron su diversidad funcional a través de inserciones, deleciones y recombinaciones de pequeños fragmentos codificados por exones primitivos de secuencias de DNA, dando origen a la diversidad de proteínas que hoy existen. Siguiendo esta tendencia, el presente trabajo desarrolla un nuevo método de generación de diversidad, el cuál es dirigido hacia la modificación completa y combinatoria de secciones (asas) de proteínas que presentan el plegamiento denominado barril TIM.

Para el desarrollo de este proyecto utilizamos como molde a la enzima fosforribosil antranilato isomerasa (PRAI) de *E. coli*, en la cuál se intercambiaron 13 diferentes asas catalíticas procedentes de otras proteínas que también poseen el plegamiento de barril TIM. Además se mutaron a saturación los dos últimos residuos que preceden a cada asa y que están ubicados en las hojas beta, los cuales apuntan hacia el sitio activo del barril. Se evaluó la viabilidad de las librerías generadas mediante fusión de las mismas a la enzima CAT (cloranfenicol acetil transferasa) como reportero de plegamiento. La fusión a CAT de las distintas librerías demostró tener desde un 25% hasta un 75% de proteínas plegadas, lo cual representa una gran población de proteínas que pueden ser analizadas en búsqueda de nuevas funciones.

La gran variabilidad encontrada bajo presión de selección de plegamiento en los sitios mutados a saturación nos indica la gran flexibilidad del sitio activo del barril TIM. Todas las proteínas analizadas sin fusión a CAT demostraron ser solubles y mantener estructura mediante un análisis por dicroísmo circular, demostrando que este sistema de fusión a CAT es un buen reportero de plegamiento para barriles TIM. La estructura de la enzima PRAI demostró que esta proteína es capaz de soportar la inserción del sitio Cre-lox y además la inserción de diferentes asas manteniendo su estructura.

# 1. INTRODUCCION

## 1.1 El mundo proteico

Todos los organismos vivos, ya sean bacterias, levaduras, animales o vegetales, comparten muchas características comunes. Una característica distintiva es que están constituidos por células y todas las células tienen una composición similar de ácidos nucleicos, carbohidratos, lípidos y proteínas. En los animales superiores existen “ciudades de células” que se agrupan para formar tejidos y órganos, mientras que en las bacterias una célula constituye la unidad de vida. Las células forman un punto intermedio en la escala de complejidad biológica y se caracterizan por poseer un compartimiento acuoso rodeado por una membrana y dentro de él, una infinidad de compuestos químicos, pero si tomamos una célula para analizar su contenido molecular, observaremos que más de la mitad de su peso seco está compuesto de proteínas [1].

El ADN tiene toda la información necesaria para generar una célula entera, pero tiene poca influencia directa sobre procesos celulares, contrario a las proteínas, las cuales dirigen una amplia gama de procesos celulares y esencialmente la maquinaria metabólica de la célula, determinan su forma y estructura, actúan como las principales herramientas de reconocimiento molecular y constituyen los mejores catalizadores biológicos [2]. Cada proteína está formada por una secuencia de aminoácidos específica, la cual determina su personalidad química, así como su estructura tridimensional; esta variabilidad hace posible la enorme versatilidad de las propiedades químicas observadas en las distintas proteínas, y posiblemente explica porqué durante la evolución se han seleccionado éstas en lugar de las moléculas de ARN para catalizar la mayoría de las reacciones celulares.

Debido a su versatilidad funcional podemos nombrar a las proteínas como “las herramientas moleculares de la vida” y por ello es importante entender la relación entre su estructura y su función, así como el proceso de su evolución, para obtener información que pueda ser utilizada en el laboratorio con el objetivo de diseñar proteínas con nuevas funciones o con actividades preestablecidas [3, 4].

Durante el siglo pasado e inicios del presente surgieron diversas técnicas moleculares tanto in vivo como in vitro para el estudio de las proteínas y principalmente de las enzimas, pero básicamente el esfuerzo en entender su función se ha centrado en diseñar mutaciones en la secuencia primaria de los genes que codifican para la proteína, es decir, cambiar un determinado aminoácido por otro y analizar el efecto producido tanto en su función como en su estructura [5]. Estudios de este tipo han llevado al establecimiento de dos estrategias generales en la ingeniería de proteínas: el diseño racional y la evolución dirigida. Estas dos estrategias siguen el principio básico de llevar a cabo mutaciones en la secuencia del gen de la proteína y analizar las

consecuencias de las mismas. La primera estrategia implica un conocimiento previo de la estructura tridimensional, mientras que la segunda no lo requiere.

## **1.2 Diseño de proteínas y Evolución dirigida de proteínas**

Hoy en día la industria biotecnológica (farmacológica, química y alimentaria) demanda el diseño de nuevas proteínas, principalmente de enzimas y anticuerpos [6, 7]. Esto provoca la necesidad de obtener enzimas que exhiban características mejoradas como: estabilidad, actividad catalítica, acción hacia diferentes sustratos y diferentes mecanismos químicos de reacción [6, 8]. Con ello, los ingenieros de proteínas se ven en la necesidad de mejorar la gama de proteínas existentes, pero además, de encontrar nuevas enzimas con diferentes características como distintos mecanismos de reacción o una actividad hacia una amplia variedad de sustratos [9]. En este sentido, en los últimos años se han encontrado nuevos genes en diferentes metagenomas los cuáles codifican para nuevas proteínas [10-12] y también la secuenciación de nuevos organismos ha arrojado una amplia gama de proteínas nunca antes descritas [13, 14]. Otra opción para el desarrollo de nuevas actividades enzimáticas, es el diseño de actividades catalíticas de novo, es decir, la modificación de proteínas existentes para que realicen funciones diferentes a las que naturalmente llevan a cabo y para ello se utiliza el diseño y la evolución dirigida de proteínas [15].

El diseño racional permite modificar específicamente una determinada región de una proteína, pero para ello es necesario tener información previa sobre la química de la reacción que realiza, la manera en la que une al sustrato, así como conocer su estructura tridimensional [16]. Generalmente las modificaciones realizadas mediante esta estrategia se establecen en la cercanía del sitio activo de la enzima [16].

La evolución dirigida de proteínas surge como una alternativa al diseño racional y se enfoca principalmente a la obtención de nuevas funciones basadas en la generación de variabilidad al azar en el gen de la proteína de interés y la posterior selección de la actividad deseada [17]. En términos generales, la evolución dirigida imita el proceso de la evolución natural, es decir, utiliza el principio de mutación-selección [8], aunque se le considera dirigida porque se seleccionan únicamente aquellas variantes cuyas propiedades son las más cercanas a las características deseadas. Básicamente, la evolución dirigida implica ciclos iterativos de mutación y/o recombinación de la secuencia de ADN que codifica para nuestra proteína, obteniendo bancos de variantes, los cuáles son acoplados a un eficiente método de selección de la nueva función o un método de búsqueda de actividad (“tamizaje”) para encontrar a la(s) mutante(s) con las características deseadas [18].

Diversos métodos son utilizados por la evolución dirigida [19], pero en general todos se basan en los siguientes pasos:

1. Elección del gen de interés, el cuál codifica la proteína con la actividad que deseamos modificar.
2. Selección del método con el que generaremos la variabilidad en la secuencia de ADN de nuestra proteína (mutagénesis). (Ver tabla 1)
3. Muestreo de la actividad deseada en la nueva proteína, mediante la detección de una señal química (tamizaje) o por complementación de la nueva función en una célula (selección).
4. Elección de las variantes ganadoras que posean la actividad buscada.
5. De ser necesario se repite el proceso para obtener en cada vuelta mejores variantes mutagenizando solamente las variantes ganadoras.

Mediante la evolución dirigida, las modificaciones de los aminoácidos pueden ubicarse en cualquier lugar de la secuencia de la proteína y por lo tanto también de su estructura, pero sólo serán seleccionados aquellos cambios que ayudaron a conferir la nueva función [20]. Actualmente la evolución dirigida constituye un método considerablemente utilizado en el diseño de enzimas y ha revolucionado a la ingeniería de proteínas, aunque presenta algunas limitaciones [8, 21-25].

Los trabajos realizados mediante evolución dirigida han demostrado la existencia de dos limitantes [19]. Por ejemplo, para una proteína con una longitud de 200 aminoácidos podríamos tener  $20^{200}$  combinaciones posibles de los aminoácidos de su secuencia, pero ni siquiera los átomos del universo nos alcanzarían para generar este número de combinaciones, cuyo número es de  $10^{77}$  átomos [26, 27]; además es imposible analizar números tan elevados de variantes, debido a que solamente podemos muestrear, en el mejor de los casos, una variabilidad de  $10^{13}$  y esto es utilizando técnicas de despliegue en ribosomas [28, 29], pero comúnmente cuando contamos con métodos de selección, analizamos combinaciones de hasta  $10^9$ , que es la eficiencia de transformación de *E. coli*, y por lo tanto es el número de variantes de la proteína que podemos analizar en forma individual, lo cuál nos lleva a sujetarnos a diseñar una variabilidad de un tamaño que pueda ser debidamente explorado en búsqueda de las características deseadas [30].

El tema de la generación de diversidad en la secuencia proteica, es un tema que merece una sección aparte debido a que es necesario el razonamiento de cuál método es el adecuado a

utilizar para lograr encontrar los cambios deseados dentro de la variabilidad generada y poder obtener un número de variantes que pueda ser explorado con las técnicas actuales.

La segunda limitante es el desarrollo de métodos de selección de las nuevas variantes, es decir, cómo vamos a poder observar qué variante tiene la actividad que estamos buscando. Para ello, tenemos que diseñar métodos que permitan ligar la propiedad buscada u objetivo con algún requerimiento fisiológico de la célula o alguna señal in vivo o in vitro, de tal manera que sólo aquellas células que sobrevivan o presenten alguna señal posible de monitorear contarán con la propiedad deseada y podrán así ser aisladas [8].

Los últimos trabajos en ingeniería de proteínas han demostrado que la forma más poderosa y exitosa de hacer modificaciones en una proteína es combinando las técnicas del diseño racional con las técnicas de evolución dirigida, por ejemplo, se realizan primero mutaciones específicas analizadas por diseño racional y posteriormente se somete a la proteína a ciclos iterativos de evolución dirigida [31, 32].

### 1.3 Generación de diversidad

Como se mencionó anteriormente, debido a que la secuencia de una proteína es producto del gen que la codifica, los cambios (mutaciones) se realizan en la secuencia de dicho gen, generando así diversidad de la secuencia original [30, 33]. De manera general podemos agrupar a estos métodos en tres categorías. La primer categoría consiste en los métodos de generación de diversidad al azar, los cuáles insertan mutaciones al azar en toda la secuencia del gen, un ejemplo lo constituye la PCR mutagénica. La segunda categoría incluye los métodos de diversidad a posiciones específicas del gen, por ejemplo el diseño de oligonucleótidos para mutar un determinado aminoácido por otro(s) aminoácido(s) codificados en el oligonucleótido. La tercera categoría incluye a los métodos de recombinación de genes, que puede ser de genes encontrados de forma silvestre o de genes diseñados sintéticamente, tal es el caso de la técnica del “shuffling” o más propiamente, barajeo de genes. Estas tres categorías agrupan un número creciente de técnicas para llevar a cabo la mutagénesis, pero las técnicas más exitosas son resumidas en la tabla 1.

Categoría	Técnica	Procedimiento
	1.1 Mutagénesis Química	Utilización de ácido nitroso para sustitución de bases en el ADN [34]
	1.2 Inserción de Nucleótidos	Se insertan nucleótidos al azar en la secuencia del gen [35]



	1.3 Cepa mutadora	Esta cepa posee defectos en la reparación del ADN por lo que en cada replicación acumula mutaciones [36-40]
1. Métodos al azar	1.4 PCR con análogos de nucleótidos	Se utilizan análogos de los nucleótidos normales para generar mutaciones en el gen de interés [41, 42]
	1.5 PCR Mutagénica	Se incrementan los rangos de error de la ADN polimerasa alterando su ambiente normal de trabajo, insertando mutaciones [43-45]
	1.6 PCR de inserción o deleción (RID)	Inserción al azar o deleción de bases dentro de una secuencia [46, 47]
2. Métodos directos	2.1 Mutagénesis a saturación	PCR mediante la que se sustituyen dos o tres bases de uno o varios codones en la secuencia del gen [48]
	2.2 Mutagénesis de pentapéptido	Se utiliza un transposón para insertar secuencias de bases al azar [49, 50]
	2.3 Mutagénesis de deleción	Utiliza una nucleasa para borrar una secuencia específica de un gen [51]
	2.4 Oligonucleótidos degenerados	Se utilizan oligonucleótidos degenerados sintetizados químicamente para extender una PCR [52-55]
	2.5 Síntesis de Trinucleótidos	Se sintetizan oligonucleótidos con codones distintos por una mezcla de nucleótidos [56]
	2.6 Fmoc- Trinucleotido	Síntesis de oligonucleótidos con codones al azar [57]
3. Métodos de Recombinación	3.1 “Shuffling”	Recombinación de alelos de genes [58, 59]
	3.2 StEP	Recombinación de secuencias de polinucleótidos mediante varios ciclos de alineamiento [60-63]
	3.3 RACHITT	Recombinación de genes parentales por solapamiento y extensión de fragmentos de los genes parentales [64, 65]
	3.4 ITCHY	Recombinación de genes distintos para generar híbridos [66-69]
	3.5 SCHEMA	Recombinación de pequeños fragmentos de proteínas homólogas, basado en un algoritmo computacional [70-72]

Tabla 1. Técnicas más utilizadas en evolución dirigida de proteínas.

Nuestro trabajo propone una nueva manera de generar diversidad en las proteínas, mediante una combinación de diseño racional y la posterior selección de plegamiento. Todo esto llevado a cabo con nuevas técnicas moleculares que implican el intercambio de asas a una proteína molde en búsqueda de nuevas funciones enzimáticas.

#### **1.4 Migración catalítica: el reto**

La evolución dirigida de proteínas en combinación con el diseño racional ha mostrado avances significativos en la última década [31]. Sólo por hacer mención a algunos ejemplos, se ha logrado aumentar la actividad catalítica de enzimas existentes en la naturaleza [70-73]; se ha incrementado la estabilidad de proteínas a altas temperaturas y/o distintos pH's [74, 75] y se han obtenido cambios de especificidad de sustrato, cofactores y enantioselectividad [32, 76, 77]. Sin embargo, no se ha logrado uno de los grandes retos de la ingeniería de proteínas: esto es, lograr una migración catalítica que implique un cambio de reacción enzimática [3]. Cambiarle la función a una determinada proteína por otra función, se ha consagrado como una barrera difícil de cruzar. Aunque pequeños pasos se han dado y comienzan a surgir trabajos interesantes al respecto, aún en estos pequeños destellos de la ingeniería de proteínas, no se ha podido igualar, mucho menos mejorar la actividad catalítica de las nuevas enzimas con respecto a la actividad mostrada por las enzimas naturales [76, 78].

En el transcurso del desarrollo del presente trabajo, Park y colaboradores publicaron un trabajo en el cual aplicaron una combinación de diseño racional y evolución dirigida. Ellos utilizaron un esquema similar al nuestro, el cual implica el corte e inserción de secciones de asas, seguido de PCR mutagénica, logrando el cambio de actividad de una glioxilasa II hacia una beta-lactamasa. Aunque este es un avance significativo, el mecanismo de reacción entre estas dos enzimas es el mismo y la nueva actividad está lejos de ser similar a la actividad de la proteína silvestre [76]. Este ejemplo constituye uno de los máximos avances de la ingeniería de proteínas respecto al cambio de función enzimática, pero la nueva función no pudo ser mejor o igual que la actividad presente en la proteína silvestre, a pesar de realizar ciclos iterativos de evolución dirigida para tratar de aumentar la actividad. Es difícil identificar cuáles mutaciones acopladas son necesarias para adquirir la función deseada, esto conduce a la necesidad de diseñar nuevos métodos y estrategias de generación de diversidad, como por ejemplo, el intercambio de asas catalíticas entre diferentes proteínas.

Una de las limitaciones que se puede intuir es que difícilmente se logrará un cambio de función realizando solamente mutaciones puntuales, por esta razón, la nueva tendencia para la modificación de proteínas implica realizar cambios más drásticos en la secuencia de la proteína de interés, lo que posiblemente nos llevaría a obtener nuevas funciones enzimáticas, objetivo

que no se ha cumplido a través de mutaciones puntuales que caen en un espacio de secuencia pequeño [79, 80]. Es decir, una serie de cambios combinatorios en un espacio de secuencia mayor, que nos conduzcan a obtener mutaciones sinérgicas y no sólo a mutaciones puntuales [32]. En este sentido una técnica llamada “family shuffling” por su nombre en inglés o barajeo de familias, permite realizar cambios estructurales drásticos pero tiene dos inconvenientes. Por un lado, las proteínas en estudio necesitan tener como mínimo un 58% de identidad de secuencia, y realmente pocas familias de proteínas la tienen, y por otro, la recombinación se limita a aquellas zonas de más alta identidad a nivel de secuencia de ADN, por lo que este método queda lejos de ser utilizado de una manera global [81].

También se ha propuesto una nueva técnica llamada “Block Shuffling” [80]. Dicha técnica plantea la combinación de distintos segmentos, entre distintas proteínas para generar nuevas funciones y está basada principalmente en que las proteínas adquirieron su diversidad funcional por la combinación de pequeños pedazos de proteínas, codificados por antiguos exones, en las primeras etapas de la evolución proteica [82]. De manera similar se han desarrollado algoritmos computacionales que predicen fragmentos estructurales que pueden ser intercambiados en una proteína sin afectar su integridad estructural permitiendo con ello la combinación de segmentos estructurales o funcionales entre proteínas [71, 72, 83].

Esta nueva visión de la ingeniería de proteínas abre la posibilidad de cambiar realmente de función a una proteína, al explorar espacios de secuencia no explorados por las técnicas tradicionales, permitiendo así salir del cuello de botella en el que actualmente se encuentra detenida [84, 85]. Siguiendo esta tendencia, el presente trabajo desarrolla un nuevo método de generación de diversidad, el cuál es dirigido hacia la modificación completa y combinatoria de secciones (asas) de proteínas que presentan el plegamiento denominado barril TIM. Este método permite incluir la diversidad biológica que ya ha sido probada naturalmente y a la par propone una nueva estrategia de exploración del espacio de secuencia mediante intercambio de asas catalíticas en barriles TIM [86].

### **1.5 El plegamiento de barril TIM ( $\beta/\alpha$ )<sub>8</sub>**

El plegamiento de barril TIM es un plegamiento que ha llamado la atención para hacer ingeniería de proteínas debido principalmente a la versatilidad de reacciones enzimáticas que puede realizar, catalizando funciones completamente distintas, que incluyen 5 de las 6 clases primarias de enzimas definidas por la Comisión de Enzimas (EC). Dichas funciones son: oxidoreductasas, transferasas, liasas, hidrolasas e isomerasas [87, 88].

La mayor parte de las proteínas que tienen este plegamiento son enzimas, encontrándose 147 distintos tipos de enzimas, 889 enzimas que presentan este dominio y la división del mismo en 17 familias [89, 90]. De todas las estructuras de proteínas conocidas, este plegamiento es el más encontrado en el "Protein Data Bank" (PDB), constituyendo alrededor del 10% [91]. La primera proteína en la cual fue encontrado este dominio fue la enzima triosa fosfato isomerasa de pollo cuya estructura fue resuelta en 1975. A partir de entonces a este plegamiento se le denominó barril TIM o también conocido como barril  $(\beta/\alpha)_8$  [92]. El barril  $(\beta/\alpha)_8$  consiste de 8 repeticiones de unidades  $(\beta/\alpha)$ , es decir, 8 hojas beta plegadas formando el centro del barril (interior), rodeadas estructuralmente por 8 alfa hélices que constituyen el exterior del barril (Ver fig. 1). Las hojas  $\beta$  y las  $\alpha$  hélices son nombradas consecutivamente a partir del extremo amino terminal de la proteína, siendo así  $\beta 1$ - $\beta 8$  y  $\alpha 1$ - $\alpha 8$ ; las asas que conectan las hojas  $\beta$  hacia las hélices  $\alpha$  son denominadas asas  $\beta/\alpha$  o también llamadas "asas catalíticas", tales como el asa  $\beta/\alpha 1$ , que conecta a la  $\beta 1$  con la  $\alpha 1$  y así consecutivamente; y las asas que conectan de las hélices  $\alpha$  hacia las hojas  $\beta$  se denominan asas  $\alpha/\beta$  [89]. Las asas  $\beta/\alpha$  tienen longitudes muy variables y con una tendencia a ser más largas que las asas  $\alpha/\beta$  localizadas en la parte inferior del barril. En algunas enzimas, las asas  $\beta/\alpha$  llegan a formar dominios completos, por ejemplo el asa  $\beta/\alpha 3$  de la piruvato cinasa constituido por 113 aminoácidos, formando un dominio extra del barril [93].



Figura 1. Estructura tridimensional de la enzima triosa fosfato isomerasa (TIM), pdb: 1tim. En color rojo se pueden observar las hélices alfa que rodean a las hojas beta de color amarillo, las cuales se encuentran en el interior del barril.

A pesar de la gran diferencia en reacciones catalíticas, los sitios activos de todos los barriles TIM están localizados en el extremo carboxilo terminal de las hojas beta (parte superior del barril), implicando que la geometría del sitio activo sea formada por estos y por los residuos de las 8 asas  $\beta/\alpha$  (ver figura 2) [89]. Además, muchas de las familias de estas enzimas requieren

la unión de cofactores o metales divalentes para llevar a cabo la catálisis, que generalmente se unen en la cara carboxilo terminal del barril, es decir, la cara catalítica. También en otros barriles como la triosa fosfato isomerasa, el sitio activo es formado por la dimerización de dos monómeros a través de sus caras catalíticas [90]. Las características anteriormente mencionadas muestran que si queremos generar distintas actividades tendríamos entonces que modificar la cara catalítica del barril, donde se encuentran las 8 asas  $\beta/\alpha$ , conservando el interior del barril en donde los residuos son sensibles a la sustitución y producen proteínas no estables [92]. Además, se ha observado que las asas  $\alpha/\beta$  tienen una mayor contribución a la estabilidad [94-97]. Esto nos hace pensar que es posible cambiar actividades catalíticas mediante el intercambio de asas, sin comprometer la estabilidad del barril [91, 98].

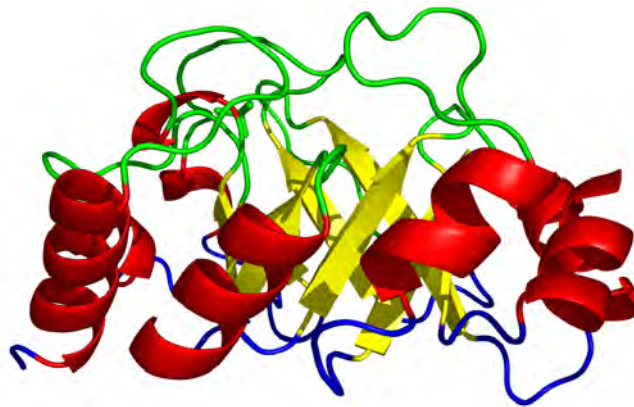


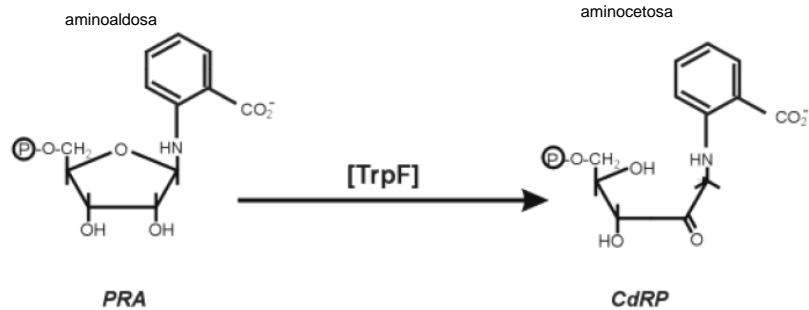
Figura 2. Vista lateral de la enzima fosforribosil antranilato isomerasa de *E. coli*. En color verde podemos apreciar las asas catalíticas del barril y en color azul las asas de estabilidad del dominio.

## 1.6 Nuestra enzima modelo: fosforribosil antranilato isomerasa (PRAI)

La enzima fosforribosil antranilato isomerasa (PRAI por sus siglas en inglés) es una proteína que posee el plegamiento de barril TIM, está codificada por el gen *TrpF* dentro del operón de ruta de biosíntesis del triptófano y cataliza el tercer paso en la ruta de síntesis del triptófano a partir del corismato. En este paso la PRAI cataliza una reacción de isomerización conocida como rearreglo de Amadori en donde una aminoaldosa es convertida en una aminocetosa [99]. (Ver esquema 1).

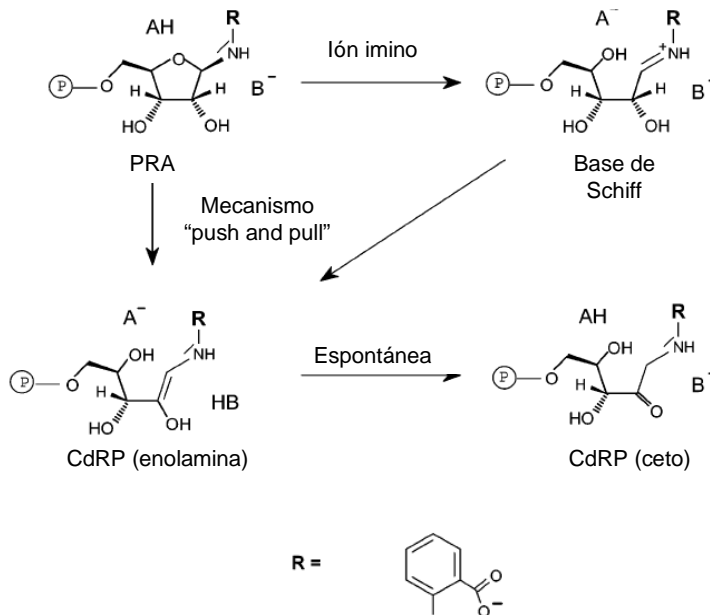
El fundamento químico de la reacción que cataliza PRAI es una reacción ácido-base donde los residuos Cys7 y Asp126 actúan como los residuos catalíticos. Se propone que el residuo Asp126 (ácido general) protona el oxígeno del anillo de la furanosa produciendo un intermediario tipo “base de Schiff” y el residuo Cys7 (base general) abstrae un protón del átomo C2’ de la ribosa, dando como resultado la apertura del anillo de la ribosa y produciendo la forma enolamina de CdRP. Alternativamente esta enolamina también puede ser formada en un solo

paso, mediante un mecanismo “push and pull” en donde la base y el ácido general actúan simultáneamente. La subsecuente formación de la enolamina de CdRP a su forma ceto ocurre espontáneamente [100]. Ver esquema 2.



Esquema 1. Rearreglo de amadori catalizado por PRAI, donde el sustrato N-(5'-fosforribosil)-antranilato (PRA) es convertido a 1-(o-carboxifenilamino)-I-deosirribulosa 5-fosfato (CdRP).

En *E. coli* esta enzima forma parte de una cadena peptídica bifuncional, es decir PRAI se encuentra fusionada con IGPS (indol glicerol fosfatos sintasa) que es la enzima que lleva a cabo el siguiente paso en la ruta de síntesis del triptófano. Sin embargo, cabe resaltar que PRAI de *E. coli* se ha podido expresar como un solo dominio (separada de la fusión con IGPS) y ha mostrado ser igualmente activa que en la fusión [101].



Esquema 2. Posible mecanismo de reacción catalizado por PRAI. El ácido (AH), la base (B-) y el sustrato PRA (R). Tomado de Henn-Sax [100].

Recientemente en nuestro grupo de investigación se ha construido una variante de PRAI, la cuál es denominada PRAI wt Cre-lox (Ver figura 3). Esta nueva proteína tiene insertado

entre la hélice 4 y la hoja beta 5 un segmento de 21 aminoácidos (linker LoxP511), cuya secuencia de ADN codificante constituye un sitio de reconocimiento para una proteína de tipo Cre-lox recombinasa. Esta variante fue construida para diseñar grandes bancos de variantes de este barril TIM, mediante la recombinación del gene por mitades separadas, lográndose con este sistema analizar bancos de un tamaño de hasta  $10^9$  variantes [102]. Esto permite ampliar el espacio de secuencia a analizar para los bancos que se construyan utilizando como molde este gene modificado. Además la nueva variante de PRAI generada demostró ser capaz de complementar la función de PRAI en una cepa de *E. coli* que porta una delección de este gen, lo que demuestra que la proteína sigue manteniendo el plegamiento de barril TIM [102].



Figura 3. Modelo de la enzima PRAI wt Cre-lox. En color verde podemos apreciar la inserción del sitio Cre-lox. Tomado de Saab-Rincón [102].

Nuestro proyecto propone que el diseño de bancos de asas en la enzima PRAI wt Cre-lox nos permitirá obtener una variabilidad de nuevas proteínas, que puedan ser probadas en búsqueda de nuevas funciones enzimáticas. Con ello uno de los primeros pasos a analizar consiste en investigar si el intercambio de asas en nuestro barril TIM por las asas de otros barriles le permite mantener la estructura y para ello utilizaremos diversos reporteros de plegamiento utilizados por la evolución dirigida, de los cuáles proporciono una breve introducción en la siguiente sección.

### 1.7- Reporteros de plegamiento en evolución dirigida

La utilización de métodos de evolución dirigida para seleccionar proteínas solubles, nos permite separar de un banco de variantes a las proteínas que se encuentran en forma soluble y con ello se facilita el análisis funcional y estructural de las mismas. Para ello diversos métodos están disponibles y no requieren información estructural o funcional de las proteínas a analizar. Una de las estrategias más utilizadas es la fusión de las proteínas de interés a reporteros de plegamiento. En esta estrategia el reportero de plegamiento posee una función o actividad

biológica que pueda detectarse y con ello poder ser seleccionada. Dos de los métodos más utilizados son la fusión a la proteína verde fluorescente (GFP) y la fusión a la proteína cloranfenicol acetil transferasa (CAT) (Ver figura 4).

La proteína verde fluorescente (GFP) es comúnmente utilizada como reportero de plegamiento, de tal manera que la proteína a analizar es expresada como el dominio amino terminal en fusión con GFP, de este modo la fluorescencia transmitida depende de la actividad del reportero y esta a su vez se correlaciona con la estabilidad de la proteína fusionada [103]. Por ejemplo, las células que expresen la GFP en fusión con otra proteína que este bien plegada provocará un efecto detectable de fluorescencia, lo que las separa de las células que no la expresen (las cuáles tienen fusionadas a proteínas insolubles).

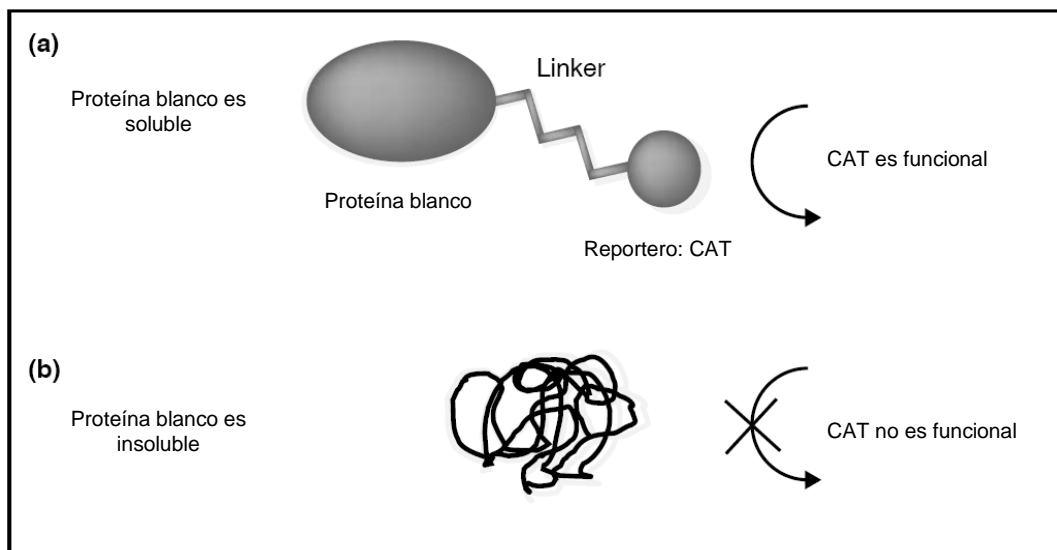


Figura 4. Principio de selección de proteínas solubles por fusión a la proteína cloranfenicol acetil transferasa (CAT) como reportero de plegamiento. La proteína cloranfenicol acetil transferasa es fusionada a la proteína blanco a través de un linker flexible. (a) Cuando la proteína blanco es soluble, la proteína CAT se pliega sin problema y es funcional y por lo tanto la célula presenta resistencia (crecimiento) en un medio suplementado con el antibiótico cloranfenicol, debido a que la función de CAT es degradar el cloranfenicol. (b) Cuando la proteína blanco es insoluble o tiene problemas para plegarse, la función de CAT es afectada y por lo tanto la célula no presenta resistencia (crecimiento) en un medio suplementado con cloranfenicol. De esta manera el ensayo sólo depende de la actividad del reportero (CAT). Imagen modificada del artículo original de Waldo [104].

La GFP es usada solamente como monitor de la producción de la proteína en estudio en forma soluble, la cuál subsecuentemente es expresada sin fusión a GFP para corroborar su solubilidad. Aplicando este método de fusión a GFP, Hecht y colaboradores encontraron variantes más solubles de la proteína  $\alpha/\beta 42$ , la cuál es precursora de Alzheimer [105].



Por otro lado, Davidson y colaboradores demostraron que la sobrevivencia de células de *E. coli* en medio suplementado con el antibiótico cloranfenicol es correlacionado con la solubilidad de distintas proteínas expresadas como el amino terminal en fusión con la proteína cloranfenicol acetil transferasa (CAT). De esta manera el crecimiento de células en medio que contiene niveles progresivos de cloranfenicol provee una presión de selección de proteínas plegadas (Ver figura 4). Este método les permitió encontrar variables solubles de la integrasa del virus VIH a partir de un banco de integrasas insolubles [106].

Otro trabajo es el de Arnold y colaboradores, quienes usaron el método de fusión a CAT para seleccionar proteínas solubles de un banco de variantes de citocromos P450, con el objetivo de separarlas de la membrana y hacerlas más solubles [107]. Esta aproximación fue un primer filtro para analizar en su ensayo de actividad enzimática solamente las proteínas seleccionadas como solubles y con ello tener un menor número de variantes a muestrear. Esta selección también les ayudó a eliminar todas las variantes que no eran completamente traducidas (con marco de lectura cambiado).

Recientemente Sterner utilizó el sistema de fusión a CAT para seleccionar proteínas con el plegamiento de barril TIM formadas a partir de mitades procedentes de diferentes barriles, obteniendo proteínas quiméricas más estables con mitades diferentes [108].

La identificación de proteínas quiméricas en forma soluble mediante la fusión a CAT ilustra un ejemplo de la aplicación satisfactoria de las técnicas de evolución dirigida para seleccionar proteínas solubles dentro de grandes bancos de variantes de proteínas. En el presente trabajo utilizamos este sistema para seleccionar proteínas solubles dentro de los diferentes bancos de asas intercambiadas en la enzima de PRAI wt Cre-lox.

## 2. ANTECEDENTES

### 2.1 Ingeniería de asas en barriles

Existen diversos trabajos de ingeniería de asas en barriles TIM, por ejemplo, mediante la manipulación de los residuos de las asas catalíticas de la enzima triosa fosfato isomerasa se ha logrado incrementar su termoestabilidad conservando prácticamente intacta la catálisis [109]. También se flexibilizó su cara catalítica, de tal forma que lograron incrementar la actividad de la enzima termófila hasta igualar la actividad de su homóloga mesófila [110], además acortando éstas asas se ha podido cambiar la dinámica de unión del sustrato [111-116]. Por otro lado, a una variante monomérica de esta enzima pero sin actividad, le fue recuperada su función mediante modificaciones en las asa  $\beta/\alpha 2$  [117]. Se han logrado cambios de especificidad mediante la manipulación de las asas en diversas enzimas con este plegamiento [118-121] así como cambios de estereo-selectividad hacia muestras racémicas de distintos sustratos [122-124]. Pero intercambio de asas entre barriles TIM aún no han sido reportados.

Recientemente se publicó un trabajo en el cuál cambiaron una sección del asa 6 de PRAI (Ser385-Gln387) por una secuencia de un epítipo de unión a un anticuerpo, dicho epítipo es 11 aminoácidos más grande que el asa 6 de PRAI. Con este trabajo demostraron mediante dicroísmo circular que la enzima mantenía prácticamente intacto su plegamiento pero había perdido la función, proponiendo que esta enzima puede ser un buen molde de los barriles  $(\beta/\alpha)_8$  para inserción de asas en el diseño de nuevas proteínas [125]. Este trabajo apoya la posible división del plegamiento del barril TIM en su cara catalítica y su cara de estabilidad, discutido en secciones anteriores

Miriam Altamirano y colaboradores llevaron a cabo un intercambio de asas entre PRAI e IGPS pero dicho trabajo fue retractado, debido a que la nueva función que ellos atribuían a la nueva proteína era producto de una contaminación por la enzima silvestre que normalmente llevaba a cabo dicha función [126, 127].

Los trabajos mencionados anteriormente son los más relacionados con el intercambio de asas en barriles, pero ninguno ha aplicado la idea de realizarlo entre enzimas que poseen este plegamiento. Existen muchos trabajos de intercambio de asas en otros plegamientos, tal es el caso de las inmunoglobulinas. En este tipo de proteína se ha observado que es posible cambiar su especificidad y mantener su estructura mediante el intercambio de diferentes asas, así como la aleatorización de las mismas [128]. Algunos ejemplos serán discutidos en la siguiente sección.

## 2.2 Ingeniería de asas en otros plegamientos

En la literatura encontramos diversos ejemplos que muestran intercambio de asas en distintos plegamientos al barril TIM. Un caso que llama la atención es el de una quimiona, a la cuál le intercambian un asa por la de una pepsina. Dicha modificación no afectó la estabilidad de la enzima, pero disminuyó la afinidad por su sustrato y se observó una disminución de su actividad enzimática [129, 130].

Por otra parte, estudios realizados sobre un citocromo C mostraron que las regiones de asas denominadas omega, pueden ser intercambiadas por asas homólogas, conservando su actividad con respecto a la proteína original [131]. Además por intercambio de asas se ha logrado cambiar la especificidad por el sustrato en distintos plegamientos sin afectar su estabilidad [132, 133], por ejemplo, al sustituirle una asa a la lisozima, se encontró una disminución de su actividad pero se conservó su estructura [134]. Todos estos trabajos muestran la flexibilidad de las asas de distintas proteínas para ser cambiadas sin comprometer la estabilidad del plegamiento y con ello permitir el diseño de nuevas proteínas, potencialmente con actividades de novo.

Los anticuerpos son indiscutiblemente la herramienta molecular más poderosa en la ciencia biomédica para generar unión hacia distintas moléculas o antígenos [135]. La alta diversidad generada en las regiones variables (CDR) hace posible la gran diversidad de unión a sus correspondientes antígenos. Esta diversidad puede ser generada por métodos de bancos sintéticos o por ejemplo, utilizando el sistema inmune de un animal [136].

Las regiones CDR están compuestas básicamente por 6 asas, y dentro de la ingeniería de anticuerpos ha sido extensivamente utilizado el llamado "loop grafting" o injerto de asas, que básicamente es un proceso mediante el cuál se translocan asas de un anticuerpo a otro para trasladar los sitios de unión o especificidad del antígeno [137]. Este injerto de asas es utilizado principalmente en la humanización de los anticuerpos, proceso mediante el cuál un anticuerpo humano recibe las regiones de otro anticuerpo generado comúnmente en un animal, como un ratón o un caballo [138]. Dicho anticuerpo es producido en un animal contra un determinado antígeno, posteriormente se lleva a cabo la translocación de las asas que poseen dicha especificidad (CDR) a un anticuerpo humano, permitiendo así una mejor aceptación de este anticuerpo híbrido por el sistema inmune del humano [138]. De esta manera se mantiene el molde del anticuerpo y sólo se insertan las regiones variables, conformadas en su mayoría por asas, que confieren la especificidad a determinadas moléculas [137].

Existen muchos ejemplos de éxito de la humanización de anticuerpos [139], pero en muchos otros casos es necesario un paso de aleatorización de todo el anticuerpo híbrido para generar la flexibilidad y el correcto acomodo de las nuevas asas dentro del nuevo molde [140]. Generalmente ésta aleatorización es producida mediante técnicas como PCR mutagénica y es analizada por selección mediante despliegue en fagos [136].

Además de las inmunoglobulinas, se han estudiado otro tipo de plegamientos capaces de tener una gran diversidad de unión a moléculas específicas, tal es el caso de las lipocalinas [141, 142], ankirinas [143], fibronectinas tipo III [144] y “affibodies” [128, 145, 146]. En estas nuevas proteínas el injerto de asas también ha resultado positivo en el desarrollo de unión a determinados antígenos, tal es el caso de la proteína neocarzinostatina (NCS). En esta proteína se insertaron las regiones CDR1 y CDR3 correspondientes a una inmunoglobulina (antilisozima) y se logró el cambio de especificidad hacia la lisozima [147].

El intercambio de asas en la ciencia biomédica es una nueva estrategia para la generación de diversos sitios de unión a diferentes moléculas en distintas proteínas molde, manteniendo la estructura [135]. Con ello se ve favorecida la posibilidad de generar bancos de asas en distintos plegamientos para buscar nuevas funciones enzimáticas.

Recordemos que en el barril  $(\beta/\alpha)_8$  las asas  $\beta/\alpha$  son las responsables de la versatilidad funcional del plegamiento, por lo cuál nuestro trabajo plantea la idea de utilizar el mismo andamiaje estructural para modificar éstas asas y con ello abrir la posibilidad de catalizar reacciones muy diversas sobre un mismo molde [94, 148]. Con esto surgen situaciones interesantes: ¿Qué pasaría si trasladamos toda la constelación catalítica de una enzima con este plegamiento a otra enzima con distinta actividad pero con el mismo plegamiento? ¿Si se realiza una combinatoria de asas encontraremos distintos mecanismos de reacción enzimática? ¿Lograremos un cambio de actividad o un cambio de especificidad? ¿O simplemente esta proteína mantendrá su estructura al cambiarle asas?

Responder a estas preguntas resulta atractivo. Desarrollar nuevas actividades catalíticas diseñadas en una enzima molde con el plegamiento de barril TIM, abre un paso para revolucionar la industria química, alimentaria y farmacéutica, y aporta un gran conocimiento para avanzar en el entendimiento sobre de la estructura y función de las proteínas y poder diseñar con ello proteínas de novo.

El presente trabajo lleva a cabo un acercamiento semiracional, para modificar la cara catalítica de nuestra enzima molde, la fosforribosil antranilato isomerasa de *E. coli* (llamada

también PRAI o TrpF) que ha sido modificada en nuestro laboratorio y que ahora posee un sitio Cre-lox [102, 149, 150]. Se construyó un banco de distintas asas de diferentes enzimas sobre esta enzima molde, y además los residuos de las hojas beta que preceden a estas asas fueron mutados a saturación con el objetivo de ayudar a que la nueva asa se acomode en la estructura y también para cambiar la composición de aminoácidos en esta región que será la zona de unión a nuevos ligandos.

Se determinará el porcentaje de plegamiento en los bancos construidos mediante la fusión al reportero de plegamiento de la proteína cloranfenicol acetil transferasa (CAT), y posteriormente se analizarán las clonas sin fusión a CAT. Solamente serán intercambiadas a PRAI wt Cre-lox las asas de las posiciones 2, 4 y 6, por ser éstas las más relacionadas con la catálisis y la unión a sustrato de nuestras enzimas en estudio.

Las asas que serán transplantadas provienen de distintas proteínas que tienen el plegamiento de barril ( $\beta/\alpha$ )<sub>8</sub> y corresponden a las siguientes enzimas: dihidrodipicolinato sintasa, aldolasa clase II, mandelato racemasa, subunidad alfa de la triptófano sintetasa, adenosin deaminasa, tiamino fosfato sintasa, ureasa y porfobilinógeno sintasa.

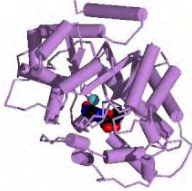



### 2.3 Enzimas en estudio

Como se mencionó en el párrafo anterior, para el presente estudio fueron seleccionadas 8 enzimas que poseen el plegamiento de barril TIM, de las cuáles fueron tomadas las respectivas asas que serán insertadas en la enzima molde (PRAI wt Cre-lox). Los criterios generales utilizados para la selección de estas enzimas fueron:

- (a) Para el caso en el que las proteínas estuvieran formadas por más de una subunidad proteica (dímero, trímero, etc), la actividad enzimática no debía depender de una interacción por el lado del sitio activo y por lo tanto por las asas catalíticas.
- (b) Las proteínas seleccionadas no debían utilizar cofactores para llevar a cabo su actividad enzimática.
- (c) Dentro de las enzimas seleccionadas deben de estar representadas todas las actividades catalíticas de los barriles TIM (oxidoreductasas, transferasas, liasas, hidrolasas e isomerasas).
- (d) Todas las enzimas deben de estar presentes en *E. coli*

Estos criterios fueron seleccionados de tal forma que la adquisición de las nuevas funciones enzimáticas por la enzima molde no dependiera de variables extras como el pegado de algún cofactor o la necesidad de oligomerizar por el sitio activo. Por otro lado, nos interesa que las enzimas seleccionadas presenten todas las actividades catalíticas de los barriles TIM para ser buscadas en la enzima molde y además que *E. coli* posea estas enzimas para poder estudiarlas en esta bacteria.

En la tabla 2 se muestran los datos de la estructura de cada enzima así como el número asignado por la comisión de enzimas (EC).

Nombre de la enzima	Código de la estructura	Número de EC
Adenosin deaminasa	 2ada	3.5.4.4
Porfobilinógeno sintasa	 116s	4.2.1.24
Dihidrodipicolinato sintasa	 1dhp	4.2.1.52
Tiamina fosfato sintasa	 1xi3	2.5.1.3

---

Subunidad alfa de la Triptófano  
sintetasa



1xc4

2.5.1.3

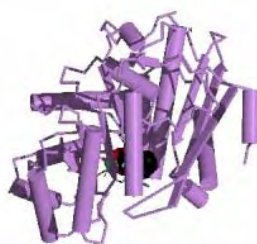
Fructosa bifosfato aldolasa



1dos

4.1.2.13

Mandelato racemasa



1mdr

5.1.2.2

Ureasa



1fwe

3.5.1.5

---

Tabla 2. Nombre, estructura y EC de las enzimas utilizadas en el presente estudio.

### 3. HIPÓTESIS

La inserción de asas de otros barriles ( $\beta/\alpha$ )<sub>8</sub> en la enzima fosforribosil antranilato isomerasa, además de las mutaciones en los últimos residuos de las hojas beta que preceden a estas asas debe de generar un gran número de variantes estables que permitan la búsqueda de nuevas actividades catalíticas.

#### 3.1 OBJETIVO GENERAL

Diseñar y construir diferentes bancos de asas de la posición 4 de diversos barriles TIM en el molde de la enzima fosforribosil antranilato isomerasa (PRAI wt Cre-lox), para el análisis de la estabilidad de su estructura.

#### 3.2 OBJETIVOS ESPECIFICOS

- Análisis de la estructura tridimensional de las 8 enzimas de interés: dihidrodipicolinato sintasa, aldolasa clase II, mandelato racemasa, subunidad alfa de la triptófano sintetasa, adenosin deaminasa, tiamino fosfato sintasa, ureasa y porfobilinogen sintasa; para la ubicación estructural de las 8 asas catalíticas de la posición 4 en estos barriles TIM.
- Diseño y síntesis de los oligonucleótidos para cada una de las 8 asas correspondientes a la posición 4 y sus adaptadores universales, así como los sitios donde llevarán variación a saturación.
- Construcción de los 8 bancos con las nuevas asas en el molde de PRAI wt Cre-lox.
- Secuenciar una muestra de los bancos construidos para evaluar el porcentaje de la correcta construcción de las variantes con las diferentes asas insertadas.
- Fusión de los bancos al reportero de plegamiento de CAT y análisis de porcentaje de plegamiento de cada uno de los bancos.
- Secuenciación de una muestra obtenida bajo presión de selección de plegamiento para análisis de variabilidad en los sitios variables.
- Análisis de la expresión de proteínas fusionadas mediante “western blot” utilizando anticuerpos contra CAT.
- Sobreexpresión de manera soluble de diferentes variantes con las nuevas asas sin fusión a CAT.



- Purificación de diferentes variantes para llevar a cabo un análisis estructural mediante dicroísmo circular.

## **4. METODOLOGIA**

La estrategia experimental la podemos dividir en cuatro partes. La primera comprende el diseño y construcción de cada asa de las diferentes enzimas en estudio en la posición del asa 4 de la secuencia del gen de PRAI, así como el análisis de dichas construcciones mediante secuenciación. Este primer punto permitió construir los 8 bancos del gen de PRAI. En la segunda parte procedimos a fusionar cada uno de estos bancos al gen de CAT (reportero de plegamiento) y con ello pudimos determinar si nuestros bancos tenían variantes plegadas. Además de los ocho bancos construidos en este trabajo, también fusionamos otros tres diferentes bancos con diferentes asas en la posición 2 y dos bancos con diferentes asas correspondientes a la posición 6, que habían sido construidas previamente en el laboratorio. También en esta sección se diseñaron y construyeron los controles para el sistema de selección de plegamiento, de tal manera que se fusionaron diferentes proteínas reportadas anteriormente como proteínas no plegadas. En la tercer parte, determinamos el porcentaje de variantes plegadas de cada banco y analizamos las secuencias de diferentes variantes crecidas en presencia de cloranfenicol (presión de selección de plegamiento). Posteriormente llevamos a cabo un análisis de "western blot" para demostrar la existencia de nuestras variantes fusionadas con CAT en la fracción soluble de los cultivos celulares. En la última sección de la metodología se realizó la sobreexpresión en la fracción soluble de diferentes variantes sin fusión a CAT. Además, se purificaron 4 variantes y fueron sometidas a un análisis de dicroísmo circular para determinar si aún mantenían estructura secundaria. La síntesis de oligonucleótidos así como las secuencias de DNA fueron realizadas por las Unidades de Síntesis y de Secuenciación, respectivamente del Instituto de Biotecnología de la UNAM.

### **4.1 Diseño de oligonucleótidos para la inserción de asas**

Para cada asa se sintetizaron dos oligonucleótidos, uno correspondiente a la cadena de ADN codificante y el otro a la cadena no codificante. Cada oligo codificante fue diseñado para que en su extremo 3' poseyera una citosina que quedaría desapareada; mientras que el oligo no codificante fue diseñado para tener en su extremo 3' una adenina que también quedaría desapareada. Con ello, al hibridar los dos oligos de cada asa obtenemos un cassette de DNA con una adenina colgando en un extremo y una citosina colgando en el otro extremo (Figura 5). Usando esta estrategia fueron sintetizados los 16 oligos correspondientes a las 8 asas de las proteínas en estudio.

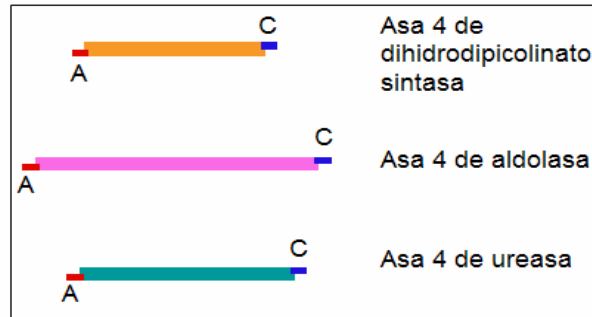


Figura 5. Diseño de oligonucleótidos que codifican para cada una de las asas.

El siguiente paso fue diseñar los conectores que unirán cada asa con el gen de PRAI. El primer conector se denominó amino y corresponde a la región de la beta 4 de PRAI, el segundo conector se denominó carboxilo y corresponde a la región de la hélice 4 de la enzima. Para el conector amino se diseñaron dos oligos, uno codificante (posee en su extremo 3' una timina) y otro no codificante. Los dos oligos llevan dos sitios de mutagénesis, estos sitios corresponden a las posiciones Q79 y H81. La posición Q79 se diseñó con la combinación de nucleótidos NNS, mientras que la posición H81 con la combinación NNT. En estas posiciones de mutagénesis la letra N significa que puede estar presente cualquiera de las 4 bases posibles, es decir, A (adenina), T (timina), C (citosina), G (guanina); y la letra S representa a citosina y a guanina. Estos dos aminoácidos posicionados en la beta 4 apuntan hacia el sitio activo por lo tanto están en la región que deseamos variar. Para el conector carboxilo se sintetizaron otros dos oligos, uno correspondiente a la hebra codificante y el otro a la no codificante (el cuál tiene diseñado en el extremo 3' una guanina). (Ver figura 6).

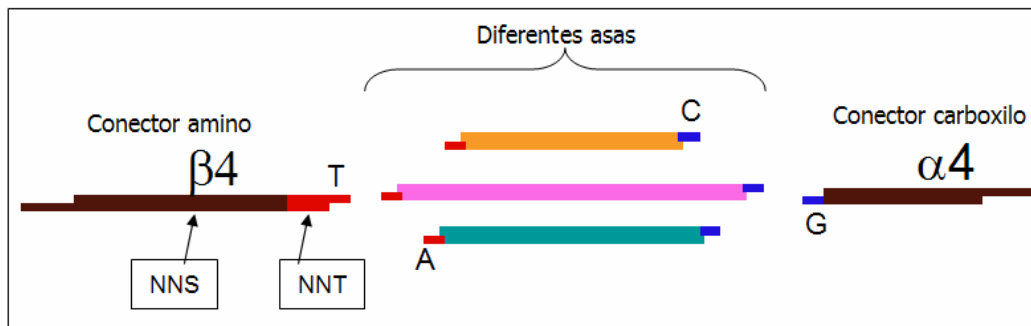


Figura 6. Diseño de oligos correspondientes a los conectores, para insertar las diferentes asas 4 al molde de PRAI.

Las secuencias y el nombre de los oligonucleotidos utilizados para la construcción de los bancos se encuentran en el anexo 1.

## 4.2 Inserción de asas en la secuencia del gen de PRAI

Todos los oligos anteriormente mencionados fueron purificados por electroforesis en gel de poliacrilamida (ver protocolo 1 anexo 2) y posteriormente fosforilados selectivamente (ver protocolo 3). De cada asa se fosforilaron los dos oligos, mientras que del conector amino el codificante y del conector carboxilo el no codificante. La fosforilación fue realizada para agregar un fosfato al extremo 5' de cada oligo (el cual es necesario en el proceso de ligación). Posteriormente estos oligos fueron hibridados (ver protocolo 4), es decir, de cada asa se hibridó el oligo correspondiente a la hebra codificante con el oligo de la hebra no codificante. De esta manera obtuvimos DNA de doble cadena para cada una de las asas y los dos conectores. El siguiente paso fue unir cada asa con los conectores y para ello por cada asa se hicieron dos reacciones de ligación aprovechando el diseño de las bases salientes de los oligos (ver protocolo 5), de tal manera que por un lado ligamos cada asa con el conector amino (hibridando la timina del conector amino con la adenina del asa) y por otro lado la misma asa con el conector carboxilo (hibridando la citosina del asa con la guanina del conector amino) tal como se observa en la figura 7.

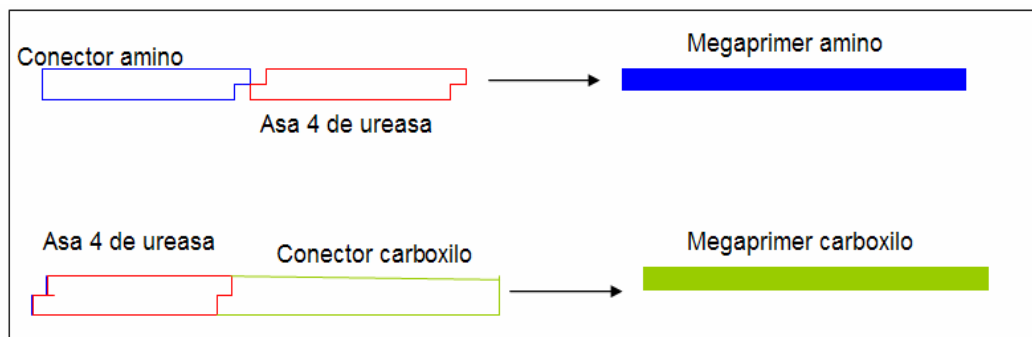


Figura 7. Proceso de ligación de cada asa con su conector para obtener los “megaprimers”.

Los productos de estas dos ligaciones fueron purificados independientemente por electroforesis en gel de poliacrilamida (ver protocolo 5), y posteriormente utilizados como “megaprimers”, para insertar cada asa en el gen de PRAI por medio de PCR.

Se hicieron dos reacciones de PCR, una para armar el fragmento del gene correspondiente al extremo amino de PRAI y la otra para armar el fragmento del gene correspondiente a la mitad carboxilo de PRAI (ver protocolo 6). En ambas reacciones se utilizó como templado el plásmido pDAN5-PRAI (que tiene clonado el gen de PRAI). Para la primer PCR usamos el oligo Hind3-Bivi y el megaprimer del conector amino; el oligo Hind3-Bivi esta diseñado para agregar en el inicio de la secuencia un sitio de restricción HindIII, un Shine-Delgarno (sitio de unión al ribosoma) y un codón de inicio del gen de PRAI. En la segunda PCR

usamos el oligo Nhe1-Bivi y el megaprimer del conector carboxilo; el oligo Nhe1-Bivi hibrida en el codón de paro de PRAI e introduce un sitio de restricción NheI. río abajo del gene. Ver Figura 8.

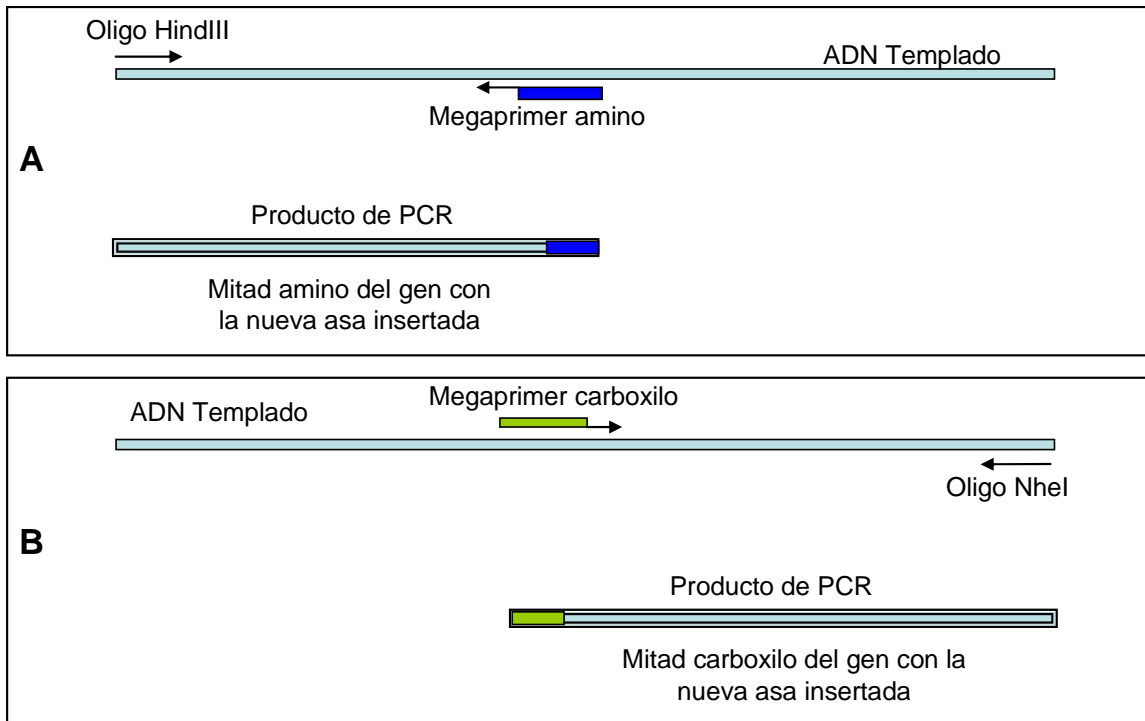


Figura 8. (A) Reacción de PCR para amplificar la mitad amino de cada gen con distinta asa. (B) Reacción de PCR para amplificar la mitad carboxilo de cada gen con distinta asa.

Los productos de las reacciones anteriores corresponden a cada una de las dos mitades del gen de PRAI, pero ahora tienen en la posición del asa 4, la secuencia de la nueva asa (determinada por el megaprimer utilizado). Cada uno de los productos de PCR de las 8 asas fue purificado por electroforesis en gel de agarosa al 1% utilizando el "kit" de Roche ("High pure PCR product purification Kit").

Después de haber amplificado por PCR las 2 mitades del gen de PRAI con la nueva asa, procedimos a armarlo por extensión del "megaprimer" mediante otra PCR (ver protocolo 7). Para esta reacción se utilizó como templado las mitades amplificadas anteriormente, aprovechando la estrategia de que las dos mitades hibridan en la región del asa insertada. Además se usaron los oligos Hind3-Bivi y Nhe1-Bivi para amplificar el producto así obtenido. Ver figura 9.

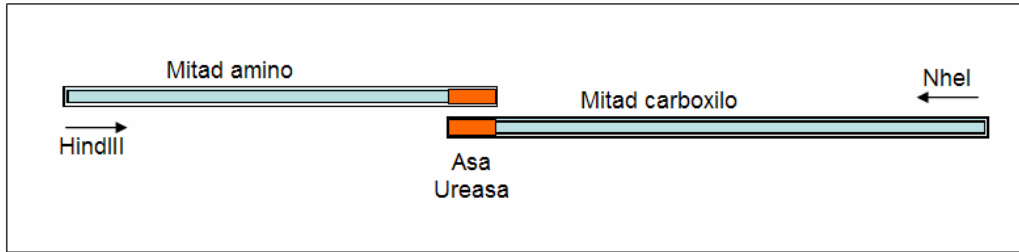


Figura 9. Reacción de PCR para la construcción de todo el gen de PRAI con la nueva asa insertada, a partir de las mitades amino y carboxilo.

De esta manera se construyeron los ocho genes (bancos) que corresponden a cada asa de las proteínas en estudio. Posteriormente el producto de PCR de cada uno de los 8 bancos fue purificado por medio de electroforesis en gel de agarosa al 1% utilizando el “kit” de Roche, digerido con las enzimas de restricción HindIII y NheI (ver protocolo 8) y ligado al vector pDAN5-PRAI (ver protocolo 10), el cuál estaba previamente digerido con las mismas enzimas de restricción y purificado por medio de electroforesis en gel de agarosa al 1% utilizando el “kit” de Roche (ver protocolo 9). Los productos de estas ligaciones se precipitaron con butanol (ver protocolo 2) y se utilizaron para transformar células electrocompetentes XL1-Blue (ver protocolo 11), las cuáles después de una hora de recuperación a 37°C se platearon e incubaron por 12 horas a 37° C en cajas de LB (ver medios) con ampicilina (Amp) ya que el vector pDAN5-PRAI porta el gen de resistencia a ampicilina. De cada transformación se cuantificó el tamaño de la ligación mediante el conteo de colonias crecidas en las anteriores cajas. El resto del volumen de cada transformación se llevó a 5 ml de LB (Amp) líquido y se dejó crecer a 37° C por 12 horas. Posteriormente se extrajo el plásmido utilizando el “kit” de Roche para aislamiento de plásmidos (“High Pure Plasmid Isolation Kit”). De esta manera cada plásmido constituye el banco de genes de cada asa.

### 4.3 Secuenciación de los bancos

De cada ligación se picaron aproximadamente 20 colonias y se crecieron en 5 ml. de LB (Amp) líquido por 12 horas a 37 °C. Posteriormente se purificó el plásmido utilizando el “kit” de Roche para aislamiento de plásmidos (“High Pure Plasmid Isolation Kit”). Todos los plásmidos purificados se mandaron a secuenciar para analizar la variabilidad presente en las posiciones que estamos mutando a saturación (Q79 y H81) y además se evaluó la calidad de la construcción de los bancos, es decir, si existe redundancia hacia un codón o aminoácido en particular, si existen clonas que posean la secuencia silvestre de PRAI o la presencia de secuencias mal construidas, producto del proceso de armado y ligación del gen.

#### 4.4 Fusión de los bancos al gen del reportero de plegamiento (CAT)

Para fusionar los bancos a CAT modificamos el vector pDAN5-PRAI. En primer término eliminamos de este vector un sitio de restricción BamHI, el cuál interfería porque CAT fusionada a PRAI esta clonada con este sitio de restricción. Esta delección del sitio de restricción fue hecha con el protocolo de síntesis del plásmido completo (ver protocolo 12) usando los oligos pdanCodBAM- y pdanNoCodBam-. En este nuevo vector subclonamos la fusión de PRAI con CAT usando los sitios de restricción HindIII al inicio del gen de PRAI y NheI al final del gen de CAT (ver figura 10). Para ello se realizó una reacción de PCR sobre el vector PT4 (el cual posee la fusión del gen de PRAI silvestre con el de CAT y que fue construido en el laboratorio del Dr. Lorenzo Segovia en el IBT-UNAM) con el oligo Hind3-Bivi (hibrida al inicio del gen de PRAI y posee el sitio de restricción HindIII) y el oligo NheI-rev-CAT (hibrida en el final del gen de CAT y posee el sitio de restricción NheI). El producto de esta PCR fue purificado por medio de electroforesis en gel de agarosa al 1% y digerido con las enzimas de restricción HindIII y NheI, y ligado al vector anteriormente construido, el cuál estaba previamente digerido con las mismas enzimas de restricción y purificado por medio de electroforesis en gel de agarosa al 1% utilizando el "kit" de Roche. El producto de esta ligación fue precipitado con butanol y se utilizó para transformar células electrocompetentes XL1-Blue, las cuales después de una hora de recuperación a 37°C se platearon e incubaron por 12 horas a 37° C en cajas de LB (ver medios) con ampicilina (Amp). Posteriormente se picaron 3 colonias y se inocularon cultivos de 5ml de LB (Amp) líquido que se dejaron crecer en agitación 12 horas a 37°C. De estos cultivos se purificó el plásmido con el "kit" de Roche y se analizó en este nuevo vector la existencia de la fusión PRAI-CAT, mediante digestión con las enzimas de restricción HindIII, NheI y BamHI y posteriormente se mandó secuenciar para analizar la posible existencia de mutaciones no deseadas dentro de los genes clonados.

Este nuevo vector (denominado pDAN5-PRAI-CAT) quedó diseñado para que al digerirlo con las enzimas HindIII y BamHI se libere el gen de PRAI y quede CAT disponible para fusionarse con el gen que sea clonado con estos sitios de restricción. Ver Figura 10.

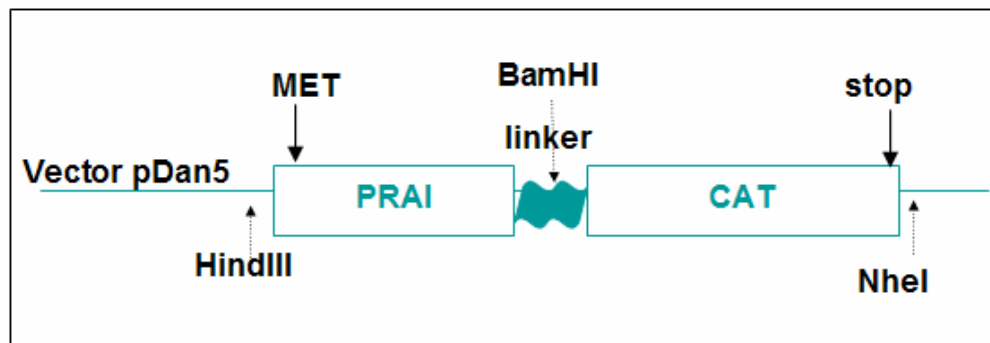


Figura 10. Diseño de la fusión del gen de PRAI con el gen de CAT en el vector pDAN5-PRAI-CAT

De esta manera todos los bancos fueron subclonados a este vector denominado pDAN5-PRAI-CAT. Para subclonarlas se realizó una PCR sobre cada uno de los bancos (ver protocolo 13), utilizando los oligos Hind3-Bivi y BamHI sinSTOP. El oligo BamHI sinSTOP posee un sitio de restricción BamHI, e hibrida sobre el final del gen de PRAI pero elimina el codón de paro existente en este gen dejando de esta manera a PRAI y a CAT como un solo gen y por lo tanto un solo polipéptido. Cada producto de PCR fue purificado por medio de electroforesis en gel de agarosa al 1% utilizando el "kit" purificación de Roche ("High pure PCR product purification Kit") y digerido con las enzimas HindIII y BamHI (ver protocolo 8). Por otro lado el vector pDAN5-PRAI-CAT fue digerido con las enzimas de restricción HindIII y BamHI y purificado por medio de electroforesis en gel de agarosa al 1% utilizando el "kit" purificación de Roche (ver protocolo 8). Cada producto de PCR digerido fue puesto en una ligación con el vector pDAN5-PRAI-CAT previamente digerido y purificado. El producto de la ligación se precipitó con butanol y se utilizó para transformar células electrocompetentes XL1-Blue, las cuáles después de una hora de recuperación a 37°C se platearon e incubaron por 12 horas a 37° C en cajas de LB con ampicilina (ver medios). Se verificó el tamaño de los bancos por conteo de colonias y se aseguró fueran mayor a 2000 variantes para tener representadas todas las variantes de cada banco.

#### **4.5 Construcción de controles para el sistema de selección de plegamiento**

Como se mencionó en la introducción, si CAT se pliega tendrá función y por lo tanto la célula va a presentar resistencia en presencia de cloranfenicol, de esta manera si la proteína que esta fusionada a CAT no se pliega, CAT tampoco lo hará y por lo tanto no habrá resistencia a cloranfenicol y la colonia no crecerá. Para la construcción de los controles negativos seleccionamos dos proteínas que normalmente no se pliegan. La primera proteína fue un anticuerpo construido por el Dr. Joel Osuna Quintero (IBT-UNAM), y caracterizado como una proteína insoluble. Para clonar este anticuerpo al vector de fusión a CAT (pDAN5-PRAI-CAT) procedimos a hacer una reacción de PCR usando como templado el plásmido que tiene el anticuerpo (pT4-Vh-papa) con el oligo HIND3-VH-pDAN y el oligo BamHI-SS-VH-pda (ver protocolo 13). El oligo HIND3-VH-pDAN esta diseñado para agregar en el inicio de la secuencia un sitio de restricción HindIII, un Shine-Dalgarno (sitio de unión al ribosoma) y un codón de inicio del gen del anticuerpo, mientras que el oligo BamHI-SS-VH-pda hibrida sobre el final del gen del anticuerpo pero elimina el codón de paro existente en este gen permitiendo de esta manera que el anticuerpo quede fusionado a CAT como un solo gen. El producto obtenido de esta PCR fue purificado por electroforesis en gel de agarosa al 1% utilizando el "kit" de Roche y posteriormente digerido con las enzimas de restricción HindIII y BamHI.

Nuestro segundo control negativo fue una permutación circular de PRAI previamente reportada como insoluble [151]. Los oligos fueron diseñados acorde al artículo publicado. La



variante seleccionada se denomina  $\beta 4/3$ . Esta es una permutación en donde la beta 4 de PRAI queda ahora como el extremo amino y la hélice alfa 3 como el extremo carboxilo de la permutación. Para la amplificación del nuevo segmento amino terminal del gen se utilizaron en una PCR los oligos HindD3 permutaci y el oligo p-Linker-B PRAI. Por el otro lado el nuevo extremo carboxilo fue amplificado en otra PCR con los oligos Bamh1 Permutaci y p-Linker-A PRAI, ambas amplificaciones se hicieron sobre el Vector (pUC18 PRAI WT) (ver protocolo 13). Ambos productos de PCR fueron purificados por medio de electroforesis en gel de agarosa al 1%. Los dos productos de las anteriores reacciones de PCR fueron usados como templado para la siguiente PCR, en presencia de los oligos HindD3 permutaci y Bamh1 Permutaci (ver protocolo 14). Este producto de PCR fue purificado por medio de electroforesis en gel de agarosa al 1% y posteriormente digerido con las enzimas de restricción HindIII y BamHI.

Un tercer control negativo fue la fusión de PRAI con CAT pero con dos codones de paro al término del gen de PRAI, de tal manera que CAT no pueda transcribirse y por lo tanto no debe presentar resistencia en presencia de cloranfenicol. Para la construcción de esta variante, llevamos a cabo una PCR sobre el templado que tiene la fusión de PRAI y CAT (plásmido pDAN5-PRAI-CAT) con los oligos Hind3-Bivi y BamH1 PRAI (ver protocolo 13). El oligo BamH1 PRAI fue diseñado de tal manera que insertara dos codones de paro al final de PRAI y además tuviera el sitio de restricción BamH1. Este producto de PCR fue purificado por medio de electroforesis en gel de agarosa al 1% y posteriormente digerido con las enzimas de restricción HindIII y BamHI.

Cada uno de los tres productos de PCR anteriormente construidos fue ligado con nuestro vector pDAN5-PRAI-CAT previamente digerido con las enzimas de restricción HindIII y BamH1 y purificado por electroforesis en gel de agarosa al 1%. El producto de cada ligación fue precipitado con butanol y transformado en células electrocompetentes XL1-Blue, las cuáles después de una hora de recuperación a 37°C se platearon e incubaron por 12 horas a 37° C en cajas de LB (ver medios) con ampicilina. Posteriormente se aisló plásmido con el “kit” de Roche de 3 colonias de cada ligación y fue digerido para corroborar la presencia del gen clonado en dichos plásmidos. Se seleccionó un plásmido de cada ligación que tuviera insertado el gen clonado y se mando secuenciar para corroborar la correcta secuencia de las variantes clonadas.

#### **4.6 Análisis del porcentaje de plegamiento de los bancos**

Los plásmidos de cada banco fusionado a CAT se transformaron en células electrocompetentes MC1061 $\Delta$ thiE. Después de una hora de recuperación a 37 ° C se les midió la OD<sub>600</sub>, para calcular el número de células y normalizar el cultivo de tal manera que en un mililitro de cultivo, el número de células fuera igual a  $1 \times 10^9$ . De allí se hicieron diluciones hasta

obtener 300 células en un volumen de 60  $\mu$ l. De este volumen se utilizaron 20  $\mu$ l y fueron depositados en forma de gota en una caja de LB sólido con ampicilina y otra gota de 20  $\mu$ l en una caja de LB sólido con ampicilina y cloranfenicol, estas cajas fueron incubadas a 30° C por 18 horas. Se hicieron 3 transformaciones independientes de cada banco y fueron goteadas de la forma explicada anteriormente. Posteriormente se analizó el número de colonias crecidas en presencia y ausencia de cloranfenicol determinando de esta manera el porcentaje de plegamiento. Es decir, se contaron el número de colonias crecidas en la gota depositada en ausencia de cloranfenicol y se compararon con el número de colonias en la gota crecida en presencia de cloranfenicol. Por ejemplo, supongamos que para la gota de un banco, contamos 100 colonias en ausencia de cloranfenicol y 80 colonias en presencia de cloranfenicol, esto significa que el 80% de las proteínas de ese banco esta plegado. Además este porcentaje se normalizó por el número de secuencias silvestres encontradas bajo presión de selección de plegamiento (crecidas en presencia de cloranfenicol).

#### **4.7 Secuenciación bajo presión de selección**

Cada banco fue transformado en células electrocompetentes MC1061 $\Delta$ thiE, las cuáles después de una hora de recuperación a 37°C se platearon e incubaron por 12 horas a 37° C en cajas de LB con ampicilina y cloranfenicol. De cada banco se picaron alrededor de 20 colonias y se inoculó un cultivo de 5 ml de LB líquido (Amp/Cm-20 (g/ml) por cada colonia, se dejaron crecer en agitación por 12 horas a 37 °C. Posteriormente se aisló plásmido con el “kit” de Roche y se mandaron secuenciar para analizar que aminoácidos se estaban seleccionando en las posiciones variables (Q79 y H81) bajo presión de selección de plegamiento. Con ello se comparó la frecuencia de aparición de residuos en presencia de presión de selección con respecto a las secuencias obtenidas en ausencia de dicha selección.

#### **4.8 Análisis de “Western blot”**

El siguiente paso de nuestra estrategia experimental fue analizar si la actividad (resistencia a cloranfenicol) presentada por nuestras variantes se debía a la presencia de la fusión de nuestras proteínas con CAT o si esta actividad era producto de la producción de CAT sin estar fusionada a nuestras proteínas. Para ello procedimos a realizar un análisis de “western blot” utilizando un anticuerpo contra la enzima cloranfenicol acetil transferasa (CAT). Se llevó a cabo el análisis para tres bancos, uno del asa 2, otro del asa 4 y finalmente otro del asa 6, de las cuáles se analizó la presencia de las fusiones en la fracción soluble.

Se transformó independientemente el plásmido de cada uno de los 3 bancos descritos anteriormente en la cepa MC1061 $\Delta$ thiE. Después de una hora de recuperación a 37°C fueron plateadas en cajas de LB sólido con ampicilina y se incubaron por 12 horas a 37 ° C. De cada

banco se picaron 10 colonias y con cada colonia se inoculó un cultivo de 5 ml de LB (Amp) líquido y se dejó crecer en agitación por 12 horas a 37 ° C. Posteriormente se extrajo el plásmido de la manera descrita anteriormente y se mandó secuenciar. También de cada colonia se inoculó otro cultivo de 5ml de medio super rico (Amp) líquido para el análisis por “western blot”. Estos cultivos se incubaron por 15 horas a 18°C y posteriormente fueron colectados por centrifugación y resuspendidos en “buffer” de lisis para ser sonicados (ver protocolo 15). Del cultivo sonificado se separó la fracción soluble de la insoluble por centrifugación y se corrió un gel de poliacrilamida al 13% para separar las proteínas presentes en la fracción soluble. Posteriormente se transfirieron las proteínas de este gel a una membrana de nitrocelulosa y se llevó a cabo el proceso para detectar la presencia de CAT a través del anticuerpo que reconoce a CAT (ver protocolo 16). De cada banco se determinó el porcentaje de plegamiento reportado por la presencia de fusión en el “western blot”.

#### **4.9 Sobreexpresión en la fracción soluble de las variantes sin fusión a CAT**

El siguiente paso fue demostrar que las proteínas reportadas como plegadas por la fusión a CAT seguían siendo solubles al quitarles CAT. Para ello seleccionamos 5 proteínas diferentes (de 5 bancos diferentes de asas) que mostraron resistencia en presencia de cloranfenicol. Cada uno de los 5 genes fueron subclonados del vector donde estaban en fusión con CAT (Pdan5-PRAI-CAT) al vector de expresión Pet28b+ (Invitrogen).

En cada una de las reacciones de PCR se utilizó como templado el respectivo plásmido con fusión a CAT y se utilizaron los oligos Pet28NcoI y Pet28HINDIII (ver protocolo 13). El oligo Pet28NcoI hibrida en la región del codon de inicio de PRAI e incluye un sitio de restricción NcoI, mientras que el oligo Pet28HINDIII hibrida en la región terminal del gen de PRAI, inserta una cola de 6 histidinas (“His-Tag”), dos codones de paro y un sitio de restricción HindIII. Cada producto de PCR fue purificado mediante electroforesis en gel de agarosa al 1% con la utilización del “kit” de Roche. Posteriormente cada producto fue digerido con las enzimas de restricción HindIII y NcoI y ligado al vector Pet28b+ previamente digerido con las mismas enzimas y purificado por electroforesis en gel de agarosa al 1% con la utilización del “kit” de Roche. El producto de cada ligación fue precipitado con butanol y transformado en células electrocompetentes XL1-Blue, las cuáles después de una hora de recuperación a 37°C se platearon e incubaron por 12 horas a 37° C en cajas de LB con kanamicina (Km). Posteriormente se picaron dos colonias de cada banco y fueron inoculadas en cultivos de 5ml de LB (Km) líquido por 12 horas a 37°C. A estos cultivos se les extrajo el plásmido y se mandaron secuenciar para asegurar que cada uno de los 5 genes de las variantes de PRAI estaba en correcto marco de lectura y no poseían mutaciones extras. Cada uno de los plásmidos (Pet28 con las mutantes de asa) fue transformado en células calcio competentes de la cepa de expresión Rosetta 2 y después de una hora de recuperación a 37 ° C

fueron plateados en cajas de LB sólido suplementadas con kanamicina y cloranfenicol. Posteriormente se llevó a cabo la sobreexpresión de cada una de las variantes siguiendo el protocolo 17. La fracción soluble de cada variante fue utilizada para ser purificada mediante afinidad a níquel usando la estrategia descrita en el protocolo 18. Se analizó en que concentración de imidazol encontramos de manera más pura a nuestras proteínas mediante un gel de poliacrilamida al 13% (SDS-Page).

#### **4.9.1 Purificación de proteínas**

Fueron sobreexpresadas 4 proteínas para su posterior purificación (una variante del asa 2, una variante del asa 4 y una variante del asa 6, además de la enzima PRAI Cre-lox WT) (ver protocolo 19 y 20). La fracción soluble de cada proteína fue purificada inicialmente por columna de níquel usando el equipo de FPLC. La fracción eluida en donde se encontró nuestra proteína fue posteriormente inyectada a una columna de permeación en gel sephacryl S200.

#### **4.9.2 Dicroísmo circular**

Para llevar a cabo el análisis de dicroísmo circular (CD) de nuestras mutantes fue necesario dializar cada una de nuestras proteínas por medio de membrana de diálisis contra un "buffer" 10mM de fosfatos con 1 mM de EDTA y 1 mM de  $\beta$ -mercaptoetanol. Finalmente cada proteína fue ajustada a una concentración final de 0.41  $\mu$ g/ $\mu$ l en un volumen de 500  $\mu$ l. Este volumen fue depositado en la celda del equipo de dicroísmo circular y fueron tomados los respectivos espectros.

## 5. RESULTADOS

### 5.1 Inserción de las asas en la secuencia del gen de PRAI

Todos los bancos fueron construidos usando el procedimiento descrito en la sección de metodología. En la figura 11 se muestra el resultado de las hibridaciones de algunos pares de oligos y las ligaciones de un asa con sus respectivos conectores.

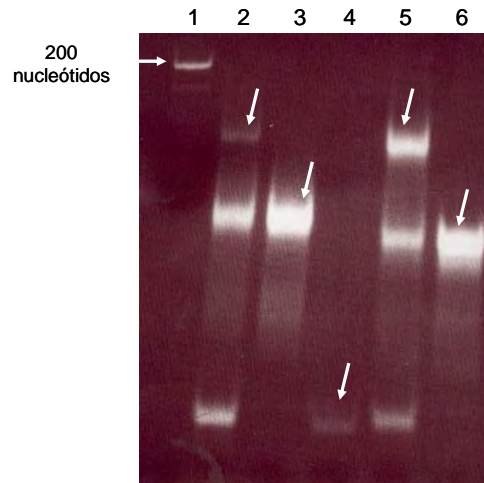


Figura 11. Electroforesis en gel de acrilamida de las hibridaciones y ligaciones de oligos. **1**-Marcador de peso molecular (200 nucleótidos), **2**-Producto de la ligación del conector amino con el asa de la ureasa (102 nucleótidos), **3**-Hibridación del conector amino (66 nucleótidos), **4**-Hibridación del asa de la ureasa (36 nucleótidos), **5**-Producto de la ligación del conector carboxilo con el asa de la ureasa (97 nucleótidos), **6**-Hibridación del conector carboxilo (61 nucleótidos)

Como se observa en la figura 11, la eficiencia de la ligación no es del 100%, esto debido a que aún existe hibridación de los conectores y del asa que no pudo ligarse. Aún así, la eficiencia de esta ligación fue suficiente para obtener cada uno de los “primers”, que sirvieron para la amplificación por PCR de la mitad amino y de la mitad carboxilo del gen de PRAI con cada una de las 8 diferentes asas.

En la figura 12 observamos los productos de la reacción de amplificación por PCR correspondientes a cada una de las mitades del gen para el banco de la ureasa. De manera similar se obtuvieron los productos de PCR con los “primers” de las demás asas.

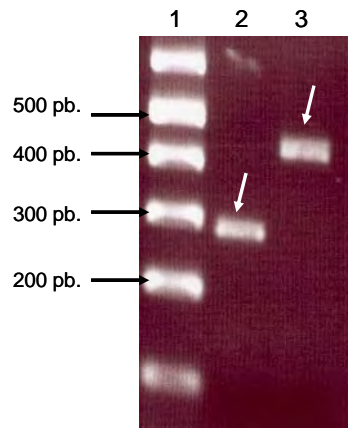


Figura 12. Electroforesis en gel de agarosa de los fragmentos amino y carboxilo del gen de PRAI con la inserción del asa de la ureasa. 1- Marcador de peso molecular, 2- Mitad amino con la inserción del asa de la ureasa (280 pb.), 3- Mitad carboxilo con la inserción del asa de la ureasa (430 pb.).

Posteriormente se armó cada uno de los bancos en una tercera reacción de PCR, donde a partir de las dos mitades de cada asa se amplificó completamente cada banco. Ver figura 13.

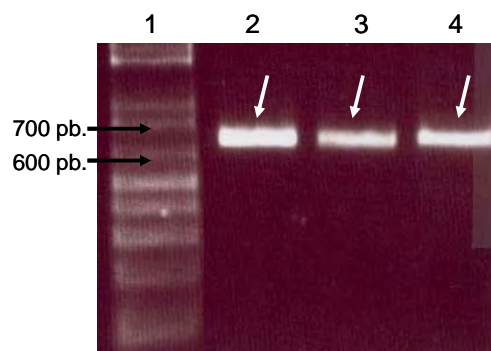


Figura 13. Electroforesis en gel de agarosa de la amplificación por PCR para armar los diferentes bancos a partir de sus correspondientes mitades. 1- Marcador de peso molecular, 2- Banco con el asa de la ureasa, 2- Banco con el asa de la aminolevulínico sintasa, 3- Banco de la subunidad alfa de la triptófano sintasa.

De la manera anteriormente detallada se obtuvieron todos los bancos y los resultados de la eficiencia de construcción se muestran en la tabla 3. La eficiencia se determinó por el número de clonas correctamente construidas (en correcto marco de lectura y que no sean silvestres) entre el número de clonas secuenciadas de cada banco. Los bancos del asa 4 deben tener un tamaño mayor a 512 variantes, que constituyen el número de combinaciones posibles para los dos sitios que estamos mutando. Es decir, 32 combinaciones para el sitio NNS (Q79) (N= 4 nucleótidos y S= 2 nucleótidos) y 16 combinaciones para el sitio NNT (H81), las cuáles al ser combinadas tienen un tamaño de 512 variantes. Las eficiencias de construcción van desde el 87% hasta el 100% (Ver tabla 3). Esto significa que en el peor de los casos, de 10 clonas secuenciadas por lo menos 8.7 están bien construidas o no son silvestres. Además el tamaño de los bancos construidos sobrepasó por mucho a las 512 variantes esperadas, permitiéndonos tener bien representada a cada banco.

Librería	Tamaño de la librería	Eficiencia de construcción
Ureasa	24,200	95.8
Porforribinogen sintasa	25,400	95.8
Mandelato racemasa	350,000	87.5
Aldolasa	210,000	95.8
Adenosin deaminasa	382,000	100
Dihidrodipicolinato sintasa	98,000	100
Tiamino sintasa	312,000	100
Sub. A de triptófano sintasa	450,000	100

Tabla 3. Tamaño de los bancos construidos del asa 4

En la figura 14 podemos observar la estructura tridimensional de PRAI, con los sitios que fueron mutados tanto del asa 4, como del asa 2 y del asa 6. Los sitios modificados a saturación nos muestran que además de cambiar las asas de PRAI por otras, estamos cambiando la estructura del sitio activo, lo cual era una pieza fundamental para el desarrollo de este proyecto.

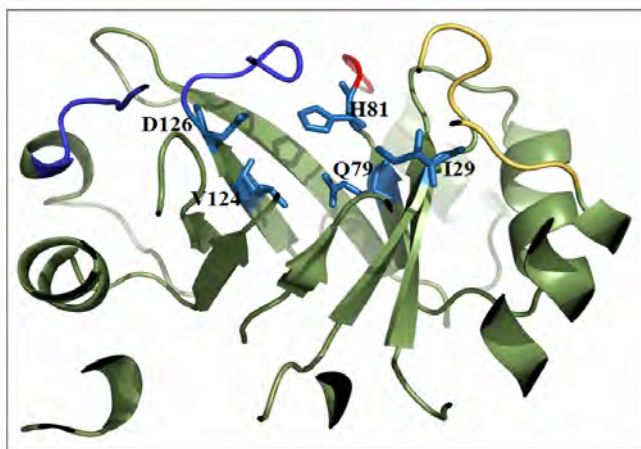


Figura 14. Estructura tridimensional de PRAI señalando los sitios mutados para las asas 2, 4 y 6. Del asa 2 (I29) del asa 4 (Q79 y H81) y del asa 6 (Y124 y D126).

Las diferentes asas insertadas en el molde de PRAI constituyen secuencias muy variables en tamaño y composición fisicoquímica (tabla 4). Esto nos muestra la gran diversidad de asas insertadas y por lo tanto las diferentes proteínas quiméricas obtenidas.

Enzima y actividad enzimática	Secuencia del asa insertada	Longitud en aminoácidos
<i>Asa β/α 2</i>		
Mandelato racemasa (isomerasa)	GYPAL	5
Aldolasa clase II (liasa)	SNGGASFIAGKGVKSDVPQ	19
Ureasa (hidrolasa)	GGTGPAAGTHATTCTPG	17
Fosforribosil antranilato isomerasa PRAI (isomerasa)	VATSPRCVN	9
<i>Asa β/α 4</i>		
Dihidrodipicolinato sintasa (liasa)	PYYNRPS	7
Tiamina fosfato sintasa (transferasa)	LGQEDLH	7
Mandelato racemasa (isomerasa)	EPTLEHD	7
Aldolasa clase II (liasa)	DLSEES	6
Subunidad alfa de Triptofano sintasa (liasa)	DVPVQQS	7
Ureasa (hidrolasa)	EDWGAT	6
Adenosin deaminasa (hidrolasa)	GDELGFPGSLF	11
Porfobilinogen sintasa (liase)	AAMDG	5
Fosforribosil antranilato isomerasa PRAI (isomerasa)	GNEE	4
<i>Asa β/α 6</i>		
Dihidrodipicolinato sintasa (liasa)	TGNL	4
Porfobilinogen sintasa (liasa)	PAGAY	5
Fosforribosil antranilato isomerasa PRAI (isomerasa)	NGQGGSGQRFD	11

Tabla 4. Datos de las asas insertadas en PRAI Cre-lox WT.

## 5.2 Secuencias de los bancos sin selección

Los resultados de la secuenciación de cada banco se muestran en la tabla 5. Esta tabla muestra los aminoácidos codificados en los 8 bancos del asa 4, para el sitio NNS (Q79) y para el sitio NNT (H81).



Número de la clona	Ureasa		Porfobilinogen sintasa		Adenosin deaminasa		Tiamina fosfato sintasa		Mandelato racemasa		Dihidrodipicolinato sintasa		Subunidad alfa de triptófano sintasa		Aldolasa clase II	
	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81
1	L	T	S	R	R	G	D	P	D	N	V	Y	T	P	R	P
2	L	S	S	C	C	I	D	C	D	D	K	T	G	I	R	D
3	L	P	V	Y	R	A	I	P	P	P	E	R	Stop	Y	L	V
4	I	S	V	V	G	Y	I	V	P	C	P	S	K	T	C	N
5	I	T	V	A	V	A	P	P	I	P	S	L	I	S	V	D
6	R	P	C	P	H	V	P	V	I	V	G	V	V	P	V	F
7	R	S	C	H	W	G	P	C	I	N	R	V	S	V	S	Y
8	R	L	T	V	M	R	H	F	I	I	D	A	K	P	S	I
9	R	T	T	I	Y	G	H	S	I	L	V	S	Y	S	H	P
10	P	H	T	D	F	R	M	N	E	H	F	S	A	V	T	N
11	G	N	L	N	F	Y	M	C	H	N	T	C	V	S	T	I
12	F	P	L	R	I	H	R	C	H	F	Y	T	I	F	T	F
13	V	F	L	S	L	G	R	H	M	C	L	C	D	D	P	S
14	V	R	I	I	K	C	L	F	Q	F	W	S	H	D	N	S
15	Y	I	I	S	L	N	L	A	Q	A	P	V	H	S	N	F
16	Y	R	H	N	A	D	V	R	K	P	P	C	F	S	I	N
17	A	Q	H	R	H	P	G	D	K	H	K	N	W	I	I	A
18	A	I	P	V	S	I	G	R	G	H	G	I	R	D	I	L
19	S	A	P	Y			A	P	V	R					K	A
20	S	V	K	F			S	N	S	N					Y	A
21	S	S	R	H			S	P	W	D					G	T
22	K	R	Y	I			S	I							G	V
23	K	P	Y	V			S	S							G	L
24							W	N								

Tabla 5. Secuencias obtenidas sin presión de selección para los residuos Q79 y H81 de los diferentes bancos del asa 4.

Además se muestran los resultados de las secuencias para los bancos del asa 2 (tabla 6) y del asa 6 (tabla 7), previamente construidas por otros miembros de nuestro grupo de investigación.

Número de la clona	Ureasa	Mandelato racemasa	Aldolasa clase II
	I29	I29	I29
1	L	L	T
2	V	A	S
3	T	W	L
4	T	W	L
5	F	W	T
6	K	W	S
7	S	D	H
8	D	D	L
9	L	G	L
10	T	L	L
11	I	W	L
12	I	N	N
13	I	R	G
14	W	R	R
15	G	G	G
16	L	D	I
17	S	D	G
18	R	P	A
19		F	
20		W	

Tabla 6. Secuencias obtenidas sin presión de selección para el residuo I29 de los diferentes bancos del asa 2.

Número de la clona	Phorofobilinogen sintasa		Dyhidrodipicolinato sintasa	
	V124	D126	V124	D126
1	V	Y	T	S
2	N	L	M	S
3	P	L	I	C
4	A	Y	V	C
5	P	H	T	G
6	R	C	Q	I
7	V	D	R	I
8	R	A	L	C
9	P	T	I	R
10	T	F	S	T
11	Y	C	L	L
12	D	S	P	V
13	A	D	V	L
14	N	P	S	S
15	W	G	S	L
16	G	G	G	T
17	S	A	A	S
18	N	T	E	F
19			E	V
20			L	R
21			T	H
22			N	T
23			S	F

Tabla 7. Secuencias obtenidas sin presión de selección para los residuos V124 y D126 de los diferentes bancos del asa 6.

Los resultados de las secuencias muestran la existencia de variabilidad en los sitios mutados, además dicha variabilidad esta en acuerdo con la diseñada teóricamente para la combinación de bases de cada sitio mutado, es decir, el sesgo observado para determinados residuos es debido a que ese residuo tiene un mayor número de codones para una determinada combinación de bases. Es importante señalar que ninguna combinación fue encontrada dos veces para cada banco, lo cual era de esperarse debido a la pequeña población muestreada. Si hubiese existido una combinación repetida podría ser indicativa de un sesgo en el banco hacia esos codones.

Este análisis de la distribución de residuos encontrados para cada asa se comparó con los residuos encontrados en las mismas posiciones pero bajo presión de selección de plegamiento y con ello se identificaron residuos que son indispensables para mantener el plegamiento de nuestras proteínas quiméricas después de la inserción de una determinada asa.

### 5.3 Fusión a CAT y análisis de los controles para el sistema de plegamiento

La viabilidad de los bancos construidos fue evaluada por la fusión de los mismos a CAT como reportero de plegamiento [106]. Para fusionar nuestros bancos al gen de CAT llevamos a cabo modificaciones en el plásmido pDAN5-PRAI, eliminando un sitio de restricción BamHI que interfería con la clonación debido a que PRAI fusionada a CAT está clonada con este sitio de restricción. Ya eliminado este sitio del vector, clonamos la fusión PRAI-CAT procedente del plásmido pT4, la cual fue amplificada mediante PCR y clonamos en este nuevo vector (pDAN5-

PRAI-CAT) todos los bancos de tal manera que quedaran en fusión con CAT. Se fusionaron los ocho bancos construidos en este trabajo, y además, tres bancos del asa 2 y dos bancos del asa 6.

Para validar nuestro sistema de selección de plegamiento fusionamos a CAT dos proteínas previamente reportadas como insolubles, las cuáles constituyen nuestros controles negativos de selección de plegamiento. Por un lado fusionamos un anticuerpo (donado por el Dr. Joel Osuna IBT-UNAM) y por otro lado una permutación circular de PRAI [151, 152].

En la tabla 8 podemos observar los resultados del crecimiento de colonias de la cepa MC1061 $\Delta$ thiE transformadas con diferentes fusiones a CAT y crecidas en diferentes concentraciones de cloranfenicol.

Nombre de la muestra	Número de colonias por placa en diferentes concentraciones de cloranfenicol						
	0 μg/ml	5 μg/ml	10 μg/ml	15 μg/ml	20 μg/ml	40 μg/ml	100 μg/ml
Anticuerpo fusionado a CAT	846	813	150	0	0	0	0
Permutación circular fusionada a CAT	921	895	123	0	0	0	0
Cepa MC1061-Thie	0	0	0	0	0	0	0
PRAI fusionada a CAT	953	967	942	956	940	925	903
PRAI con codon de paro entre la fusión con CAT	985	0	0	0	0	0	0

Tabla 8. Crecimiento de colonias de diferentes proteínas fusionadas a CAT. Todas las placas de LB estaban suplementadas con ampicilina y la correspondiente concentración de cloranfenicol.

A partir de la tabla 8 podemos observar que nuestros controles negativos (proteínas insolubles) ya no crecen en la concentración de 15 μg/ml de cloranfenicol, por lo tanto, determinamos que esta es la mínima concentración de cloranfenicol para reportar proteínas plegadas. Es decir, lo que crece por encima de esta concentración corresponde a proteínas plegadas o menos insolubles, por ejemplo, la fusión de PRAI con CAT. Por debajo de esta concentración los resultados son falsos positivos debido a que nuestros controles negativos a 10 μg/ml de cloranfenicol muestran crecimiento y por lo tanto resistencia, siendo reportadas como solubles, cuando nosotros sabemos que no lo son. Basados en estos resultados decidimos utilizar la concentración de 20 μg de cloranfenicol para la selección de plegamiento de nuestros bancos (ver figura 15).

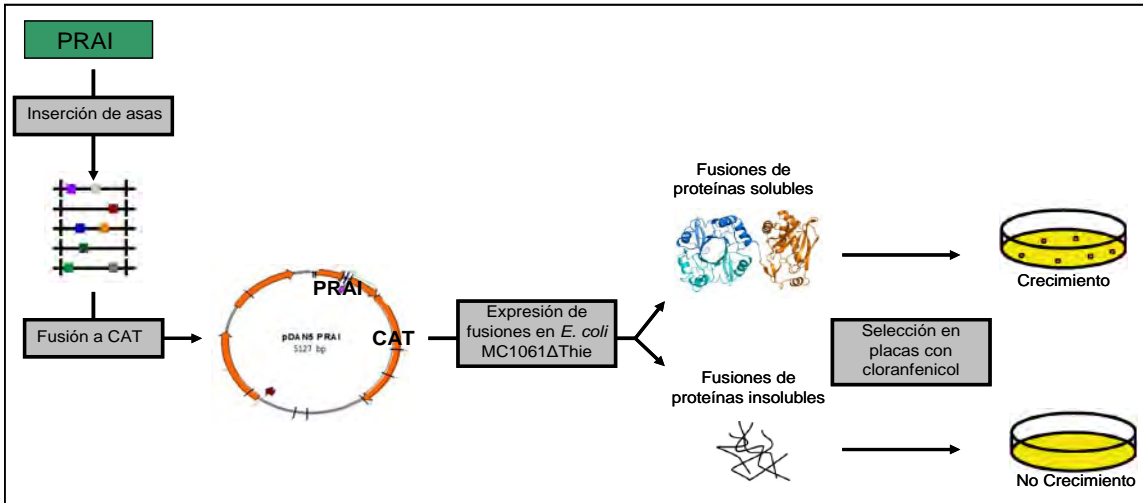
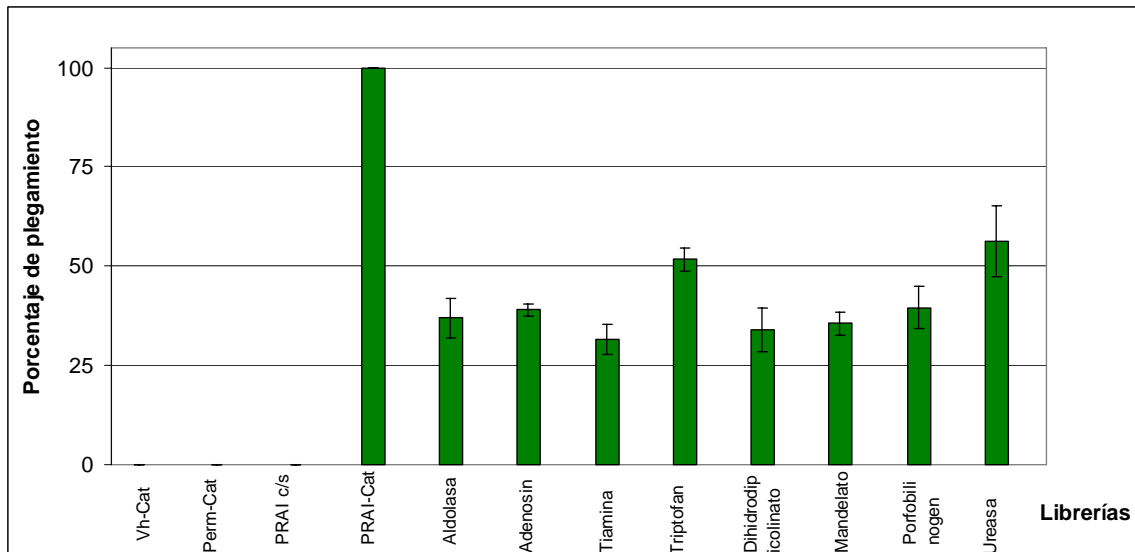
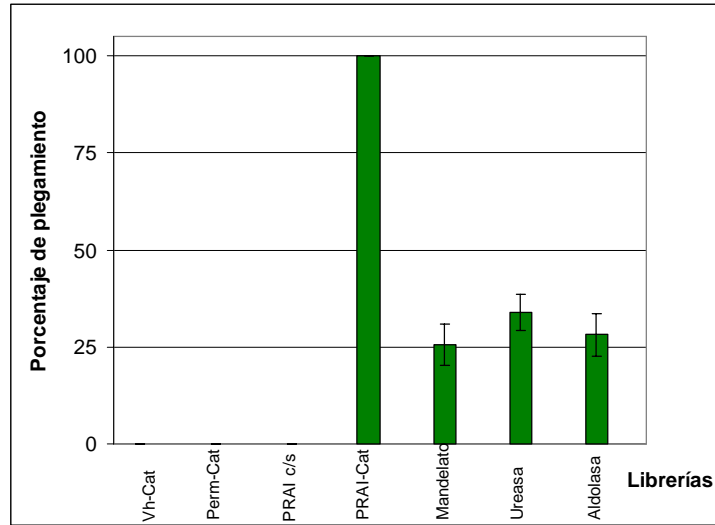


Figura 15. Esquema general del proceso de selección de plegamiento mediante crecimiento en presencia de cloranfenicol.

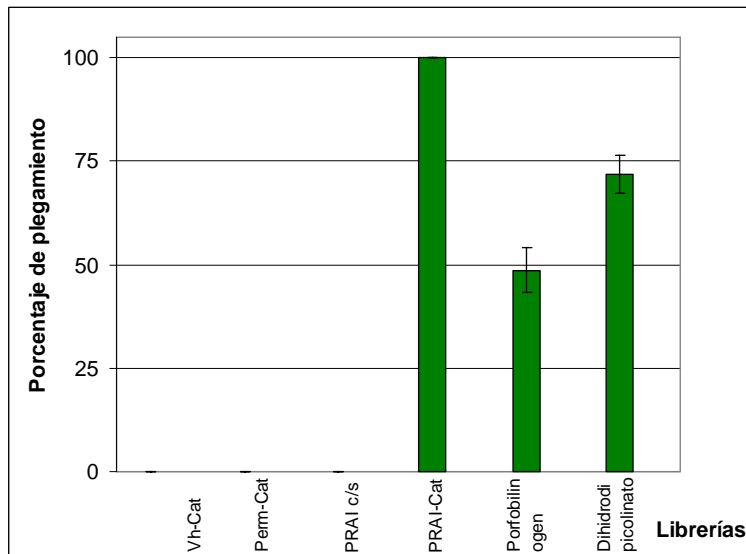
El porcentaje de plegamiento para cada banco fue calculado por el número de colonias crecidas en presencia de cloranfenicol entre el número de colonias crecidas en ausencia de este antibiótico (ver figura 15). De esta manera obtuvimos los distintos porcentajes de plegamiento para cada banco de las asas 2, 4 y 6. Además estos porcentajes se normalizaron por el número de colonias en las que su secuencia correspondía a la enzima silvestre. En las gráficas 1, 2 y 3 se muestran los resultados para las tres asas y sus respectivos bancos.



Gráfica 1. Porcentajes de plegamiento para los 8 bancos del asa 4.



Gráfica 2. Porcentajes de plegamiento para los 3 bancos del asa 2.



Gráfica 3. Porcentajes de plegamiento para los 2 bancos del asa 6.

En las gráficas 1, 2 y 3 podemos observar una ligera diferencia de porcentaje de plegamiento para los diferentes bancos, pero cada asa se comporta de manera diferente en lo que respecta al porcentaje de plegamiento y no existe una relación directa entre el tamaño o la composición de residuos del asa con el porcentaje de plegamiento o la posición de la inserción del asa. Los porcentajes de plegamiento obtenidos son bastante buenos porque nos indican que tenemos desde el 25% hasta el 70% de proteínas plegadas. Lo que abre la posibilidad de encontrar proteínas nuevas con actividades potencialmente novedosas, porque al menos mantienen la estructura.

## 5.4 Secuencias obtenidas bajo presión de selección

Para analizar qué residuos se estaban fijando en las posiciones variables con la presión de selección de plegamiento, mandamos a secuenciar clonas de cada asa, de los diferentes bancos, pero seleccionadas en presencia de 20 µg/ml de cloranfenicol. Los resultados de la variabilidad obtenida en los sitios mutados se muestran en las tablas 9, 10 y 11.

Número de la clona	Ureasa		Porfobilinogen sintasa		Adenosin deaminasa		Tiamina fosfato sintasa		Mandelato racemasa		Dihidrodipicolinato sintasa		Triptófano sintetasa		Aldolasa clase II	
	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81
1	G	R	P	S	K	S	K	G	I	H	V	G	S	A	G	H
2	Q	H	I	I	R	P	R	T	F	S	K	N	K	G	N	P
3	P	H	P	V	I	R	S	F	M	N	E	I	I	R	Q	N
4	<u>S</u>	<u>T</u>	V	P	S	C	N	D	V	L	V	T	F	G	H	L
5	<u>D</u>	<u>T</u>	I	N	V	H	S	I	H	Y	F	D	I	D	S	A
6	L	Y	<u>H</u>	<u>Y</u>	L	H	R	L	<u>T</u>	<u>Y</u>	T	Y	H	L	I	R
7	C	P	L	S	M	R	L	S	W	I	L	H	N	P	K	G
8	K	P	T	N	<u>S</u>	<u>R</u>	R	N	C	P	D	Y	Q	N	L	G
9	<u>S</u>	<u>T</u>	T	G	N	R	S	L	R	S	S	S	G	H	R	D
10	A	T	F	R	R	Y	S	N	<u>P</u>	<u>N</u>	<u>R</u>	<u>V</u>	G	D	T	N
11	<u>D</u>	<u>T</u>	C	S	K	H	V	V	W	A	P	R	T	Y	G	T
12	R	D	S	H	G	N	P	H	N	D	Y	D	V	H	R	A
13	<u>N</u>	<u>P</u>	Y	S	C	H	L	N	<u>P</u>	<u>N</u>	T	P	H	N	Y	D
14	M	A	E	C	R	C	C	N	P	T	R	P	L	L	H	V
15	T	Y	T	Y	A	V	C	D	W	S	A	N	L	D	W	T
16	<u>N</u>	<u>P</u>	<u>H</u>	<u>Y</u>	K	D	M	C	<u>T</u>	<u>Y</u>	A	C	Y	C	R	H
17	M	H	G	A	Q	T	R	S	N	A	V	I	G	S	A	P
18	P	T	S	T	<u>S</u>	<u>R</u>	R	G	P	S	<u>R</u>	<u>V</u>	I	H	G	L
19	L	H	N	T	P	D			Y	I	P	I	E	H	K	T
20			E	R							S	A	Y	P		

Tabla 9. Secuencias obtenidas con presión de selección para los residuos Q79 y H81 de los diferentes bancos del asa 4. En color naranja y subrayados, las combinaciones de residuos que se repiten

Número de la clona	Ureasa	Mandelato racemasa	Aldolasa clase II
	I29	I29	I29
1	S	R	T
2	S	R	T
3	S	R	T
4	S	R	T
5	S	R	T
6	I	R	T
7	I	R	T
8	F	A	T
9	A	A	T
10	A	A	T
11	A	A	T
12	F	G	S
13	F	D	N
14	V	D	G
15	L	D	N
16	L	D	M
17	Q	W	C
18	K	V	S

Tabla 10. Secuencias obtenidas con presión de selección para el residuo I29 de los diferentes bancos del asa 2.

Número de la clona	Porfobilinogen sintasa		Dihidrodipicolinato sintasa	
	V124	D126	V124	D126
1	V	G	<u>T</u>	<u>C</u>
2	D	H	T	I
3	T	T	L	H
4	F	T	R	T
5	K	I	C	C
6	H	C	C	Y
7	Y	L	A	L
8	L	R	G	T
9	E	C	P	T
10	Y	W	<u>T</u>	<u>C</u>
11	L	P	K	F
12	P	F	S	V
13	<u>T</u>	<u>C</u>	V	L
14	L	A	R	P
15	<u>T</u>	<u>C</u>	P	I
16	R	V	L	P
17	P	H	F	L
18	P	R	S	T

Tabla 11. Secuencias obtenidas con presión de selección para los residuos V124 y D126 de los diferentes bancos del asa 6. En color naranja y subrayados, las combinaciones de residuos que se repiten.

Uno de los primeros resultados obtenidos de las tablas anteriores, son las combinaciones repetidas de secuencias para algunos bancos (Ver tablas 12 y 13).

Diferentes librerías del asa 4									
Ureasa		Porfobilinogen sintasa		Adenosin deaminasa		Mandelato racemasa		Dihidrodipicolinato sintasa	
Q79	H81	Q79	H81	Q79	H81	Q79	H81	Q79	H81
<u>S</u>	<u>T</u>	<u>H</u>	<u>V</u>	<u>S</u>	<u>R</u>	<u>T</u>	<u>V</u>	<u>R</u>	<u>V</u>
<u>D</u>	<u>T</u>					<u>P</u>	<u>N</u>		
<u>N</u>	<u>P</u>								

Tabla 12. Secuencias obtenidas por duplicado para las posiciones Q79 y H81 en diferentes bancos del asa 4 de PRAI

Diferentes librerías del asa 6			
Porfobilinogen sintasa		Dihidrodipicolinato sintasa	
V124	D126	V124	D126
<u>T</u>	<u>C</u>	<u>T</u>	<u>C</u>

Tabla 13. Secuencias obtenidas por duplicado para las posiciones V124 y D126 en diferentes bancos del asa 6.

La probabilidad de encontrar 2 veces la misma clona dentro del tamaño de nuestro banco es de  $3.8 \times 10^{-6}$ , lo que nos habla de que existe una selección para esta combinación de residuos en esas posiciones específicas, ya que al azar es imposible encontrar la misma clona dos veces repetida en las 18 clonas que secuenciamos (como lo demuestra el hecho de que

cuando secuenciamos sin presión de selección no observamos este tipo de combinaciones de residuos).

Para los bancos del asa 4 encontramos que algunas asas realmente muestran una preferencia para la combinación de residuos específicos, mientras que otras no. Por otro lado, para los dos bancos del asa 6, se encontró la misma combinación de residuos (T-C). La preferencia por ciertas combinaciones de aminoácidos encontrada para el asa 6 independientemente del asa insertada, aporta información sobre el distinto comportamiento de la estructura del barril, dependiendo de la posición de inserción de las asas, por ejemplo para el asa 4 no encontramos combinaciones en común para todas las asas insertadas, como se observa para el asa 6, a pesar de haber insertado sólo dos asas.

Otro resultado interesante, es que en ninguna clona, de ninguna asa encontramos los residuos que PRAI tiene normalmente en estas posiciones, lo que puede implicar que estos aminoácidos no son los óptimos en el contexto de un barril con asas diferentes, o bien, que estos residuos pudieran haber sido seleccionados por un papel más funcional que estructural en la enzima PRAI silvestre.

Para determinar qué residuos se estaban seleccionando positiva o negativamente comparamos las frecuencias de aparición de cada residuo bajo presión de selección con respecto a las frecuencias de aparición en ausencia de presión de selección, de esta manera también se normaliza por la diferente representación de codones para cada aminoácido. Las diferencias entre las dos frecuencias fueron analizadas para cada sitio de mutagénesis y para cada asa. Estas diferencias las podemos observar en las gráficas 4-8 y son señalados en color rojo los residuos que con un 95% de confianza demuestran estar siendo seleccionados positivamente o negativamente.



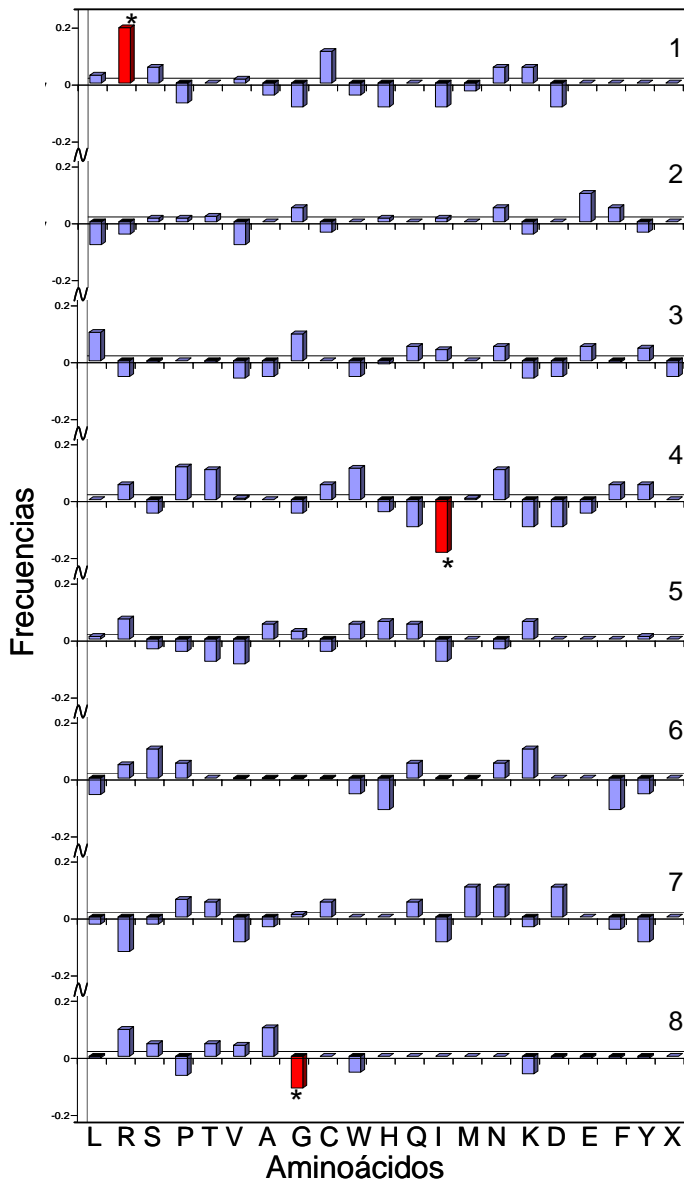


Gráfico 4. Diferencias entre frecuencias de aparición para el residuo Q79 (NNS), de los residuos obtenidos en presencia de selección con respecto a los residuos sin selección para las diferentes aspas 4. En color rojo, los residuos que tienen una selección estadísticamente significativa. 1- tiamina fosfato sintasa; 2- porfolibinogen sintasa; 3- subunidad alfa de la triptófano sintetasa; 4- mandelato racemasa; 5- aldolasa clase II; 6- adenosin desaminasa; 7- ureasa; 8- dihidrodipicolinato sintasa.

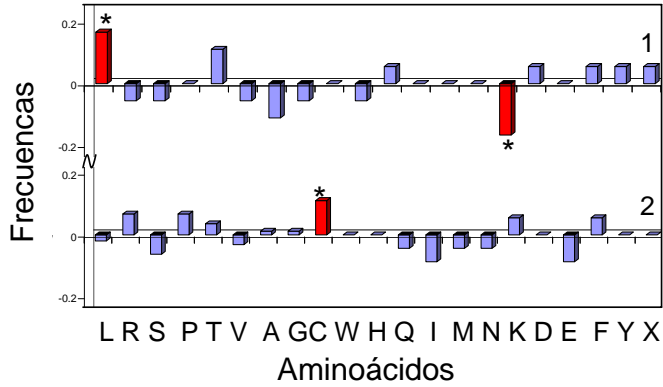


Gráfico 5. Diferencias entre frecuencias de aparición para el residuo V124 (NNS), de los residuos obtenidos en presencia de selección con respecto a los residuos sin selección para las diferentes aspas de la posición 6. En color rojo, los residuos que tienen una selección estadísticamente significativa. 1- porfobilinógeno sintasa; 2- dihidrodipicolinato sintasa.

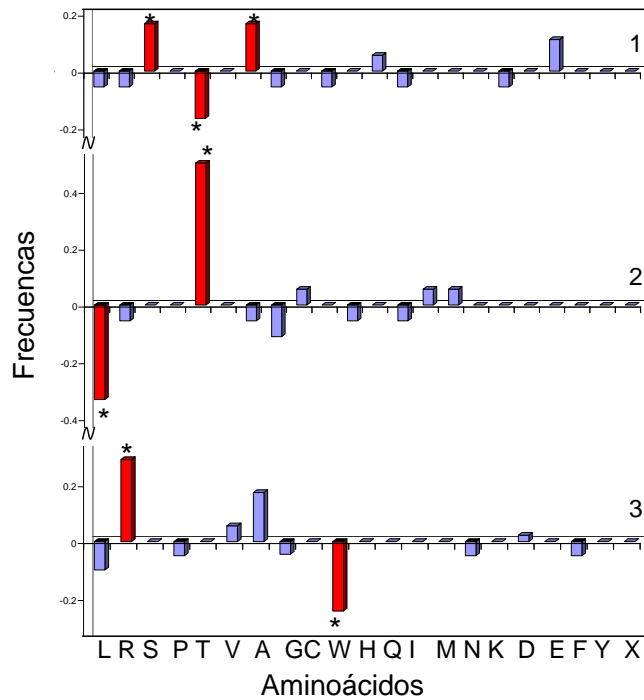


Gráfico 6. Diferencias entre frecuencias de aparición para el residuo I29 (NNS), de los residuos obtenidos en presencia de selección con respecto a los residuos sin selección para las diferentes aspas de la posición 6. En color rojo, los residuos que tienen una selección estadísticamente significativa. 1- ureasa; 2- aldolasa; 3- mandelato racemasa.

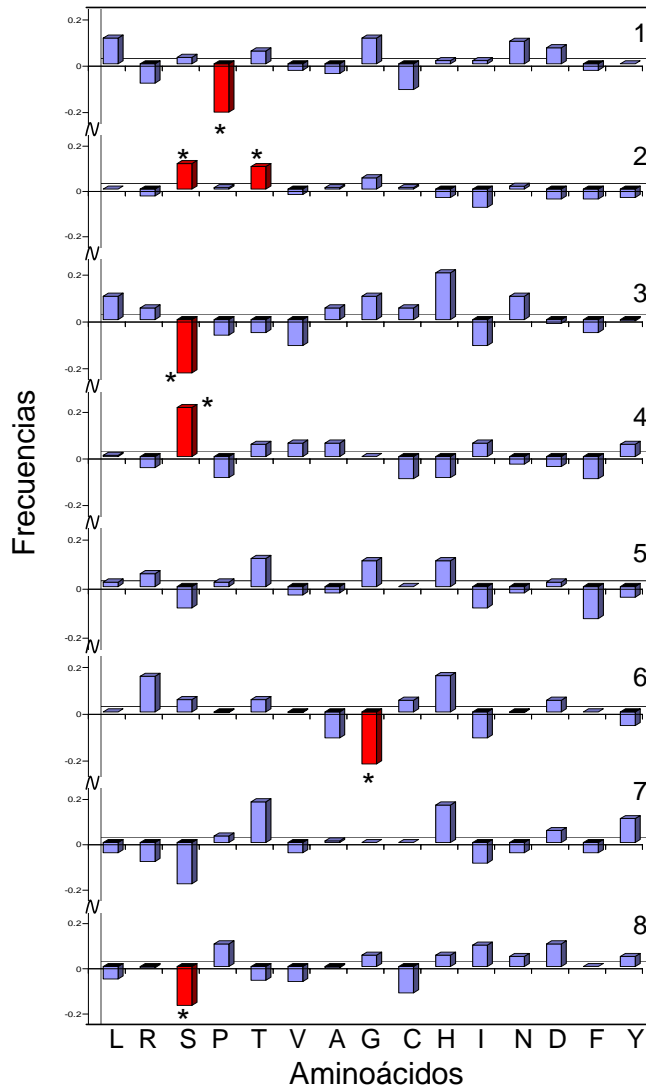


Gráfico 7. Diferencias entre frecuencias de aparición para el residuo H81 (NNT), de los residuos obtenidos en presencia de selección con respecto a los residuos sin selección para las diferentes asns 4. En color rojo, los residuos que tienen una selección estadísticamente significativa. 1- tiamina fosfato sintasa; 2- porfolibinogen sintasa; 3- subunidad alfa de la triptófano sintetasa; 4- mandelato racemasa; 5- aldolasa clase II; 6- adenosin desaminasa; 7- ureasa; 8- dihidrodipicolinato sintasa.

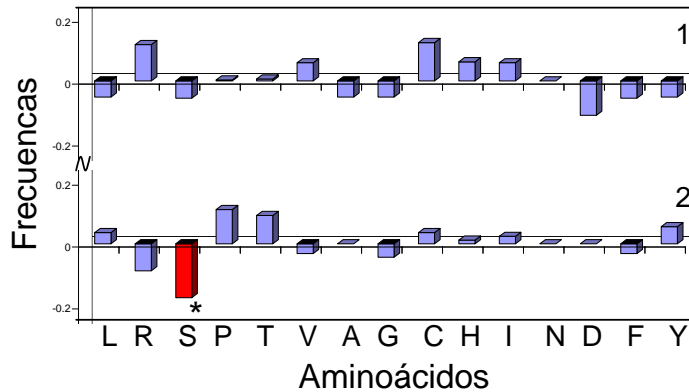


Gráfico 8. Diferencias entre frecuencias de aparición para el residuo D126 (NNT), de los residuos obtenidos en presencia de selección con respecto a los residuos sin selección para las diferentes asas de la posición 6. En color rojo, los residuos que tienen una selección estadísticamente significativa. 1- porfobilinógeno sintasa; 2- dihidrodipicolinato sintasa.

Los resultados en forma resumida de los residuos que están presentando una selección de plegamiento se muestran en las tablas 14, 15 y 16.

Enzima de la cuál proviene el asa	Q79	H81
Tiamina fosfato sintasa	R (+)	P (-)
Porfobilinogen sintasa	---	S (+), T (+)
Sub. Alfa de triptófano sintetasa	---	S (-)
Mandelato racemasa	I (-)	S (+)
Aldolasa clase II	---	---
Adenosin deaminasa	---	G (-)
Ureasa	---	---
Dihidrodipicolinato sintasa	G (-)	S (-)

Tabla 14. Selección de residuos para las posiciones Q79 y H81 del asa 4. Selección positiva (+), selección negativa (-).

Podemos observar para los bancos del asa 4 que el residuo seleccionado en cada posición depende directamente del asa insertada (tabla 14). Por ejemplo, los bancos del asa de la porfobilinógeno sintasa y la mandelato racemasa presentan una selección positiva por serina, mientras que este residuo es incompatible en el contexto de las asas correspondientes de la subunidad alfa de triptófano sintetasa y de la dihidrodipicolinato sintetasa. Ninguna de estas asas tienen parecido alguno entre sí, ni de tamaño, ni de características fisicoquímicas, lo cual corrobora la selección dependiendo del asa.

<b>Enzima de la cuál proviene el asa</b>	<b>V124</b>	<b>D126</b>
Porfobilinogen sintasa	L (+), K (-)	---
Dihidrodipicolinato sintasa	C (+)	S (-)

Tabla 15. Selección de residuos para las posiciones V124 y D126 del asa 6. Selección positiva (+), selección negativa (-).

Para los bancos del asa 6 (tabla 15), observamos un comportamiento similar al del asa 4, es decir los residuos seleccionados positiva o negativamente dependen del asa insertada en esa posición y no del sitio de inserción. Pero si existen combinaciones de residuos que son seleccionadas, aunque los residuos de cada posición por sí solos no hayan sido seleccionados. Por ejemplo, para los dos bancos del asa 6 encontramos que se selecciona la combinación (T-C), mientras que estos residuos no se encuentran seleccionados de manera independiente.

<b>Enzima de la cuál proviene el asa</b>	<b>I29</b>
Ureasa	S (+), A (+), T (-)
Aldolasa clase II	L (-), T (+)
Mandelato racemasa	R (+), W (-)

Tabla 16. Selección de residuos para la posición I29 del asa 2. Selección positiva (+), selección negativa (-).

Para los bancos del asa 2 observamos el mismo comportamiento que para las demás asas, es decir el residuo seleccionado depende directamente del asa insertada. Por ejemplo, mientras que para el asa insertada de la ureasa, la treonina se selecciona negativamente, para el asa de la aldolasa clase II se selecciona positivamente.

Por otro lado el asa de la aldolasa clase II presentó una alta presión de selección de treonina con una frecuencia de aparición por arriba del 0.4. Siendo esta asa la única que tuvo una fuerte presión a un residuo en particular.

## 5.5 Análisis *in vitro* del plegamiento (“Western-blot”)

Para confirmar los resultados obtenidos por el método de selección *in vivo* a través de la resistencia a cloranfenicol, llevamos a cabo un análisis de “western blot” con un anticuerpo contra CAT de diferentes bancos para analizar la presencia *in vitro* de las proteínas fusionadas a CAT en la fracción soluble del extracto celular. En la figura 16 mostramos el resultado de los 3 diferentes “western blot” para tres diferentes bancos.

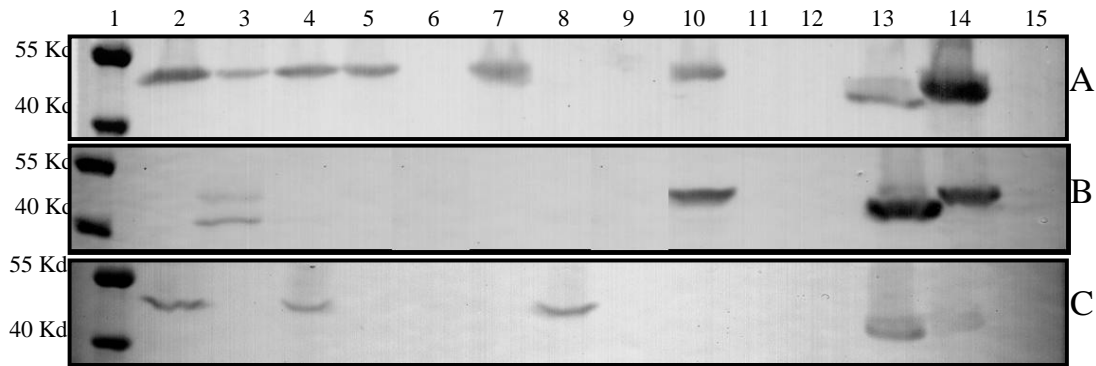


Figura 16. Análisis de “western blot” para diferentes bancos. (A) Banco 4 de ureasa. (B) Banco 2 de mandelato racemasa. (C) Banco 6 de porfobilinógeno sintasa. Carriles: 1- Marcador de peso molecular, 2-11 diez colonias tomadas al azar de cada banco, 12- CAT, 13- PRAI WT en fusión con CAT, 14- clona seleccionada de cada banco en presencia de cloranfenicol (20 ug/ml), 15 PRAI Cre-lox WT  
Pesos moleculares: CAT = 22 Kda; PRAI = 23 Kda; Fusión = 45 Kda.

Es importante señalar que la banda correspondiente a PRAI WT en fusión con CAT de la figura 16 se observa por debajo del peso molecular de las diferentes variantes seleccionadas, debido a que las variantes fueron construidas en la PRAI que tiene insertado el Cre-Lox, lo cuál aumenta su tamaño. En la sección de anexos (anexo 4) pueden observarse las imágenes completas de cada “western blot”.

Las clonas utilizadas para el “western blot” fueron secuenciadas (ver tabla 17 y 18) y con ello se determinó el porcentaje de plegamiento reportado por la presencia de nuestras quimeras en fusión con CAT. Los resultados de las secuencias para las 10 colonias del asa 4 del banco de la ureasa demostraron que las 10 clonas tenían una secuencia en buen marco de lectura y correspondían a nuestras proteínas quiméricas fusionadas a CAT, por lo tanto, como observamos presencia de fusión en 6 de las 10 clonas, esto significa que tenemos un 60% plegado en ese banco por el análisis de “western blot”.

Para el caso del banco del asa 2 de mandelato racemasa sólo 7 secuencias corresponden a nuestras quimeras, por lo tanto como se observa sólo la fusión de 2 clonas, esto representa el 28% de proteínas plegadas para ese banco. Las otras 3 secuencias resultaron ser monstruos, productos del proceso de construcción del gen con el asa insertada.

Del banco del asa 6 de porfobilinógeno sintasa sólo 7 secuencias corresponden a nuestras quimeras, por lo tanto como sólo se observa la fusión de 3 clonas, esto representa el 43 % de plegamiento. Las secuencias de las 3 clonas faltantes de las 10 analizadas para esta asa, mostraron ser genes mal construidos.

Las secuencias de las clonas encontradas en la fracción soluble en forma de fusión de los diferentes bancos se ilustran en la tabla 17, mientras que las secuencias de las clonas en donde no se observó una fusión con CAT se ilustran en la tabla 18.

Secuencias de las diferentes librerías					
Nombre de la clona	Asa 4 de ureasa	Asa 6 de porfobilinogen sintasa		Asa 2 de mandelato racemasa	
	<b>Q79</b>	<b>H81</b>	<b>V124</b>	<b>D126</b>	<b>I29</b>
1	H	H	A	L	W
2	R	S	A	Y	R
3	L	S	V	H	
4	S	N			
5	G	A			
6	N	H			

Tabla 17. Secuencias de las diferentes posiciones mutadas que presentaron fusión a CAT de manera soluble.

Secuencias de las diferentes librerías					
Nombre de la clona	Asa 4 de ureasa	Asa 6 de porfobilinogen sintasa		Asa 2 de mandelato racemasa	
	<b>Q79</b>	<b>H81</b>	<b>V124</b>	<b>D126</b>	<b>I29</b>
1	N	H	H	L	N
2	V	G	N	L	F
3	P	N	X	H	F
4	H	V	T	H	

Tabla 18. Secuencias de las diferentes posiciones mutadas que no presentaron fusión a CAT

Como había de esperarse, ninguna secuencia de las clonas que no presentaron fusión a CAT se encontró en las secuencias obtenidas bajo presión de selección en cloranfenicol (ver tablas 9, 10 y 11). Esto corrobora que el sistema de selección de plegamiento funciona correctamente para seleccionar variantes plegadas en fusión a CAT.

Es interesante señalar, que de las 7 secuencias correspondientes a las construcciones del asa 2, dos secuencias correspondieron a las clonas observadas en la fracción soluble (ver cuadro 16). De las otras 5 secuencias restantes, 3 secuencias fueron las que se encontraron sin fusión a CAT (ver cuadro 17) y las otras 2 que tampoco mostraron fusión a CAT resultaron tener el marco de lectura cambiado y por ello no selecciono algún residuo.

## 5.6 Sobreexpresión de variantes sin fusión a CAT

Para demostrar que nuestras variantes permanecían solubles cuando les quitábamos la fusión de CAT, seleccionamos 5 variantes seleccionadas en presencia de cloranfenicol y procedimos a hacer las construcciones correspondientes para sobreexpresarlas y a purificarlas de la fracción soluble mediante afinidad a níquel. En la figura 17 podemos observar los resultados de las proteínas encontradas en la fracción soluble, después del proceso de purificación.

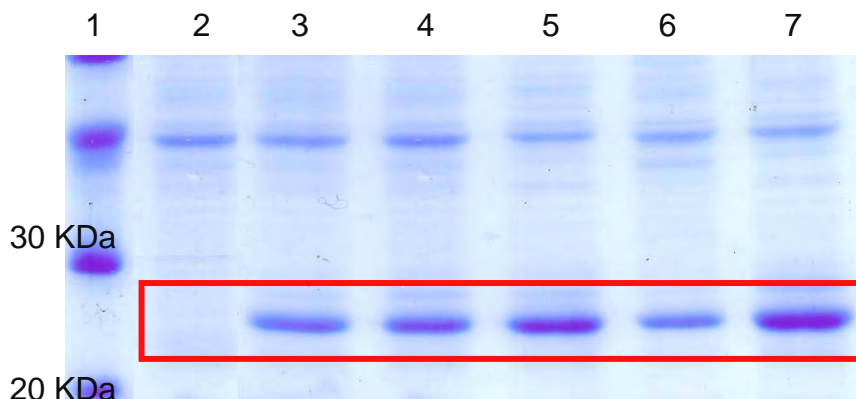


Figura 17. Fotografía de un gel de SDS con las fracciones solubles de las siguientes proteínas: 1- Marcador de peso molecular; 2- Cepa con plásmido vacío, 3- PRAI Cre-lox WT; 4- PRAI con el asa 4 de ureasa; 5- PRAI con el asa 6 de la subunidad alfa de la triptófano sintetasa; 6- PRAI con el asa 6 de la porfobilinógeno sintasa; 7- PRAI con el asa 2 de aldolasa clase II. El recuadro de color rojo muestra a las proteínas de interés.

Como podemos observar en la figura 17, obtuvimos a nuestras 4 variantes en forma soluble en el extracto celular, lo cual es interesante porque tenemos proteínas que tienen nuevas asas insertadas y demuestran estar en forma soluble. Esto habla de la capacidad del plegamiento de la enzima PRAI wt Cre-lox para soportar la inserción de asas manteniendo su estructura. Además estas 4 variantes fueron seleccionadas al azar y ninguna fue insoluble, esto refuerza la idea de que lo reportado en los demás bancos son ciertamente proteínas solubles. Pero para obtener un análisis más fino de los cambios estructurales por la inserción de asas en PRAI, procedimos a purificar 4 variantes para ser analizadas mediante dicroísmo circular.

## 5.7 Purificación de proteínas para análisis de dicroísmo circular.

Se purificaron 4 variantes, dichas variantes son: PRAI wt Cre-lox, una variante con el asa 4 de la subunidad alfa de la triptófano sintetasa, una variante con el asa 2 de la aldolasa clase II y una última variante con el asa 6 de porfobilinógeno sintasa.

La proteína PRAI Cre-lox WT fue purificada primero por una columna de níquel y posteriormente por exclusión molecular. En la figura 18 (A) se observa el cromatograma resultado de la purificación por la columna de níquel.



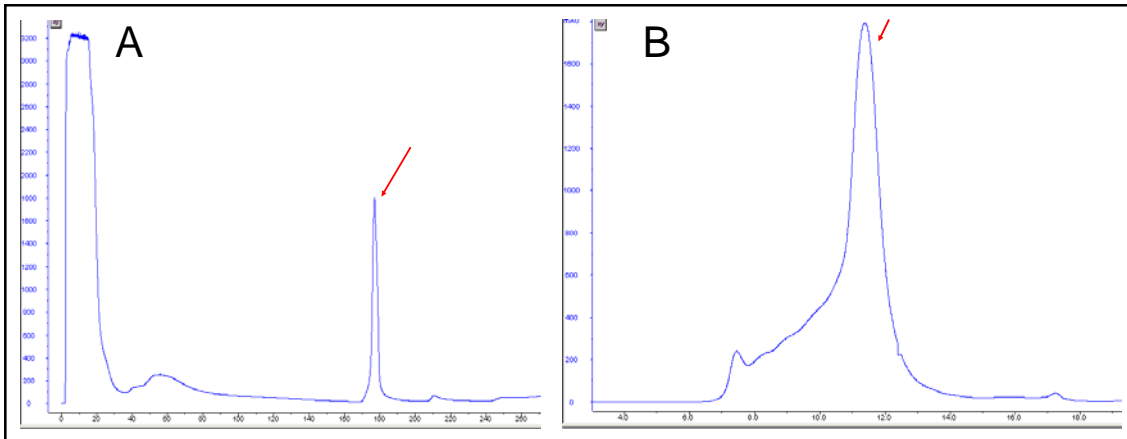


Figura 18. Cromatogramas de purificación de PRAI Cre-lox WT. (A) Purificación por afinidad a níquel. (B) Purificación por permeación en gel en la columna Superosa 12. En las líneas azules se indica la absorbancia a 280 nm, correspondiente a la concentración de proteína. Los picos señalados con una flecha en color rojo corresponden a la proteína PRAI Cre-lox WT.

Observamos que nuestra proteína se unió muy ligeramente a la columna debido a que eluyó con una baja concentración de imidazol (25 mM) y por lo tanto se encuentra aún contaminada con otras proteínas. Por lo cual procedimos a purificarla por medio de una columna de exclusión molecular. Para ello utilizamos una columna de permeación en gel de baja definición (Superosa 12). El resultado de la purificación por esta columna lo podemos observar en la figura 18 (B).

Debido a que nuestra proteína todavía presentaba otras proteínas que la contaminaban, procedimos a llevar a cabo un tercer paso de purificación utilizando otra columna de exclusión molecular de alta definición (Sephacryl S200). Los resultados obtenidos de esta purificación se muestran en la figura 19 (A). Como resultado obtuvimos finalmente a nuestra proteína con un nivel alto de pureza (ver figura 20).

Ya estandarizado el proceso de purificación para la proteína silvestre (PRAI Cre-lox WT), procedimos a purificar las demás variantes, primero por afinidad a níquel y posteriormente por exclusión molecular. La purificación por afinidad a níquel la llevamos a cabo por medio de “batch” utilizando la columna empacada con sefarosa para flujo rápido de Amersham (ver protocolo 18). De esta manera colectamos de cada mutante las fracciones correspondientes a los “buffers” A y B (10 y 20 mM de imidazol respectivamente) debido a que ya sabíamos que en estas fracciones eluye nuestra proteína. Posteriormente la concentramos y las inyectamos en la columna Sephacryl S200. Los cromatogramas de los resultados obtenidos por la columna de exclusión molecular para las diversas variantes se muestran en la figura 19 secciones B, C y D.

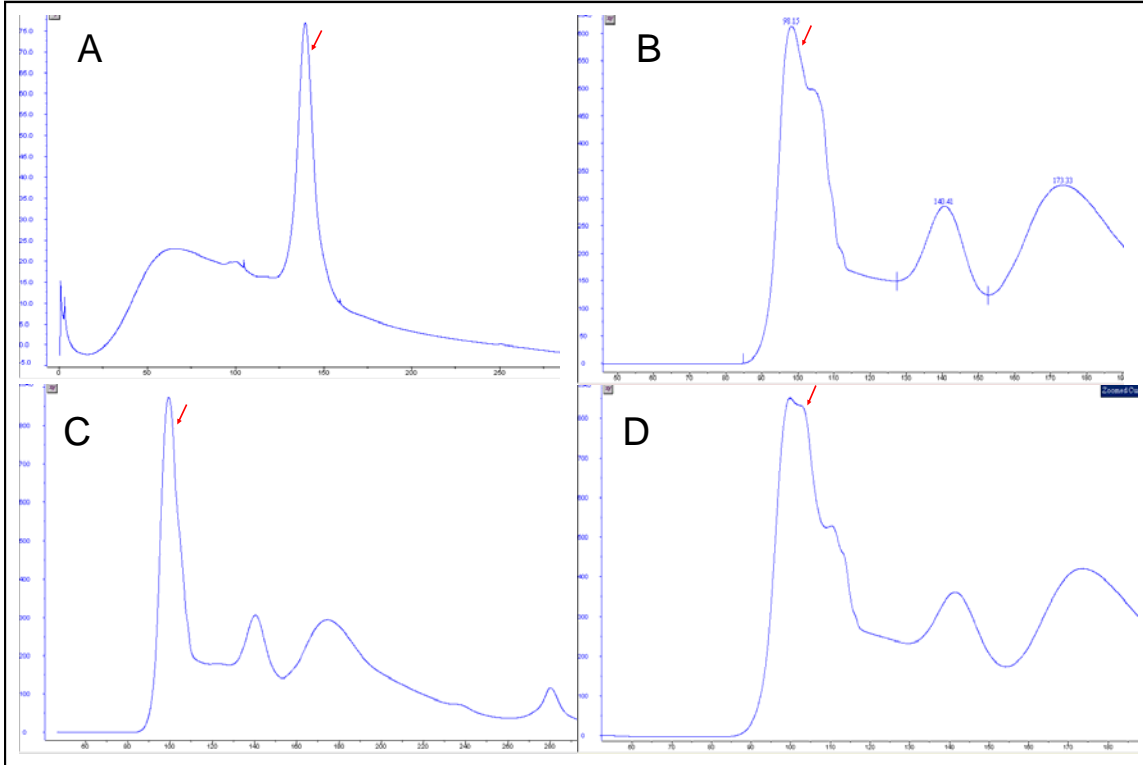


Figura 19. Cromatogramas de purificación de PRAI y variantes en la columna Sephacryl S200. (A) PRAI Cre-Lox wt (B) variante con el asa 4 de la subunidad alfa de la subunidad alfa de la triptófano sintetasa. (C) variante con el asa 2 de la aldolasa clase II (D) variante con el asa 6 de la porfobolinogeno sintetasa. En línea azul se indica la absorbancia a 280 nm, correspondiente a la concentración de proteína. El pico señalado con una flecha de color rojo corresponde a las respectivas proteínas.

Se pudo observar que todas las proteínas eluyeron en el mismo volumen de “buffer” en exclusión molecular y los cromatogramas fueron muy similares. Posterior a esta purificación procedimos a analizar por un gel de acrilamida (“SDS-PAGE”) la pureza y la presencia de nuestras proteínas. En la figura 24 podemos observar a las cuatro proteínas después de todos los pasos de purificación y que son finalmente las que utilizaremos para el análisis por dicroísmo circular.

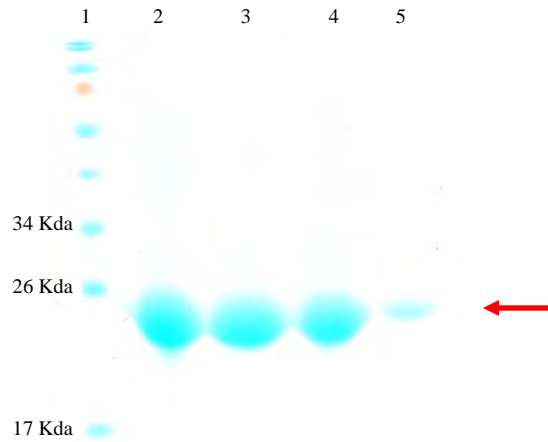


Figura 20. Gel de SDS, donde se muestran nuestras proteínas al final de todos los pasos de purificación. La flecha roja indica la presencia de nuestras proteínas. Carril 1-Marcador de peso molecular. Carril 2-Variante de PRAI Cre-Lox con el asa 6 de la porfobilinógeno sintasa. Carril 3-Variante de PRAI Cre-Lox con el asa 4 de la subunidad alfa de la triptófano sintetasa. Carril 4-Variante de PRAI Cre-Lox con el asa 2 de la aldolasa clase II. Carril 5-PRAI Cre-lox WT.

## 5.8 Análisis de dicroísmo circular

Los resultados de los espectros obtenidos del dicroísmo circular se muestran en la figura 21 y representan el promedio de seis mediciones para cada una de las variantes.

Como resultado de los espectros de dicroísmo circular, podemos observar que si bien, existe un cambio estructural en nuestras variantes con respecto a la proteína silvestre, dicho cambio no es muy drástico. Estos espectros tomados en el UV-lejano muestran que la variante que conserva más su estructura con respecto a la proteína silvestre es la que tiene insertada el asa 4 de la subunidad alfa de la triptófano sintetasa mientras que la variante que sufre los mayores cambios estructurales es la proteína que tiene insertada el asa 6 de porfobilinógeno sintasa. Mientras que la proteína que tiene insertada el asa 2 de la aldolasa clase II mostró las menores perturbaciones estructurales. Pero básicamente todos los espectros son superponibles entre sí.

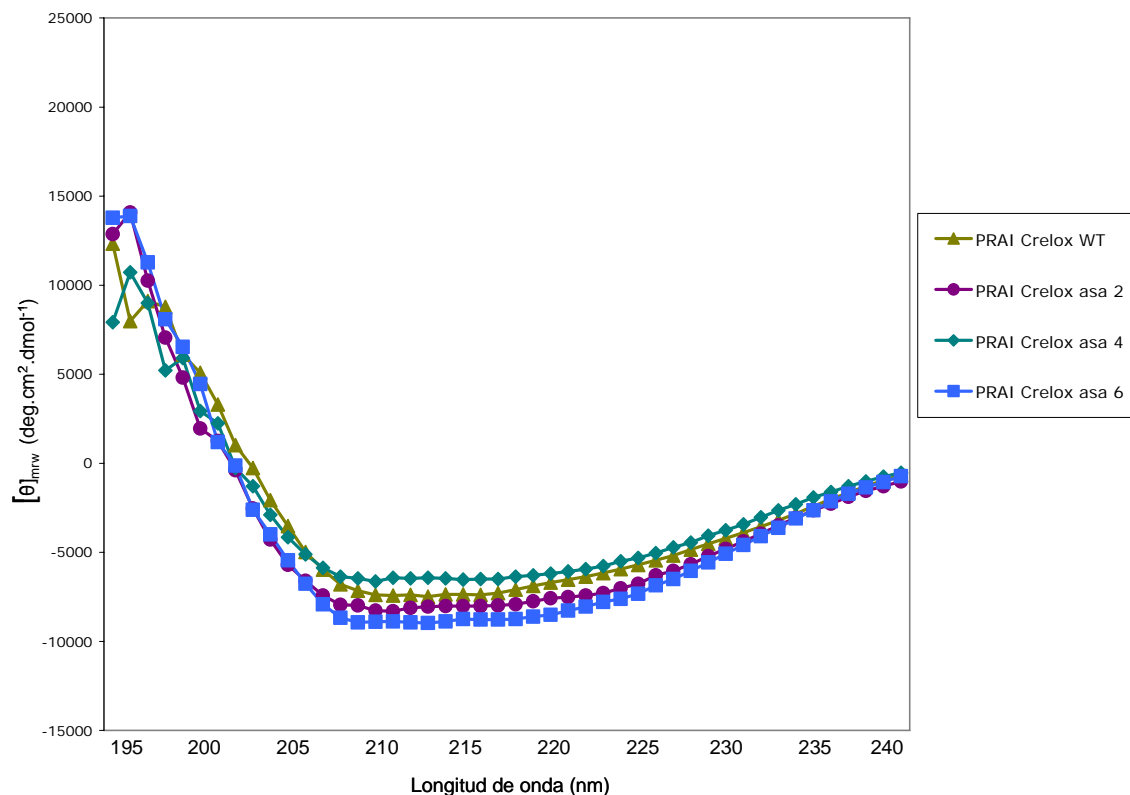


Figura 21. Análisis por dicroísmo circular (CD) en el UV-lejano. La proteína silvestre es denominada (PRAI Cre-lox WT), la variante que tiene insertada el asa 2 de la aldolasa clase II (PRAI Cre-lox asa 2), la variante que tiene insertada el asa 4 de la subunidad alfa de la triptófano sintetasa (PRAI Cre-lox asa 4) y la variante que tiene insertada el asa 6 de porfobilinógeno sintasa (PRAI Cre-lox asa 6).

Con los datos obtenidos del dicroísmo circular procedimos a realizar una predicción de estructura secundaria utilizando el programa CDSSTR. Con este programa obtuvimos los resultados presentados en la tabla 19.

	% de alfa hélices	% de hojas beta	% desordenado	% Total
PRAI Crelox WT	14	35	50	99
PRAI Crelox asa 2	28	28	44	100
PRAI Crelox asa 4	14	38	49	101
PRAI Crelox asa 6	27	26	47	100

Tabla 19. Predicción de estructura secundaria para cada una de las variantes de PRAI

La mutante con la inserción del asa 4 de la subunidad alfa de la triptófano sintetasa, es la que menos cambios estructurales produjo al ser comparada con el contenido de hojas beta y alfa hélices de la enzima silvestre (PRAI Cre-lox WT). La mutante del asa 2 aumentó considerablemente el contenido de alfa hélices y disminuyó el contenido de hojas beta, así como

el porcentaje de estructura desordenada. Interesantemente en la estructura tridimensional de la enzima aldolasa clase dos (PDB: 1dos) el asa dos se encuentra formando una hélice, lo que podría explicar el aumento en contenido de alfa hélices para esta mutante, esto abre la posibilidad de que en PRAI esta asa este adoptando también la estructura de alfa hélice.

Por otro lado, la mutante del asa 6 también aumento su contenido de alfa hélices y disminuyó la proporción de hojas beta lo que nos indica un cambio conformacional mayor al provocado por la inserción del asa 4. Curiosamente el asa insertada es muy pequeña (solo está compuesta por 5 aminoácidos) comparada con el asa 6 de PRAI que está compuesta por 12 residuos (ver tabla 4). En la estructura de la enzima porfobilinogen sintasa (PDB: 1l6s) esta asa no posee ningún tipo de estructura secundaria. En este sentido no explicaría porque en PRAI provoca un aumento en la composición de alfa hélices, aunque probablemente esté contribuyendo a que la hélice 6 de PRAI (que sólo tiene una vuelta de hélice) se propague, formando así una hélice más larga y ello explique el aumento de este tipo de estructura. Sin embargo, esta ganancia en contenido helicoidal es a expensas del contenido de hebras  $\beta$  y no de proteína no estructurada.

El asa 4 de la subunidad alfa de la triptófano sintetasa fue la que menos cambios estructurales ocasionó. Probablemente debido a que tiene una estructura, longitud y secuencia muy parecida al asa 4 de PRAI y por eso se acomodó estructuralmente igual en esta posición. Además cabe resaltar que hay fuertes evidencias estructurales de la relación evolutiva entre estas dos proteínas [98].

## 6. DISCUSION

La distribución de residuos encontrada en las posiciones mutadas a saturación sin presión de selección muestra estar de acuerdo con el número de codones representados en el diseño de los oligonucleótidos para cada aminoácido. Lo cual es un indicativo de que los bancos no están sesgados hacia determinados residuos como resultado del proceso de construcción de los mismos o de una mala síntesis de los oligonucleótidos.

El diseño de los controles negativos para el método de selección de plegamiento nos permitió evaluar la veracidad del mismo como reportero de proteínas plegadas, además esto nos permitió analizar el comportamiento de las diferentes fusiones en distintas cepas de *E. coli*, lo que nos llevó a evaluar con mayor precisión la cantidad de proteínas plegadas de cada banco, así como la concentración adecuada de cloranfenicol para obtener proteínas plegadas.

Inicialmente probamos el sistema de fusión a CAT en la cepa XL1-Blue, pero en esta cepa obtuvimos muchos falsos positivos. Por ejemplo, la fusión de PRAI con CAT pero con codones de paro entre la fusión crecía hasta altas concentraciones de cloranfenicol, lo que significaba que CAT se estaba transcribiendo a pesar de estar en un marco de lectura incorrecto. También, las construcciones de las proteínas insolubles en fusión con CAT mostraban crecer en presencia de altas concentraciones de cloranfenicol. Además, obteníamos de nuestros bancos que sólo alrededor del 1% de la población estaba de forma plegada. Esto nos llevó a probar diversas cepas y encontramos que la cepa MC1061 $\Delta$ thiE aportaba los mejores resultados. En esta cepa todos los controles funcionaban de manera correcta por lo que procedimos a hacer todos nuestros análisis de plegamiento utilizando esta cepa. Probablemente las diferencias genéticas entre estas dos cepas ocasionaban que una funcionara de manera correcta mientras que otra no funcionara adecuadamente. Recientemente Sterner y colaboradores publicaron un trabajo donde utilizaron el sistema de fusión a CAT para obtener proteínas solubles de grandes bancos, y curiosamente el 95% de las variantes que ellos seleccionan como plegadas (en presencia de 350 u  $\mu$ g/ml de cloranfenicol) son falsos positivos [108]. Analizando más a fondo en su trabajo nos dimos cuenta que ellos utilizaron la cepa XL1-Blue para seleccionar plegamiento, lo que corrobora los resultados obtenidos por nosotros en la misma cepa. De tal manera que al cambiar de cepa obtuvimos muy buenos resultados y sin ningún falso positivo, de hecho todas nuestras proteínas seleccionadas bajo este sistema y expresadas sin CAT fueron solubles, contrastando los resultados obtenidos en su trabajo.

En PRAI los residuos que preceden a cada asa y que en este trabajo fueron mutados a saturación están altamente conservados, probablemente debido a que en este barril tienen un compromiso más funcional que estructural, esto debido a que de todos los bancos construidos

sólo encontramos una clona (del asa 4 de la ureasa, ver tabla 8) que tenía la combinación de residuos silvestres en las posiciones mutadas. Además en la diversidad de proteínas plegadas que tenemos encontramos una gran variabilidad de residuos en estas posiciones, aunado a ello fue posible el intercambio de asas totalmente diferentes a las asas silvestres manteniéndose la estructura de PRAI.

Los porcentajes de plegamiento observados para los diferentes bancos varían incluso entre asas de la misma posición, lo que hace imposible correlacionar este porcentaje con características generales de las asas insertadas, como tamaño o composición fisicoquímica.

Interesantemente una predicción de estructura secundaria con el algoritmo PSIPRED [153], a partir de la secuencia de la proteína que tiene el asa 2 de mandelato racemasa, predijo que ciertamente la región que ganaba hélices alfa era precisamente el asa insertada. Además esta asa en su contexto original esta formando precisamente una hélice, esto sugiere que el asa intercambiada probablemente conservó su estructura en este nuevo molde proteico.

## 7. CONCLUSIONES

Durante nuestro proyecto de investigación, construimos variantes de PRAI Cre-lox WT a las cuales les reemplazamos las asas 2, 4 y 6 por las mismas asas de otras proteínas que conservan el plegamiento de barril TIM, pero que realizan distintas funciones enzimáticas, además mutamos a saturación las posiciones que preceden a cada asa para mejorar su inserción en el molde de PRAI.

La inserción de sitios de mutagénesis “bisagras” permitió el correcto posicionamiento de cada asa, debido a que bajo presión de selección de plegamiento solo una clona del asa 4 de la ureasa mostró conservar los residuos que la proteína silvestre mantenía en esas posiciones y además cada asa seleccionó diferentes residuos para adaptarse mejor a la estructura de PRAI.

El método de selección genética de plegamiento y el análisis de “western blot” muestran valores en porcentaje de plegamiento muy similares para los bancos analizados lo que valida nuestro sistema para obtener proteínas plegadas.

La fusión a CAT de los distintos bancos demostró tener desde un 25% hasta un 75% de proteínas plegadas, lo cual representa una gran población de proteínas que pueden ser analizadas en búsqueda de nuevas funciones.

La gran variabilidad encontrada en los sitios mutados a saturación nos indica la gran flexibilidad del sitio activo del barril TIM, lo que sugiere que muy probablemente el barril TIM está evolucionado para permitir grandes cambios en la cara catalítica permitiendo así la gran diversidad de funciones que hoy encontramos en este plegamiento.

Todas las proteínas analizadas sin fusión a CAT demostraron ser solubles y mantener estructura, demostrando que este sistema de fusión a CAT es un buen reportero de plegamiento para barriles TIM. La estructura de la enzima PRAI demostró que este barril es capaz de soportar la inserción del sitio Cre-lox y además la inserción de diferentes asas manteniendo su estructura.

El alto porcentaje de proteínas plegadas producto del intercambio de asas en la cara catalítica de PRAI así como la amplia variabilidad encontrada en los residuos mutados a saturación confirman la versatilidad del plegamiento del barril TIM para dirigir la mutagénesis a estas áreas específicas. La mutagénesis dirigida de esta forma, nos permite ampliar la exploración del espacio de secuencia de una proteína, sin alterar la integridad del plegamiento, lo cual es imposible de explorar por las estrategias tradicionales de mutagénesis generando una



nueva herramienta de evolución dirigida para investigar nuevas actividades catalíticas en barriles TIM.

Este trabajo sugiere que las asas catalíticas de los barriles TIM están diseñadas y han sido evolucionadas dentro del contexto del barril TIM para ser intercambiadas, soportando la idea de que en las primeras etapas de evolución, la recombinación de pequeños fragmentos como lo constituyen las asas jugaron un papel importante en la diversificación de funciones en barriles TIM manteniendo la estructura del barril pero variando las asas, además permitiendo la inserción de dominios más grandes en estas asas.

Nuestro trabajo fundamenta la teoría de que ancestralmente existían proteínas homólogas que tenían diferentes secuencias entre ellas precisamente en las asas y que mediante recombinación se fueron intercambiando las asas (que eran las regiones no homólogas entre sí) diversificando las funciones de las proteínas originales, tal como recientemente ha sido sugerido por Tawfik, como un mecanismo para diversificar funciones enzimáticas [86] y como lo demuestra nuestro trabajo dicho acercamiento puede ser aplicado en barriles TIM.

Este trabajo constituye el primer trabajo reportado donde se involucre una combinatoria de asas en barriles TIM proponiendo una nueva estrategia de exploración de secuencia en búsqueda de nuevas actividades catalíticas.

## 8. ANEXOS

### Anexo 1. Secuencias de oligonucleótidos

Tabla 20. Tabla de oligonucleótidos que codifican las asas y los conectores

Nombre	Secuencia (5'---3')	Asa correspondiente
L4MaRCod	GAG CCG ACG CTT CAA CAC GAC	Asa 4 codificante de mandelato racemasa
L4MaRNoCod	TC GTG TTG AAG CGT CGG CTC A	Asa 4 no codificante de mandelato racemasa
L4ThiCod	TTG GGG CAG GAA GAT TTG CAC	Asa 4 codificante de tiamina sintasa
L4ThiNoCod	TG CAA ATC TTC CTG CCC CAA A	Asa 4 no codificante de tiamina sintasa
L4AldCod	GAC CTG TCT GAA GAA TCC	Asa 4 codificante de aldolasa
L4AldNoCod	GA TTC TTC AGA CAG GTC A	Asa 4 no codificante de aldolasa
L4UreCod	GAA GAC TGG GGA GCT ACC	Asa 4 codificante de ureasa
L4UreNoCod	GT AGC TCC CCA GTC TTC A	Asa 4 no codificante de aldolasa
L4NHCod2	TTA TCG CTG GCG GCA GTG NNS CTG NNT	Conector amino codificante
L4NHNoCod	NN CAG SNN CAC TGC CGC CAG CGA TAA CAC CTT AGC TTT G	Conector amino no codificante
L4ADACod	GGT GAT GAA CTT GGT TTC CCG GGA AGT CTG TTC	Asa 4 codificante de adenosin desaminasa
L4ADANoCod	AA CAG ACT TCC CGG GAA ACC AAG TTC ATC ACC A	Asa 4 no codificante de adenosin desaminasa
L4AlevCod	GCC GCG ATG GAC GGC	Asa 4 codificante de aminolevulínico sintasa
L4AlevNoCo	CC GTC CAT CGC GGC A	Asa 4 no codificante de aminolevulínico sintasa
L4DiPCod	CCT TAC TAC AAT CGT CCG TCC	Asa 4 codificante de dihydrodipicolinato sintasa
L4DiPNoCod	GA CGG ACG ATT GTA GTA AGG A	Asa 4 no codificante de dihydrodipicolinato sintasa
L4TrpSCod	GAT-GTG-CCA-GTT-GAA- GAG-TCC	Asa 4 codificante de la subunidad alfa de la triptófano sintasa
L4TrpSNoCo	GAC-TCT-TCA-ACT-GGC-ACA- TCA	Asa 4 no codificante de la subunidad alfa de la triptófano sintasa
L4CONoCod	CAG-CGT-ATC-GAT-ATA-CAG- CTG-G	Conector carboxilo no codificante
L4COCodbis	CAG-CTG-TAT-ATC-GAT-ACG- CTG-CGT-GAA-GCT-CTG- CCA-GCA	Conector carboxilo codificante

Tabla 21. Tabla de oligonucleótidos utilizados en el proyecto.

<b>Nombre</b>	<b>Secuencia (5'---3')</b>
Hind3-Bivi	A TGA CAG TCC GAA GCT TCA GGA GGG GTG TTG ATG
Nhe1-Bivi	ATT GGT TTG CCG CTA GCT CAT TAA TAT GCG CG
Pet28HINDIII	CTG CGA AAG CTT TCA TTA GTG GTG GTG GTG GTG GTG GGA TCC ATA TGC GCG CAG
Pet28NcoI	TAT ATA CCA TGG GGG AGA ATA AGG TAT GTG GC
pdanCodBAM-	TGG CTT TAA TGA GGG TCC ATT CGT TTG TGA
pdanNoCodBAM-	TCA CAA ACG AAT GGA CCC TCA TTA AAG GCA
NheI rev CAT	T GCC GCT AGC TTA GTG GTG GTG GTG GTG GTG AGA TCT TTA
BamHI sinSTOPPR	T GCC GGA TCC ATA TGC GCG CAG CGT CTG GAA A
p-Linker-A PRAI	TAC GAT CCC TCT GGC GGG CTC GAG AAT AAA GTA TGT
p-Linker-B PRAI	CCC GCC AGA GGG ATC GTA ATA TGC GCG CAG CGT C
Bamh1 Permutaci	T GCC GGA TCC GTG ATC CGA TAA CAC CTT AGC TTT G
HinD3 Permutaci	AG TCC GAA GCT TCA GGA GGG GTG TTG ATG CGC GAT CTG GCA GTC CAA CTG CAT GG
HIND3-VH-pDAN	AG TCC GAA GCT TCA GGA GGG GTG TTG ATG GAC GTC CAG CTC CAG CAG TCT GGA
BAMH1-SS-VH-pda	T GCC GGA TCC ACT AGT GAC AGT GAC CAG AGT ACC TTG
BamHI PRAI	T GCC GGA TCC TCA TTA ATA TGC GCG CAG CGT C

## **Anexo 2. Protocolos de Biología Molecular**

### **Protocolo 1- Purificación de oligonucleótidos**

La purificación de oligos se llevo a cabo en geles de poliacrilamida al 20%, 8M de Urea. Cada gel fue corrido a 200 volts en una fuente de poder Bio-Rad por aproximadamente 2 horas, posteriormente se visualizaron las bandas de DNA correspondientes al tamaño de cada oligo con una lámpara de luz ultravioleta. Estas bandas son cortadas e introducidas en un tubo Eppendorf de 1.5 ml y se les realizó el siguiente procedimiento:

- Agregar 1 ml de "buffer" de elusión para oligos e incubar a 80°C con agitación por 4 horas
- Centrifugar por 5 minutos a 14,000 RPM y transferir el sobrenadante a un tubo eppendorf limpio
- Secar en Savant hasta reducir el volumen a aproximadamente 300  $\mu$ l.
- Precipitar el oligo con n-butanol (ver protocolo 2)
- Resuspender el oligo con 100  $\mu$ l de agua grado milliQ

Después de la purificación del oligo se tomaron 2  $\mu$ l para verificar su pureza y concentración en un gel delgado (0.75 mm) de poliacrilamida 20%, 8M Urea.

### **Protocolo 2- Precipitación con n-Butanol**

- Agregar n-Butanol en volumen 1:1 con el DNA
- Agitar fuertemente hasta homogenizar el butanol y el medio acuoso en el que esta el DNA
- Centrifugar por 10 min. a 14,000 RPM
- Extraer el n-butanol con pipeta
- Secar en Savant por 5 minutos
- Resuspender la pastilla con 100  $\mu$ l de agua grado milliQ

### **Protocolo 3- Fosforilación de oligonucleótidos**

Se fosforilaron 800 pmoles de cada oligo, en un volumen final de 100  $\mu$ l. La reacción de fosforilación es la siguiente:

- X  $\mu$ l de oligonucleótido purificado
- 10  $\mu$ l de "buffer" PNK (Polynucleotide Kinase, New England Biolabs ® de un stock 10X)
- 15  $\mu$ l ATP (de un stock 10mM)
- 2  $\mu$ l de Kinasa (New England Biolabs ®)
- Completar el volumen con agua grado milliQ

Esta reacción se incubó por una hora a 37°C. Posteriormente se inactivó la enzima incubando la reacción por 10 minutos a 65°C.

### **Protocolo 4- Hibridación de oligonucleótidos**

En un tubo Eppendorf de 1.5 ml. mezclamos 800 pmol. de cada uno de los dos oligos complementarios (uno correspondiente a la hebra codificante y el otro a la no codificante) anteriormente fosforilados. Colocamos el tubo en un baño de agua hirviendo y posteriormente retiramos el agua de la fuente de calor y dejamos enfriar hasta llegar a temperatura ambiente. Después de este proceso nuestros dos oligos están ahora en doble cadena de DNA (ya hibridados).

### **Protocolo 5- Ligación de asas con sus respectivos conectores y purificación de las mismas**

Las reacciones de ligación de cada asa con su respectivo conector se llevaron a cabo en un volumen de 100  $\mu$ l y fueron preparadas de la manera siguiente:

Para la ligación del asa con el conector amino:

- X  $\mu$ l del conector amino hibridado (800 pmol)
- X  $\mu$ l del asa hibridada (800 pmol)
- 20  $\mu$ l de "buffer" de ligasa de Invitrogen ® (de un stock 5x)

- 10  $\mu$ l ATP (de un stock 10mM)
  - 3  $\mu$ l de ligasa de Fermentas  $\text{\textcircled{R}}$
  - Completar el volumen con agua grado milliQ
- Para la ligación del asa con el conector carboxilo:
- X  $\mu$ l del conector carboxilo hibridado (800 pmol)
  - X  $\mu$ l del asa hibridada (800 pmol)
  - 20  $\mu$ l de "buffer" de ligasa de Invitrogen  $\text{\textcircled{R}}$  (de un stock 5x)
  - 10  $\mu$ l ATP (de un stock 10mM)
  - 3  $\mu$ l de ligasa de Fermentas  $\text{\textcircled{R}}$
  - Completar el volumen con agua grado milliQ

Cada una de estas reacciones se incubó por 12 horas a 16 $^{\circ}$  C. Posteriormente cada ligación fue purificada en un gel nativo de Acrilamida al 10%, después de correr el gel a 200 volts con una fuente de poder de BioRad por aproximadamente 2 horas, procedimos a teñir el gel con bromuro de etidio y se cortó la banda correspondiente al tamaño para cada una de las ligaciones. Las bandas cortadas se colocaron en un tubo Eppendorf $\text{\textcircled{R}}$  con 1 ml de "buffer" de Elusión de Oligos y se incubaron a 65 $^{\circ}$ C en agitación por 3 horas. Posteriormente se centrifugó cada tubo por 5 minutos a 14,000 RPM y se transfirió el sobrenadante a un tubo limpio. Dicho tubo fue puesto a secar en Savant hasta que el volumen se redujo a aproximadamente 300  $\mu$ l. Se procedió a precipitar el oligo con n-butanol hasta observarse la formación de una pastilla. Ya formada la pastilla se lava con 1 ml de etanol absoluto frío (-20  $^{\circ}$ C). Posteriormente, se centrifugaron a 14,000 rpm por 10 minutos y se removió cuidadosamente el sobrenadante, dejándose la pastilla que posteriormente se puso a secar en Savant. Finalmente, se resuspendió la pastilla con 100  $\mu$ l de agua grado milliQ.

#### **Protocolo 6- Reacción de amplificación por PCR del extremo amino y carboxilo terminal del gen de PRAI**

La reacción de PCR para amplificar el extremo amino de PRAI fue hecha en un volumen de 100  $\mu$ l utilizando los siguientes reactivos:

- 1  $\mu$ l Templado: plásmido pDAN5-PRAI (stock de 50 ng/ $\mu$ l)
- 1  $\mu$ l Oligo Hind3-Bivi (stock de 25 pmol/ $\mu$ l)
- 10  $\mu$ l Producto de la ligación del asa con el conector amino (100 ng)
- 10  $\mu$ l "buffer" de PCR para *Pfu* polimerasa (10X)
- 10  $\mu$ l dNTP's (2.5 mM)
- 5  $\mu$ l Enzima *Pfu* polimerasa (purificada en el laboratorio)
- 63  $\mu$ l Agua grado milliQ

La reacción de PCR para amplificar el extremo carboxilo de PRAI fue hecha en un volumen de 100  $\mu$ l utilizando los siguientes reactivos:

- 1  $\mu$ l Templado: plásmido pDAN5-PRAI (stock de 50 ng/ $\mu$ l)
- 1  $\mu$ l Oligo NheI-Bivi (stock de 25 pmol/ $\mu$ l)
- 10  $\mu$ l Producto de la ligación del asa con el conector carboxilo (100 ng)
- 10  $\mu$ l "buffer" de PCR para *Pfu* polimerasa (10X)
- 10  $\mu$ l dNTP's (2.5 mM)
- 5  $\mu$ l Enzima *Pfu* polimerasa (purificada en el laboratorio)
- 63  $\mu$ l Agua grado milliQ

Las condiciones de las dos reacciones de PCR fueron: primer ciclo de desnaturalización 94 $^{\circ}$ C 5 min., 30 ciclos con las siguientes condiciones: 94 $^{\circ}$ C 1 min., 60 $^{\circ}$ C 1:30 min., 72 $^{\circ}$ C 1 min. y una extensión final de 72 $^{\circ}$ C 10 min. Todas las reacciones de PCR se hicieron en un termociclador iCycler de Bio-Rad $\text{\textcircled{R}}$ . Cada uno de estos productos de PCR fue purificado por electroforesis en gel de agarosa al 1% utilizando el "kit" de Roche ("High pure PCR product purification Kit") y resuspendido en 100  $\mu$ l de agua grado milliQ.

### **Protocolo 7- Reacción de amplificación por PCR para armar el gen de PRAI a partir de las mitades amino y carboxilo**

Esta reacción de PCR fue hecha para armar cada uno de los 8 genes con distintas asas insertadas y se usó como templado las mitades amplificadas para cada banco. Se hizo en un volumen de 100  $\mu$ l utilizando los siguientes reactivos:

- 15  $\mu$ l Mitad amino de cada asa (400 ng)
- 15  $\mu$ l Mitad carboxilo de cada asa (400 ng)
- 1  $\mu$ l Oligo Hind3-Bivi (stock de 25 pmol/ $\mu$ l)
- 1  $\mu$ l Oligo NheI-Bivi (stock de 25 pmol/ $\mu$ l)
- 10  $\mu$ l "buffer" de PCR para *Taq* polimerasa (10X)
- 10  $\mu$ l dNTP's (2.5 mM)
- 4  $\mu$ l MgCl<sub>2</sub> (50 mM)
- 5  $\mu$ l Enzima *Taq* polimerasa (purificada en el laboratorio)
- 39  $\mu$ l Agua grado milliQ

En esta reacción de PCR se agregaron los oligos Hind3-Bivi y NheI-Bivi hasta el décimo ciclo. Por lo cuál fueron 10 ciclos sin oligos y 20 ciclos con oligos. Las condiciones utilizadas fueron: primer ciclo de desnaturalización 94°C 5 min., 30 ciclos con las siguientes condiciones: 94°C 1 min., 60°C 1:30 min., 72°C 1:30 min. y una extensión final de 72°C 10 min. Todas las reacciones de PCR se hicieron en un termociclador iCycler de Bio-Rad®. Cada uno de estos productos de PCR fue purificado por electroforesis en gel de agarosa al 1% utilizando el "kit" de Roche ("High pure PCR product purification Kit") y resuspendido en 100  $\mu$ l de agua grado milliQ.

### **Protocolo 8- Reacción de digestión de productos de PCR**

Las reacciones de digestión para los productos de PCR se hicieron en un volumen de 100  $\mu$ l y utilizando los siguientes reactivos:

- 70  $\mu$ l del producto de PCR (1.5  $\mu$ g)
- 1  $\mu$ l de enzima de restricción Hind III (New England Biolabs)
- 1  $\mu$ l de enzima de restricción NheI (New England Biolabs)
- 1  $\mu$ l de BSA (New England Biolabs)
- 10  $\mu$ l de "buffer" 2 (New England Biolabs 10X)
- 17  $\mu$ l de agua grado milliQ

Las enzimas utilizadas dependen del sitio de restricción con el que nos interesa cortar el DNA, así como el "buffer" a utilizarse que será el indicado para cada enzima.

Esta reacción fue incubada por 12 horas a 37° C y posteriormente purificada, cortando la banda correspondiente al tamaño del gen en un gel de agarosa al 1% y se extrajo el DNA usando el "kit" ("High PCR Purification Isolation" de Roche).

### **Protocolo 9- Reacción de digestión de plásmido**

El vector pDAN5-PRAI se sometió a una reacción de restricción con las enzimas HindIII y NheI. Para la reacción de digestión se adicionó el volumen necesario de vector purificado correspondiente a 1  $\mu$ g, 10  $\mu$ l de "buffer" 2 de New England Biolabs (10x), 1  $\mu$ l de BSA (100x), 1.5 $\mu$ l de cada enzima de restricción y se completó el volumen con agua (grado milliQ) para tener un volumen final de 100  $\mu$ l. esta reacción fue incubada por 12 horas a 37°C. Esta digestión fue posteriormente corrida en un gel de agarosa al 1% y se cortó la banda que corresponde al tamaño del vector sin el inserto. Posteriormente se extrajo el DNA del gel de agarosa utilizando el "kit" de Roche "High PCR Purification Isolation de Roche".

Las enzimas de restricción utilizadas, dependen de los sitios de restricción que se van a usar para clonar.

### **Protocolo 10- Reacción de ligación de vector con inserto**

La relación molar de inserto: vector fue de 3:1, respectivamente. Se hizo un control negativo de esta reacción para asegurar la eficiencia de la digestión del vector, el cual no debe recircularizarse en ausencia de inserto. Las reacciones se llevaron a un volumen final de 100  $\mu$ l de la manera siguiente:

- X  $\mu$ l de inserto previamente digerido (0.7  $\mu$ g)
- X  $\mu$ l de vector previamente digerido (0.7  $\mu$ g)
- 10  $\mu$ l de "buffer" de ligación (Fermentas 10X)
- 2  $\mu$ l de ligasa (Fermentas 1U/ $\mu$ l)
- X  $\mu$ l de agua grado milliQ

Esta reacción fue incubada por 12 horas a 16°C. Posteriormente fue purificada mediante precipitación con butanol, secada en Savant por 5 minutos y resuspendida en 10  $\mu$ l de agua grado milliQ.

### **Protocolo 11- Transformación de células electrocompetentes**

Se transformaron 5  $\mu$ l de plásmido (1  $\mu$ g de DNA) en 50  $\mu$ l de células electrocompetentes XL1-Blue usando celdas de 0.1 cm y un electroporador Gene Pulser, todos ellos de Bio-Rad. Cada transformación se recuperó en 980  $\mu$ l de medio SOC, 10  $\mu$ l de Glucosa (2M) y 10  $\mu$ l de MgCl<sub>2</sub> (1M) a 37°C en agitación por 1 hora. Posteriormente se platearon 10  $\mu$ l, 50  $\mu$ l y 100  $\mu$ l del cultivo celular recuperado en cajas de medio LB sólido suplementado con el antibiótico al que posee resistencia el plásmido.

### **Protocolo 12- Síntesis del plásmido completo mediante PCR**

Esta es una reacción de PCR en donde se amplifica completamente el plásmido insertando las mutaciones o cambios diseñados en los oligos utilizados. La reacción así como sus componentes fueron ajustados a un volumen final de 100  $\mu$ l y los reactivos utilizados son los siguientes:

- 1  $\mu$ l de vector pDAN5-PRAI (25 ng)
- 1  $\mu$ l de oligo pdanCodBAM- (25 pmol/ $\mu$ l)
- 1  $\mu$ l de oligo pdanNoCodBAM- (25 pmol/ $\mu$ l)
- 5  $\mu$ l "buffer" de PCR para *Pfu* polimerasa (stock 10X)
- 5  $\mu$ l dNTP's (2.5 mM)
- 5  $\mu$ l de *Pfu* polimerasa (purificada en el laboratorio)
- 2  $\mu$ l de *Taq* polimerasa (purificada en el laboratorio)
- 30  $\mu$ l de agua grado milliQ

Las condiciones para esta reacción fueron: primer ciclo de desnaturalización 95°C 5 min., 16 ciclos con las siguientes condiciones: 95°C 1 min., 60°C 1 min., 68°C 6 min. y una extensión final de 72°C 10 min.

Posteriormente esta reacción fue llevada a un volumen final de 100  $\mu$ l con "buffer" de la enzima de restricción DpnI y agua grado milliQ, además se adicionó 1  $\mu$ l de la enzima DpnI y se incubó por una hora a 37°C. Esta reacción fue precipitada con butanol y secada en Savant por 5 min. y resuspendida en 10  $\mu$ l de agua y fueron transformados 4  $\mu$ l en células electrocompetentes XL1-Blue.

### **Protocolo 13- Reacción general de PCR**

Esta reacción fue utilizada para amplificar segmentos de DNA de manera normal. Es decir, se utilizó para hacer subclonaciones de un vector a otro y lo único que cambia en las diferentes reacciones son los oligos usados para la amplificación. Los reactivos a utilizarse quedan de la siguiente manera:

- 1  $\mu$ l Templado (generalmente plásmido en un concentración de 50 ng/ $\mu$ l)
- 1  $\mu$ l Oligo 1 (stock de 25 pmol/ $\mu$ l)
- 1  $\mu$ l Oligo 2 (stock de 25 pmol/ $\mu$ l)

- 10  $\mu$ l "buffer" de PCR para *Pfu* polimerasa (10X)
- 10  $\mu$ l dNTP's (2.5 mM)
- 5  $\mu$ l Enzima *Pfu* polimerasa (purificada en el laboratorio)
- 72  $\mu$ l Agua grado milliQ

Las condiciones de las dos reacciones de PCR fueron: primer ciclo de desnaturalización 94°C 5 min., 30 ciclos con las siguientes condiciones: 94°C 1 min., 60°C 1:30 min., 72°C 1 min. y una extensión final de 72°C 10 min. Todas las reacciones de PCR se hicieron en un termociclador iCycler de Bio-Rad®. Cada uno de estos productos de PCR fue purificado por electroforesis en gel de agarosa al 1% utilizando el "kit" de Roche ("High pure PCR product purification Kit") y resuspendido en 100  $\mu$ l de agua grado milliQ.

#### **Protocolo 14- Reacción de amplificación por PCR para armar el gen de la permutación circular de PRAI**

Esta reacción de PCR fue hecha para armar la permutación circular de PRAI. Se hizo en un volumen de 100  $\mu$ l utilizando los siguientes reactivos:

- 15  $\mu$ l Mitad producto de la amplificación amino terminal del gen (300 ng)
- 15  $\mu$ l Mitad producto de la amplificación carboxilo terminal del gen (300 ng)
- 1  $\mu$ l Oligo Hind3 permutaci (stock de 25 pmol/ $\mu$ l)
- 1  $\mu$ l Oligo BamHI permutaci (stock de 25 pmol/ $\mu$ l)
- 10  $\mu$ l "buffer" de PCR para *Taq* polimerasa (10X)
- 10  $\mu$ l dNTP's (2.5 mM)
- 4  $\mu$ l MgCl<sub>2</sub> (50 mM)
- 5  $\mu$ l Enzima *Taq* polimerasa (purificada en el laboratorio)
- 39  $\mu$ l Agua grado milliQ

En esta reacción de PCR se agregaron los oligos Hind3 permutaci y BamHI permutaci hasta el décimo ciclo. Por lo cuál fueron 10 ciclos sin oligos y 20 ciclos con oligos. Las condiciones utilizadas fueron: primer ciclo de desnaturalización 94°C 5 min., 30 ciclos con las siguientes condiciones: 94°C 1 min., 60°C 1:30 min., 72°C 1:30 min. y una extensión final de 72°C 10 min. Todas las reacciones de PCR se hicieron en un termociclador iCycler de Bio-Rad®. Cada uno de estos productos de PCR fue purificado por electroforesis en gel de agarosa al 1% utilizando el "kit" de Roche ("High pure PCR product purification Kit") y resuspendido en 100  $\mu$ l de agua grado milliQ.

#### **Protocolo 15- Expresión de proteínas para el análisis por "western blot" de fusiones a CAT**

- Inocular 5 ml de medio rico (Amp/Cm) con una colonia de células transformadas de la cepa MC1061 $\Delta$ thiE con el plásmido de la fusión pDAN 5-PRAI-CAT.
- Incubar por 12 horas a 37°C a 250 rpm
- Inocular de nuevo un tubo de 5 ml de medio rico (Amp/Cm) con 200  $\mu$ l del cultivo anterior
- Incubar por 15 horas a 20 °C a 250 rpm
- Centrifugar los cultivos en tubos eppendorf de 2 ml a 4000 rpm por 5 min a 4 °C
- Desechar sobrenadante
- Resuspender en 200  $\mu$ l de "buffer" PBS suplementado con Complete™ (inhibidor de proteasas)
- Sonicar incubando en un baño de hielo con punta chica 5 veces por 10 segundos cada vez, cuidar de no tocar las paredes del tubo ni hacer espuma
- Centrifugar a 11,000 rpm por 20 min a 4°C
- Transferir el sobrenadante a un tubo limpio lo cual constituye la fracción soluble
- Resuspender el pellet en 100  $\mu$ l de "buffer" PBS-Complete™ con 1% de SDS, incubar 10 minutos a 95°C. Esta es la fracción insoluble
- Conservar las muestras a 4 °C para analizarlas por SDS-PAGE.



El uso de IPTG para inducir la producción de proteína en esta cepa provocó que todas nuestras variantes se fueran a cuerpos de inclusión, por lo que decidimos expresar sin IPTG.

#### **Protocolo 16- Análisis por “western-blot” (“semi dry transfer”)**

- Preparar el gel de acrilamida al 13% (SDS-Page)
- Correr a 150 volts por aproximadamente 2 horas. Tener cuidado de que no baje el “buffer” nunca de nivel
- Antes de que las muestras salgan del gel preparar la cámara de transferencia
- Retirar el gel, siempre humectando con agua y depositarlo en una bandeja con agua
- Cortar de 6 papeles filtro y una membrana de nitrocelulosa del tamaño del gel. Remojar los papeles y la membrana en “buffer” de transferencia por 5 minutos.
- Colocar 3 papeles filtro en el ánodo de la cámara, luego poner el gel con SDS y posteriormente la membrana de nitrocelulosa y encima los tres papeles filtro restantes. Cerrar la cámara y colocar un peso fuerte para asegurar contacto. Correr una hora y media a 80mA.
- Después de la transferencia tomar la membrana y ponerla en un recipiente
- Bloquear la membrana con 20 ml de TBST + 1mg de leche (leche Svelties Nestlé). Dejar por 12 horas a 4°C
- Lavar 3 veces con TBST por 10 minutos cada vez.
- Incubar con el primer anticuerpo anti-CAT por una hora. Este anticuerpo está en relación 1:3000 en “buffer” TBST con 0.1% de leche. Incubar por una hora en agitación constante.
- Lavar 3 veces con TBST por 10 minutos cada vez.
- Incubar con el segundo anticuerpo (anti-Dig) acoplado a fosfatasa alcalina. Este anticuerpo esta en relación 1:5000 en “buffer” TBST con 0.1% de leche. Incubar por una hora en agitación constante
- Lavar dos veces con TBST
- Reveler con aproximadamente un mililitro de BCIP/NBT- blue (SIGMA)
- Parar la reacción de revelado con EDTA/agua 0.005 molar
- Enjuagar con agua grado milliQ
- Secar en medio de papel absorbente y digitalizarlo para su análisis

#### **Protocolo 17- Sobreexpresión de proteínas en Vector pET28**

- Inocular 5 ml de medio LB (Km/Cm) con una colonia de células transformadas de la cepa Rosetta 2 con el plásmido pET28
- Incubar por aproximadamente 5 horas a 37 °C a 280 rpm (hasta alcanzar una OD600 = 0.5)
- Incubar nuestro cultivo en hielo por 10 minutos
- Agregar al cultivo 1mM de IPTG (Isopropil B-D-1-tiogalactopiranosido) para inducir la expresión de nuestra proteína.
- Incubar nuestro cultivo recién inducido en hielo por 10 minutos
- Incubar los tubos en agitación (280 rpm) por 14 horas a 20 °C
- Centrifugar los cultivos en tubos eppendorf de 2 ml a 4000 rpm por 5 min a 4 °C
- Desechar sobrenadante
- Resuspender la pastilla del cultivo en 400 µl de “buffer” de sonicación (fosfatos 10 mM, pH 7.5, (0.5mM EDTA pH 7.7, 50mM de NaCl, 5% glicerol, 0.1 mM DTT, 0.1 mM de PMSF y 2.5 mg de lizosima).
- Sonicar en hielo con punta chica 6 veces por 10 segundos cada vez, cuidar de no tocar las paredes del tubo ni hacer espuma
- Centrifugar a 11,000 rpm por 20 min a 4 °C
- Transferir el sobrenadante a un tubo limpio. Esta constituye la fracción soluble
- Resuspender el pellet en 100 µl de “buffer” PBS-Complete con 1% de SDS, incubar 10 minutos a 95 °C. Esta es la fracción insoluble.
- Conservar las muestras a 4°C para analizarlas por SDS-PAGE.

### **Protocolo 18- Purificación por afinidad a resina de níquel (en pequeña escala)**

1. Agregar a un tubo eppendorf 200  $\mu$ l de soporte de sefarosa (fast flow de Amersham). Agregar 900  $\mu$ l de agua para remover el etanol en el que estaba guardada para su conservación.
2. Centrifugar 3 minutos a 4 °C a una velocidad de 4000 rpm
3. Extraer la fracción soluble (agua) de la columna de sefarosa
4. Repetir el paso 2 y 3
5. Agregar 900  $\mu$ l de una solución de 10 mg/ml de  $\text{NiCl}_2$  en 50mM de NaAcO pH 4.0.
6. Incubar en agitación por 10 minutos a temperatura ambiente
7. Centrifugar por 3 minutos a 4 °C a 4000 rpm
8. Remover el sobrenadante ( $\text{NiCl}_2$ ) de la columna
9. Agregar 900  $\mu$ l de agua e incubar en agitación por 3 minutos
10. Centrifugar 3 minutos a 4 °C a una velocidad de 4000 rpm
11. Remover el agua de la columna
12. Agregar a la columna, la fracción soluble de la sobreexpresión de nuestra proteína e incubar a 4 °C en agitación por 30 minutos.
13. Centrifugar 4 minutos a 4 °C a 4000 rpm y desechar sobrenadante
14. Agregar a la columna 700  $\mu$ l de "buffer" A e incubar a 4 °C en agitación por 5 minutos
15. Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será la fracción de proteínas eluidas en nuestro "buffer" A)
16. Agregar a la columna 700  $\mu$ l de "buffer" B e incubar a 4 °C en agitación por 5 minutos
17. Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será nuestra fracción de proteínas eluidas en nuestro "buffer" B)
18. Agregar a la columna 700  $\mu$ l de "buffer" C e incubar a 4 °C en agitación por 5 minutos
19. Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será nuestra fracción de proteínas eluidas en nuestro "buffer" C)
20. Agregar a la columna 900  $\mu$ l de "buffer" de elución e incubar a 4 °C en agitación por 15 minutos
21. Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será nuestra fracción de proteínas eluidas en nuestro "buffer" de elución)
22. Analizar en un SDS-Page al 13 % las diferentes fracciones eluidas para ubicar en que fracción (gradiente de imidazol) eluye nuestra proteína de interés de manera más pura y eficiente

### **Protocolo 19- Sobreexpresión de las proteínas con interés de purificar**

- Inocular 5 ml de medio LB (Km/Cm) con una colonia de células transformadas de la cepa Rosetta 2 con el plásmido pET28
- Incubar por 12 horas a 37 °C a 280 rpm
- Con el cultivo anterior inocular un matraz de un litro de LB (Amp/Cm) e incubar a 37 °C en agitación (280 rpm) por aproximadamente 5 horas (hasta alcanzar una  $\text{OD}_{600} = 0.7$ )
- Ya obtenida la densidad óptica especificada anteriormente poner el cultivo en hielo por 15 minutos
- Agregar al cultivo IPTG (Isopropil B-D-1-tiogalactopiranosido) a una concentración final de 1 mM para inducir la expresión de nuestra proteína.
- Incubar nuestro cultivo recién inducido en hielo por 15 minutos
- Posteriormente incubar a 20 °C por 15 horas en agitación (280 rpm)
- Centrifugar los cultivos a 4000 rpm por 5 min a 4 °C y desechar el sobrenadante
- Resuspender la pastilla del cultivo en 25 ml de "buffer" de sonicación (fosfatos 10 mM, pH 7.5, (0.5mM EDTA pH 7.7, 50mM de NaCl, 5% glicerol, 0.1mM DTT, 0.1 mM de PMSF y 2.5 mg de lizosima).
- Sonicar en hielo con punta grande 6 veces por 20 segundos cada vez, cuidar de no tocar las paredes del tubo ni hacer espuma
- Centrifugar a 17,000 rpm por 20 min a 4°C
- Transferir el sobrenadante a un tubo falcon limpio y esa es la fracción soluble que será transferida a la columna de níquel para la purificación de la proteína

### **Protocolo 20- Purificación por afinidad a resina de níquel (a gran escala)**

- 1- Agregar 3 ml de soporte de sefarosa (fast flow de Amersham) en un tubo falcon de 50 ml. Agregar 10 ml de agua para remover de la columna el etanol en el que estaba guardada.
- 2- Centrifugar 3 minutos a 4 °C a una velocidad de 4000 rpm
- 3- Extraer la fracción soluble (agua y etanol) de la columna de sefarosa
- 4- Repetir el paso 2 y 3
- 5- Agregar 8 ml de una solución de 10 mg/ml de NiCl<sub>2</sub> en 50 mM de NaAcO pH 4.0.
- 6- Incubar en agitación por 10 minutos a temperatura ambiente
- 7- Centrifugar por 3 minutos a 4°C a 4000 rpm y remover el sobrenadante (NiCl<sub>2</sub>) de la columna
- 8- Agregar 8 ml de agua e incubar en agitación por 3 minutos a temperatura ambiente y centrifugar por 3 minutos a 4°C a una velocidad de 4000 rpm
- 9- Remover el agua de la columna
- 10- Agregar a la columna, la fracción soluble de la sobreexpresión de nuestra proteína e incubar a 4 °C en agitación por 30 minutos.
- 11- Centrifugar 4 minutos a 4 °C a 4000 rpm y desechar sobrenadante
- 12- Agregar a la columna 8 ml de "buffer" A e incubar a 4 °C en agitación por 5 minutos
- 13- Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será nuestra fracción de proteínas eluidas en nuestro "buffer" A)
- 14- Agregar a la columna 8 ml de "buffer" B e incubar a 4 °C en agitación por 5 minutos
- 15- Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será nuestra fracción de proteínas eluidas en nuestro "buffer" B)
- 16- Agregar a la columna 8 ml de "buffer" C e incubar a 4 °C en agitación por 5 minutos
- 17- Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será nuestra fracción de proteínas eluidas en nuestro "buffer" C)
- 18- Agregar a la columna 13 ml de "buffer" de elución e incubar a 4 °C en agitación por 15 minutos
- 19- Centrifugar por 3 minutos a 4 °C a una velocidad de 4000 rpm y tomar sobrenadante en un tubo (esta será nuestra fracción de proteínas eluidas en nuestro "buffer" de elución)
- 20- Analizar en un SDS-Page al 13 % las diferentes fracciones eluidas para ubicar en que fracción (gradiente de imidazol) eluye nuestra proteína de interés de manera más pura y eficiente

## **Anexo 3. Soluciones**

### **1. LB (1 litro)**

- 10 g de triptona.
- 10 g de NaCl
- 5 g de extracto de levadura
- Completar con agua miliQ a 1 lt.

Para medio sólido

- Agregar 15 g de bacto agar

### **2- PBS 1X (1 litro)**

- 1.44 g Na<sub>2</sub>HPO<sub>4</sub> (8mM)
- 0.24 g KH<sub>2</sub>PO<sub>4</sub> (1.5 mM)
- 0.2 g KCl (3 mM)
- 8 grde NaCl (137 mM)

Ajustar el pH a 7.4

### **3- SOC**

- 20 g de triptona
- 5 g de extracto de levadura de cerveza
- 10 ml de NaCl 1M
- 2.5 ml KCl 1M

Completar con agua milliQ a 1 lt.

### **4- Ampicilina**

- 200 µg de ampicilina/ml
- Agua como solvente
- Para esterilizarse se filtra
- Guardar a -20 °C

### **5- “Buffer” de sonicación**

- 10 mM fosfato de potasio pH 7.6
- 0.5 mM EDTA pH 7.7
- 50 mM NaCl
- 5% glicerol
- 0.1 mM DTT (agregar en fresco)
- 0.1 mM de PMSF
- 1 mg/ml de lizosima

### **6- “Buffer” de purificación para permeación en gel**

- 10 mM fosfato de potasio pH 7.6
- 1 mM de EDTA
- 0.4 mM DTT

### **7- “Buffer” A (50 ml)**

- 50 mM fosfato de potasio pH 7.6
- 1 M de NaCl
- 10 mM de imidazol

### **8- “Buffer” B (50 ml)**

- 50 mM fosfato de potasio pH 7.6
- 1 M de NaCl
- 20 mM de imidazol

### **9- “Buffer” C (50 ml)**

- 50 mM fosfato de potasio pH 7.6

- 100 mM de NaCl
- 100 mM de imidazol

**10- “Buffer” de elución**

- 50 mM fosfato de potasio pH 7.6
- 100 mM de NaCl
- 500 mM de imidazol

**11- “Buffer” de regeneración de columna de níquel**

- 10 mg/ml  $\text{NiCl}_2$
- 50 mM NaAc pH 4.0

#### Anexo 4. Figuras del análisis de “western blot”

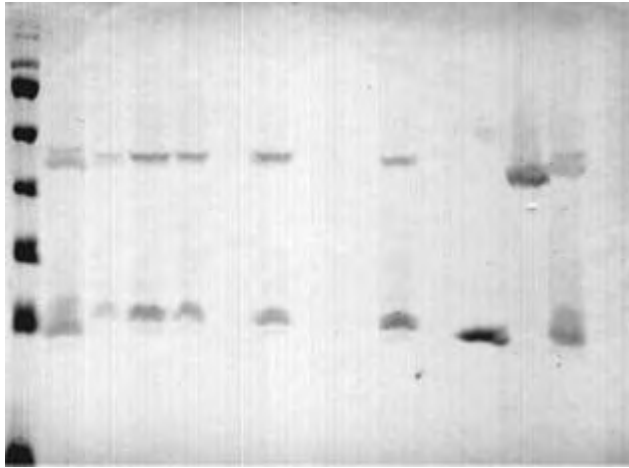


Figura 22. Análisis para el banco del asa 4 de ureasa

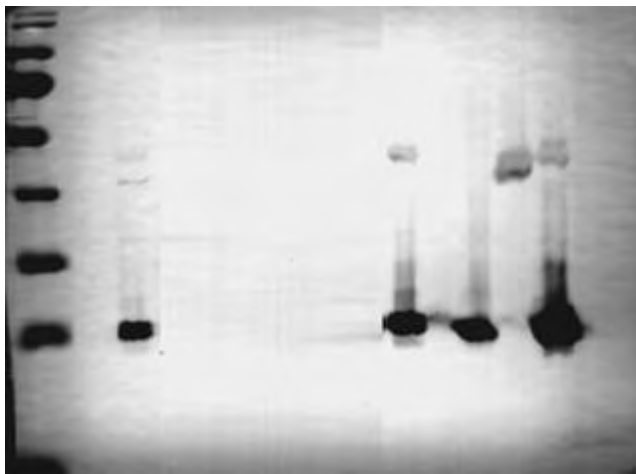


Figura 23. Análisis para el banco del asa 2 de mandelato racemasa

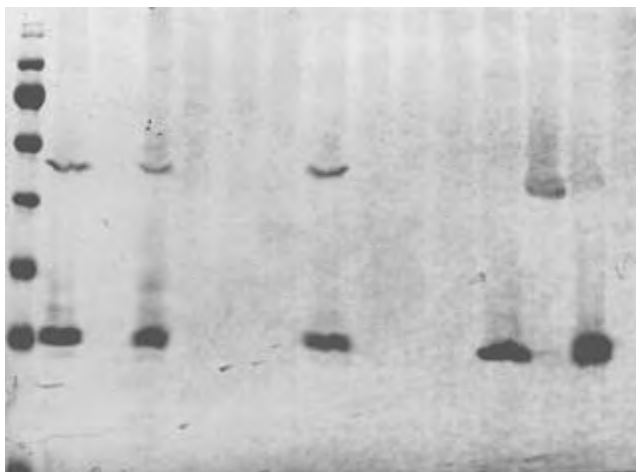


Figura 24. Análisis para el banco del asa 6 de porfobilinógeno sintasa

## 9. BIBLIOGRAFÍA

1. Alberts, B.L., J. et al., *Molecular Biology of the Cell*. Fourth Edition ed. 2002. 3-85.
2. Wolfenden, R. and M.J. Snider, *The depth of chemical time and the power of enzymes as catalysts*. *Acc Chem Res*, 2001. **34**(12): p. 938-45.
3. Farinas, E.T., *Directed evolution approaches for protein engineering*. *Comb Chem High Throughput Screen*, 2006. **9**(4): p. 235-6.
4. Kyte, J., *Mechanism in protein chemistry*, ed. Garland. 1995.
5. Guo, H.H., J. Choe, and L.A. Loeb, *Protein tolerance to random amino acid change*. *Proc Natl Acad Sci U S A*, 2004. **101**(25): p. 9205-10.
6. Vasserot, A.P., et al., *Optimization of protein therapeutics by directed evolution*. *Drug Discov Today*, 2003. **8**(3): p. 118-26.
7. Rubin-Pitel, S.B. and H. Zhao, *Recent advances in biocatalysis by directed enzyme evolution*. *Comb Chem High Throughput Screen*, 2006. **9**(4): p. 247-57.
8. Cohen, N., et al., *In vitro enzyme evolution: the screening challenge of isolating the one in a million*. *Trends Biotechnol*, 2001. **19**(12): p. 507-10.
9. Johannes, T.W. and H. Zhao, *Directed evolution of enzymes and biosynthetic pathways*. *Curr Opin Microbiol*, 2006. **9**(3): p. 261-7.
10. Schloss, P.D. and J. Handelsman, *Metagenomics for studying unculturable microorganisms: cutting the Gordian knot*. *Genome Biol*, 2005. **6**(8): p. 229.
11. Tringe, S.G. and E.M. Rubin, *Metagenomics: DNA sequencing of environmental samples*. *Nat Rev Genet*, 2005. **6**(11): p. 805-14.
12. Lorenz, P. and J. Eck, *Metagenomics and industrial applications*. *Nat Rev Microbiol*, 2005. **3**(6): p. 510-6.
13. Venter, J.C., et al., *Environmental genome shotgun sequencing of the Sargasso Sea*. *Science*, 2004. **304**(5667): p. 66-74.
14. Ferrer, M., F. Martinez-Abarca, and P.N. Golyshin, *Mining genomes and 'metagenomes' for novel catalysts*. *Curr Opin Biotechnol*, 2005. **16**(6): p. 588-93.
15. Douthwaite, J. and L. Jermutus, *Exploiting directed evolution for the discovery of biologicals*. *Curr Opin Drug Discov Devel*, 2006. **9**(2): p. 269-75.
16. Schmitzer, A.R., F. Lepine, and J.N. Pelletier, *Combinatorial exploration of the catalytic site of a drug-resistant dihydrofolate reductase: creating alternative functional configurations*. *Protein Eng Des Sel*, 2004. **17**(11): p. 809-19.
17. Orenca, M.C., et al., *Predicting the emergence of antibiotic resistance by directed evolution and structural analysis*. *Nat Struct Biol*, 2001. **8**(3): p. 238-42.
18. Taylor, S.V., P. Kast, and D. Hilvert, *Investigating and Engineering Enzymes by Genetic Selection*. *Angew Chem Int Ed Engl*, 2001. **40**(18): p. 3310-3335.
19. Neylon, C., *Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution*. *Nucleic Acids Res*, 2004. **32**(4): p. 1448-59.
20. Yuan, L., et al., *Laboratory-directed protein evolution*. *Microbiol Mol Biol Rev*, 2005. **69**(3): p. 373-92.
21. Schmidt-Dannert, C., *Directed evolution of single proteins, metabolic pathways, and viruses*. *Biochemistry*, 2001. **40**(44): p. 13125-36.
22. Eijsink, V.G., et al., *Directed evolution of enzyme stability*. *Biomol Eng*, 2005. **22**(1-3): p. 21-30.
23. Chatterjee, R. and L. Yuan, *Directed evolution of metabolic pathways*. *Trends Biotechnol*, 2006. **24**(1): p. 28-38.
24. Jurgens, C., et al., *Directed evolution of a (beta alpha)8-barrel enzyme to catalyze related reactions in two different metabolic pathways*. *Proc Natl Acad Sci U S A*, 2000. **97**(18): p. 9925-30.
25. Leemhuis, H., et al., *New genotype-phenotype linkages for directed evolution of functional proteins*. *Curr Opin Struct Biol*, 2005. **15**(4): p. 472-8.
26. Koonin, E.V., Y.I. Wolf, and G.P. Karev, *The structure of the protein universe and genome evolution*. *Nature*, 2002. **420**(6912): p. 218-23.

27. Burley, S.K. and J.B. Bonanno, *Structuring the universe of proteins*. Annu Rev Genomics Hum Genet, 2002. **3**: p. 243-62.
28. Hanes, J. and A. Pluckthun, *In vitro selection and evolution of functional proteins by using ribosome display*. Proc Natl Acad Sci U S A, 1997. **94**(10): p. 4937-42.
29. Hanes, J., et al., *Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries*. Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14130-5.
30. Jermutus, L. and J. Pelletier, *Creating and evaluating protein diversity*. Curr Opin Biotechnol, 2001. **12**(4): p. 331-3.
31. Chica, R.A., N. Doucet, and J.N. Pelletier, *Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design*. Curr Opin Biotechnol, 2005. **16**(4): p. 378-84.
32. Peimbert, M. and L. Segovia, *Evolutionary engineering of a beta-Lactamase activity on a D-Ala D-Ala transpeptidase fold*. Protein Eng, 2003. **16**(1): p. 27-35.
33. Wong, T.S., D. Zhurina, and U. Schwaneberg, *The diversity challenge in directed protein evolution*. Comb Chem High Throughput Screen, 2006. **9**(4): p. 271-88.
34. Sweasy, J.B. and L.A. Loeb, *Detection and characterization of mammalian DNA polymerase beta mutants by functional complementation in Escherichia coli*. Proc Natl Acad Sci U S A, 1993. **90**(10): p. 4626-30.
35. Dube, D.K., et al., *Selection of new biologically active molecules from random nucleotide sequences*. Gene, 1993. **137**(1): p. 41-7.
36. Greener, A., M. Callahan, and B. Jerpseth, *An efficient random mutagenesis technique using an E. coli mutator strain*. Mol Biotechnol, 1997. **7**(2): p. 189-95.
37. Bornscheuer, U.T., J. Altenbuchner, and H.H. Meyer, *Directed evolution of an esterase: screening of enzyme libraries based on pH-indicators and a growth assay*. Bioorg Med Chem, 1999. **7**(10): p. 2169-73.
38. Alexeeva, M., et al., *Deracemization of alpha-methylbenzylamine using an enzyme obtained by in vitro evolution*. Angew Chem Int Ed Engl, 2002. **41**(17): p. 3177-80.
39. Nguyen, A.W. and P.S. Daugherty, *Production of randomly mutated plasmid libraries using mutator strains*. Methods Mol Biol, 2003. **231**: p. 39-44.
40. Bessette, P.H., et al., *Construction of designed protein libraries using gene assembly mutagenesis*. Methods Mol Biol, 2003. **231**: p. 29-37.
41. Zaccolo, M., et al., *An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues*. J Mol Biol, 1996. **255**(4): p. 589-603.
42. Zaccolo, M. and E. Gherardi, *The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase*. J Mol Biol, 1999. **285**(2): p. 775-83.
43. Cadwell, R.C. and G.F. Joyce, *Randomization of genes by PCR mutagenesis*. PCR Methods Appl, 1992. **2**(1): p. 28-33.
44. Cadwell, R.C. and G.F. Joyce, *Mutagenic PCR*. PCR Methods Appl, 1994. **3**(6): p. S136-40.
45. Cirino, P.C., K.M. Mayer, and D. Umeno, *Generating mutant libraries using error-prone PCR*. Methods Mol Biol, 2003. **231**: p. 3-9.
46. Murakami, H., T. Hohsaka, and M. Sisido, *Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs*. Nat Biotechnol, 2002. **20**(1): p. 76-81.
47. Murakami, H., T. Hohsaka, and M. Sisido, *Random insertion and deletion mutagenesis*. Methods Mol Biol, 2003. **231**: p. 53-64.
48. Miyazaki, K. and F.H. Arnold, *Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function*. J Mol Evol, 1999. **49**(6): p. 716-20.
49. Hayes, F., B. Hallet, and Y. Cao, *Insertion mutagenesis as a tool in the modification of protein function. Extended substrate specificity conferred by pentapeptide insertions in the omega-loop of TEM-1 beta-lactamase*. J Biol Chem, 1997. **272**(46): p. 28833-6.
50. Hayes, F. and B. Hallet, *Pentapeptide scanning mutagenesis: encouraging old proteins to execute unusual tricks*. Trends Microbiol, 2000. **8**(12): p. 571-7.



51. Pikkemaat, M.G. and D.B. Janssen, *Generating segmental mutations in haloalkane dehalogenase: a novel part in the directed evolution toolbox*. Nucleic Acids Res, 2002. **30**(8): p. E35-5.
52. Ward, B. and T. Juehne, *Combinatorial library diversity: probability assessment of library populations*. Nucleic Acids Res, 1998. **26**(4): p. 879-86.
53. Palfrey, D., M. Picardo, and A.V. Hine, *A new randomization assay reveals unexpected elements of sequence bias in model 'randomized' gene libraries: implications for biopanning*. Gene, 2000. **251**(1): p. 91-9.
54. Bergquist, P.L., R.A. Reeves, and M.D. Gibbs, *Degenerate oligonucleotide gene shuffling (DOGS) and random drift mutagenesis (RNDM): two complementary techniques for enzyme evolution*. Biomol Eng, 2005. **22**(1-3): p. 63-72.
55. Gibbs, M.D., K.M. Nevalainen, and P.L. Bergquist, *Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling*. Gene, 2001. **271**(1): p. 13-20.
56. Virnekas, B., et al., *Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis*. Nucleic Acids Res, 1994. **22**(25): p. 5600-7.
57. Gaytan, P., et al., *Combination of DMT-monomer and Fmoc-trinucleotide phosphoramidites in oligonucleotide synthesis affords an automatable codon-level mutagenesis method*. Chem Biol, 1998. **5**(9): p. 519-27.
58. Stemmer, W.P., *DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution*. Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10747-51.
59. Stemmer, W.P., *Rapid evolution of a protein in vitro by DNA shuffling*. Nature, 1994. **370**(6488): p. 389-91.
60. Vanhercke, T., et al., *Reducing mutational bias in random protein libraries*. Anal Biochem, 2005. **339**(1): p. 9-14.
61. Zhao, H., et al., *Molecular evolution by staggered extension process (StEP) in vitro recombination*. Nat Biotechnol, 1998. **16**(3): p. 258-61.
62. Aguinaldo, A.M. and F. Arnold, *Staggered extension process (StEP) in vitro recombination*. Methods Mol Biol, 2002. **192**: p. 235-9.
63. Aguinaldo, A.M. and F.H. Arnold, *Staggered extension process (StEP) in vitro recombination*. Methods Mol Biol, 2003. **231**: p. 105-10.
64. Coco, W.M., *RACHITT: Gene family shuffling by Random Chimeragenesis on Transient Templates*. Methods Mol Biol, 2003. **231**: p. 111-27.
65. Coco, W.M., et al., *DNA shuffling method for generating highly recombined genes and evolved enzymes*. Nat Biotechnol, 2001. **19**(4): p. 354-9.
66. Lutz, S., et al., *Creating multiple-crossover DNA libraries independent of sequence identity*. Proc Natl Acad Sci U S A, 2001. **98**(20): p. 11248-53.
67. Ostermeier, M., J.H. Shim, and S.J. Benkovic, *A combinatorial approach to hybrid enzymes independent of DNA homology*. Nat Biotechnol, 1999. **17**(12): p. 1205-9.
68. Griswold, K.E., et al., *Evolution of highly active enzymes by homology-independent recombination*. Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10082-7.
69. Ostermeier, M. and S. Lutz, *The creation of ITCHY hybrid protein libraries*. Methods Mol Biol, 2003. **231**: p. 129-41.
70. Silberg, J.J., J.B. Endelman, and F.H. Arnold, *SCHEMA-guided protein recombination*. Methods Enzymol, 2004. **388**: p. 35-42.
71. Otey, C.R., et al., *Functional evolution and structural conservation in chimeric cytochromes p450: calibrating a structure-guided approach*. Chem Biol, 2004. **11**(3): p. 309-18.
72. Meyer, M.M., et al., *Library analysis of SCHEMA-guided protein recombination*. Protein Sci, 2003. **12**(8): p. 1686-93.
73. Castle, L.A., et al., *Discovery and directed evolution of a glyphosate tolerance gene*. Science, 2004. **304**(5674): p. 1151-4.
74. Solbak, A.I., et al., *Discovery of pectin-degrading enzymes and directed evolution of a novel pectate lyase for processing cotton fabric*. J Biol Chem, 2005. **280**(10): p. 9431-8.

75. Wang, T., et al., *Directed evolution for engineering pH profile of endoglucanase III from Trichoderma reesei*. *Biomol Eng*, 2005. **22**(1-3): p. 89-94.
76. Park, H.S., et al., *Design and evolution of new catalytic activity with an existing protein scaffold*. *Science*, 2006. **311**(5760): p. 535-8.
77. Joerger, A.C., S. Mayer, and A.R. Fersht, *Mimicking natural evolution in vitro: an N-acetylneuraminase lyase mutant with an increased dihydrodipicolinate synthase activity*. *Proc Natl Acad Sci U S A*, 2003. **100**(10): p. 5694-9.
78. Dwyer, M.A., L.L. Looger, and H.W. Hellinga, *Computational design of a biologically active enzyme*. *Science*, 2004. **304**(5679): p. 1967-71.
79. Williams, G.J., A.S. Nelson, and A. Berry, *Directed evolution of enzymes for biocatalysis and the life sciences*. *Cell Mol Life Sci*, 2004. **61**(24): p. 3034-46.
80. Tsuji, T., M. Onimaru, and H. Yanagawa, *Towards the creation of novel proteins by block shuffling*. *Comb Chem High Throughput Screen*, 2006. **9**(4): p. 259-69.
81. Cramer, A., et al., *DNA shuffling of a family of genes from diverse species accelerates directed evolution*. *Nature*, 1998. **391**(6664): p. 288-91.
82. Gilbert, W., S.J. de Souza, and M. Long, *Origin of genes*. *Proc Natl Acad Sci U S A*, 1997. **94**(15): p. 7698-703.
83. Endelman, J.B., et al., *Site-directed protein recombination as a shortest-path problem*. *Protein Eng Des Sel*, 2004. **17**(7): p. 589-94.
84. Cipolla, L., *Combinatorial libraries of biocatalysts: application and screening*. *Comb Chem High Throughput Screen*, 2004. **7**(2): p. 101-14.
85. Bolon, D.N., C.A. Voigt, and S.L. Mayo, *De novo design of biocatalysts*. *Curr Opin Chem Biol*, 2002. **6**(2): p. 125-9.
86. Tawfik, D.S., *Biochemistry. Loop grafting and the origins of enzyme species*. *Science*, 2006. **311**(5760): p. 475-6.
87. Nagano, N., E.G. Hutchinson, and J.M. Thornton, *Barrel structures in proteins: automatic identification and classification including a sequence analysis of TIM barrels*. *Protein Sci*, 1999. **8**(10): p. 2072-84.
88. Hocker, B., *Directed evolution of (beta/alpha)(8)-barrel enzymes*. *Biomol Eng*, 2005. **22**(1-3): p. 31-8.
89. Wierenga, R.K., *The TIM-barrel fold: a versatile framework for efficient enzymes*. *FEBS Lett*, 2001. **492**(3): p. 193-8.
90. Nagano, N., C.A. Orengo, and J.M. Thornton, *One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions*. *J Mol Biol*, 2002. **321**(5): p. 741-65.
91. Hocker, B., et al., *Stability, catalytic versatility and evolution of the (beta/alpha)(8)-barrel fold*. *Curr Opin Biotechnol*, 2001. **12**(4): p. 376-81.
92. Banner, D.W., et al., *Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data*. *Nature*, 1975. **255**(5510): p. 609-14.
93. Larsen, T.M., et al., *Structure of rabbit muscle pyruvate kinase complexed with Mn<sup>2+</sup>, K<sup>+</sup>, and pyruvate*. *Biochemistry*, 1994. **33**(20): p. 6301-9.
94. Gromiha, M.M., et al., *Locating the stabilizing residues in (alpha/beta)<sub>8</sub> barrel proteins based on hydrophobicity, long-range interactions, and sequence conservation*. *Proteins*, 2004. **55**(2): p. 316-29.
95. Luger, K., et al., *Correct folding of circularly permuted variants of a beta/alpha barrel enzyme in vivo*. *Science*, 1989. **243**(4888): p. 206-10.
96. Urfer, R. and K. Kirschner, *The importance of surface loops for stabilizing an eightfold beta/alpha barrel protein*. *Protein Sci*, 1992. **1**(1): p. 31-45.
97. Soberon, X., P. Fuentes-Gallego, and G. Saab-Rincon, *In vivo fragment complementation of a (beta/alpha)(8) barrel protein: generation of variability by recombination*. *FEBS Lett*, 2004. **560**(1-3): p. 167-72.
98. Henn-Sax, M., et al., *Divergent evolution of (beta/alpha)<sub>8</sub>-barrel enzymes*. *Biol Chem*, 2001. **382**(9): p. 1315-20.
99. Hommel, U., M. Eberhard, and K. Kirschner, *Phosphoribosyl anthranilate isomerase catalyzes a reversible amadori reaction*. *Biochemistry*, 1995. **34**(16): p. 5429-39.

100. Henn-Sax, M., et al., *Two (beta/alpha)<sub>8</sub>-barrel enzymes of histidine and tryptophan biosynthesis have similar reaction mechanisms and common strategies for protecting their labile substrates*. *Biochemistry*, 2002. **41**(40): p. 12032-42.
101. Eberhard, M., et al., *Indoleglycerol phosphate synthase-phosphoribosyl anthranilate isomerase: comparison of the bifunctional enzyme from Escherichia coli with engineered monofunctional domains*. *Biochemistry*, 1995. **34**(16): p. 5419-28.
102. Saab-Rincon, G., et al., *Generation of variability by in vivo recombination of halves of a (beta/alpha)<sub>8</sub> barrel protein*. *Biomol Eng*, 2005. **22**(4): p. 113-20.
103. Waldo, G.S., et al., *Rapid protein-folding assay using green fluorescent protein*. *Nat Biotechnol*, 1999. **17**(7): p. 691-5.
104. Waldo, G.S., *Genetic screens and directed evolution for protein solubility*. *Curr Opin Chem Biol*, 2003. **7**(1): p. 33-8.
105. Wurth, C., N.K. Guimard, and M.H. Hecht, *Mutations that reduce aggregation of the Alzheimer's A $\beta$ <sub>42</sub> peptide: an unbiased search for the sequence determinants of A $\beta$  amyloidogenesis*. *J Mol Biol*, 2002. **319**(5): p. 1279-90.
106. Maxwell, K.L., et al., *A simple in vivo assay for increased protein solubility*. *Protein Sci*, 1999. **8**(9): p. 1908-11.
107. Sieber, V., C.A. Martinez, and F.H. Arnold, *Libraries of hybrid proteins from distantly related sequences*. *Nat Biotechnol*, 2001. **19**(5): p. 456-60.
108. Seitz, T., et al., *Stabilisation of a (beta/alpha)<sub>8</sub>-barrel protein designed from identical half barrels*. *J Mol Biol*, 2007. **372**(1): p. 114-29.
109. Williams, J.C., et al., *Structural and mutagenesis studies of leishmania triosephosphate isomerase: a point mutation can convert a mesophilic enzyme into a superstable enzyme without losing catalytic power*. *Protein Eng*, 1999. **12**(3): p. 243-50.
110. Merz, A., et al., *Improving the catalytic activity of a thermophilic enzyme at low temperatures*. *Biochemistry*, 2000. **39**(5): p. 880-9.
111. Noble, M.E., J.P. Zeelen, and R.K. Wierenga, *Structures of the "open" and "closed" state of trypanosomal triosephosphate isomerase, as observed in a new crystal form: implications for the reaction mechanism*. *Proteins*, 1993. **16**(4): p. 311-26.
112. Joseph, D., G.A. Petsko, and M. Karplus, *Anatomy of a conformational change: hinged "lid" motion of the triosephosphate isomerase loop*. *Science*, 1990. **249**(4975): p. 1425-8.
113. Sampson, N.S. and J.R. Knowles, *Segmental motion in catalysis: investigation of a hydrogen bond critical for loop closure in the reaction of triosephosphate isomerase*. *Biochemistry*, 1992. **31**(36): p. 8488-94.
114. Sampson, N.S. and J.R. Knowles, *Segmental movement: definition of the structural requirements for loop closure in catalysis by triosephosphate isomerase*. *Biochemistry*, 1992. **31**(36): p. 8482-7.
115. Alvarez, M., et al., *Triose-phosphate isomerase (TIM) of the psychrophilic bacterium Vibrio marinus. Kinetic and structural properties*. *J Biol Chem*, 1998. **273**(4): p. 2199-206.
116. Norledge, B.V., et al., *Modeling, mutagenesis, and structural studies on the fully conserved phosphate-binding loop (loop 8) of triosephosphate isomerase: toward a new substrate specificity*. *Proteins*, 2001. **42**(3): p. 383-9.
117. Saab-Rincon, G., et al., *Different strategies to recover the activity of monomeric triosephosphate isomerase by directed evolution*. *Protein Eng*, 2001. **14**(3): p. 149-55.
118. Cheon, Y.H., et al., *Manipulation of the active site loops of D-hydantoinase, a (beta/alpha)<sub>8</sub>-barrel protein, for modulation of the substrate specificity*. *Biochemistry*, 2004. **43**(23): p. 7413-20.
119. Wymer, N., et al., *Directed evolution of a new catalytic site in 2-keto-3-deoxy-6-phosphogluconate aldolase from Escherichia coli*. *Structure (Camb)*, 2001. **9**(1): p. 1-9.
120. Franke, D., G.A. Sprenger, and M. Muller, *Easy access to (R,R)-3,4-dihydroxy-3,4-dihydrobenzoic acid with engineered strains of Escherichia coli*. *ChemBiochem*, 2003. **4**(8): p. 775-7.
121. Fong, S., et al., *Directed evolution of D-2-keto-3-deoxy-6-phosphogluconate aldolase to new variants for the efficient synthesis of D- and L-sugars*. *Chem Biol*, 2000. **7**(11): p. 873-83.

122. Cho, C.M., A. Mulchandani, and W. Chen, *Bacterial cell surface display of organophosphorus hydrolase for selective screening of improved hydrolysis of organophosphate nerve agents*. Appl Environ Microbiol, 2002. **68**(4): p. 2026-30.
123. Chen-Goodspeed, M., et al., *Structural determinants of the substrate and stereochemical specificity of phosphotriesterase*. Biochemistry, 2001. **40**(5): p. 1325-31.
124. Chen-Goodspeed, M., et al., *Enhancement, relaxation, and reversal of the stereoselectivity for phosphotriesterase by rational evolution of active site residues*. Biochemistry, 2001. **40**(5): p. 1332-9.
125. Patrick, W.M. and J.M. Blackburn, *In vitro selection and characterization of a stable subdomain of phosphoribosylanthranilate isomerase*. Febs J, 2005. **272**(14): p. 3684-97.
126. Altamirano, M.M., et al., *Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold*. Nature, 2000. **403**(6770): p. 617-22.
127. Altamirano, M.M., et al., *Retraction. Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold*. Nature, 2002. **417**(6887): p. 468.
128. Binz, H.K., P. Amstutz, and A. Pluckthun, *Engineering novel binding proteins from nonimmunoglobulin domains*. Nat Biotechnol, 2005. **23**(10): p. 1257-68.
129. Nugent, P.G., et al., *Protein engineering loops in aspartic proteinases: site-directed mutagenesis, biochemical characterization and X-ray analysis of chymosin with a replaced loop from rhizopuspepsin*. Protein Eng, 1996. **9**(10): p. 885-93.
130. Albert, A., et al., *Protein engineering aspartic proteinases. Site-directed mutagenesis, biochemical characterisation, and X-ray analysis of chymosins with substituted single amino acid substitutions and loop replacements*. Adv Exp Med Biol, 1998. **436**: p. 169-77.
131. Fetrow, J.S., T.S. Cardillo, and F. Sherman, *Deletions and replacements of omega loops in yeast iso-1-cytochrome c*. Proteins, 1989. **6**(4): p. 372-81.
132. Hoyne, P.A., et al., *Properties of an insulin receptor with an IGF-1 receptor loop exchange in the cysteine-rich region*. FEBS Lett, 2000. **469**(1): p. 57-60.
133. Ma, B.J., et al., *Reciprocal transfer of class I MHC allele specificity between activating Ly-49P and Ly-49W receptors by exchange of beta 4-beta 5 loop residues*. J Immunol, 2003. **171**(10): p. 5337-44.
134. Pickersgill, R., et al., *Making a small enzyme smaller; removing the conserved loop structure of hen lysozyme*. FEBS Lett, 1994. **347**(2-3): p. 199-202.
135. Presta, L.G., *Selection, design, and engineering of therapeutic antibodies*. J Allergy Clin Immunol, 2005. **116**(4): p. 731-6; quiz 737.
136. Jung, S., A. Honegger, and A. Pluckthun, *Selection for improved protein stability by phage display*. J Mol Biol, 1999. **294**(1): p. 163-80.
137. Ewert, S., A. Honegger, and A. Pluckthun, *Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering*. Methods, 2004. **34**(2): p. 184-99.
138. Haruyama, H., et al., *Humanization of the mouse anti-Fas antibody HFE7A and crystal structure of the humanized HFE7A Fab fragment*. Biol Pharm Bull, 2002. **25**(12): p. 1537-45.
139. Juraja, S.M., et al., *Engineering of the Escherichia coli Im7 immunity protein as a loop display scaffold*. Protein Eng Des Sel, 2006. **19**(5): p. 231-44.
140. Binz, H.K. and A. Pluckthun, *Engineered proteins as specific binding reagents*. Curr Opin Biotechnol, 2005. **16**(4): p. 459-69.
141. Schlehuber, S. and A. Skerra, *Lipocalins in drug discovery: from natural ligand-binding proteins to "anticalins"*. Drug Discov Today, 2005. **10**(1): p. 23-33.
142. Skerra, A., *Lipocalins as a scaffold*. Biochim Biophys Acta, 2000. **1482**(1-2): p. 337-50.
143. Binz, H.K., et al., *High-affinity binders selected from designed ankyrin repeat protein libraries*. Nat Biotechnol, 2004. **22**(5): p. 575-82.
144. Koide, A., et al., *The fibronectin type III domain as a scaffold for novel binding proteins*. J Mol Biol, 1998. **284**(4): p. 1141-51.
145. Nord, K., et al., *Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain*. Nat Biotechnol, 1997. **15**(8): p. 772-7.
146. Eklund, M., et al., *Anti-idiotypic protein domains selected from protein A-based affibody libraries*. Proteins, 2002. **48**(3): p. 454-62.

147. Nicaise, M., et al., *Affinity transfer by CDR grafting on a nonimmunoglobulin scaffold*. Protein Sci, 2004. **13**(7): p. 1882-91.
148. Kallenbach, N., *Breaking open a protein barrel*. Proc Natl Acad Sci U S A, 2001. **98**(6): p. 2958-60.
149. Sterner, R., et al., *Phosphoribosyl anthranilate isomerase from Thermotoga maritima is an extremely stable and active homodimer*. Protein Sci, 1996. **5**(10): p. 2000-8.
150. Hennig, M., et al., *Crystal structure at 2.0 Å resolution of phosphoribosyl anthranilate isomerase from the hyperthermophile Thermotoga maritima: possible determinants of protein stability*. Biochemistry, 1997. **36**(20): p. 6009-16.
151. Akanuma, S. and A. Yamagishi, *Identification and characterization of key substructures involved in the early folding events of a (beta/alpha)<sub>8</sub>-barrel protein as studied by experimental and computational methods*. J Mol Biol, 2005. **353**(5): p. 1161-70.
152. Akanuma, S., et al., *A detailed unfolding pathway of a (beta/alpha)<sub>8</sub>-barrel protein as studied by molecular dynamics simulations*. Proteins, 2005. **58**(3): p. 538-46.
153. McGuffin, L.J., K. Bryson, and D.T. Jones, *The PSIPRED protein structure prediction server*. Bioinformatics, 2000. **16**(4): p. 404-5.