

129

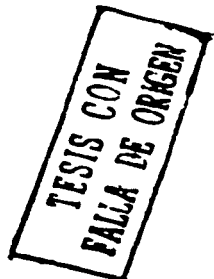
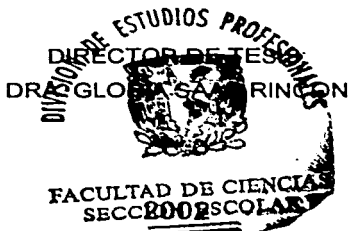


UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

RECOMBINACION *IN VIVO* DE FRAGMENTOS DE LA ENZIMA FOSFORRIBOSIL ANTRANILATO ISOMERASA (PRAI)

T E S I S
QUE PARA OBTENER EL TITULO DE
B I O L O G O
P R E S E N T A:
EUGENIO MANCERA RAMOS





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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



UNIVERSIDAD NACIONAL
MEXICANA

DRA. MARÍA DE LOURDES ESTEVA PERALTA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito: **Recombinación in vivo**
 de fragmentos de la enzima Fosforribosil Antranilato Isomerasa (PRAI).

realizado por **Eugenio Mancera Ramos**

con número de cuenta **9754609-6** , quien cubrió los créditos de la carrera de:

Biología

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis

Propietario **Dra. Gloria Saab Rincón**

Gloria Saab Rincón

Propietario **Dr. Lorenzo Segovia Forcella**

Segovia

Propietario **Dr. Alexander de Luna Fors**

Alexander de Luna Fors

Suplente **Dr. Xavier Soberón Mainero**

Xavier Soberón Mainero

Suplente **Dr. Víctor Valdés López**

Víctor Valdés López

FACULTAD DE CIENCIAS
U.N.A.M.

Consejo Departamental de Biología

Juan Manuel Rodríguez Chávez

M.en C. Juan Manuel Rodríguez Chávez



DEPARTAMENTO
DE BIOLOGIA

A mis abuelos

A Kasia

A mis padres

Agradezco sinceramente:

A la Dra. Gloria Saab Rincón, directora de esta tesis, por toda la atención y esfuerzo dedicados al proyecto.

A todos los integrantes del grupo del Dr. Xavier Soberón y compañeros del laboratorio.

A mis amigos de este rumbo Rocío, Paula, Argel, Diego, Arlene, Ángel, Blanca, Alonso, Mariana y los compañeros del taller.

A la Dra. Patricia Joseph por su apoyo.

A la Sra. Juana y a Pancho.

Al M. en C. Eugenio Bustos López, de la Unidad de Síntesis de Macromoléculas del IBT, por el suministro de oligonucleótidos.

A la B. Maricela Olvera Rodríguez y al B. René Hernández Vargas, de la Unidad de Secuenciación del IBT, por la secuenciación de DNA.

Índice

Resumen	1
Introducción	2
Barriles α/β	7
Barriles α/β en fragmentos.....	12
Recombinación <i>in vivo</i>	14
Enzimas modelo.....	23
Fosforribosil antranilato isomerasa.....	24
Tirosafosfato isomerasa.....	27
Objetivo	29
Objetivo particular	29
Metodología	30
Construcción de la proteína completa y de las mitades por separado en pDAN5.....	30
Recombinación <i>in vivo</i>	35
Análisis de la recombinación.....	37
Utilización del sistema en librerías.....	38
Barajeo de librerías.....	39
Protocolos.....	41
Medios y soluciones.....	47
Resultados	49
Construcciones y complementación de las funciones de las proteínas.....	49
Construcción y complementación de PRAI.....	49
Construcción y complementación de TIM.....	52

Recombinación <i>in vivo</i> de PRAI.....	55
Recombinación <i>in vivo</i> de los fagémidos portadores de las mitades de PRAI para formar una enzima PRAI activa.....	55
Proporciones de los fagémidos recombinantes	59
Patrones de restricción de los fagémidos recombinados.....	61
Utilización del sistema en librerías.....	63
Discusión.....	67
Complementación de las funciones de las enzimas con el péptido LOXP511 insertado.....	67
Recombinación <i>in vivo</i> de fragmentos de PRAI	71
Recombinación <i>in vivo</i> de fragmentos de PRAI para formar una enzima activa.....	71
Proporciones de los genotipos de los fagémidos recombinados.....	73
Utilización del sistema de recombinación <i>in vivo</i> en librerías.....	76
Conclusiones.....	80
Bibliografía.....	82

Resumen

La evolución dirigida de proteínas presenta dos pasos limitantes, uno de los cuales es la generación de variabilidad. A pesar de que se ha puesto gran interés en el desarrollo de mejores métodos, la eficiencia de transformación celular sigue siendo un importante cuello de botella, restringiendo el tamaño de las librerías a menos de 10^9 variantes. En la generación de librerías de anticuerpos este problema ha sido resuelto utilizando la recombinación *in vivo* sitio específica. En esta estrategia se utiliza la recombinación para combinar de nuevo y repetidamente las diferentes parejas de cadenas pesadas y cadenas ligeras de una librería previamente construida. Lo anterior es posible al insertar un sitio reconocido por la Cre recombinasa (sitio *loxP*) entre las cadenas a recombinar, el cual es traducido en un péptido conector. En el presente trabajo se describe la implementación de un sistema de recombinación *in vivo* para la fosforribosil antranilato isomerasa (PRAI) de *Escherichia coli*, una enzima citoplasmática de un solo dominio barril α/β . Basándose en estudios previos de corte, se generó una enzima activa *in vivo* con el sitio *loxP* insertado entre la $\alpha 4$ y $\beta 5$. Posteriormente se formó una proteína completa y activa por medio de recombinación, a partir de la integración de dos de sus fragmentos clonados de manera independiente en el vector de recombinación. La eficiencia del proceso demuestra que el sistema puede ser una buena herramienta en la generación de grandes librerías de proteínas para la búsqueda de nuevas propiedades enzimáticas.

Introducción

Las proteínas conforman uno de los cuatro principales grupos de biomacromoléculas. En gran medida las células están hechas de proteínas, las cuales constituyen más de la mitad del peso seco de la célula; "cada una de las propiedades que caracteriza a un organismo vivo está afectada por proteínas" (Creighton T. E., 1993). A pesar de que los ácidos nucleicos son los que almacenan la información requerida en los diferentes procesos de los seres vivos, las proteínas son el aparato que ejecuta esta información. La expresión de los genes no sólo resulta en la síntesis de proteínas sino que también depende de éstas. Las proteínas determinan la forma y estructura celular, sirven como instrumentos de reconocimiento y catálisis molecular, y almacenan y transportan una gran variedad de partículas, desde electrones hasta macromoléculas. La gran cantidad de reacciones químicas de los organismos vivos depende de las proteínas, ya que por sí solas estas reacciones no podrían ocurrir lo suficientemente rápido para sustentar la vida. Las enzimas (proteínas catalíticas) aumentan enormemente la tasa de reacción de estos fenómenos químicos. El intervalo de funciones que las proteínas cumplen es enorme, desde enzimas hasta algunas proteínas que son simplemente estructurales formando, por ejemplo, el citoesqueleto de las células y los tendones, huesos, uñas y pelo de los animales (Alberts B. D. *et al.*, 1983; Creighton T. E., 1993; Lodish H. A. *et al.*, 2000).

Por lo anterior, las proteínas han sido y son uno de los principales objetos de estudio de la bioquímica. Es importante conocer y entender su estructura y función, así como entender el proceso mediante el cual las proteínas modernas han evolucionado a partir de precursores más simples (Höcker B. *et al.*, 2001^b). Además, su entendimiento no sólo es fundamental para la ciencia básica sino también para la ciencia aplicada debido a su valor económico. En las últimas décadas el estudio de las proteínas, y en especial de las enzimas, ha sido revolucionado por la ingeniería de proteínas. A través del uso de la tecnología del ácido desoxidoribonucleico (DNA) recombinante y de la química de los ácidos nucleicos ha sido posible confeccionar las proteínas. Se ha logrado la producción de grandes cantidades de proteínas antes inaccesibles, así como la rápida determinación de su estructura primaria a través de la secuenciación del gen que las codifica y la modificación de sus estructuras por medio de la mutación del gen correspondiente (Fersht A., 1985). En la reciente expansión de la biotecnología, la estrategia anterior juega un papel fundamental en la

búsqueda de nuevas propiedades protéicas de aplicación para la industria química, farmacéutica y de alimentos (Cohen N. *et al.*, 2001), lográndose que esta búsqueda no esté limitada a la tradicional exploración de proteínas ya existentes en los organismos naturales. La primera aproximación se dio con la aplicación de estrategias de diseño racional, las cuales comprenden la modificación específica de la proteína objetivo basándose en la teoría general y el conocimiento previo de ésta. Para lo anterior se necesita contar con información estructural de alta calidad, con el fin de poder generar modelos moleculares que permitan obtener las conformaciones estructurales que más probablemente adquirirá la proteína, después de que se le realicen las diferentes modificaciones (Jermutus L. *et al.*, 2001). Sin embargo, la limitada habilidad para predecir los requerimientos estructurales de las proteínas hace que esta estrategia sea difícil y poco fructificante (Schmidt-Dannert C., 2001).

Como una alternativa a la estrategia racional, y más que nada enfocada a la obtención de nuevas propiedades aplicables de las proteínas, recientemente surgió la estrategia conocida como evolución dirigida. La idea general fue primeramente utilizada por Spiegelman en ácidos nucleicos (Mills D.R. *et al.*, 1967). Posteriormente, el grupo de Arnold, como uno de los pioneros, aplicó los principios de la teoría de evolución molecular creada por Eigen (1984) y Kauffman (1993), para la creación de enzimas mejoradas (Chen K. *et al.*, 1993). En términos generales la estrategia imita el proceso de evolución molecular natural de mutación – selección, aunque se le considera dirigida porque se seleccionan únicamente aquellas variantes cuyas propiedades son más cercanas a las requeridas para la aplicación de interés. El grupo de Arnold demostró que varios ciclos iterativos de este proceso, es decir mutación azarosa seguida por muestreo de las librerías creadas, arrojaban enzimas más activas y estables en tiempos considerablemente cortos (Schmidt-Dannert C., 2001). Actualmente la evolución dirigida es una metodología muy usada en el diseño de enzimas y que ha revolucionado la ingeniería de proteínas (Jermutus L. *et al.*, 2001, Schmidt-Dannert C., 2001). Lo anterior debido a que puede ser usada para la confección de proteínas con cierto fin, ya sea una aplicación en ciencia básica o práctica, incluso ante la ausencia de información estructural detallada y falta de conocimiento del sistema (Taylor S. V. *et al.*, 2001). Sin embargo, la evolución dirigida, más que desplazar al enfoque racional lo ha complementado.

La cantidad de trabajos donde se reporta la modificación de cierta proteína por evolución dirigida es considerable a la fecha (tabla 1) y aumenta a pasos agigantados (Schmidt-Dannert C., 2001). Los trabajos más comunes tienen que ver con el mejoramiento de la estabilidad y actividad enzimática, aunque también existen trabajos de cambio de requerimientos de sustrato de las enzimas y enantioselectividad, además de que se prevé que esta estrategia puede ser útil para mejorar la expresión heteróloga de proteínas funcionales, modificar los requerimientos de activadores y cofactores, o mejorar la resistencia a condiciones oxidativas o modificaciones químicas (Schmidt-Dannert C., 2001). Es importante mencionar que la evolución dirigida ha mostrado ser de gran utilidad en el diseño de proteínas dirigido a investigar la relación función – estructura, revelando aspectos que hubieran sido difíciles o imposibles de predecir racionalmente.

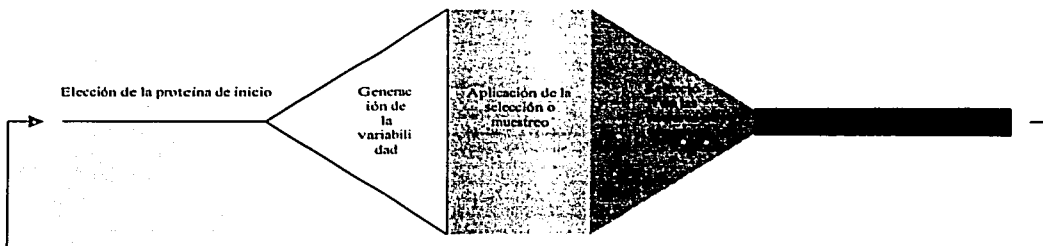


Figura 1. Pasos de la evolución dirigida. El ancho de las figuras representa de manera general el número de variantes involucradas en cada paso.

Los métodos de los distintos autores tienen variedades, pero en general consisten en los pasos ilustrados en la figura 1. Los trabajos hasta ahora realizados han revelado que son dos los pasos limitantes en la evolución dirigida: la generación de variabilidad (mutación) y el aislamiento de las mejores variantes (selección o muestreo). Por lo tanto, éstos pasos son los que se encuentran en la mira para el desarrollo de nuevas herramientas que permitan mejorar la estrategia. Hasta ahora la forma más eficiente de aislar a las mejores variantes han demostrado ser los esquemas de selección *in vivo*, ya que permiten la búsqueda exhaustiva en librerías grandes, de hasta 10^{10} variantes (Taylor S. V. *et al.*, 2001).

Tabla 1. Algunos trabajos de evolución dirigida

Enzima objetivo	Función objetivo	Referencia
kanamicina nucleotidiltransferasa	termoestabilidad	Liao et al. (1986)
subtilisina E	actividad en solventes orgánicos	(1993); Arnold & Chen (1994)
β -lactamasa	actividad hacia nuevos sustratos	Stemmer (1994)
subtilisina BPN ⁺	estabilidad en la ausencia de Ca ²⁺	Strausberg et al. (1995)
para-nitrobenzil esterasa	actividad hacia pNB ésteres; actividad en solventes orgánicos	Moore & Arnold (1996); Moore et al. (1997); Arnold & Moore (1998)
thymidine kinase	especificidad por sustrato (terapia genica)	Illick et al. (1996)
β -galactosidasa	actividad hacia nuevos sustratos; especificidad por sustrato	Zhang et al. (1997)
Subtilisina E	nivel de expresión; actividad en solventes orgánicos	You & Arnold (1996)
06-alkylguanine-DNA alquiltransferasa	protección contra agentes alquilantes (terapia genica)	Christians & Loebe (1996)
Ruta de desintoxicación por arsénito	Resistencia a arsénico	Cramer et al. (1997)
aminoacil-tRNA sintetasa	aminoacilación de un tRNA modificado	Liu et al. (1997)
aspartato aminotransferasa	actividad hacia amino y 2-oxo ácidos β -ramificados	Yano et al. (1997)
lipasa	capacidad para lavado	Okkels et al. (1997)
lipasa	enantioselectividad en hidrólisis de p-nitrofenil 2-metildecanoate	Retz et al. (1997)
lipasas	actividad hacia p-nitrofenil ésteres de cadena larga	van Kampen et al. (1998)
pNB esterasa	termoestabilidad	Giver et al. (1998)
esterasa	enantioselectividad de hidrólisis de un éster 3-hidroxi estéricamente impedido	Bornscheuer et al. (1998)
subtilisina E	termoestabilidad	Zhao & Arnold (1999)
Subtilisina E	termoestabilidad	Zhao et al. (1998)
H. tentus subtilisina	nivel de expresión (actividad total de la enzima secretada)	Naki et al. (1998)
subtilisin BPN ⁺	actividad a 10 ⁴ °C	Taguchi et al. (1998)
3-isopropilalate deshidrogenasa	termoestabilidad	Akanuma et al. (1998)
cefalosporinasas	actividad hacia moxalactama	Cramer et al. (1998)
corismato mutasa	conversión a enzima monomérica (solubilidad)	MacBeath et al. (1998)
bifenil dioxygenasas	degradación de bifenilos policlorinados (PCBs)	Kumamaru et al. (1998)
FLP recombinasa recombinación <i>in vivo</i>	eficiencia a elevadas temperaturas en <i>E. coli</i> y en células mamíferas; termoestabilidad <i>in vitro</i>	Buchholz et al. (1998)
endonucleasa EcoRV	Extensión del sitio de reconocimiento	Lanio et al. (1998)
citocromo P450cam	actividad incrementada en la ruta de peróxido shunt, hacia naltalene	Joo et al. (1999)
mioglobina	actividad peroxidasa	Wan et al. (1998)
aspartato aminotransferasa	especificidad por sustrato	Oue et al. (1999)
TEM-1 β -lactamasa	actividad hacia cefotaxima	Zaccolo et al. (1999)
glutation transferasa	especificidad por sustrato	Hansson et al. (1999)
bifenil dioxygenasa	rango extendido de sustratos	Brahmann et al. (1999)
Coprinus cinereus peroxidase	estabilidad a peróxido, termoestabilidad	Cherry et al. (1999)
β -glucuronidase	retención de función después de entrecruzamiento con glutaraldehido	Matsumura et al. (1999)
celulasa I of <i>B. stearothermophilus</i>	estabilidad	Matsumura et al. (1999)
subtilisina S41	Termoestabilidad mejorada	Miyazaki & Arnold (1999)
hidantoinasa	enantioselectividad + actividad invertida total	Miy et al. (2000)
subtilisinas con varias propiedades	mejoramiento en la actividad	Ness et al. (1999)
esterasa	enantioselectividad	Henke & Bornscheuer (1999)
timidina kinasa	especificidad por sustrato	Christians et al. (1999)
catecol 2,3-dioxygenasas	termoestabilidad	Kikuchi et al. (1999)
mdol-3-glicerol-fosfato sintasa	conferir nueva actividad catalítica	Altamirano et al. (2000)
<i>B. stearothermophilus</i> lactato deshidrogenasa	eliminar requerimiento de fructosa 1,6 bisfosfato (FBP)	Allen & Holbrook (2000)
kanamicina nucleotidil transferasa	termoestabilidad	Hoseki et al. (1999)
<i>B. stearothermophilus</i> amidasa	incremento de expresión en <i>E. coli</i>	Cheung & Oriol (2000)
phospholipase A1	termoestabilidad	Song & Rhee (2000)
Peroxidasa de rabano	Incremento de actividad/expresión en <i>S. cerevisiae</i>	Morawski et al. (2000)
L-2-hidroxiacido deshidrogenasa	requerimiento de co-factor (fructosa 1,6 bisfosfato)	Allen et al. (2000)
litoene desaturasa y licopeno ciclasa	rutas para nuevos carotenoides (especificidad por sustrato y de reacción)	Schmidt-Dannert et al. (2000)
citocromo P450 Bm-3	especificidad por sustrato	Li et al. (2000)

Tabla 1. Algunos trabajos de evolución dirigida (Obtenida de Saab-Rincón, comunicación personal).

En éstos se liga la propiedad objetivo con algún requerimiento fisiológico de la célula o alguna señal *in vivo*, de tal manera que sólo aquellas células que sobrevivan o presenten la señal contarán con la propiedad deseada. Sin embargo, lo anterior no siempre es posible y muchas veces presenta serias limitantes (para profundizar en este tema ver Cohen N. *et al.*, 2001; Schmidt-Dannert C., 2001). Por el otro lado, la forma más simple de crear variabilidad es por mutación puntual al azar, usando por ejemplo la reacción en cadena de la polimerasa (PCR) mutagénica. Esta herramienta es el típico punto de inicio en los trabajos de evolución dirigida, donde comúnmente se realizan ciclos iterativos de mutación azarosa seguida de selección, con el fin de que se acumulen más mutaciones benéficas. La tasa de mutación utilizada es generalmente no mayor a un aminoácido por proteína, debido a las limitantes de muestreo y a que con tasas mayores gran parte de la librería resulta no funcional. Esta herramienta ha probado ser muy eficiente para el mejoramiento de la estabilidad o actividad de las enzimas, sin embargo no es adecuada si el cambio en el péptido requiere la modificación simultánea de más de un aminoácido (Schmidt-Dannert C., 2001).

De manera complementaria a la mutación puntual al azar, se utiliza la recombinación de los genes que resultan del primer método. Esto permite la combinación aditiva de mutaciones benéficas y la eliminación de mutaciones deletéreas. El método más utilizado para la recombinación es el *shuffling* de DNA (*DNA shuffling*) desarrollado por Stemmer (Stemmer W. P. *et al.*, 1994^a) y las variantes que de ahí han derivado. Éstas metodologías incluso han sido utilizadas independientemente de la mutación puntual al azar para recombinar genes homólogos, buscando generar variabilidad en lo que se conoce como *shuffling* de familia (*family shuffling*) (Cramer A. *et al.*, 1998). La ventaja de éste es que el punto de partida son proteínas naturales que ya son viables y cuya secuencia es el resultado de un largo proceso evolutivo natural evitándose, entre otras cosas, las mutaciones deletéreas que pudieran surgir en la mutación al azar. Se han desarrollado otros métodos para hacer más eficiente la recombinación como el RACHITT desarrollado por Coco *et al.* (2001). A pesar de sus aportaciones, la limitante de este tipo de métodos es que dependen del reensamblaje por PCR de fragmentos cortos azarosos generados a partir de genes homólogos, lo que demanda altos niveles de identidad de secuencia entre los fragmentos a reensamblar y, por lo tanto, de los genes a ser recombinados (Schmidt-

Dannert C., 2001). En general, los métodos con los que se cuenta actualmente están limitados a dos eventos mutacionales, sustitución de bases y recombinación homóloga, y por lo tanto sería de gran utilidad para la evolución dirigida el desarrollo de sistemas que permitieran la recombinación de genes no homólogos, es decir, no dependientes de la secuencia. Esto permitiría, entre otras cosas, la recombinación de plegamientos estructurales homólogos con poca identidad de secuencia, así como de motivos y dominios estructurales no homólogos, procesos que en la naturaleza han sido y son importantes eventos evolutivos. Algunos métodos como el ITCHY (Ostermeier M. *et al.* 1999) o el SHIPREC (Sieber V. *et al.* 2001) empiezan a ser desarrollados en esta dirección (Schmidt-Dannert C., 2001). Y se espera que el desarrollo en paralelo de métodos para el muestreo de las variantes y generación de variabilidad impulse la evolución dirigida (Cohen N. *et al.*, 2001). Este es el contexto en el cual es desarrollado el presente trabajo.

Barriles α/β

En el campo de la ingeniería de proteínas, últimamente se ha puesto gran atención en el grupo de los barriles α/β , barriles TIM o barriles $(\beta\alpha)_8$. Este grupo de proteínas se caracteriza por tener un dominio de alrededor de 250 aminoácidos, conformado por ocho estructuras β -plegadas, rodeadas por siete u ocho estructuras α -hélices (Wierenga R. K., 2001). Cada una de las β -plegadas interiores está conectada a una α -hélice exterior, por lo que se puede considerar que es una estructura $(\alpha/\beta)_8$ (Farber G. *et al.*, 1990). Las ocho β -plegadas forman en el interior una red de puentes de hidrógeno que definen la estructura del dominio. La β -plegada 1 forma puentes de hidrógeno medulares con sus β -plegadas vecinas, la 2 y la 8, y así de manera correspondiente. Usualmente hay sólo una α -hélice entre cada par de β -plegadas vecinas, siendo la región más hidrofóbica la zona entre las β -plegadas y α -hélices, y no la cavidad del barril (Wierenga R. K., 2001). Por otro lado, aunque este motivo estructural es llamado barril, su forma usual es elíptica y no circular, donde la posición del eje mayor varía. La altura del barril también varía considerablemente de proteína a proteína (Reardon D. *et al.*, 1995). Adicionalmente, las enzimas que presentan el dominio α/β , comúnmente tienen otras estructuras que preceden, interrumpen (en las asas que conectan las β -plegadas a las α -hélices) o siguen al barril. La posición de estas

estructuras, en combinación con la forma del barril, ha sido utilizada para dividir estas proteínas en diferentes familias. La proteína arquetipo del barril α/β , y la primera en la que se describió la estructura, es la tiroso fosfato isomerasa (TIM), de donde viene el nombre de barril TIM (figura 2) (Banner *et al.*, 1975; Farber G. *et al.*, 1990).

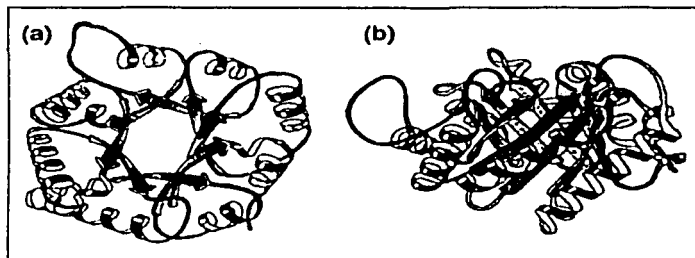


Figura 2. Uno de los primeros dibujos de una de las unidades barril α/β de la enzima triosa fosfato isomerasa (TIM), primera enzima cristalizada con esta estructura. (a) Vista superior, (b) vista de costado, perpendicular al eje (Obtenida de Farber G. *et al.*, 1990).

El grupo de proteínas que tienen dominios barril α/β es muy amplio; se calcula que aproximadamente 10% de las enzimas pueden contener esta estructura, siendo probablemente la estructura enzimática más común. En la actualidad, la base de datos de estructuras de proteínas SCOP distingue 23 superfamilias de barriles TIM (para mayor información de las familias consultar Farber G. *et al.*, 1990; Reardon D. *et al.*, 1995; Copley R. *et al.*, 2000; Nagano N. *et al.*, 2002). Las proteínas con este dominio son principalmente enzimas: enzimas metabólicas y glicohidrolasas extracelulares, más que nada (Bork P. *et al.*, 1995). Los barriles α/β son enzimas versátiles que actúan como oxidoreductasas, transferasas, liasas, hidrolasas o isomerasas (Höcker B. *et al.*, 2001^b) e incluyen algunas de las enzimas más eficientes (Wierenga R. K., 2001). Los residuos del sitio activo se encuentran en el carboxilo (C) terminal de las β -plegadas y en las asas (*loops*) que las unen con la α -hélice contigua, en lo que se conoce como la cara catalítica del barril. Por el contrario, las asas que unen a las α -hélices con las β -plegada en la cara

contraria del barril son esenciales para la estabilidad de la enzima. Esta compartimentalización poco común es atractiva desde el punto de vista de la ingeniería de proteínas, ya que permite pensar en la posibilidad de modificar una propiedad sin alterar la otra (Höcker B. *et al.*, 2001^b).

A pesar de las similitudes en su estructura, los barriles α/β realizan una amplia gama de actividades funcionales lo que, junto a su poca relación a nivel de estructura primaria, ha motivado un estudio intensivo de su evolución. Además las características anteriores las señala como buenas candidatas para ser utilizadas en ingeniería de proteínas (Bork P. *et al.*, 1995). De manera general, se han propuesto dos caminos evolutivos del grupo: el convergente y el divergente. Es posible que el motivo estructural sea particularmente estable y que muchas proteínas ancestrales, con diferentes orígenes, hayan evolucionado hacia esta forma en un proceso de convergencia. La evidencia principal de esta ruta es la falta de homología en las secuencias entre los miembros de la familia α/β . Alternativamente, todas las proteínas α/β podrían ser descendientes de un ancestro común y la falta de homología entre las secuencias podría ser debida a la antigüedad de la divergencia, considerando que la estructura tridimensional evoluciona más lento que la estructura primaria (Reardon D. *et al.*, 1995). Para sustentar esta propuesta se han utilizado diferentes evidencias, como la posición similar del sitio activo y otros caracteres estructurales (Farber G. *et al.*, 1990). Cada uno de los miembros de las 23 superfamilias de barriles α/β presentadas en la base de datos de estructuras de proteínas SCOP, tienen un probable ancestro común y un estudio reciente realizado por Copley R. *et al* (2000) provee evidencia estadística confiable que indica que al menos 12 de las 23 superfamilias comparten un ancestro común. En el trabajo más reciente realizado al respecto se propone agrupar a los barriles TIM en 21 familias y se proveen evidencias que sugieren un ancestro común para 17 de estas 21 familias (Nagano N. *et al.*, 2002). En la actualidad, se puede decir que un proceso de evolución divergente a partir de un ancestro común explica mejor la información disponible, en comparación con un proceso de evolución convergente a una estructura estable (Reardon D. *et al.*, 1995). Incluso, se proponen los genes *hisA* y *hisF*, que pertenecen al operón de biosíntesis de histidina, como los posibles ancestros comunes. Lo anterior es debido a un singular patrón interno en la secuencia y estructura de estas proteínas en el que se observaba una repetición de la mitad del tamaño de la molécula, lo

que sugiere un fenómeno de duplicación/fusión de un gen ancestral. Se propone entonces un modelo evolutivo en el que, la primera duplicación génica de la secuencia de un medio barril ancestral fue seguida por fusión génica y mutación adaptativa de las secuencias iniciales idénticas de los medios barriles, para formar un barril α/β completo con un amplio espectro funcional (figura 3) (Fani R. *et al.*, 1994, Thoma R. *et al* 1998, Lang D. *et al* 2000). Recientemente se logró la expresión por separado de ambas mitades del gen *hisF* de *Thermotoga maritima* en *Escherichia coli*, observándose que cada mitad es una proteína plegada *in vitro*, aunque ambas forman cuerpos de inclusión *in vivo* y no son funcionales (Höcker B. *et al.* 2001^a). Además, de modo inesperado se ha encontrado similitud estructural y de secuencia de aminoácidos entre el medio barril HisF-C y miembros del plegamiento tipo falvodoxina, sugiriéndose que el medio barril es un dominio estructural que evoluciona independientemente (Höcker B. *et al.*, 2002). Sin embargo, los únicos casos de enzimas barril α/β donde se ha encontrado evidencia de este fenómeno son los genes *hisA* y *hisF*. Incluso, análisis detallado de estas proteínas indica que posiblemente el barril completo es el resultado evolutivo de repetidos eventos de duplicación y fusión de un dominio menor que un medio barril $(\beta\alpha)_4$, es decir, $(\beta\alpha)_2$ o $(\beta\alpha)_1$ (Nagano N. *et al.*, 2002; Brill M. *et al.*, 2002).

Dicho lo anterior, es claro que la evolución de las proteínas barril α/β sigue siendo un punto polémico. Aún cuando es posible que con el descubrimiento de nuevas proteínas y la resolución de más estructuras tridimensionales la historia evolutiva de la familia pueda llegar a aclararse, el surgimiento de nuevos datos también podría complicarla aún más.

Sin embargo, las características tanto evolutivas como estructurales y funcionales antes mencionadas hacen que las enzimas barriles α/β sean excelentes candidatos para ser utilizadas en ingeniería de proteínas y evolución dirigida. Son varios los trabajos de ingeniería de proteínas realizados con estas enzimas, dentro de los cuales cabe destacar la conversión de la función de *hisA* hacia *trpF* que implica el cambio de especificidad de sustrato (Jürgens C. *et al.*, 2000). Este es de gran importancia debido a que refleja uno de los dos principales procesos de divergencia evolutiva en enzimas, el cambio de especificidad manteniendo el mecanismo catalítico (Gerlt J. A. *et al.*, 2001). Todo indica que el barril α/β en el futuro será una de las opciones de plegamiento que permitirá la

creación de enzimas con actividades catalíticas totalmente nuevas (Wierenga R. K., 2001, Wilson Miles E. *et al.*, 2000).

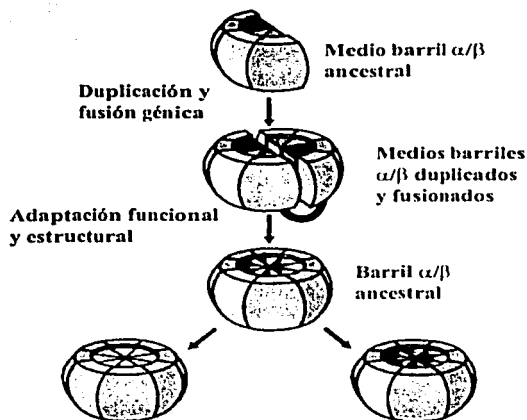


Figura 3. Modelo evolutivo de duplicación-fusión de un medio barril ancestral dando lugar al barril completo ancestral (Obtenido de Lang D. *et al.*, 2000).

El entendimiento de las estrategias y mecanismos de la naturaleza en la evolución de enzimas parece ser fundamental al proveer la información necesaria para la confección de proteínas. Existe suficiente evidencia que sugiere que muchas proteínas, si no es que todas, evolucionaron a partir de la integración de genes, ya sea homólogos o no. Esto puede ser resultado del barajeo de exones y/o de mecanismos genéticos específicos como los rearrreglos de transposones o la recombinación sitio-específica (Riechmann L. *et al.*, 2000, Gerlt J. A. *et al.*, 2001, Wilson Miles E. *et al.*, 2000). Es de especial interés para este estudio la idea, surgida principalmente a partir de la observación de la simetría de la estructura y de las evidencias de posibles fracciones de barriles independientes ancestrales, de trabajar con unidades menores que un dominio barril $(\beta\alpha)_x$ (Gerlt J. A. *et al.*, 2001, Wilson Miles E. *et al.*, 2000). Lo anterior permitiría aumentar la variabilidad de las proteínas con las que se trabaja y combinar péptidos de diversos orígenes, en la búsqueda

de nuevas propiedades enzimáticas de los dominios proteicos barril α/β . Las evidencias de la ocurrencia de este tipo de eventos en la naturaleza apuntan a la factibilidad de esta idea.

Barriles α/β en fragmentos

A pesar de que la mayor parte de los estudios en los que se trabaja con fragmentos de barriles α/β no están enfocados a la creación de variabilidad y conjugación de péptidos de orígenes diferentes en la búsqueda de nuevas propiedades enzimáticas sino a cuestiones de plegamiento y evolución, el conocimiento obtenido en éstos bien pueden ser utilizado como base teórica para el primer fin. Por lo mismo, en estos trabajos no se le ha dado gran importancia a que la proteína resultante de la unión de los fragmentos sea funcional, es decir, que complemente la función, lo cual es un punto esencial para el objetivo de crear nuevas propiedades catalíticas.

En uno de los primeros trabajos realizados con el fin de estudiar los procesos de plegamiento, la subunidad α de la enzima triptofano sintasa (α TS) (un barril α/β) fue sometida a proteólisis parcial, produciéndose dos fragmentos contiguos α -1 y α -2 (α -1: de la β 1 a la β 6 y α -2: de la α 6 a la α 8) que pueden reensablarse en un complejo activo *in vitro* (Miles E. W. *et al.*, 1982). Lo anterior sirvió, en cierta forma, como antecedente para que un modelo de plegamiento parecido se probara en la enzima fosforribosil antranilato isomerasa (PRAI) de *Saccharomyces cerevisiae*. El barril α/β fue cortado entre el sexto y el séptimo módulo $\beta\alpha$, y cada fragmento resultante [$(\beta\alpha)_6$ y $(\beta\alpha)_2$] fue expresado de manera separada para probar su capacidad de adoptar la conformación nativa de manera autónoma. De igual manera que con la α TS, los fragmentos *in vitro* se rearreglan dando lugar a una proteína casi totalmente activa, aunque *in vivo* la función no se complementa y los péptidos forman cuerpos de inclusión (Eder J. *et al.*, 1992). Las observaciones de estudios de plegamiento con PRAI de *E. coli* son consistentes con estos resultados (Jasanoff A. *et al.*, 1994). En otro trabajo parecido, con el propósito de investigar cómo interfieren las modificaciones estructurales con la estabilidad de las proteínas, se modificó una unidad $\beta\alpha$ de la triosa fosfato isomerasa (TIM) de *E. coli*, una típica proteína barril α/β , asumiendo que el barril α/β pseudo-simétrico puede ser dividido en 8 motivos sucesivos formados por

un asa, una β plegada, un asa y una α hélice. El octavo de estos motivos de la TIM de *E. coli* fue remplazado por el motivo correspondiente de la enzima de pollo, logrando una proteína híbrida activa *in vitro* y con una estabilidad comparable a la de la TIM silvestre (Mainfroid M., *et al.*, 1993). Con la misma enzima pero de músculo de pollo, y con el objetivo de estudiar las implicaciones evolutivas de la organización génica en exones e intrones sobre los péptidos codificantes, se cortó la enzima por separado en tres diferentes límites exón/intrón y los complejos se analizaron catalíticamente. Se encontró que la expresión de ambas porciones de los tres diferentes cortes *in vivo* complementa la función, aunque de manera pobre (las colonias tardan 5 días en crecer). Lo mismo sucede al medir la actividad *in vitro*, pues hay actividad detectable de TIM pero mucho más baja que aquella de la enzima silvestre (75 000 veces más baja) (Bertolaet B. *et al.*, 1995). Posteriormente, y de nuevo con la subunidad α de la triptofano sintasa (α TC) de *E. coli*, se produjeron 8 diferentes fragmentos amino – terminal del barril α/β , los cuales fueron caracterizados por dicroísmo circular (DC) y resonancia magnética nuclear. Los 8 fragmentos de 105 hasta 214 residuos, que contienen sucesivamente más elementos de estructura secundaria, se hicieron por mutagénesis con codones de paro. Aunque no se midió la actividad enzimática de ninguno de los fragmentos, el estudio arrojó resultados interesantes: principalmente se encontró que fragmentos grandes (> 100 residuos) se pueden comportar como unidades de plegamiento autónomas y que la ocurrencia de unidades $\beta\alpha\beta$, que incluso son menores, se ve favorecida (Zitzewitz J. A. *et al.*, 1999). Con el afán de investigar la propuesta de que los dominios protéicos han evolucionado a través del ensamblaje y/o intercambio combinatorio de fragmentos polipeptídicos más pequeños, recientemente se fusionó el DNA codificante de la mitad amino – terminal del barril α/β de la proteína CspA con DNA genómico fragmentado de *E. coli* y las quimeras fueron clonadas para ser desplegadas en fagos filamentosos. Los péptidos plegados fueron seleccionados por proteólisis. Cuatro de los péptidos se expresaron como proteínas solubles, formando monómeros y exhibiendo desdoblamiento cooperativo. Incluso uno de ellos es más estable que la proteína CspA completa. Estos resultados tienen importantes implicaciones importantes tanto para la evolución de las proteínas a partir de péptidos independientes, como para la posibilidad de crear nuevas proteínas por medio de este proceso (Riechmann L. *et al.*, 2000). Como se mencionó en la sección anterior, el gen *hisF* de *T. maritima* también ha sido dividido en

dos fragmentos $[(\beta\alpha)_{1-4}]$ y $(\beta\alpha)_{5-8}]$, los cuales al ser conjugados *in vitro* complementan la función (Höcker B. *et al.* 2001^a).

Con el propósito de utilizar barriles α/β en fragmentos para ingeniería de proteínas y específicamente para evolución dirigida, nuestro grupo también ha trabajado con complementación de fragmentos de PRAI de *E. coli*. La enzima fue cortada en tres diferentes posiciones en las asas inferiores, entre las β -plegadas y las α -hélices, para obtener dos fragmentos distintos en cada corte: un corte entre $\alpha 2$ y $\beta 3$ $[(\beta\alpha)_2]$ y $(\beta\alpha)_6]$, uno entre $\alpha 6$ y $\beta 7$ $[(\beta\alpha)_6]$ y $(\beta\alpha)_2]$ y por último uno entre $\alpha 4$ y $\beta 5$ $[(\beta\alpha)_4]$ y $(\beta\alpha)_4]$. Al expresar como policistrón cada pareja de fragmentos, sólo los fragmentos que resultaron del corte entre $\alpha 4$ y $\beta 5$ $[(\beta\alpha)_4]$ y $(\beta\alpha)_4]$ complementan la función *in vivo*. Cuando se expresan estos últimos fragmentos $[(\beta\alpha)_{1-4}]$ y $(\beta\alpha)_{5-4}]$ en vectores diferentes, la complementación disminuye notoriamente (las colonias tardan entre 3 y 4 días en crecer) (Fuentes P., 2001). Al añadir un cierre (*zipper*) de leucina como puente de unión entre las mitades producto de este corte expresadas en vectores diferentes, se logra la complementación total de la función (crecimiento igual a la PRAI silvestre) (Saab-Rincón G., resultados no publicados).

Todas estas evidencias, además de dar buenas bases y direcciones para trabajar con fragmentos de barriles α/β , refuerzan la factibilidad de la idea.

Recombinación *in vivo*

Con el propósito de trabajar con fragmentos de barriles α/β y buscando una herramienta que permita combinar o barajar estos fragmentos de manera más eficiente que su simple expresión por separado en una misma célula (como en Fuentes P., 2001) se llegó a la recombinación *in vivo*. Se ha sugerido que la recombinación entre exones ha sido un fenómeno clave en la evolución natural de las proteínas y, por lo tanto, también se ha propuesto que esta estrategia puede ser utilizada en la evolución artificial de estas biomoléculas (Fisch I. *et al.*, 1996). Al parecer esta estrategia tiene la ventaja de permitir la combinación y asociación de genes pre-transcripcionalmente, además de que al llevarse a cabo *in vivo*, no tiene las limitantes del proceso de transformación por electroporación que suelen tener los sistemas de recombinación *in vitro*, tales como el barajeo de genes (*gene shuffling*) y sus variantes. La recombinación sitio específica puede ser inducida cuando se

introducen moléculas de ácidos nucleicos con los sitios reconocidos por las recombinasas específicas a una bacteria que expresa la recombinasa que actúa en esos sitios. La mayor parte de los sistemas de recombinación sitio específico en bacterias caen dentro de dos grandes familias de enzimas que realizan procesos genéticos similares como la integración y escisión de genomas de bacteriófagos, y la regulación de la expresión génica y el número de copias de un plásmido. Una de estas es la de las integrasas o tirosin recombinasas que cuenta con más de 100 miembros, mientras que la otra es la que comprende a las enzimas resolvasas (Sblattero D. *et al.*, 2001). La recombinación *in vivo* para generar variabilidad en péptidos ha sido utilizada, más que nada, en la generación de grandes bibliotecas de anticuerpos, ya sean Fab's o scFvs (Waterhouse P. *et al.*, 1993; Griffiths A. D. *et al.*, 1994; Geoffroy E. *et al.*, 1994; Davies J. *et al.*, 1995; Tsurishita N. *et al.*, 1996; Zahra D. G. *et al.*, 1999; Sblattero D. *et al.*, 2000), existiendo sólo un trabajo donde se utiliza la recombinación *in vivo* con péptidos pequeños, que no son anticuerpos, pero sí ligandos (Fisch I. *et al.*, 1996). Las recombinasas utilizadas en todos los trabajos anteriores pertenecen a la familia de las integrasas, y son la lambda recombinasa y la Cre recombinasa. El sistema de la lambda recombinasa aprovecha el mecanismo utilizado por el fago lambda para integrarse al genoma bacteriano durante su ciclo de replicación. En éste intervienen cuatro diferentes proteínas (Int, Xis, IHF y FIS), cuyas interacciones determinan la dirección de la recombinación, y dos diferentes sitios *att* (*attP* y *attB*) uno en el fago y otro en la bacteria. Estos sitios, después de la primera fase, la recombinación integrativa, son convertidos a otros dos sitios *att* diferentes (*attL* y *attR*) necesarios para la segunda y última fase de recombinación escisiva (para más detalles ver Landy A., 1989). De esta manera, flanqueando los segmentos de DNA a recombinar con las combinaciones apropiadas de sitios *att*, en la presencia de las proteínas apropiadas, se puede dirigir la recombinación según sea requerida (Sblattero D. *et al.*, 2001). Este sistema, que ha sido el menos empleado, fue utilizado con anticuerpos monoclonales de chimpancé anti-HIV gp160 Ab (Geoffroy E. *et al.*, 1994). Por medio del sistema, se asoció irreversiblemente el plásmido y el fagémido, portadores de la cadena ligera y la cadena pesada del anticuerpo, respectivamente. Los vectores correctamente recombinados fueron seleccionados por medio de un marcador genético creado únicamente después de un evento correcto de

recombinación. Sin embargo, con este sistema no ha sido creada alguna librería funcional de tamaño confiable.

Por el otro lado, el sistema de la Cre recombinasa, también denominado *loxP/cre*, es probablemente el sistema de recombinación sitio específico más usado y no sólo en la creación de grandes librerías de anticuerpos. El sistema es una herramienta importante en la manipulación *in vivo* de genomas eucariotes. Por medio de la expresión controlada de la Cre recombinasa, se ha logrado aplicar estrategias de desactivación y activación génica en células mamíferas, así como crear ratones transgénicos (Siegel R. W. *et al.*, 2001). Este sistema fue descubierto al aislar un fragmento del genoma del bacteriófago P1 necesario para la recombinación (*locus* de *recombinación* por lo tanto *lox*) y un gen cuyo producto causa la *recombinación* (por lo tanto *cre*) (Sblattero D. *et al.*, 2001). Este sistema puede estar involucrado en varios procesos *in vivo*, entre los que se encuentra la partición del genoma después de la replicación. El sistema está compuesto por la proteína Cre recombinasa de 34 kDa, la cual media la recombinación entre dos sitios *loxP* de 34 pares de bases (pb) y no requiere ningún factor accesorio. El sitio *loxP* está constituido por dos repeticiones invertidas de 13 pb que flanquean una secuencia central asimétrica de 8 pb, dándole al sitio una apariencia de direccionalidad (figura 4). La recombinación mediada por la Cre recombinasa de dos sitios *loxP* alineados escinde el DNA intermedio, mientras que la misma recombinación, pero entre sitios *loxP* invertidos, resultara en la inversión del mismo DNA (Siegel R. W. *et al.*, 2001). La estructura cristalográfica ha revelado que a cada una de las repeticiones invertidas de 13 pb, junto con 2 bases de la parte central, se une un monómero Cre y la recombinación involucra la formación de una estructura o unión tipo-Holliday (Guo F. *et al.*, 1997; Gopaul D. N. *et al.*, 1998). El intercambio de hebras comienza después de que se producen cortes de sólo una hebra en sitios específicos dentro de cada centro de 8 pb, por medio de una tirosina conservada en dos subunidades Cre, generando nucleófilos que completan el intercambio de un juego de hebras. La estructura de tipo-Holliday intermediaria es resuelta hacia los productos recombinantes después del corte e intercambio del par restante de hebras de DNA por el segundo par de subunidades Cre (figura 5) (Guo F. *et al.*, 1997; Gopaul D. N. *et al.*, 1998; Siegel R. W. *et al.*, 2001). Diferentes estudios han revelado que las mutaciones en la región central del sitio *loxP* tienen efectos profundos en la recombinación. Únicamente los pares de sitios que tienen

identidad en la región central de 8 pb son aptos para recombinar entre sí (Sblattero D. *et al.*, 2001). Se han encontrado cinco sitios que tienen alguna mutación en la parte central del sitio que recombinan con un 100% de eficiencia con sitios idénticos, pero con mínima eficiencia con otros sitios *loxP* diferentes (Siegel R. W. *et al.*, 2001). Uno de los sitios más comúnmente usados es el *loxP511*, que tiene la substitución de una A por una G en la séptima base del segmento central (Sblattero D. *et al.*, 2001). Actualmente se cuenta con métodos que permiten estudiar las eficiencias de recombinación *in vivo* entre todas las combinaciones de sitios *loxP*, lo que abre nuevas posibilidades para la explotación del sistema (Siegel R. W. *et al.*, 2001).

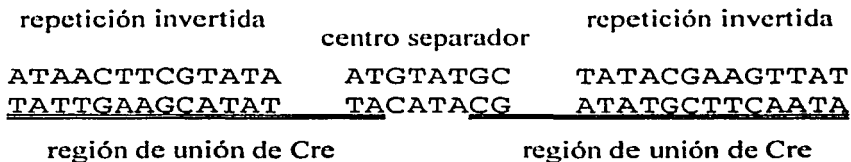


Figura 4. Secuencia del sitio *loxP* silvestre (Sblattero D. *et al.* 2001).

El sistema *loxP/cre* fue usado por primera vez para la creación de una librería de anticuerpos, con el fin de sobrepasar las limitaciones que representan los procesos de ligación y transformación para el tamaño de la librería. Se elaboró un sistema modelo en el que se echa mano de la infección combinatoria y la recombinación *in vivo* para la generación de la librería (Waterhouse P. *et al.*, 1993). En éste se transforma *E. coli* con un repertorio de cadenas pesadas clonadas en un plásmido, para después ser infectadas con un repertorio de cadenas ligeras clonadas en un fago (infección combinatoria). Dentro de la célula, ambas cadenas son fusionadas en un mismo vector por recombinación *in vivo*, ya que el repertorio de cadenas pesadas está flanqueado por dos sitios *loxP* diferentes (*loxPwt* y *loxP511*) y el repertorio de cadenas ligeras está acompañado por una cadena pesada nula idéntica en todos los fagos pero también flanqueada por los mismos sitios *loxP* (*loxPwt* y

loxP511), idéntica en todos los fagos pero también flanqueada por los mismos sitios *loxP* (*loxPwt* y *loxP511*). Al ocurrir un evento de recombinación entre los sitios *loxPwt* y otro

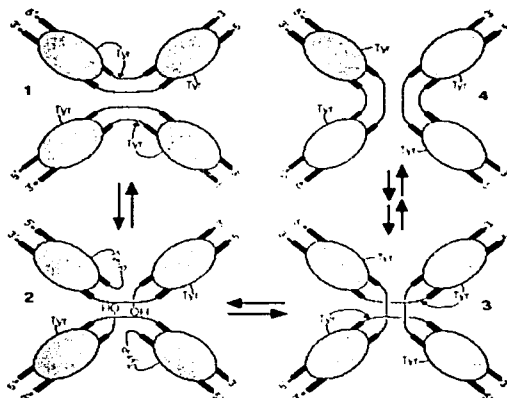


Figura 5. Mecanismo molecular de la reacción de recombinación realizada por la Cre recombinasa entre dos sitios *loxP* homólogos (Obtenida de Guo F. *et al.*, 1997).

entre los sitios *loxP511* del plásmido y el fago, la cadena pesada nula, presente en el fago con la cadena ligera, es intercambiada por las diferentes cadenas pesadas del repertorio clonado en el plásmido (figura 6). Dado que cada célula transformada con un plásmido con la cadena pesada es infectada por un fago portador de una cadena ligera, y la recombinación es reversible, al final de la recombinación cada célula contiene una nueva combinación funcional de VH/VL, así como tres no funcionales (figura 6). Debido a que la infección es extremadamente eficiente y casi todas las células en un cultivo pueden ser infectadas, la diversidad combinatoria de fragmentos Fab puede ser tan grande como el número de células en el cultivo (Waterhouse P. *et al.*, 1993).

Subsecuentemente, este sistema modelo fue aplicado a la creación de una librería combinatoria de Fab altamente diversa (Griffiths A. D. *et al.*, 1994). Para esto se prepararon dos librerías, una de cadenas pesadas de 10^8 clonada en el plásmido y

flanqueada por los dos sitios *loxP* heterólogos antes mencionados, y la segunda de cadenas ligeras de 8×10^5 clonada en el fago que contiene la cadena pesada nula flanqueada por los mismos sitios *loxP* heterólogos. Estas fueron transformadas e infectadas, respectivamente, en un cultivo de *E. coli*, para por último introducir la Cre recombinasa infectando con fago P1. Por el número de células cultivadas, se estimó que la biblioteca resultante contiene cerca de 6.5×10^{10} diferentes anticuerpos desplegados en fago. A pesar de sus ventajas, este sistema tiene algunos problemas: (i) el sistema es reversible, por lo que la librería resultante estará contaminada con combinaciones no funcionales; (ii) el uso de fagos filamentosos hace que el sistema sea poco estable, resultando en la delección del anticuerpo; (iii) la recombinación entre múltiples vectores en una sola célula no es posible; y (iv) los productos de la recombinación son finales, sin que exista la posibilidad de que sean recombinados nuevamente (Sblattero D. *et al.*, 2001).

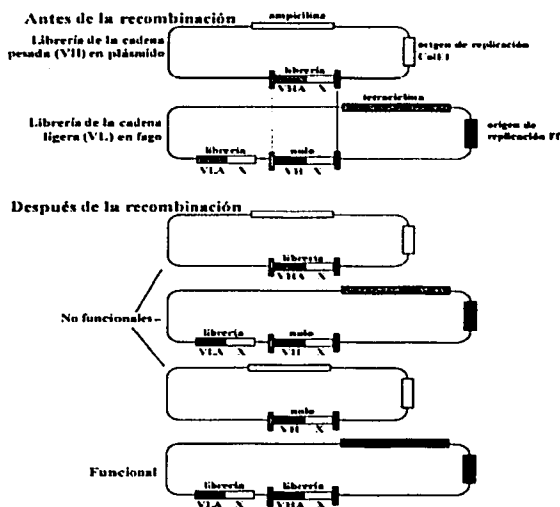


Figura 6. Recombinación usando el sistema de dos vectores. Como punto de partida se tiene el plásmido que contiene la librería de la cadena pesada (VH) acompañada de un gen X y un fago que contiene la librería de la cadena ligera (VL) acompañada de un gen X. Después de la recombinación se obtienen la combinación funcional VHA/VLA, además de los vectores no funcionales (Obtenida de Sblattero D. *et al.*, 2001).

Posteriormente, el sistema ha sido utilizado en varios estudios con algunas variantes. Primero se utilizó en la generación de un gran repertorio de cadenas pesadas de humano camellizadas (Davies J. *et al.*, 1995). Esto se logró por medio de la integración de un sitio *loxP* en la secuencia codificante de la cadena pesada. El repertorio del extremo 5' de la cadena pesada, en el que el asa H2 de la cadena fue remplazada por el sitio *loxP*⁵¹¹, fue clonado en el fago aceptor, mientras que el repertorio del extremo 3' fue clonado en el plásmido donador entre los sitios *loxP* correspondientes. Después de la recombinación se restableció el dominio completo de la cadena pesada observándose que éste se pliega como un típico dominio inmunoglobulino. El sitio *loxP*^{wt} también fue utilizado como péptido conector con un sistema *loxP/cre* similar al descrito por Waterhouse *et al.* (1993), pero uniendo la cadena pesada y la cadena ligera de un anticuerpo. Lo último al desarrollarse nuevos vectores para la recombinación *in vivo* de cadenas pesadas y ligeras de inmunoglobulinas (Tsurushita N. *et al.*, 1996). Al igual que en el caso anterior, se demostró que el péptido conector no perturba la asociación entre ambas cadenas. Este sistema fue utilizado para recombinar un scFv sencillo cuando fue probado el sistema, pero no ha sido utilizado para construir librerías.

Como se mencionó al inicio, existe un único estudio donde se utiliza el sistema *loxP/cre* para la creación de grandes librerías de péptidos que no son anticuerpos. En éste se desarrolla una especie de estrategia de barajeo de exones, construyéndose dos repertorios de péptidos pequeños de 10 residuos al azar. Uno de los repertorios fue clonado en un plásmido flanqueado por los sitios *loxP* heterólogos y el otro en un fago que contiene una región adyacente también flanqueada por los mismos sitios *loxP*, para después, por infección combinatoria y recombinación, armar una librería combinatoria de péptidos de más de 10¹¹ miembros diferentes. La parte más innovadora de este estudio es que el sitio *loxP* necesario entre ambos péptidos para la recombinación fue colocado en un intrón autoeditable (*self-splicing*) del grupo I que es funcional en bacterias, por lo que el sitio fue escindido después de la recombinación, evitándose así posibles problemas de plegamiento (Fisch I. *et al.*, 1996). El repertorio final fue seleccionado para actividades de unión a un anticuerpo monoclonal y a dos enzimas.

Con el afán de mejorar algunas de las limitaciones que tiene el sistema desarrollado por Waterhouse *et al.*, (1993) recientemente fue desarrollado un sistema *loxP/cre*

genéticamente más estable y que permite seleccionar las clonas que se recombinaron correctamente (Zahra D. G. *et al.*, 1999). La primera limitación se superó utilizando un sistema de fagémido en lugar de uno de fago, mientras que la selección se logró por medio de un marcador genético parecido al utilizado en el sistema de la lambda recombinasa: el gen de resistencia a cloranfenicol es colocado por enfrente de su promotor cuando ocurre la recombinación de manera correcta.

El último de los sistemas desarrollados de cre recombinación *in vivo*, a diferencia de todos los anteriores incluyendo el de la lambda recombinasa, utiliza sólo un vector en lugar de dos (Sblattero D. *et al.*, 2000). Este sistema hace a un lado las ideas de que una bacteria no puede ser infectada por más de un fago o fagémido y de que diferentes plásmidos con un mismo origen de replicación no pueden coexistir dentro de una misma célula, aprovechando de mejor manera los potenciales de la recombinación *in vivo*. De esta forma se pretende que la recombinación dentro de una célula no se lleve a cabo únicamente entre dos moléculas de DNA, como en los sistemas anteriores, sino entre muchas. El sistema también está basado en la idea de una librería de cadenas pesadas y ligeras separadas por dos sitios *loxP* no homólogos pero, en lugar de clonar cada librería en diferentes vectores, se clonan en un mismo fagémido y la recombinación es empleada para una mayor combinación de éstas (Sblattero D. *et al.*, 2001). La librería es infectada con una alta multiplicidad de infección (MOI) para asegurar que cada célula sea infectada por múltiples fagémidos y así sea posible la múltiple recombinación entre éstos (figura 7). En el estudio, primero se analizaron los seis posibles marcos de lectura de los sitios *loxP* silvestre y *loxP511*, encontrando una traducción del sitio *loxP511* (ITSYNNVYYTKL) que tiene sólo un aminoácido básico, no tiene codones de paro y es la menos hidrofóbica, optimizando el péptido conector. Su capacidad de funcionar como tal fue probada con diferentes anticuerpos. Con este sitio *loxP511*, se construyó un fagémido (pDAN5) (figura 8) en el que se clonó una librería de anticuerpos preconstruida (7×10^7), con la cadena ligera fuera de los sitios *loxP* y la cadena pesada dentro. Dicha librería clonada en el nuevo vector fue infectada con una alta multiplicidad de infección en una cepa de *E. coli* que expresa de manera constitutiva la Cre recombinasa, con el fin de llevar a cabo la recombinación y aumentar la variabilidad de la librería. El sistema de recombinación fue probado con la complementación de un cierto anticuerpo clonado en el mismo vector, pero cada cadena por

separado; posteriormente fue utilizado para la creación de la amplia librería antes mencionada. Además, se determinó el número de fagémidos que entran en una sola célula, así como el grado de recombinación que hay entre éstos. Con lo anterior, y con el número de células utilizadas, se estimó que la librería final obtenida fue de alrededor de 3×10^{11} , mientras que si todas las posibles combinaciones de genes estuvieran presentes, la librería podría llegar a ser hasta 10 veces mayor. Se espera que todos los anticuerpos dentro de la librería sean funcionales, además de que el procedimiento de recombinación puede ser repetido indefinidamente. Curiosamente, el sistema no sufre de incompatibilidad de plásmidos, ya que se observa que la reacción de recombinación alcanza el equilibrio sin que haya ningún plásmido específico presente en exceso sobre los otros (Sblattero D. *et al.*, 2000).

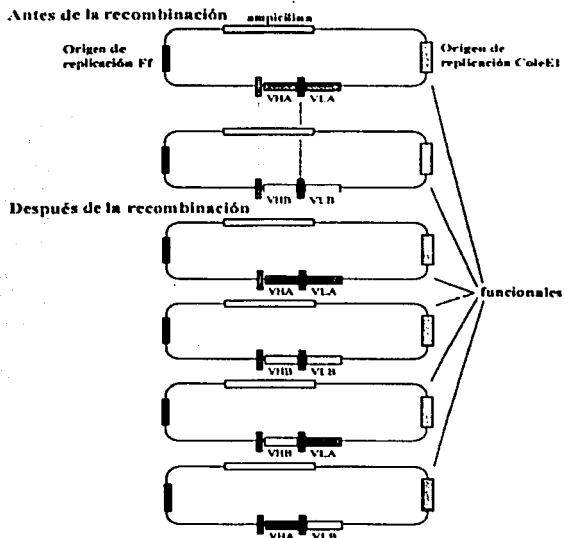


Figura 7. Recombinación con el sistema de un solo vector. Una sola bacteria es infectada por dos fagémidos o más conteniendo las combinaciones originales funcionales VHA/VLA y VHB/VLB. Después de la recombinación se forman las 4 posibles combinaciones, todas dan anticuerpos funcionales.

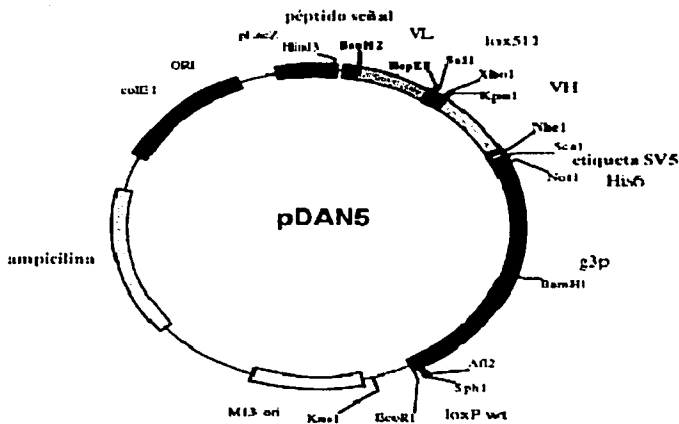


Figura 8. Mapa del vector pDAN5 construido por Sblattero D. *et al.* (2000) para la recombinación de anticuerpos en el sistema de un solo vector (Obtenida de Sblattero D. *et al.*, 2000).

De esta manera, se ha sugerido que el sistema puede ser utilizado para la creación de grandes librerías de cualquier proteína en la que pueda ser insertada una secuencia de recombinación sitio específica (Sblattero D. *et al.*, 2001). Por todas estas características, el citado sistema de recombinación *in vivo* desarrollado por Sblattero D. y Bradbury A. (2000) es muy atractivo como herramienta para trabajar con fragmentos de barriles $\alpha\beta$.

Enzimas modelo

Las proteínas barriles $\alpha\beta$ utilizadas como modelo en el presente estudio son la fosforribosil antranilato isomerasa (PRAI) de *E. coli* y la tirosafosfato isomerasa (TIM) de *Trypanosoma brucei*. Como se mencionó anteriormente, se han realizado estudios con ambas enzimas donde se caracteriza su comportamiento después de fragmentarlas y en

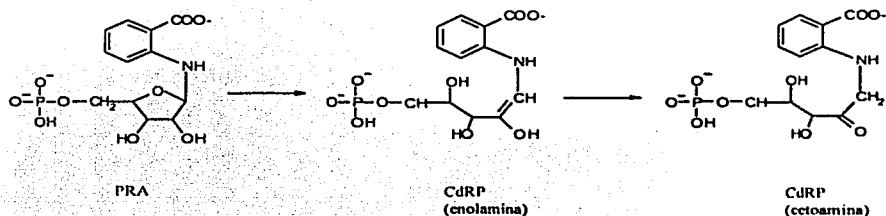
ambos casos se ha logrado la complementación de su función al expresarlas *in vivo* (Jörg E. *et al.*, 1992, Bertolaet B. *et al.*, 1995, Fuentes P., 2001). Además se trata de enzimas bien caracterizadas tanto estructural como funcionalmente. Otro punto importante a favor es que nuestro grupo de investigación ha trabajado con estas enzimas y, por lo mismo, se tiene experiencia en su manejo, además de que se cuenta con los genes clonados y con varios sistemas de selección.

Fosforribosil antranilato isomerasa de Escherichia coli (ePRAI)

La enzima N-(5'Fosforribosil)-antranilato isomerasa (PRAI), codificada por el gen *TrpF* en el operón de biosíntesis de triptofano, es un pequeño barril α/β (figura 9) que cataliza el tercer paso en la biosíntesis de triptofano a partir de corimato (corimato-antranilato-PRA-CdRP-IGP-indol-triptofano) en algunas plantas y microorganismos (Jasanoff A. *et al.*, 1994). En este paso, la PRAI rearrregla el N-(5'Fosforribosil)-antranilato (PRA) a 1-[(2-carboxipentil)amino]-1-deoxiribulosa 5-fosfato (CdRP) en una reacción intramolecular redox conocida como rearreglo de Amadori. Esta reacción involucra pasos elementales de transferencia de protones y posiblemente lleva a una eniolamida como primer producto (esquema 1) (Hommel U. *et al.*, 1995). La estructura cristalográfica de la proteína refinada a 2 Å revela que ésta difiere de un plegamiento $(\beta\alpha)_n$ típico en la presencia de un giro helicoidal adicional en el asa que conecta la β plegada 8 con la α hélice 8, así como en la degeneración de su α hélice 5, la cual está alterada por el residuo P356 y no adopta la geometría helicoidal (Jasanoff A. *et al.*, 1994). El sitio activo se encuentra en el C-terminal del barril conformado por las asas que conectan la β plegada 1 con la α hélice 1 ($\beta 1\alpha 1$) y la que conecta la β plegada 6 con la α hélice 6 ($\beta 6\alpha 6$), además de un sitio de unión a fosfato formado por las asas $\beta 7\alpha 7$ y $\beta 8\alpha 8$ (Altamirano M. M. *et al.*, 2000, Wilmanns M. *et al.*, 1992). Los residuos catalíticos importantes son la Cys7 y la Asp126, ubicados correspondientemente en cada una de las asas catalíticas primeramente mencionadas (Wilmanns M. *et al.*, 1992, Jürgens C. *et al.*, 2000).

La organización estructural de PRAI difiere ampliamente entre los microorganismos. En *Saccharomyces cerevisiae* es un monómero, mientras que en *E. coli*, con la que tiene 28% de identidad de secuencia, se presenta como una bienzima

covalentemente unida a la indol-3-glicerol-fosfato sintasa (IGPS) que cataliza el paso subsecuente en la síntesis de triptofano. A pesar de esto se ha demostrado que ePRA1 y eIGPS son enzimas catalíticamente independientes. Lo anterior ha sido señalado por la estructura cristalográfica (Priestle S. P. *et al.*, 1987), además de que han sido separadas a nivel genético y expresadas como proteínas monofuncionales (Eberhard M. *et al.*, 1995). Las enzimas monofuncionales son proteínas estables, monoméricas con una actividad catalítica completa. Así mismo no se encontró ninguna asociación entre ambas al manejarlas juntas, ni *in vitro* ni *in vivo*. Lo anterior permite trabajar con una enzima monomérica independiente.



Esquema 1. Reacción catalizada por la PRA1.

Evolutivamente, la PRA1 ha sido relacionada con enzimas de la vía de biosíntesis de histidina. Las búsquedas en la Base de Datos de Proteínas (PDB) demuestran que TrpF tiene relaciones estructurales distantes, pero significativas, con HisA y HisF, siendo la segunda proteína más parecida en ambos casos (Lang D. *et al.*, 2000). Así mismo, éstos hallazgos han sido reforzados por los resultados del trabajo de ingeniería de proteínas en el que se generan variantes de HisA con actividad de TrpF (Jürgens C. *et al.*, 2000). Esto último sugiere que ambas enzimas pudieron haber evolucionado a partir de una misma enzima ancestral. Esta hipótesis ha sido reforzada por comparaciones entre los sitios activos y mecanismos catalíticos de HisA y TrpF (Henn-Sax M. *et al.*, 2002). Cabe recordar que HisA y HisF son los barriles $\alpha\beta$ cuyos análisis de secuencia y estructura sugirieron un modelo evolutivo de duplicación-fusión de un medio barril ancestral (Fani R. *et al.*, 1994; Thoma R. *et al.*, 1998; Lang D. *et al.*, 2000).

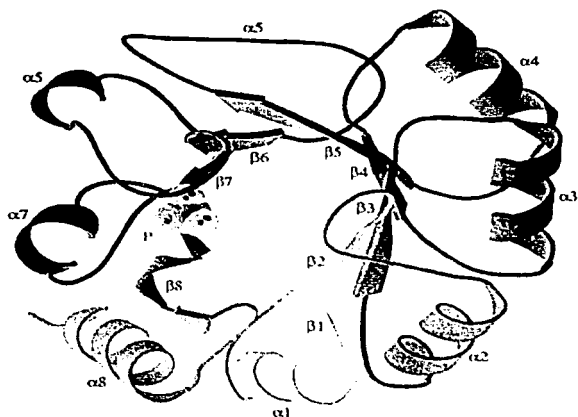


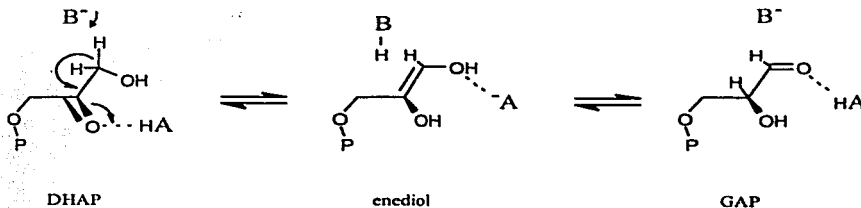
Figura 9. Fosforribosil antranilato isomerasa (PRAI) de *E. coli*.

Finalmente, es importante mencionar algunos estudios de ingeniería de proteínas con PRAI, además de los ya comentados, para destacar las posibilidades de confeccionar y, principalmente, de insertar un péptido ajeno en este barril α/β . ePRAI ha sido permutada circularmente de dos diferentes maneras, y las enzimas resultantes han sido catalíticamente activas, y prácticamente indistinguibles de la enzima original por varios criterios estructurales y espectroscópicos (Luger K. *et al.*, 1989). También le ha sido agregado un duplicado de los dos últimos módulos $\beta\alpha$ C-terminales en el N-terminal, creando un homólogo de diez módulos $\beta\alpha$ (Luger K. *et al.*, 1990). Las proteínas $(\beta\alpha)_{10}$ purificadas son enzimáticamente casi totalmente activas y los espectros de absorbancia, fluorescencia y dicroísmo circular, así como el comportamiento de desplegamiento reversible, fueron muy parecidos a los de la proteína original $(\beta\alpha)_8$. Por último, se construyó una variante de ePRAI que contiene una duplicación en serie del quinto módulo $\beta\alpha$ (Urfer R. *et al.*, 1992).

Esta variante ($\beta\alpha$), es enzimáticamente activa y por lo tanto debe tener un centro silvestre ($\beta\alpha$)₈.

Tirosafosfato isomerasa de *Trypanosoma brucei* (TbTIM).

La tirosafosfato isomerasa codificada en por gen *tpi* es una proteína antigua, ubicua, y relativamente muy conservada, que se cree que se originó antes de la división de Arquea, Eubacteria y Eucaria. Se conoce mucho de su estructura y su mecanismo catalítico, y se ha propuesto como un buen candidato para estudios de evolución, estabilidad y estructura de proteínas (Bertolaet B. L. *et al.*, 1995). Se trata de una enzima glicolítica esencial que cataliza la interconversión de dos triosas fosfato, la dihidroxiacetona fosfato (DHAP) y el D-gliceraldeído 3-fosfato (GAP), a partir de la formación intermedia de un *cis*-enadiol (esquema 2) (Knowles J. R. *et al.*, 1991). La enzima media la protonación y desprotonación necesaria para las dos enolizaciones, de DHAO a enadiol y de GAP a enadiol. Esto lo realiza de manera muy eficiente, de tal manera que la interconversión *in vivo* es catalizada a una tasa limitada únicamente por la difusión. Las partes importantes para la catálisis son los residuos catalíticos His95 en el asa $\beta4\alpha4$ y Glu167 en el asa $\beta6\alpha6$, además de la Asn11 y la Lys13, ambos del asa $\beta1\alpha1$, y el asa $\beta6\alpha6$ (residuos 166 a 176) que funciona como una especie de tapa que cambia su conformación en la presencia del sustrato y secuestra al intermediario (Knowles J. R. *et al.*, 1991, Wierenga R. K. *et al.*, 1992). Los residuos que forman parte del sitio activo están en el C-terminal de las β plegadas, tal como en un típico barril $\alpha\beta$.



Esquema 2. Reacción catalizada por la TIM.

La TIM se encuentra en forma de homodímero en todas las especies en las que se ha estudiado (Saab-Rincón G. *et al.*, 2001). La estructura cristalográfica de TIM de *Trypanosoma brucei*, refinada a 1.83 Å de resolución, muestra que es un dímero muy denso que consiste de dos subunidades idénticas barriles $\alpha\beta$ de 250 residuos (26,000 Da) sin otro dominio adicional (figura 10) (Borchert T. V. *et al.*, 1993). Se sabe que sólo el dímero intacto es completamente activo (Schliebs W. *et al.*, 1996), a pesar de que cada subunidad tiene un sitio activo independiente y la reacción no es realizada de manera simultánea en ambas. La interfase entre cada monómero es muy extensa, consistiendo más que nada en todos los residuos de las asas $\beta 1\alpha 1$ a $\beta 4\alpha 4$, aunque sólo el asa $\beta 3\alpha 3$ es la que se incrusta en una profunda cavidad entre las asas $\beta 1\alpha 1$ y $\beta 4\alpha 4$, cercana al sitio activo de la otra subunidad, estabilizando la arquitectura del sitio activo por medio de interacciones de van der Waals y puentes de hidrógeno (Borchert T. V. *et al.*, 1994). Con base en esta información, se generó una variante de TIM monomérica eliminando el asa $\beta 3\alpha 3$ y reduciendo la región hidrofóbica. Concretamente en la nueva variante, llamada monoTIM, se reemplazó un fragmento de 15 residuos de la parte interactuante del asa por otro de 8 residuos más hidrofílicos (Borchert T. V. *et al.*, 1993, 1994).

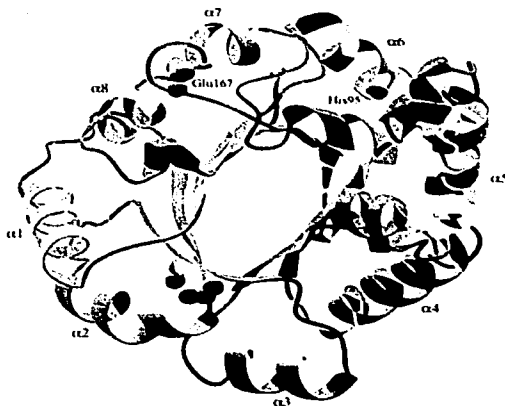


Figura 10. Triosa fosfato isomerasa monomérica (monoTIM) de *T. brucei*.

MonoTIM es una enzima monomérica soluble, muy parecida en estabilidad a la proteína silvestre y con una conformación de barril α/β , pero con sólo un milésimo de la actividad original del dímero. Así mismo, se ha demostrado que monoTIM cataliza la interconversión a través del mismo mecanismo que la enzima silvestre (Schliebs W. *et al.*, 1996). Últimamente, a través de estrategias de evolución dirigida, se han logrado variantes de monoTIM con un incremento de 11 veces en su k_{cat} y reducción de 4 veces en su K_m (Saab-Rincón G. *et al.*, 2001). De esta manera es posible trabajar con una enzima monomérica al igual que en el caso de PRAI.

Objetivo

El objetivo general es utilizar la recombinación *in vivo* para trabajar con fragmentos de dominios barriles α/β con miras a desarrollar una estrategia para la ingeniería de este prometedor grupo de proteínas. En concreto, se pretende implementar un sistema de recombinación *in vivo* que permita combinar bibliotecas preexistentes de proteínas, independientemente de su origen, para generar mayor variabilidad en la búsqueda de nuevas propiedades enzimáticas.

Objetivo particular

El objetivo particular del trabajo es formar genes de las proteínas PRAI y TIM funcionales, a partir de la integración de dos de sus fragmentos, clonados de manera independiente en el vector de recombinación (pDANS), por medio de recombinación *in vivo* con la Cre recombinasa. De esta forma, se pretende demostrar que el sistema *cre/loxP* de recombinación *in vivo* puede ser utilizado para recombinar fragmentos de barriles α/β . Además, se pretende analizar la eficiencia del proceso y probar su utilidad en librerías.

Metodología

La estrategia experimental puede ser dividida en tres partes principales. La primera comprende la construcción de las proteínas completas en el vector de recombinación pDAN5 con el sitio *loxP511* insertado y, a su vez, la construcción de cada uno de los fragmentos de las proteínas por separado en el mismo vector. El primer paso de esta etapa permitió probar el efecto de la inserción del sitio *loxP511* sobre la actividad de las proteínas, mientras que el segundo generó las construcciones necesarias para la recombinación. La segunda parte consiste en la recombinación *in vivo* de los fragmentos clonados en pDAN5 para formar proteínas completas y el análisis de la eficiencia del proceso. Finalmente, la tercera parte esta formada por la utilización del sistema en librerías para validar su utilidad.

Es importante mencionar que el sistema de recombinación a ser implementado es el desarrollado por Sblattero D. *et al.* (2001) debido a las ventajas mencionadas en la introducción.

Construcción de la proteína completa y de las mitades por separado en pDAN5 (figuras 11 y 12).

Las dos mitades de cada una de las proteínas fueron amplificadas por separado por PCR (ver protocolo 1) (figura 11) agregándoseles los sitios de restricción necesarios para la clonación en pDAN5, además de algunas secuencias necesarias para su expresión correcta. Los *primers* (sintetizados en la Unidad de Síntesis del Instituto de Biotecnología, UNAM) utilizados para la amplificación de la primera mitad ($\beta 1-\alpha 4$) (figura 13) se diseñaron para agregar en el inicio de la secuencia un sitio de restricción HindIII, un Shain-Dalgarno (sitio de unión del ribosoma) y un codón de inicio (oligos B1HindPRAI y pDATimHind), y al final de la secuencia una glicina extra y un sitio de restricción Sall (oligos A4SallPRAI y pDATimSal). Los *primers* utilizados para amplificar la segunda mitad ($\beta 5-\alpha 8$) (figura 13) se diseñaron para agregar al inicio de la secuencia un sitio XhoI y una glicina extra (oligos B5XhoIPRAI y pDANmTXho), y al final uno o dos codones de término (taa y tga) ninguno de los cuales es el ambar y un sitio de restricción NheI (oligos A8NheIPRAI y

pDATimNhe). Las glicinas extras se agregaron para proporcionar flexibilidad suficiente y con el fin de que el péptido conector codificado por el sitio *loxP511* (LOXP511) se pueda acomodar en la estructura de la proteína. Así mismo los codones de paro fueron necesarios para asegurar que la proteína g3p, requerida para el despliegue en fago y que se encuentra después del gen del anticuerpo en el vector pDAN5, no se exprese. Es importante aclarar que los sitios de restricción que se añadieron entre la proteína y el péptido LOXP511 para la clonación (*Sall* y *XhoI*) codifican para una serie de aminoácidos que aumentan el tamaño del péptido conector. De este modo los *primers* para la amplificación del final de la primera mitad y el inicio de la segunda mitad (oligos A4*Sall*PRA1, pDATim*Sal*, B5*XhoI*PRA1 y pDANmTXho) no sólo agregan glicinas en los extremos del péptido conector, sino que además agregan en el extremo 5' la secuencia que codifica para otra glicina más, una serina y una treonina (GST), y en el extremo 3' la secuencia que codifica para tres serinas y una glicina más (SSSG) (ver figura 13, 17 y 18).

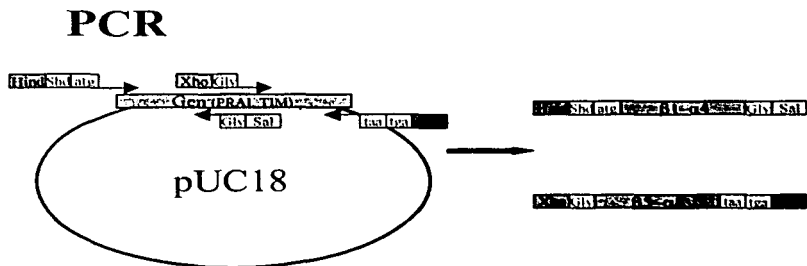


Figura 11. Amplificación de cada una de las mitades de las proteínas por PCR.

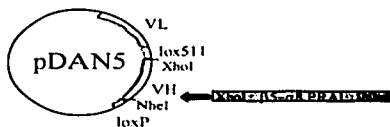
Las mitades amplificadas por PCR se purificaron con el *kit* de Roche (Boehringer-Mannheim) para purificar productos de PCR (“*High Pure PCR Product Purification Kit*”), fueron digeridas, la primera con las enzimas *HindIII* y *Sall* y la segunda con *XhoI* y *NheI* (ver protocolo 2), y fueron purificadas de nuevo por medio de electroforesis en gel de agarosa al 1% utilizando el *kit* de extracción de Qiagen (“*QIAquick Gel Extraction Kit*”),

con lo que quedaron listas para la clonación. En el caso de PRAI, se clonó primero la segunda mitad, mientras que en el caso de monoTIM primero la primera, debido a la presencia de sitios de restricción repetidos (XhoI) en el gen de PRAI que interferían con las clonaciones. Independientemente de lo anterior, el vector (pDAN5) fue digerido con las enzimas correspondientes a la mitad a clonar, purificado por medio de electroforesis en gel de agarosa al 1% con el mismo *kit* de extracción de Qiagen y puesto a ligar con la mitad correspondiente del gen (ver protocolo 3). El producto de la ligación se precipitó con butanol (ver protocolo 4) y se utilizó para transformar células JM101 electrocompetentes (ver protocolo 5), las cuales después de una hora de recuperación se platearon e incubaron 12 horas a 37° C en cajas de LB con ampicilina (Amp) (ver medios), ya que pDAN5 porta el gen de resistencia a ampicilina. Algunas de las colonias obtenidas se crecieron en LB Amp líquido, se purificó el plásmido utilizando el *kit* de Roche (Boehringer-Mannheim) para aislamiento de plásmidos ("*High Pure Plasmid Isolation Kit*") y se probó la presencia de la mitad clonada por medio de PCR con los *primers* correspondientes (oligos B5XhoIPRAI y A8NheIPRAI para la segunda mitad de PRAI y oligos pDATimHind y pDATimSal para la primera mitad de monoTim) y electroforesis en gel de agarosa al 1%. Posteriormente se eligió uno de los vectores que sí tenía la mitad clonada, se digirió con las enzimas para la clonación de la otra mitad, se purificó por medio de electroforesis en gel de agarosa al 1% de la misma manera antes descrita y se puso a ligar con la mitad del gen a clonar (las clonaciones se esquematizan en la figura 12). La ligación se precipitó con butanol y se utilizó para transformar células, pero ahora fueron células de la cepa de selección, es decir JMB9 $\Delta TrpF$ en el caso de PRAI y VEJM101 $\Delta TrpI$ en el caso de TIM. Las células se lavaron con el medio de selección líquido (ver medios) (ver protocolo 6) y se platearon, tanto en cajas de medio de mínimo Amp como en cajas de LB Amp (control positivo), para determinar si las enzimas con el inserto complementan las funciones y por lo tanto siguen siendo activas. En el caso de PRAI el medio de selección es el Vogel Bonner (VB) y en el caso de TIM es el medio mínimo M9 (ver medios). Algunas de las colonias obtenidas se crecieron en LB Amp, se aisló el plásmido con el *kit* ya antes mencionado y se verificó que la construcción estuviera completa y no contuviera mutaciones por secuenciación (en la Unidad de Secuencia del Instituto de Biotecnología, UNAM) con otros *primers* que hibridizan en pDAN5 justo por fuera de las secuencias clonadas (oligos

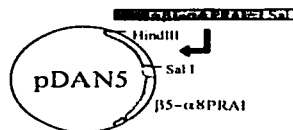
pDAN5Hind y pDAN5NheI) (figura 14). Además, se realizó una curva de crecimiento comparando el crecimiento de la cepa de selección transformada con el plásmido portador de la proteína sin la inserción del péptido conector y con la inserción.

Construcción de PRAI

1a. clonación

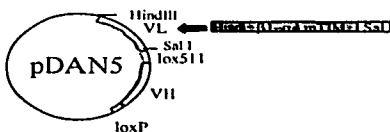


2a. clonación



Construcción de mTIM

1a. clonación



2a. clonación

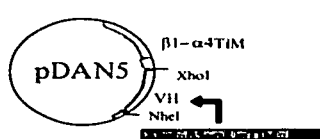
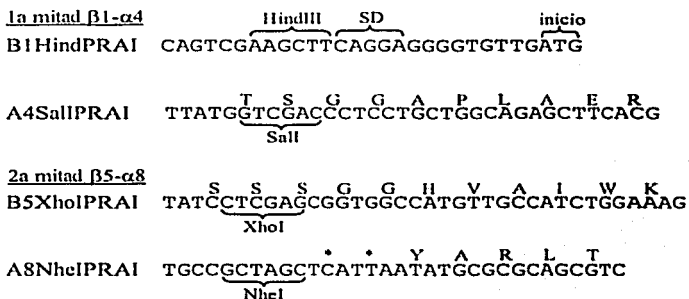


Figura 12. Clonaciones de los fragmentos de las proteínas en el fagémido pDAN5.

Al final de los pasos anteriores se obtuvieron, por lo tanto, las proteínas PRAI y monoTIM con el sitio de recombinación loxP511 insertado entre la α hélice 4 y la β plegada 5, clonadas en el vector de recombinación pDAN5. Además de pDAN5 con sólo una de las mitades, 1ª o 2ª dependiendo de la proteína, por lo que fue necesario clonar la mitad faltante en pDAN5 con el fin de tener a cada una de las mitades clonadas por separado para la recombinación. Esto último se realizó de la misma forma en que se clonaron las primeras mitades, pero en los sitios de restricción correspondientes (figura. 12). También se realizaron curvas de crecimiento comparando el crecimiento de la cepa de selección transformada con las tres diferentes construcciones: pDAN5+ β 1- α 4+ β 5- α 8, pDAN5+ β 1- α 4+VH y pDAN5+VL+ β 5- α 8.

Oligos para PRAI



Oligos para mTIM

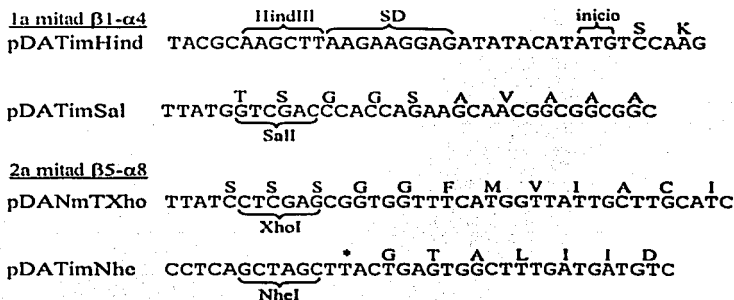


Figura 13. Oligos utilizados para la amplificación de las mitades de las proteínas por PCR para las construcciones en el vector pDAN5.

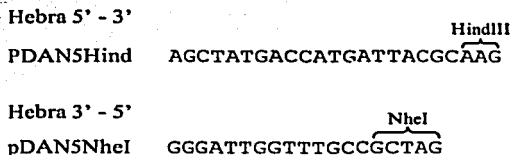


Figura 14. Oligos utilizados para secuenciar las diferentes construcciones y fagémidos.

Recombinación *in vivo*

Recombinación in vivo (figura 15). El sistema de recombinación *in vivo* de Sblattero D. *et al.* (2001) tiene la cualidad de ser un sistema desarrollado en un fagémido. Lo anterior tiene importantes ventajas en cuanto a la eficiencia de transformación celular, ya que la cantidad de células transformadas depende, casi en su totalidad, de la cantidad de células en etapa infectiva que puedan ser cultivadas. Es importante tener en cuenta este punto para las siguientes etapas experimentales.

Dadas las características del sistema, primero fue necesario transformar células F⁺ (en este caso DH5 α F⁺) con cada una de las mitades por separado, para poder obtener pDAN5 infeccioso con cada una de las mitades (figura 15). La transformación se llevó a cabo con células competentes por CaCl₂ (ver protocolo 7) debido a que solo se necesitan una o dos colonias. Las células transformadas se platearon en cajas de LB Amp y se incubaron por una noche a 37 °C. Se tomó una colonia transformada por cada una de las mitades, ambas se crecieron a 37 °C en medio rico 2xYT (ver medios) con ampicilina hasta alcanzar una densidad óptica a 600 nm (DO₆₀₀) igual a 0.3. Los fagémidos fueron rescatados con fago *helper* R408 adicionado en una proporción de 20:1 (fago *helper* :células), incubando por 4.5 horas a 30 °C (ver protocolo 8). Los fagos se obtuvieron del sobrenadante después de centrifugar a 3000 g por 15 minutos y se precipitaron con una solución 2.5M de NaCl y 20% PEG 8000, siendo resuspendidos en amortiguador PBS 1x (ver protocolo 9). Los diferentes fagos obtenidos fueron titulados, ya sea por la técnica de formación de colonias o formación de placas (ver protocolos 10 y 11). Una vez con los fagos infectivos y titulados de cada una de las mitades se procedió a realizar la recombinación (ver protocolo 12). Se creció un cultivo de la cepa que expresa la Cre recombinasa constitutivamente (Cre⁺JL03) en medio 2xYT 1% de glucosa hasta una DO₆₀₀ = 0.5, se infectó en una proporción de 100:1 (fago:células) con cada uno de los fagos portadores de cada mitad, dejando reposar por 30 minutos a 37 °C a manera de que ocurra la infección, para luego agregar ampicilina e incubar el cultivo por un periodo de entre 6 y 15 horas a 30 °C. Este tiempo de incubación permite que se lleve a cabo la recombinación. Posteriormente, el cultivo se diluyó 20:1 en el mismo medio (2xYT 1% glucosa Amp), se dejó crecer hasta una DO₆₀₀ = 0.5 y se rescató con fago *helper* R408 adicionándolo en una proporción 20:1, e incubando en agitación por 18 horas a 30 °C. Los fagos resultantes se

obtuvieron del sobrenadante después de centrifugar a 3000 g por 15 minutos y se precipitaron de la misma forma antes mencionada. A la par de estos experimentos de recombinación se realizaron, como controles negativos, dos recombinaciones en las que se infectó la cepa Cre+ sólo con los fagos de alguna de las mitades en una proporción de 200:1. Dichas recombinaciones se realizaron de la misma manera que en las que se infectó con ambas mitades.

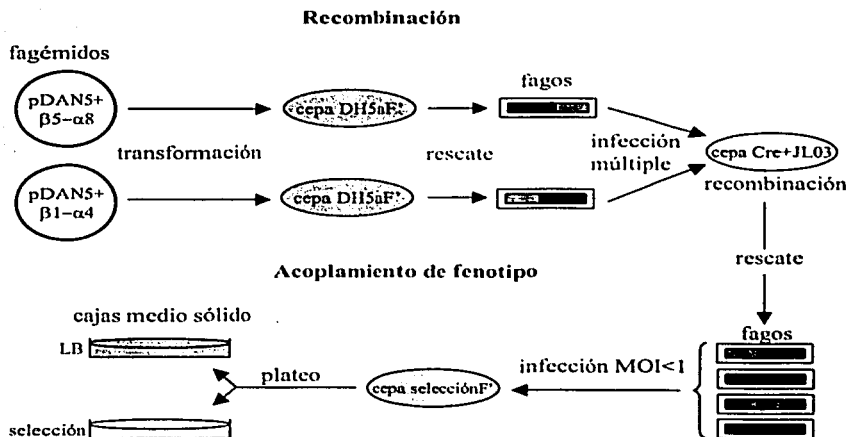


Figura 15. Esquema del procedimiento seguido para la recombinación *in vivo*.

Posteriormente se realizaron procedimientos de recombinación con diferentes tiempos de recombinación debido a que este factor puede influir en la eficiencia del proceso (Bradbury A., comunicación personal). Se probaron tiempos de 6, 9, 12 y 15 horas. Por último, cabe mencionar que se realizaron experimentos de recombinación infectando cada mitad en una proporción menor (20:1 de cada fago portador de cada mitad por célula, en lugar de 100:1) para observar el efecto de este factor.

Análisis de la recombinación (figura 15). Los fagos resultado de la recombinación se titularon con las técnicas antes mencionadas y con ellos se infectaron células de la cepa de selección, para probar si la recombinación generó genes completos de la proteína y, por lo tanto, estos complementan la función. Se creció un cultivo de células de selección F' (JMB9F') (no se cuenta con una cepa de selección de TIM que sea F') en 2xYT con tetraciclina (Tc) (ya que la cepa JMB9F' es resistente a tetraciclina) hasta una $DO_{600} = 0.5$, se infectó con los fagos en una proporción de 1:1 o menor, se dejó 30 minutos sin agitación a 37 °C, se lavaron las células con medio de selección líquido como en etapas anteriores y se platearon tanto en medio de selección suplementado con Amp como en LB suplementado con Amp (ver protocolo 13). Algunas de las colonias crecidas en el medio de selección se crecieron en medio VB Amp Tc líquido, se purificó el fagémido con el mismo *kit* de aislamiento ya usado y se probó que éste portara ambas mitades de la proteína por medio de PCR con los *primers* correspondientes (oligos B1HindPRA1 y Δ8NheIPRA1) y electroforesis en gel de agarosa al 1%. Además se secuenciaron tres representantes con los *primers* que hibridizan por fuera del gen de la proteína (oligos pDAN5Hind y pDAN5NheI). En este paso, al infectar el cultivo en una proporción menor que 1:1, se acopló el fenotipo de cada célula al genotipo de un sólo fagémido, de tal forma que se pudieran analizar los diferentes genotipos creados por la recombinación.

Además de la ocurrencia de la recombinación, se midió la eficiencia del proceso. Lo anterior se hizo por medio de dos procedimientos. Por una parte se comparó el número de colonias crecidas en medio de selección con el número de colonias crecidas en medio rico, para ver qué porcentaje de las segundas representan las primeras. Para esto fue necesario diluir entre cuatro y cinco factores las células lavadas, infectadas en una proporción menor de 1:1 del experimento anterior, para platearlas de nuevo en cajas de LB Amp Tc y medio de selección Amp Tc a manera de réplicas. Las cajas de medio rico se incubaron a 37 °C por una noche, mientras que las cajas de medio de selección se incubaron hasta por 50 horas. Después se contaron las colonias en ambas cajas y se compararon. Por el otro lado, se tomaron 100 colonias al azar de las cajas de LB Amp Tc del experimento anterior, se crecieron separadas en LB Amp Tc y se purificaron los plásmidos con el *kit* de aislamiento. Los fagémidos obtenidos se analizaron por patrón de digestión con las enzimas AlwNI y

MscI. Para lo anterior cada plásmido obtenido se digirió por separado con cada una de las enzimas y se hizo electroforesis en gel de agarosa al 1% de los fragmentos resultantes.

Por último se realizaron curvas de crecimiento de los cuatro posibles genotipos resultado de la recombinación en medio rico para ver si su producto no es tóxico para la célula, lo cual alteraría las proporciones esperadas.

Utilización del sistema en librerías (figura 16)

Se utilizaron tres diferentes mutantes deletéreas de PRAI previamente reportadas (Wilmanns M. *et al.*, 1992 y Henn-Sax M. *et al.*, 2002). La primera es la sustitución de la cisteína catalítica 7 del asa uno ($\beta 1\alpha 1$), situada en la primera mitad, por una alanina (Cys7Ala). La segunda es la sustitución del aspártico catalítico 126 del asa 6 ($\beta 6\alpha 6$), situado en la segunda mitad, por una alanina (Asp126Ala). La última, es la conjugación de estas mutantes en una doble mutante que tiene ambos residuos catalíticos sustituidos por una alanina (Cys7Ala Asp126Ala). Con las mutantes se crearon dos bibliotecas de la doble mutante contaminadas con diferentes proporciones de las mutantes sencillas, una con un milésimo de cada mutante y la otra con un cienmilésimo de cada mutante, con el fin de demostrar que, por medio de la recombinación, se podía regenerar la proteína activa sin las dos mutaciones deletéreas. Para ello se determinó la cantidad de DNA en las alícuotas de plásmidos superenrollados de cada una de las mutantes a través de absorbancia de luz UV₂₆₀ y, con base en esto, se hicieron las diluciones requeridas. Ambas bibliotecas se transformaron por electroporación en células XL1blue de la misma manera que la transformación anterior de células JM101 y JMB9 en la primera sección de la metodología, aunque en lugar de platear las células éstas fueron utilizadas para inocular medio líquido 2xYT Amp (ver protocolo 14). Se dejó crecer el cultivo hasta una $DO_{600} = 0.5$ y se rescataron los fagémidos de las bibliotecas, infectando en una proporción 20:1 con fago *helper* R408. Los fagos resultantes se obtuvieron centrifugando y precipitando con PEG de la misma forma que en pasos anteriores. Cada una de las bibliotecas se tituló con los mismos métodos y fue utilizada para infectar un cultivo de células Cre+, donde se lleva a cabo la recombinación, en una proporción de 200:1 (fagos bibliotecas:células). Después de las 12 horas de recombinación se siguió el protocolo descrito anteriormente para rescatar y

obtener los fagos (ver protocolo 14). Las bibliotecas ya recombinadas se titularon y fueron sometidas a selección en busca del fagémido que portara PRAI activa, es decir sin ninguna de las mutaciones deletéreas. Esto se llevó a cabo de la misma manera que al probar la ocurrencia de la recombinación en la etapa anterior, pero a mayor escala (ver protocolo 14). Se creció un gran cultivo de células de selección F⁺ (JMB9F⁺) (5×10^{11} células) por cada biblioteca y se infectaron por separado en una proporción <1:1. Se dejaron reposar a 37 °C, se lavaron y se inocularon en medio VB Amp Tc líquido. Se utilizó medio líquido de selección debido a la enorme cantidad de células sometidas a selección. El cultivo se dejó creciendo en agitación a 37 °C hasta que se observó crecimiento y entonces se tomó una alícuota para inocular más medio líquido VB Amp Tc. Este último paso de enriquecimiento se repitió tres veces. Al final se tomó una alícuota del último cultivo, se diluyó 10^{-4} y 10^{-5} y se plateó en cajas del mismo medio (VB Amp Tc) que se incubaron a 37 °C para obtener colonias aisladas. De las colonias obtenidas, algunas se crecieron en medio líquido VB Amp Tc, se aisló el plásmido con el *kit* y estos se mandaron a secuenciar con los *primers* que hibridizan por fuera del gen (oligos pDAN5Hind y pDAN5NheI). Así mismo se tomó otra alícuota del último cultivo de selección, se aislaron los fagémidos con el mismo *kit*, se electroporaron en la cepa de selección (JMB9) y se platearon en cajas de VB Amp para dejar crecer a 37 °C.

Barajeo de librerías. Con el fin de comparar la capacidad del método de recombinación *in vivo* de combinar mutantes raras en una librería contra la misma capacidad del método de barajeo de genes (*gene shuffling*), que es el comúnmente usado, se realizó el barajeo de genes de la misma biblioteca con las mutantes sencillas diluidas 10^{-5} . Primero, como control positivo, se realizó el barajeo de una biblioteca con las mutantes sencillas diluidas 10^{-2} y después se continuó con el barajeo de la biblioteca con las mutantes sencillas diluidas 10^{-5} . El barajeo de genes se llevó a cabo como fue descrito primeramente por Stemmer W. P. C. (1994). La librería ya barajada se clonó en el vector pDAN5 de la misma forma en que se clonaron las mitades anteriormente, aunque en los sitios HindIII y NheI directamente, y se transformó por electroporación en la cepa JMB9 plateando las células, después de lavarlas con medio mínimo para evitar contaminación con triptofano, en cajas de VB Amp. Después de incubarlas a 37 °C durante más de 50 horas se tomaron

algunas colonias, se crecieron, se purificaron los plásmidos y se secuenciaron con los *primers* que hibridizan en pDAN5 por fuera del gen (oligos pDAN5Hind y pDAN5NheI).

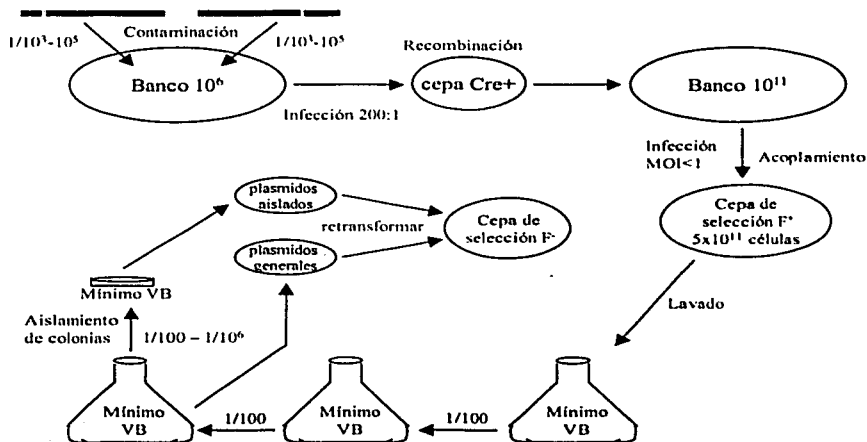


Figura 16. Esquema del procedimiento seguido en la utilización del sistema de recombinación *in vivo* en librerías.

Protocolos

1. Reacción en cadena de la polimerasa (PCR)

Para una reacción de 100 μL

- Amortiguador (10x): 10 μL
- dNTPs (2.5 mM): 8 μL
- Oligo 1 (30 – 40 pM): 1 μL
- Oligo 2: 1 μL
- Templado: 1 μL
- Polimerasa: Tac Polimerasa purificada en el laboratorio (2.5 U/ μL) 4 μL , Expand High Fidelity (Roche) 0.7 – 1 μL (se puede lograr mejor rendimiento si la enzima es agregada después del primer choque de calor)
- MgCl (50 mM): 4 μL
- Completar con H_2O a 100 μL

Condiciones del PCR (estándar utilizado en las construcciones para pDAN5)

- 1- Desnaturalización 95 °C 4 min (sin polimerasa), 1 ciclo.
- 2- Desnaturalización 95 °C 1 min; hibridación 50 – 55 °C 1 min; extensión 72 °C 2 min, 25 ciclos.
- 3- Extensión 72 °C 10 min, 1 ciclo.

2. Digestión

Para una reacción de 50 μL

- Amortiguador específico para la enzima (10x generalmente): 5 μL
- BSA (sólo si es necesario 10 o 100x generalmente): 0.5 o 5 μL
- Enzima 1: 1.5 μl (las enzimas HindIII, Sall y XhoI fueron usadas a 20 U/ μL , mientras que HincII se usó a 8 U/ μL , NheI y AlwNI* a 10 U/ μL y MscI* a 3 U/ μL)
- Enzima 2: 1.5 μL
- DNA: ~40 μL (la mayor cantidad en el volumen final)
- Si es necesario, completar con H_2O al volumen final

*al hacer las digestiones con las enzimas AlwNI y MscI se varió la concentración de enzima (desde 0.02 hasta 0.05 $\mu\text{L}/\mu\text{L}$) buscando la mejor resolución.

Condiciones de la digestión

12 horas** a 37 °C***

*para una digestión total es recomendable dejarla toda la noche, aunque en ocasiones incluso 6 hrs. son suficientes. Para una digestión parcial o cualitativa 2 hrs. son suficientes.

**es necesario ajustarse a la temperatura óptima a la que trabaja la enzima. Con las enzimas utilizadas es de 37° C.

3. Ligación

Para una ligación pequeña* de entre 100 y 400 ng de DNA final en 20 μ L

- Amortiguador (5x): 4 μ L
- Ligasa T4 (400 U/ μ l): 1 μ L
- Vector: Vng**
- Inserto: Ing = (Vng/kbV)(kbl[8])**
- Completar con H₂O a 20 μ L

*para ligaciones de bancos grandes es recomendable aumentar el tamaño de la ligación, de tal forma que se tengan entre 1 y 2 μ g de DNA final.

**la relación entre cantidad de vector en nanogramos (Vng) e inserto en nanogramos (Ing) está dada por esta ecuación, donde kbV es el tamaño del vector en kilo bases (kb) y kbl el tamaño del inserto en kb.

Condiciones de la ligación

Una noche a 16 °C***

***cuando no se requiere gran eficiencia de ligación es posible hacerla durante 2 hrs. a temperatura ambiente.

4. Precipitación con butanol

- Agregar butanol (entre 500 y 1400 μ l, dependiendo de la cantidad de DNA a precipitar de 20 hasta 100 μ l)
- Agitar fuertemente hasta que no se note la separación entre el medio acuoso en el que está suspendido el DNA y el butanol.
- Centrifugar en una centrifuga de mesa por 10 min a velocidad máxima
- Extraer cuidadosamente el butanol con pipeta
- Escurrir
- Secar en el Savant algunos minutos
- Resuspender con 5-20 μ L de H₂O

5. Transformación de células por electroporación

- Preparar un tubo pequeño de ensayo con 1 mL. de medio SOC, 10 μ l glucosa 2M y 10 μ L de una mezcla de MgCl₂ MgSO₄ 2M por cada electroporación a realizar y poner a enfriar en hielo una celda por cada electroporación.
- Voltaje de electroporación:
 - 1.8 V para celdas chicas
 - 2.5 V para celdas grandes

- Vertir el DNA sobre las células (50 μ L de células electrocompetentes cuando es una electroporación chica y hasta 200 μ L cuando es grande), mezclar con mucho cuidado tratando de no hacer burbujas y se toma sin las burbujas creadas. Se vierte \sim 1 μ L de DNA cuando es DNA superenrollado (la eficiencia es mucho mayor) y \sim 2 μ L cuando es producto de una ligación.
- Las células se vierten en la celda lentamente y sobre una de las paredes, de tal manera que la células resbalen y no formen burbujas.
- Dar unos golpecitos para que las células se distribuyan a lo largo de la celda y se pueda observar que las células hacen contacto con ambas paredes.
- Las células siempre se mantienen en hielo y se tratan con cuidado, ya que son muy delicadas.
- Antes de electroporar, se succiona el medio SOC con una pipeta pasteur estéril dejándolo listo para verterlo sobre las células después de la electroporación.
- La celda se seca bien, se pone en el electroporador y se da el pulso.
- Rápidamente se saca del electroporador, se vierte el medio SOC en la celda, se mezcla, y por último se succiona y se pone en el tubo de ensaye donde se incubarán las células a 37° C en agitación por una hora.

6. Lavado de células

- Centrifugar el cultivo de células a una velocidad de entre 4000 y 4500 rpm (alrededor de 3000 g) por entre 4 y 5 min.
- Desechar el sobrenadante bajo condiciones estériles con cuidado de no tirar células.
- Remover cuidadosamente con una pipeta el resto del medio rico.
- Resuspender las células cuidadosamente ya sea con una pipeta o movimientos circulares suaves con el mismo volumen de medio de selección o PBS que de medio rico desechado.
- Repetir la centrifugación y resuspensión dos veces más. Para el 2° y el 3er ciclo ya no es necesario remover con una pipeta el medio restante.

7. Transformación de células competentes por Ca₂Cl

Células competentes

- Crecer en 30 mL de medio LB un cultivo de las células deseadas hasta una DO₆₀₀= 0.6
- Centrifugar a 2500 rpm en frío (4°C) por 10 min.
- Desechar el sobrenadante y resuspender las células con 2 o 3 ml de Ca₂Cl 0.1M previamente enfriado (4°C)
- Añadir del mismo Ca₂Cl hasta completar 15 ml.
- Incubar en hielo por 30 min.
- Centrifugar de la misma manera que en el primer paso.
- Desechar el sobrenadante y resuspender en 1.5 mL de Ca₂Cl 0.1M frío para después añadir 123 μ l de DMSO.
- Hacer alícuotas de 200 μ L en microtubos de 1.5 mL.
- Almacenar a -80° C.

Transformación

- Agregar 5 μ L del DNA a transformar a una de las alícuotas de células competentes y agitar levemente.
- Incubar en hielo por 45 min agitando levemente cada 15 min.
- Dar un choque de calor de 90 seg a 42 °C.
- Añadir 500 μ L de LB e incubar por una hora a 37 °C.

8. Rescate de fagémidos

- Crecer un cultivo a partir de una sola colonia de la cepa con el fagémido a rescatar en 40 ml de medio 2xYT suplementado con los antibióticos necesarios (Amp en el caso de pDAN5) hasta una $DO_{600} = 0.5$.
- Añadir el fago "helper" (en este caso el R408) en una proporción de 20:1 (fago:células), considerando que en 1ml de cultivo a $DO_{600} = 0.5$ hay 5×10^8 células, y dejar reposando 30 min a 37 °C sin agitación.
- Incubar con agitación a 30 °C durante 6 a 18 horas y, por último, obtener los fagos del sobrenadante de centrifugar el cultivo a 3000 g por 15 minutos.

9. Precipitación de fagos

- Centrifugar el cultivo donde se encuentran los fagos a 3000 g por 15 min
- Tomar el sobrenadante y agregar 1/5 del volumen de una solución de PEG 8000 20% y NaCl 2M previamente enfriada (4 °C).
- Incubar a 4 °C por media hora
- Centrifugar a 4000 g por 10 min a 4 °C
- Desechar el sobrenadante y resuspender los fagos cuidadosamente con 1/10 del volumen inicial del cultivo de PBS
- Los fagos pueden guardarse a 4 °C ó para un almacenamiento más largo, se agrega a – 20% de glicerol y se mantiene a –70 °C.

10. Titulación por número de placas

- Crecer un cultivo de células F' hasta una $DO_{600} = 0.6 - 0.8$. 200 μ L por cada dilución de cada fago a titular.
- Preparar un tubo de ensaye pequeño con 3ml de agar suave (Top agar) 2xYT manteniéndolo fundido a 50 °C y una caja de 2xYT seca y entibiada por cada dilución a probar.
- Hacer una serie de diluciones del fago a titular en PBS. Es recomendable probar las diluciones de entre 10^{-7} o 10^{-8} hasta 10^{-10} o 10^{-11} , ya que se esperan títulos de entre 10^{10} y 10^{13} fagos por ml.
- Colocar en un microtubo de 1.5 ml 200 μ L de las células, agregar 100 μ L de la dilución de fago a titular y mezclar.

- Agregar la mezcla anterior a un tubo de agar suave caliente, mezclar y verter en una caja de 2xYT asegurándose de que el agar suave quede bien esparcido.
- Esperar a que el agar suave se solidifique e incubar a 37 °C por una noche.
- El título del fago se obtiene multiplicando el número de placas por el inverso de la dilución usada por 10 (#placas x 1/factor de dilución x 10).

11. Titulación por número de colonias (sólo es posible cuando el fagémido da resistencia a algún antibiótico)

- Crecer un cultivo de células F' hasta una $DO_{600} = 0.7 - 1.0$. 200 μ L por cada dilución de cada fago a titular.
- Diluir el fago de la misma manera que en el protocolo anterior, pero empezando las diluciones desde un orden de magnitud mayor (10^{-6} o 10^{-7}).
- Colocar en un microtubo de 1.5 ml 200 μ l de las células, agregar 100 μ L de la dilución de fago a titular y mezclar.
- Dejar las mezclas reposando a 37 °C sin agitación por 30 min.
- platear 100 μ L de cada dilución en cajas de LB más el antibiótico al que da resistencia el fagémido.
- Incubar las cajas por una noche a 37 °C.
- El título del fago se obtiene multiplicando el número de placas por el inverso de la dilución usada por 100 (#colonias x 1/factor de dilución x 100).

12. Recombinación *in vivo*

- Crecer la cepa Cre+ en medio 2xYT 1% de glucosa en agitación a 37 °C hasta una $DO_{600} = 0.5$. 2 mL por cada recombinación .
- Infectar los 2 mL de cultivo con el fago a recombinar en una proporción de 200:1 (hasta 20:1) (en el caso de que sean dos fagos diferentes es posible infectar cada uno en una proporción de 100:1) dejando reposar por 30 min. a 37 °C para después agregar ampicilina (o el antibiótico al que da resistencia el fago).
- Incubar el cultivo con agitación a 30 °C por el tiempo de recombinación deseado (6 – 15 hrs.).
- Después del tiempo de recombinación se diluye el cultivo en 38 ml (20:1) del mismo medio (2xYT 1% glucosa, Amp) y se crece a 37 °C con agitación hasta una $DO_{600} = 0.5$
- Se rescatan los fagémidos agregando "helper" en una proporción 20:1. Se deja reposar por 30 min. a 37 °C y se incuba por entre 6 y 18 hrs. a 30 °C.
- Obtener los fagos del sobrenadante después de centrifugar a 3000 g por 15 min. Precipitar con 2.5 M NaCl PEG 8000 centrifugando a 4000 g por 10 min.

13. Acoplamiento del genotipo del fagémido al fenotipo de la célula (prueba de la recombinación)

- Crecer la cepa de selección F' (JMB9F' Δ TrpF) en 2xYT a 37° C en agitación hasta $DO_{600} = 0.5$. 1 mL por cada recombinación a probar.
- Infectar con los fagos producto de la recombinación en una proporción menor de 1:1.
- Dejar reposar a 37 °C por 30 min sin agitación.
- Lavar las células con medio de selección líquido (VB).
- Platear, tanto en LB como en medio de selección (VB), con el antibiótico al que da resistencia el fagémido (Amp en el caso de pDAN5). Es recomendable platear diluciones del cultivo lavado debido al gran número de células. Incluso es posible diluir hasta 10^{-6} .
- Incubar a 37 °C las cajas hasta la aparición de colonias (una noche en LB y hasta 4 días en VB)

14. Recombinación de librerías

Creación de la librería primaria de fagos

- Con las células resultado de la transformación con la librería original, inocular 40 mL de medio 2xYT Amp (o el antibiótico al que el fagémido de resistencia).
- Crecer en agitación a 37 °C hasta una $DO_{600} = 0.5$
- Rescatar infectando con "helper" (R408) en una proporción 20:1, considerando que aproximadamente en 1 mL de cultivo a una DO_{600} de 1.0 hay 10^9 células.
- Dejar 30 min. a 37 °C sin agitación.
- Incubarlas entre 6 y 18 horas en agitación a 30 °C.
- Obtener los fagos del sobrenadante después de centrifugar a 3000 g por 15 min y precipitar con PEG 8000 20% y 2.5M NaCl.
- Titular.

Recombinación de la librería primaria antes creada para crear la librería secundaria

- Crecer 2 ml de la cepa Cre+ hasta una $DO_{600} = 0.5$ a 37 °C en agitación, en medio 2xYT 1% glucosa.
- Infectar el cultivo con el fago antes preparado en una proporción de 200:1.
- Dejar reposar 30 min. a 37 °C sin agitación, para después agregar ampicilina (o el antibiótico específico).
- Incubar el cultivo 12 hrs. a 30 °C en agitación.
- Agregar el cultivo a 38 mL del mismo medio (dilución 1/20 en medio 2xYT 1% glucosa Amp) y dejar crecer en agitación a 37 °C hasta una $DO_{600} = 0.5$.
- Rescatar agregando "helper" en una proporción 20:1, dejar reposar a 37 °C por 30 min.
- Incubar por 18 hrs. a 30 °C.
- Obtener los fagos del sobrenadante después de centrifugar y precipitación con PEG 8000 y 2.5M NaCl.
- Titular.

Análisis de la recombinación (los volúmenes dependerán del tamaño de la librería secundaria)

- Crecer un cultivo de células de selección en 1 L 2xYT hasta una $DO_{600} = 0.5$
- Infectar con los fagos resultado de la recombinación en una proporción de 1:1 o menor.
- Dejarlos 30 min sin agitación a 37 °C.

- Lavar las células con medio de selección líquido o PBS, pero en el último lavado en lugar de resuspender con 1 L se resuspende en 1/10 del volumen inicial de medio rico (es decir 100 ml).
- Con estas células se inocula 1 l de medio de selección con el antibiótico específico (VB Amp) y se incuba en agitación a 37 °C.

Medios de cultivo

1. LB (1 litro)

- 10 g triptona
- 10 g NaCl
- 5 g extracto de levadura
- Completar con agua miliQu a 1 L

Para medio sólido:

- 15 g bacto agar

2. 2xYT (1 litro)

- 16 g triptona
- 10 g NaCl
- 10 g extracto de levadura
- Completar con agua miliQu a 1 L

Para medio sólido:

- 15 g bacto agar

Para medio suave (Top agar):

- 6 g bacto agar

3. VB (1 litro)

- 20 ml glucosa 20%
- 20 ml casaminoácidos 20%
(esterilizados y no filtrados)
- 50 ml sales VB 20x
- 1 ml FeCl₃ 5mg/ml
- 250 µl vitamina B1 (tiamina) 10 mg/ml
- completar con agua miliQu a 1 L

Para medio sólido:

- 24g agar noble

4. M9 (1 litro)

- 10 mL glucosa 20%
- 312 µL casaminoácidos 20%
- 100 mL sales M9 10x
- 1 mL vitamina B1 1%
- 1 mL MgSO₄·7H₂O 1M
- 2 mL CaCl₂ 50 mM

Para medio sólido:

- 15 g bacto agar

5. SOC (1 litro)

- 20 g triptona
- 5 g extracto de levadura
- 10 mL NaCl 1M
- 2.5 mL KCl 1M
- Completar con agua a 1 L

Soluciones

1. Ampicilina (1000x)

- 200 µg ampicilina/mL
- Agua como solvente
- Se filtra para esterilizar
- Guardar a -20 °C

2. Tetraciclina (1000x)

- 10 µg tetraciclina/mL
- Etanol como solvente
- No es necesario esterilizar
- Guardar a -20 °C

3. Salas VB 20x (1 litro)

- $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 16.2 mM
- ácido cítrico 205.5 mM
- K_2HPO_4 1.148 M
- $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ 168.5 mM
- Ajustar a pH 7

4. Salas M9 10x

- K_2HPO_4 172.2 mM
- Na_2HPO_4 422.6 mM
- NaCl 85.5 mM
- NH_4Cl 186.9 mM

5. PBS 1x

- KH_2PO_4 1.5 mM
- Na_2HPO_4 8 mM
- NaCl 137 mM
- KCl 3 mM
- Ajustar a pH 7.4

Resultados

Construcciones y complementación de las funciones de las proteínas

Para implementar un sistema de recombinación *in vivo* con enzimas es necesario que el sitio usado para la recombinación, que se encuentra conectando los fragmentos que serán recombinados, no afecte de manera importante la conformación nativa de la enzima. Lo anterior no sólo con el fin de poder analizar la recombinación, sino más que nada con la idea de que la recombinación resulte en mayores ventajas que desventajas. En este trabajo, la forma de medir el cambio en la conformación nativa es por medio de la actividad de las enzimas, por eso es que primero se analiza el impacto de la inserción del sitio *loxP511* sobre la actividad enzimática, para después continuar con la recombinación como tal. Con el afán de causar el menor daño a la conformación de las enzimas, basándose en los antecedentes y específicamente en los análisis de corte realizados por Fuentes P. (2001), se decidió insertar el sitio *loxP511* entre la α hélice 4 y la β plegada 5 (figuras 17 y 18). El sitio de inserción es una de las asas inferiores, las cuales, como ya se mencionó, parecen contribuir más que nada a la estabilidad de la proteína y no a su actividad, ya que el sitio activo en los barriles α/β se encuentra en las asas superiores.

Construcción y complementación de PRAI. La amplificación y la clonación de ambas mitades se realizó de tal manera que el péptido conector quedara exactamente entre la prolina 97 y la alanina 98 del asa $\alpha 4$ - $\beta 5$ (figura 17). En el caso de la PRAI, primero se realizó la clonación de la segunda mitad de la proteína ($\beta 5$ - $\alpha 8$), ya que la primera mitad contiene un sitio de restricción XhoI. Lo anterior evitó que al digerir con XhoI, para clonar la segunda mitad, se eliminara una porción de la primera mitad. A pesar de ello la segunda clonación, es decir la de la primera mitad, dio algunos problemas debido a otro sitio de restricción XhoI no reportado en el vector del que se partió y que se encuentra presente en la cadena ligera del anticuerpo (VL). El sitio provoca que al utilizar XhoI para la primera clonación se cortara, además de la cadena pesada del anticuerpo, el sitio Sall junto con todo el sitio *loxP511*. La presencia del sitio XhoI adicional sólo fue evidente hasta que se secuenció una de las construcciones finales. Para eliminar el sitio fue necesario digerir previamente el vector con la enzima de restricción HincII, que corta en dos sitios que

rodean al sitio XhoI extra y no eliminan ninguna región importante del vector, y religar el vector resultante.



Figura 17. Estructura secundaria de la PRAI con la inserción del péptido conector LOXp511. ▽ indican los residuos catalíticos.

Una vez eliminados los problemas anteriores, se pudo clonar primero la segunda mitad ($\beta 5 - \alpha 8$) en la parte que se recombina del plásmido y posteriormente la primera mitad ($\beta 1 - \alpha 4$) en la parte fija del vector. Después de la segunda clonación, la de la primera mitad, se transformaron células de la cepa de selección $\Delta TrpF$ y, una vez lavadas, se platearon en medio de selección con su réplica en medio rico. Después de una noche de incubación, se observó que la función se complementa. Como se puede ver en la figura 19, el número de colonias crecidas en medio rico fue similar al de las crecidas en el medio de selección, aunque resultó difícil la estimación por la gran cantidad de colonias crecidas. Lo anterior nos indica que el péptido conector codificado en el sitio *loxP511* no afecta considerablemente la actividad de PRAI. Al alinear las secuencias obtenidas de las colonias crecidas en medio de selección con una secuencia teórica de pDANS + $\beta 1 - \alpha 4$ PRAI + $\beta 5 - \alpha 8$ PRAI se confirmó que la construcción fue correcta. Con estos plásmidos secuenciados se

retransformaron células de selección (JMB9 Δ TrpF) y se realizaron curvas de crecimiento en medio de selección. En la figura 20 se muestra que el crecimiento de la cepa de selección transformada con pDAN5 con la proteína interrumpida por el sitio *loxP511* (pDAN5 + β 1- α 4PRAI + β 5- α 8PRAI) fue muy parecido al de la cepa transformada con pDAN5 con la proteína silvestre (pDAN5 + PRAI WT).

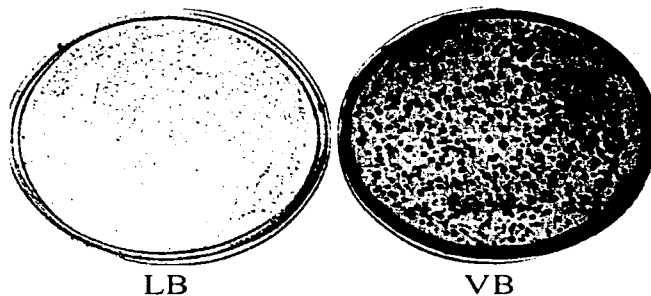


Figura 19. Cajas de medio LB Amp y VB Amp plateadas con células de la cepa JMB9 transformadas con el fagémido pDAN5+ β 1 α 4PRAI+ β 5 α 8PRAI e incubadas durante una noche.

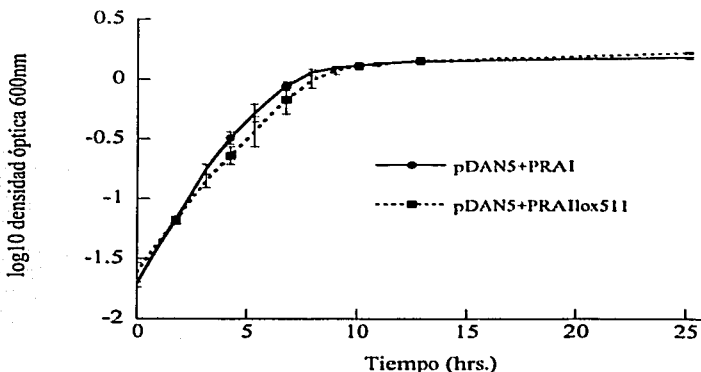


Figura 20. Curvas de crecimiento de la cepa JMB9 en medio VB Amp líquido transformada con pDAN5+PRAIwt (pDAN5+PRAI) por un lado y por el otro con pDAN5+ β 1 α 4PRAI+ β 5 α 8PRAI (pDAN5+PRAIloxP511).

A manera de controles negativos se transformaron por separado células de la cepa de selección con los vectores portadores de sólo una de las mitades (pDAN5 + β 1- α 4PRAI VH y pDAN5 + VL + β 5- α 8PRAI). Como se muestra en la figura 21, la enzima completa con el péptido conector interrumpiéndola complementa la función, mientras que las mitades de la enzima por separado no complementan en lo absoluto. Dado estos resultados, la enzima PRAI es una buena opción para probar la recombinación.

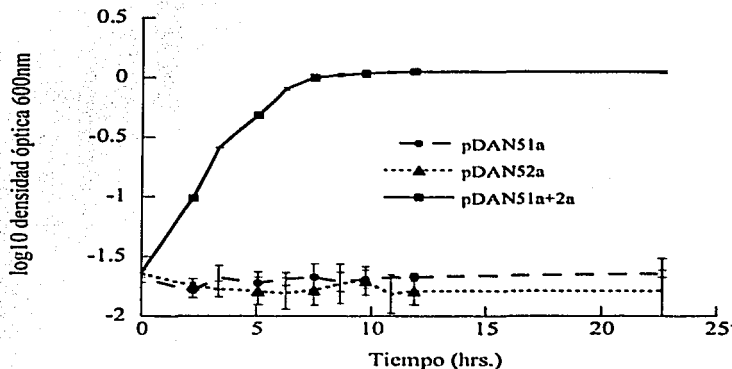


Figura 21. Curvas de crecimiento de la cepa JMB9 en medio VB Amp líquido transformada con pDAN5+ β 1 α 4PRAI+ β 5 α 8PRAI (pDAN51a+2a), con pDAN5+ β 1 α 4PRAI+VII (pDAN51a) y con pDAN5+VL+ β 5 α 8PRAI (pDAN52a).

Construcción y complementación de la TIM. De igual manera que con la PRAI, la amplificación de las dos mitades de TIM se realizó de tal forma que quedara dividida por el péptido conector exactamente entre la glicina 113 y la fenilalanina 114 (figura 18). En este caso no fue necesario clonar primero la segunda mitad, ya que el gen de MonoTIM (MTIM) no contiene algún sitio de restricción de los utilizados para la clonación. Por lo mismo no fue requerida la remoción del sitio XhoI extra encontrado en la VL del

anticuerpo, ya que como esta mitad fue extraída en un inicio con HindIII y Sall, el sitio no interfiere.

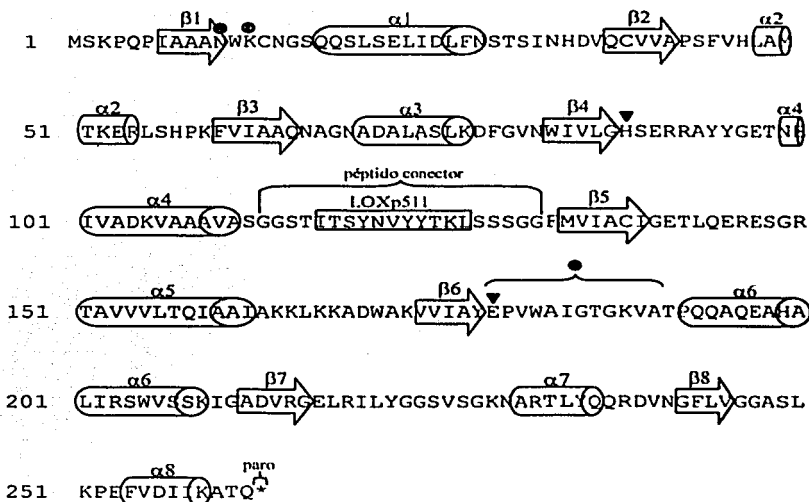


Figura 18. Estructura secundaria de la MTIM con la inserción del péptido conector LOXp511. ▼ indica los residuos catalíticos, ⊙ indica los residuos importantes para la catálisis.

Después de la segunda clonación, se transformaron células de la cepa de selección Δtpi (VEJM101) (Saab-Rincón G. *et al*, 2001) y se plataron en medio de selección Amp, con sus respectivas réplicas en medio rico Amp. En este caso se dejaron incubando a 37 °C 40 horas, ya que la variante de TIM monomérica, al ser menos activa, hace que las colonias crezcan más lento. A pesar de que las colonias en LB crecieron desde las 12 horas y en buen número, no creció ninguna colonia en el medio de selección, incluso después de 5 días de incubación (figura 22). Al parecer, la inserción del péptido conector inactiva de alguna manera la proteína. De todas formas, se obtuvo el plásmido de algunas de las colonias crecidas en medio rico y se verificó la presencia de ambas mitades, primero por medio de PCR y posteriormente secuenciando. Los alineamientos de las secuencias de estos

plásmidos con la secuencia teórica de MTIM corroboraron los resultados obtenidos por PCR, revelando que la construcción es correcta. De tal manera que no fue necesario hacer curvas de crecimiento para describir la no complementación de la función.

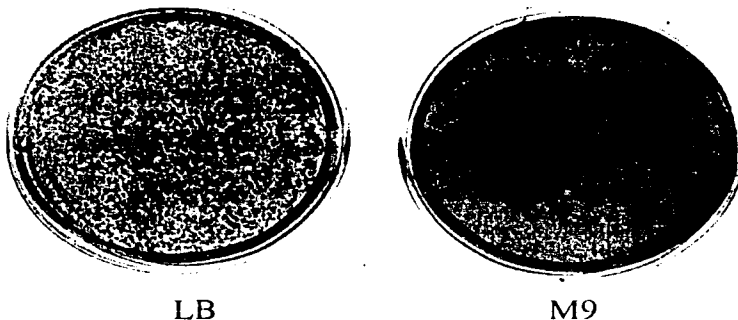


Figura 22. Cajas de medio LB Amp y M9 Amp plateadas con células de la cepa VEJMI01 transformadas con el fagémido pDAN5+ β 1 α 4 MTIM+ β 5 α 8 MTIM las cuales fueron incubadas durante una noche en el caso de LB y durante 5 días en el caso de M9.

La variante monomérica de TIM evolucionada hacia una mayor actividad es 40 veces menos activa que la proteína silvestre, a pesar de ser más activa que la MTIM original. Por lo mismo es posible que la actividad de la variante MTIM sea mucho más sensible a la inserción del péptido conector. Con base en esta idea se decidió hacer una construcción para la recombinación con TIM silvestre homodimérica. Dicha proteína se dimeriza postraduccionalmente, por lo que a nivel génico se podría trabajar de la misma forma que con los monómeros. Como TIM silvestre es varias veces (cuarenta) más activa que la variante monomérica se esperaban mejores resultados. La construcción se realizó de la misma forma que con la variante activa de MTIM y, a pesar de todo, los resultados fueron los mismos, la función no se complementó incluso después de 5 días de incubación (figura 23). La presencia de ambas mitades en el plásmido fue corroborada por PCR, usando como templado plásmidos purificados de las colonias crecidas en medio rico, y posteriormente secuenciando y alineando la secuencia con la secuencia teórica de TIMWT. Los resultados anteriores indican que ninguna de las variantes de TIM tolera una inserción

o que la posición en que se hizo la inserción no fue la correcta. Por lo mismo ninguna de las variantes permite evaluar los procesos de recombinación y por lo tanto no fueron utilizadas en las siguientes etapas.

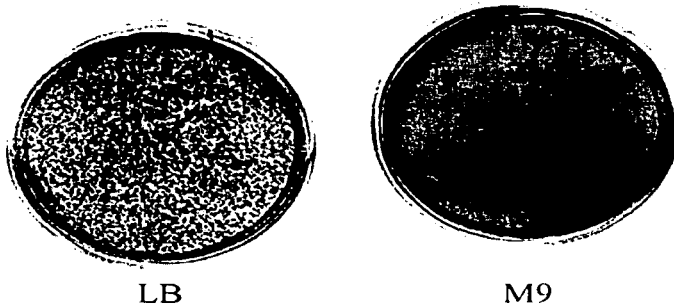
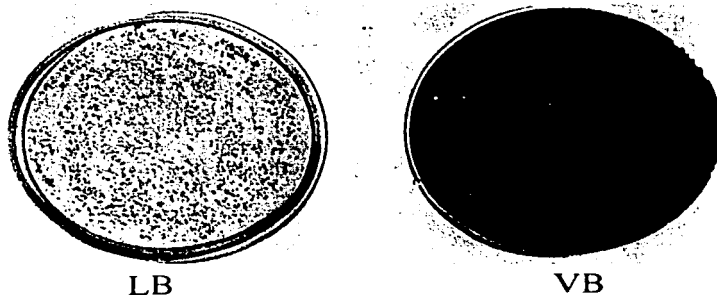


Figura 23. Cajas de medio LB Amp y M9 Amp plateadas con células de la cepa VEJM101 transformadas con el fagémido pDAN5+ β 1 α 4wtTIM+ β 5 α 8wtTIM las cuales fueron incubadas durante una noche en el caso de LB y durante 5 días en el caso de M9

Recombinación *in vivo* de PRA1

Recombinación in vivo de los fagémidos portadores de las mitades de PRA1 para formar una enzima PRA1 activa. El primer paso, como se observa en la figura 15, fue la obtención de los fagos infecciosos pDAN5, portadores por separado de cada una de las mitades de PRA1, utilizando una cepa F'. Los títulos de ambos fagos obtenidos fueron buenos; es decir, entre 1×10^{11} y 1×10^{12} fagos/ml. Con dichos fagos se infectó un cultivo de células de la cepa Cre⁺, crecidas en exceso de glucosa en una alta multiplicidad de infección (100:1 de cada fago). Este exceso es esencial para lograr una múltiple infección y que la recombinación se lleve a cabo entre múltiples fagémidos, dentro de cada célula. No se conoce totalmente la biología de la infección, pero se ha encontrado que la glucosa favorece la múltiple infección posiblemente porque detiene la maquinaria inhibitoria de infecciones posteriores (Sblattero D., *et al.*, 2000); tanto la alta proporción de infección como la glucosa parecen favorecerla.

La recombinación se dio durante 12 horas en agitación a 30 °C, temperatura a la cual se favorece la expresión de la Cre recombinasa. Los fagémidos ya recombinados y rescatados tuvieron igualmente títulos de alrededor de 1×10^{11} y 1×10^{12} . Para probar que la recombinación ocurrió con estos fagos, se infectaron células de selección $\Delta TrpF^+$ en una proporción de 1:1 o menor, para asegurar que sólo un fago infecte a una célula y por lo tanto el fenotipo de la célula reflejara el genotipo del fagémido. Después de lavarlas, se platearon en medio de selección con el antibiótico al que da resistencia pDAN5 y con su réplica correspondiente en medio rico. Después de 12 horas de incubación se observó un gran número de colonias en medio rico, mientras que las primeras colonias en el medio de selección se observaron hasta las 22 horas. Como se muestra en la figura 24, a las 40 horas de incubación también creció un gran número de colonias en el medio de selección, de tal manera que fue difícil calcular una cantidad exacta.



Figuras 24. Cajas de medio LB Amp Te y VB Amp Te plateadas con células de la cepa JMB9F⁺ infectadas con los fagos producto de recombinación las cuales fueron incubadas durante por una noche en el caso de LB y en el caso de VB durante 40 horas.

Cabe resaltar dos observaciones en relación a las colonias crecidas en medio de selección, por un lado la heterogeneidad en la velocidad de crecimiento y por otro la lentitud de la mayoría, en relación con las colonias de la cepa de selección F⁺ (JMB9) transformadas con la misma construcción, obtenidas en la primera etapa experimental. Independientemente de lo anterior, se obtuvieron los fagémidos de algunas de estas colonias escogidas al azar, aunque representando todos los fenotipos existentes,

confirmándose la presencia de ambas mitades de la proteína por PCR y secuenciando algunos representantes. El alineamiento con la secuencia teórica de pDAN5 + β 1- α 4PRAI + β 5- α 8PRAI volvió a confirmar la presencia de la proteína completa, indicando que el evento de recombinación sí ocurrió.

A su vez, como controles negativos, se realizaron recombinaciones infectando con los fagos portadores de una sola mitad de la proteína. Por un lado se infectó sólo con fagos de una mitad y por el otro sólo con fagos de la otra mitad, como se representa en la parte derecha de la figura 25.

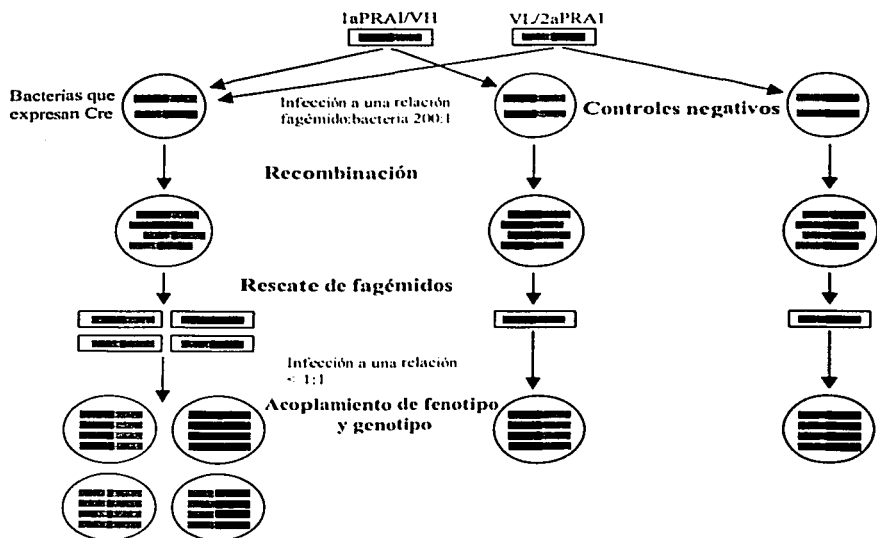


Figura 25. Proceso teórico de recombinación al infectar con los dos fágidos portadores por separado de cada una de las mitades y al infectar con sólo uno de los fágidos.

Inesperadamente, aunque a las 22 horas no creció ninguna colonia, a las 40 horas se observaron unas cuantas. Se obtuvieron los fágidos de diferentes colonias y se

analizaron por PCR. Algunos de estos plásmidos parecían tener ambas mitades, sin embargo, la mayoría tuvo sólo la mitad correspondiente o incluso ninguna. Se intentó secuenciar estos fagémidos con los *primers* que rodean al gen (pDANHind y pDANNhe), pero sólo se pudo obtener la secuencia de uno de ellos. El análisis demostró que están ocurriendo deleciones importantes en sitios extraños. Congruentemente al retransformar los plásmidos obtenidos en la cepa de selección original (JMB9 Δ TrpF) y en la cepa de selección F' (JMB9 Δ TrpF F'), no se complementa la función incluso después de 5 días de incubación. Esto indica que las primeras colonias son falsos positivos.

Tratando de aclarar algunas de estas incongruencias, se prepararon células competentes por CaCl₂ aislando sólo un fenotipo de colonia de la cepa de selección F', y se realizaron curvas de crecimiento después de transformarlas con el fagémido conteniendo el gen de la proteína completa, cada una de las mitades por separado y con el fagémido sin ninguna mitad. Como se observa en la figura 26, sólo crecieron las células transformadas con el fagémido que porta el gen de la proteína completa, las cuales además crecieron más lento que la cepa original (F') transformada con el mismo plásmido. Esto último explica la lentitud de crecimiento de las colonias de la cepa de selección F'.

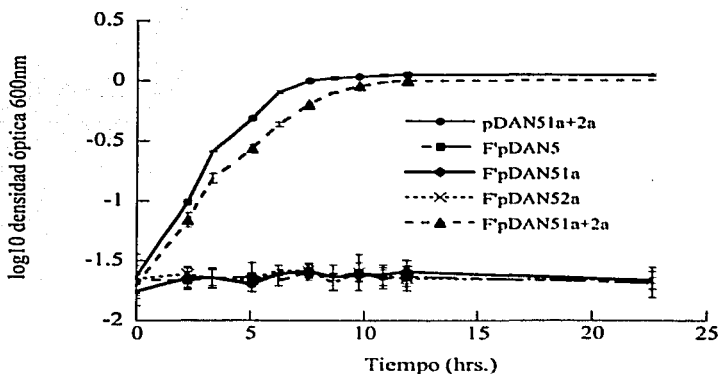


Figura 26. Curvas de crecimiento de la cepa JMB9F' en medio VB Amp Tc líquido infectada por separado con cada una de las construcciones de PRAI y comparada con la curva de crecimiento de la cepa JMB9 transformada con pDAN5+ β 1 α 4PRAI+ β 5 α 8PRAI.

Proporciones de los fagémidos recombinantes. La complementación de la función y la obtención de construcciones con ambas mitades de la PRAI sólo indica que la recombinación ocurre. La eficiencia de ésta se analizó por medio de las proporciones de los genotipos de los fagémidos ya recombinados. Como se observa en la parte izquierda de la figura 25, se esperaría que, cuando se infecta con los fagos de ambas mitades, las cuatro posibles combinaciones de fragmentos recombinados tienden a llegar a un equilibrio. Entonces si la recombinación no favorece ninguna de las combinaciones, cada uno de los cuatro posibles genotipos (los dos genotipos parentales [pDAN5 + β 1- α 4PRAI + VH y pDAN5 + VL + β 5- α 8PRAI], el genotipo con el anticuerpo completo [pDAN5 + VL + VH] y el genotipo con la proteína completa [pDAN5 + β 1- α 4PRAI + β 5- α 8PRAI]), al tener la misma probabilidad de ocurrir, representarían 25% de los fagémidos totales. Una de las maneras en que se analizaron los genotipos de los fagémidos fue a través del acoplamiento de cada genotipo al fenotipo de una célula de selección. Por lo antes mencionado, se esperaría que el número de colonias crecidas en medio de selección representase el 25% de las colonias crecidas en medio rico, ya que sólo aquellas colonias con el fagémido con ambas mitades de la proteína crecerían en medio de selección. Además, es posible que el tiempo de recombinación sea un factor determinante de la llegada al equilibrio de los cuatro genotipos, por lo que se hicieron recombinaciones variando los tiempos (6, 9, 12 y 15 horas) y se analizaron las proporciones de colonias en medio de selección. Es importante remarcar que para poder contar las colonias, debido a la gran eficiencia de transformación celular por infección, fue necesario diluir las células antes de platearlas. A diferencia de lo esperado, los porcentajes de colonias crecidas en medio de selección en todos los diferentes tiempos de recombinación estuvieron muy por debajo del 25% y fueron muy variables entre las repeticiones que se realizaron (entre 5 y 25 veces más bajos).

Sin embargo existe la posibilidad de que dichos valores tan bajos se deban no a la ineficiencia del proceso de recombinación, sino a un problema intrínseco del sistema de selección. Debido a la naturaleza del medio de selección, pueden estar creciendo un menor número de células independientemente del fagémido con el que hayan sido transformadas. Con el fin de descartar esta posibilidad se transformaron células únicamente con pDAN5 + PRAIwt, y se platearon de la misma manera en medio rico y en medio de selección. Como se sospechaba, la cantidad de colonias crecidas en medio de selección fue alrededor de 13

veces menor que la cantidad de colonias crecidas en medio rico, siendo también variable (figura 27). En esta figura se muestra además cómo la remoción de otros factores selectivos, como la adición de triptofano o la remoción del segundo antibiótico (tetraciclina), hace que el porcentaje de colonias crecidas en medio de selección aumente, aunque nunca llegue a ser del 100%.

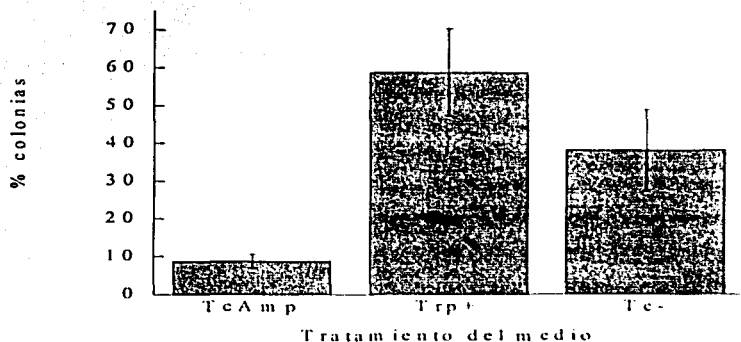


Figura 27. Gráfica de los porcentajes de colonias crecidas en medio de selección con respecto a las colonias crecidas en medio rico de la cepa JMB9F' transformada con el vector pDAN5+PRA1wt. TcAmp: VB con ambos antibióticos. Trp+: VBAmptc adicionado con triptofano y Tc-: VBAmptc sin tetraciclina.

Así entonces, tomando en cuenta esta disminución intrínseca del sistema de selección (figura 28), las medias de las proporciones de colonias obtenidas en medio de selección en los diferentes tiempos de recombinación no están tan alejadas de las esperadas, aunque la variación es considerable. Dado el error tan grande, se puede decir que no hay diferencias significativas al cambiar los tiempos de recombinación en esta escala temporal. Sin embargo, la variabilidad disminuye hacia los tiempos intermedios (12 horas), por lo que en experimentos posteriores sería recomendable usar estos tiempos de recombinación.

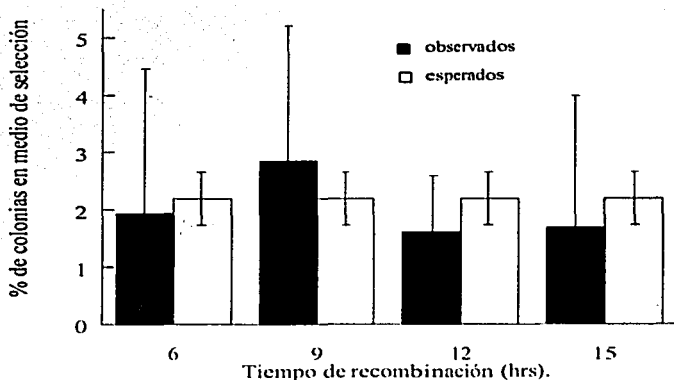


Figura 28. Gráfica de los porcentajes de colonias de la cepa JMB9F^r crecidas en VB Amp Tc con respecto a las crecidas en LB Amp Tc infectadas con los fagémidos resultado de los diferentes tiempos de recombinación. Estos porcentajes se comparan con los porcentajes de colonias esperados teóricamente después de hacer las correcciones por el efecto del medio de selección.

Para cada repetición de los experimentos de recombinación anteriores se realizaron recombinaciones con sólo una o la otra de las mitades a manera de controles negativos. Al acoplar el genotipo de los plásmidos resultantes de éstos con el fenotipo celular no se obtuvieron colonias en medio de selección.

Así mismo a través de todos los experimentos pasados se estuvieron buscando falsos positivos, como los encontrados en los controles negativos, que pudieran afectar las proporciones obtenidas por medio de secuenciación de colonias al azar. Sin embargo, los falsos positivos parecen estar relacionados con la densidad de células plateadas, ya que al utilizar bajas densidades celulares no se encontraron falsos positivos. Incluso las colonias con crecimiento muy lento (más de tres días) muestran tener la construcción correcta al ser secuenciadas.

Patrones de restricción de los fagémidos recombinados. La otra forma de analizar los genotipos resultantes de la recombinación es por medio del polimorfismo en sus

patrones de restricción. Como se observa en la figura 29, las enzimas AlwNI y MscI permiten hacer dicho análisis. Se tomaron 100 colonias al azar del experimento de acoplamiento del fenotipo al genotipo de los fagémidos resultado de la recombinación por 12 horas sin selección, se purificaron sus fagémidos y se digirieron con las enzimas anteriores por separado. La figura 30 muestra las proporciones obtenidas de los diferentes genotipos. Las proporciones de los fagémidos con la primera mitad y con ambas mitades de PRAI son cercanas al 25% esperado (33.4% y 30.5% respectivamente), pero las proporciones de pDAN5 con ambas mitades del anticuerpo y con únicamente la segunda mitad de PRAI son mucho más bajas que las esperadas (1.4% y 5.72% respectivamente). También es notorio el gran número de genotipos que parecen ser rearrreglos no esperados (monstruos). Es importante hacer notar la gran cantidad de genotipos que no pudieron ser determinados (31%), principalmente debido a la poca cantidad de plásmido con que se contaba resultado de una extracción deficiente.

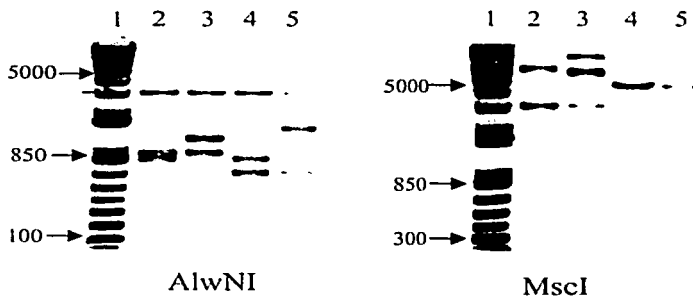


Figura 29. Electroforesis en gel de agarosa de los fragmentos de restricción de los cuatro posibles fagémidos resultantes de la recombinación digeridos con las enzimas AlwNI por un lado y MscI por el otro. 1: marcador (1Kbplus) 2: pDAN5-XhoI, 3: pDAN5+β1α4PRAI+VH, 4: pDAN5+VL+β5α8PRAI y 5: pDAN5+β1α4PRAI+β5α.

Otro posible factor que pudiera afectar las proporciones es que el producto (proteína) de alguno de los genotipos sea tóxico para la célula, impidiendo tener colonias

transformadas con este fagémido incluso en medio rico. Para eliminar esta posibilidad, se transformaron células con los cuatro genotipos posibles y se crecieron en medio rico. La figura 31 muestra que sólo la construcción pDAN5 + b1-a4 + VL tiene una menor tasa de crecimiento lo cual podría hacer que esta mitad estuviese un poco menos representada. Sin embargo, no se tuvo problema en identificar una proporción de ésta cercana a lo esperada. Así, la falta de representación puede deberse a otros fenómenos como la naturaleza recombinatoria de la cepa Cre⁺.

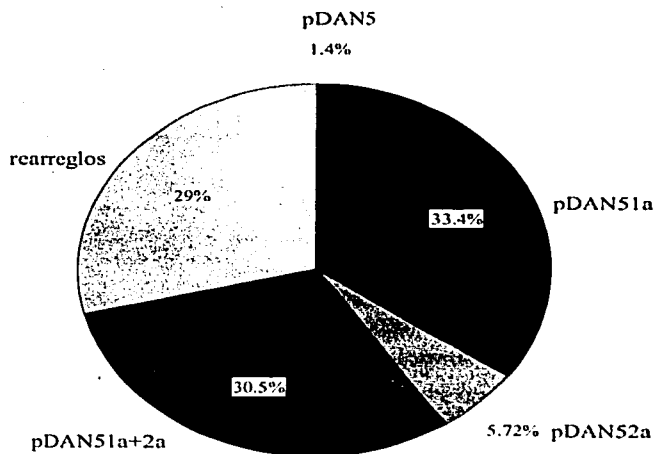


Figura 30. Porcentajes de cada uno de los genotipos de los fagémidos resultado de la recombinación determinados por patrones de restricción.

Utilización del sistema en librerías

La utilización del sistema en librerías está dirigida más que nada a probar sus alcances en cuanto a combinación de variantes poco representadas en una biblioteca. Como se explica en la introducción, un paso fundamental en la creación de variabilidad es la combinación de variantes que confieran una propiedad deseada. Entre mayor tasa de

mutagenicidad empleada mayor es la probabilidad de acumular mutaciones deletéreas, por lo que es importante poder combinar mutantes benéficas para aumentar la variabilidad muestreada sin llegar a un banco grande de genes muertos. En el mismo contexto, entre menos representadas estén las variantes en una biblioteca, será menos probable combinar las adecuadas y por lo tanto deberá ser más grande la biblioteca para incluir la combinación que dé una clona mejorada. Con el fin de demostrar la utilidad del sistema de recombinación *in vivo* en la creación de bibliotecas de enzimas, se utilizaron dos mutantes sencillas deletéreas de PRA1 y la combinación de éstas en una doble mutante deletérea. Cada una de estas mutaciones reemplaza uno de los residuos catalíticos situados en cada mitad. De esta manera, al contaminar bibliotecas de la doble mutante con cierta proporción de las mutantes sencillas y recombinarlas, podremos analizar la capacidad del sistema de combinar las mutantes sencillas hacia una proteína activa sin mutaciones. Utilizando diferentes proporciones de las mutantes sencillas en la biblioteca podremos encontrar los límites de combinación de variantes muy poco representadas del sistema.

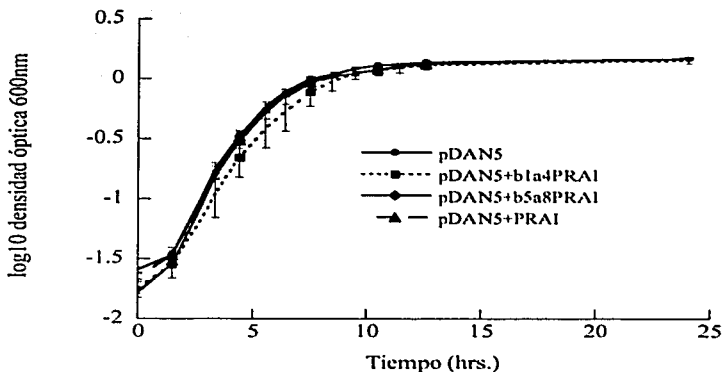


Figura 31. Curvas de crecimiento de la cepa DH5αF' en medio LB Amp líquido transformada por separado con cada uno de los posibles fágémidos resultado de la recombinación.

La primera biblioteca que se probó contiene una de cada una de las mutantes sencillas, por cada 1000 dobles mutantes. El procedimiento utilizado es el que se representa

en la figura 16. El paso que puede ser más limitante es la transformación por electroporación de la librería primaria en una cepa F' para poder obtener posteriormente fagos para la recombinación. Entre más eficiente sea la transformación de la primera librería, mayores serán las probabilidades de que las mutantes sencillas se encuentren representadas en los fagos resultantes y puedan recombinarse para dar una proteína activa. Con la primera librería se lograron transformar más de 1×10^9 células, por lo que la librería fue de un tamaño aceptable. El otro paso que puede ser limitante es el sistema de selección de la variante sin mutantes, ya que aunque la biblioteca sea muy grande, sólo será posible probar el número de variantes que puedan ser acopladas al fenotipo de una célula de selección. Por lo anterior, el sistema de selección estará limitado a la cantidad de células en estado infectivo que puedan ser crecidas. En este caso se utilizó 1 litro de cultivo a una densidad óptica de 0.5, que representa alrededor de 5×10^{11} células. Dado lo último, como máximo sólo fueron analizadas 5×10^{11} variantes. Para poder seleccionar tal cantidad de células fue necesario hacer una preselección en medio líquido, en lugar de cajas de medio sólido ya que se hubieran necesitado 50 000 cajas plateando 10 millones de células por caja.

Después de recombinar esta librería con las mutantes sencillas en una proporción de 1/1000, se lavaron y se crecieron en medio líquido. En este cultivo se observó un crecimiento obvio hasta las 24 horas de incubación. Con el fin de enriquecer la o las clonas activas se hicieron varios pasos de crecimiento y reinoculación en medio mínimo. Después de 72 horas de incubación, a través de las diferentes alícuotas de medio de selección fresco, se diluyeron las células y se plateó una pequeña alícuota en cajas para obtener colonias aisladas. Se purificó el fagémido de algunas de las colonias. Por otro lado, se purificó fagémido de 5 ml del cultivo líquido de la última alícuota de enriquecimiento sin aislar colonias (ver figura 16). Los fagémidos resultantes se retransformaron por electroporación en la cepa de selección F' y se platearon en cajas de medio VB. Se obtuvieron muchas colonias, por lo que se secuenciaron algunos representantes. Las secuencias revelaron que los fagémidos no tienen ninguna de las mutaciones deletéreas.

Con el fin de demostrar los límites del sistema se hizo una nueva biblioteca primaria, pero ahora con cada una de las mutantes sencillas en una proporción de 1/100,000 y se recombinó. La biblioteca primaria también fue mayor a 1×10^9 variantes y la selección

se llevó a cabo de la misma manera que con la biblioteca anterior. Inesperadamente se observó crecimiento ya después de una noche, sin embargo se prosiguió de la misma forma. Al final se obtuvieron los fagémidos de algunas colonias aisladas en cajas de medio sólido, así como fagémidos purificados de 5 ml del último cultivo líquido y se retransformaron en células JMB9 F⁻ (ver figura 16). Inesperadamente, ninguno de los tres fagémidos obtenidos de colonias aisladas en medio sólido al transformarse en la cepa JMB9F⁻ no complementaron la función de PRAI. Consistente con esto, sólo crecieron algunas colonias de las células que habían sido transformadas con los fagémidos purificados del medio de cultivo líquido. Al analizar las secuencias de los diferentes fagémidos se encontró que aquellos provenientes de las colonias aisladas de la cepa de selección F⁻ en cajas de medio sólido seguían teniendo alguna de las mutantes y por lo tanto eran falsos positivos. En contraste, los fagémidos obtenidos de las células de selección F⁻ después de transformarse con los fagémidos aislados directamente de la última alícuota de medio líquido no tuvieron ninguna de las mutaciones. Por lo tanto, aunque los falsos positivos dificultaron la obtención de la variante correcta, fue posible seleccionarla. En estos experimentos, el enriquecimiento en medio líquido en la cepa de selección F⁻ es utilizado como preselección ya que es necesario retransformar los plásmidos obtenidos de este paso en la cepa original de selección F⁻ y platear en medio mínimo para eliminar los falsos positivos.

A manera de controles negativos, cada una de las bibliotecas primarias, fue infectada en la cepa de selección, antes de someterlas a la recombinación, acoplando el genotipo al fenotipo, y plateada en medio de selección. Después de varias horas de incubación se obtuvieron algunas colonias las cuales resultaron ser falsas positivas ya que al retransformar el plásmido en la cepa JMB9F⁻ no hubo crecimiento.

Por último, para comparar los resultados anteriores con el método de *gene shuffling* se realizó el barajeo de la última librería, la que tiene las mutantes sencillas diluidas 1/100,000, además del barajeo de una librería que tiene las mutantes sencillas diluidas 1/100 (como control positivo). Del barajeo de la librería con las mutantes sencillas 1/100 se obtuvieron varias colonias en la cepa de selección F⁻ en las que se complementa la función y cuyos fagémidos, revelado al ser secuenciados, son portadores de PRAI sin mutaciones. Sorprendentemente, estos mismos resultados, aunque con menos colonias, se obtuvieron del barajeo de la biblioteca con las mutantes sencillas diluidas 1/100,000.

Discusión

Complementación de las funciones de las enzimas con el péptido LOXP511 insertado

El efecto de la inserción del péptido conector LOXP511 sobre la conformación nativa de las enzimas y, por lo tanto, sobre su actividad es un punto fundamental en la implementación de un sistema de recombinación *in vivo* con miras a ser utilizado en evolución dirigida de proteínas. La utilidad del sistema será cuestionable a pesar de las posibles ventajas que el mismo pueda aportar en cuanto a generación de variabilidad, si la actividad se ve mayormente afectada. Al perder la conformación activa, la recuperación de una conformación que permita restablecer, e incluso llegar a mejorar las propiedades de la enzima, será más difícil. Hasta antes de este trabajo, los efectos de la inserción del péptido conector LOXP511 sólo habían sido probados en anticuerpos, enlazando la cadena ligera con la pesada. El péptido fue diseñado con este fin dando buenos resultados (Sblattero D. *et al.*, 2000), sin embargo es poco lo que se puede deducir a partir de éstos en cuanto al posible efecto que tenga su utilización en enzimas. Probablemente, debido a las características estructurales y funcionales, los anticuerpos sean péptidos mucho más permisivos a alteraciones de esta naturaleza.

Los resultados de complementación de la función de PRAI, obtenidos con la enzima portadora del péptido conector, pueden parecer sorprendentes si son vistos fuera del contexto de los antecedentes. El tamaño del péptido conector LOXP511 (18 aa) es equivalente al de un módulo $\beta\alpha$ del barril, por lo que su inserción no es una alteración despreciable. Sin embargo, los diferentes trabajos de plegamiento con la enzima, en especial los del grupo de Kirschner D. y posteriormente los estudios de corte y utilización de cierres de leucina para unir fragmentos (Fuentes G. P. 2001; Saab-Rincón G., comunicación personal), sugerían que es posible realizar una inserción de este tipo con un efecto negativo mínimo sobre la actividad de la enzima. Los resultados aquí obtenidos apoyan la idea de que la PRAI es una enzima robusta que permite modificaciones de este tipo, ya que la proteína con la inserción complementó la función casi de igual manera que la proteína sin la inserción. A su vez, los resultados anteriores sugieren que es posible

trabajar de manera más general con barriles $\alpha\beta$ interrumpidos por este tipo de péptidos conectores en sitios equivalentes.

Por otro lado, desde un enfoque de plegamiento de la proteína, es muy interesante el sitio donde está insertado el péptido conector. Éste se encuentra en el asa que conecta la α -hélice 4 con la β -plegada 5, que es una de las asas inferiores del barril. Se ha propuesto que estas asas inferiores están relacionadas, más que nada, con la estabilidad del barril (Urfer R. *et al.*, 1992), lo que ha llevado a suponer que las asas de los barriles podrían estar divididas funcionalmente, siendo las inferiores (α - β) las importantes para la estabilidad y las superiores (β - α) las necesarias para la catálisis. Los resultados obtenidos contradicen en cierta forma lo anterior, ya que una de las asas inferiores es alterada y, a pesar de ello, la proteína sigue siendo activa y, por lo tanto, sigue teniendo la conformación de barril $\alpha\beta$ que es la forma más probable de la enzima activa. Sin embargo, habría que hacer otro tipo de pruebas para caracterizar directamente la estabilidad de la enzima. Una posibilidad remota es que una fracción del péptido conector (la primera) estuviera formando parte del barril a manera de β -plegada, de tal forma que el resto del péptido conector junto con la β -plegada 5 quedaran fuera del barril, pero en la parte superior de éste, como asa β - α , sin afectar la estabilidad. Lo anterior es sumamente improbable ya que se alteraría fuertemente la conformación nativa de la enzima, sin posibilidad de complementar la función como lo hace. Para determinar con exactitud el acomodo del péptido conector en la estructura del barril habría que hacer estudios de emisión de fluorescencia o cristalografía que podrían arrojar resultados interesantes al respecto.

La PRAI, entonces, es una excelente candidata para ser usada en el sistema de recombinación *in vivo*. El mantenimiento de la actividad después de la inserción del sitio LOXP511 permite que el sistema de complementación (o selección) con la cepa Δ TrpF, sea utilizado para el análisis del proceso de recombinación *in vivo*.

Por otro lado, los resultados de complementación de función con TIM son totalmente contrarios. La conformación de la enzima no permite la inserción del péptido conector LOXP511 en el asa inferior entre la α -hélice 4 y la β -plegada 5, ya que al hacerlo la actividad se pierde por completo, o al menos es totalmente indetectable con el sistema de complementación. Existen pocos antecedentes de modificaciones en TIM y, en realidad, no

tan relacionados a cambios de esta naturaleza. El trabajo que más información provee al respecto es el de Bertolaet B. L. *et al.* (1995), en el que la TIM de músculo de pollo es fragmentada en los diferentes límites intrón – exón, observándose una complementación débil de la función al expresar juntos los fragmentos. A pesar de este antecedente, se decidió insertar el péptido conector en el mismo lugar que en la PRAI, con base en la relación estructural de las dos enzimas. Los criterios usados para fragmentar la TIM de pollo, al ser los límites exón – intrón, podrían ser totalmente ajenos a la proteína utilizada. Además, algunos de estos límites están exactamente en estructuras secundarias, donde la inserción de un péptido podría traer consecuencias fatales a la conformación total. La idea de insertar el péptido conector en una asa que es una región más flexible, y especialmente en una asa que ya había aceptado inserciones en proteínas estructuralmente relacionadas, parecía más adecuada.

Incluso la actividad de la proteína silvestre (TIMwt), que es una de las enzimas conocidas más eficientes, se vió totalmente alterada por la inserción del péptido conector. Con la enzima monomérica existía la posibilidad de que la menor eficiencia catalítica causada por la monomerización se estuviera sumando al efecto negativo de la inserción, resultando en la inactivación total de la enzima. Sin embargo, la enzima silvestre más robusta sufrió la misma inactivación. Podría pensarse que el péptido conector interfiere con la capacidad de dimerización de la proteína silvestre. Lo último no es muy probable, ya que la inserción es en las asas inferiores, mientras que las asas que intervienen en la dimerización son las superiores.

Modelamientos estructurales teóricos de TIM con el péptido conector podrían generar información relevante para la localización de un buen sitio de inserción. Este tipo de datos, junto con la información que puedan aportar estudios de corte y modificaciones de la misma naturaleza, serían sumamente útiles para lograr una enzima activa con el péptido conector insertado. Si la elección del sitio particular de inserción en TIM hubiera estado basada en el mismo tipo de antecedentes directos que se tenían para PRAI, muy probablemente hubiera sido posible obtener una enzima activa.

Los resultados negativos de complementación con ambas variantes de TIM nos hablan de las restricciones del sistema. La inserción de un péptido conector en enzimas, cuya secuencia es necesaria para la recombinación no es una modificación trivial. Por lo

mismo, la simpleza de la implementación de un sistema de recombinación *in vivo* general, o para la mayoría de los barriles α/β , se pone en duda hasta cierto punto. Lo último independientemente de la utilidad del sistema en PRAI. Posiblemente sea necesaria bastante información tanto teórica como práctica para poder encontrar un sitio de inserción idóneo para cada barril α/β . Sin embargo, existen otras posibilidades que tienen que ver con la modificación del sistema de recombinación. Una posibilidad es la alteración de la especificidad de la Cre recombinasa, de tal forma que ésta sea capaz de reconocer y mediar la recombinación entre sitios que codifiquen para péptidos conectores con características más aceptadas por la conformación de la proteína a recombinar. Este tipo de alteraciones de especificidad han sido logradas a través de evolución dirigida de la proteína (Buchholz F. *et al.*, 2001) y podrían dar buenos resultados al hacer más general la aceptación de la inserción del péptido. Otra posibilidad es la utilización de un intrón autoeditable en procariontes como se hace en el trabajo de Fisch I. *et al.* (1996). En este sistema, el sitio *loxP* utilizado para la recombinación es colocado dentro del intrón, de tal forma que después de la recombinación se favorece la autoedición de éste, reduciéndose considerablemente el tamaño del péptido conector a un tamaño final de alrededor de 5 aminoácidos. Aunque la disminución en el tamaño del péptido representa una importante ventaja, al editar el sitio *loxP* se pierde la posibilidad de utilizar de nuevo estos genes en otro ciclo de recombinación.

Los resultados de complementación de ambas proteínas también nos dicen aspectos importantes de las características generales del grupo de los barriles α/β y especialmente de las diferencias entre estos dos integrantes de la familia, la PRAI y la TIM. Como se observa, los conocimientos obtenidos de los estudios previos de la PRAI no pudieron ser extrapolados a la TIM, a pesar de pertenecer a la misma familia. Esto nos recuerda que los barriles α/β , a pesar de su probable homología y sus múltiples características en común, sobre todo estructurales, tienen propiedades particulares distintas, incluso estructurales. Por otro lado, situando a ambas enzimas en un contexto evolutivo, se podría pensar que la PRAI es mucho más permisiva a la inserción en la mitad de la proteína, debido al parecido de ésta con un barril ancestral resultado de la dimerización de un medio barril. La PRAI es la enzima más parecida a HisA y HisF, que son las proteínas en las que se encontró la

evidencia para postular el modelo de la evolución de un barril completo a partir de un dímero de medios barriles ancestrales. Posiblemente la estructura de la PRAI mantenga más características de un dímero de medios barriles α/β que la TIM, lo cual permite la inserción en este sitio. La TIM, al divergir del barril α/β ancestral, perdió las características que permiten la inserción en la PRAI.

Recombinación *in vivo* de fragmentos de PRAI

Recombinación in vivo de fragmentos de PRAI para formar una enzima activa. La forma en que se analizó la recombinación *in vivo* fue a través de la integración de PRAI completa, y por lo tanto activa, a partir de dos mitades clonadas por separado en pDANS por medio de recombinación. De esta manera, al acoplar el fenotipo de células de selección con el genotipo de los vectores producto de la recombinación, sólo se observa crecimiento de colonias en medio de selección si el proceso de recombinación se llevó a cabo, es decir, si se integró la proteína activa. Los resultados demuestran que esto es exactamente lo que está ocurriendo. Sin embargo, a diferencia de los resultados obtenidos con el crecimiento de las colonias de la cepa de selección F' en la primera parte experimental, las colonias tardan más en crecer y muestran tamaños muy heterogéneos, a pesar de que todas las células provienen de una sola colonia JMB9F'. Independientemente de su tamaño, todos los vectores de las colonias analizadas contenían ambas mitades de la proteína en orden correcto. La lentitud de crecimiento podría esperarse, ya que los productos génicos del plásmido F, a pesar de estar integrados al cromosoma, representan una carga metabólica extra para la célula. Además, el proceso de infección por el que se transforman las células también podría causar un retardamiento en el crecimiento celular. Estos factores son totalmente independientes de la actividad enzimática del péptido portado por el vector con que se les haya infectado. Lo anterior es corroborado por las curvas de crecimiento mostradas en la figura 26, donde se puede ver que el crecimiento de las células de la cepa de selección F' es más lento que el de las células de la cepa F', a pesar de haber sido transformadas con el mismo vector. Por el otro lado, la heterogeneidad de tamaños es difícil de explicar. Tratando de eliminarla, se intentó aislar un fenotipo de colonias por medio de la repetida esparción de la cepa F' original, pero el resultado siempre fue el mismo.

Parecería ser que una característica intrínseca de esta cepa JMB9F' es la variabilidad. Dos meras especulaciones al respecto, ajenas a las características intrínsecas de la cepa, son primero, que el momento exacto en que se lleva a cabo la infección varía dentro del cultivo celular una vez que se añaden los fagos, haciendo variar el inicio del crecimiento, y segundo, que la disponibilidad de fago *helper* hace que el ciclo del fagémido como fago siga, alterando el crecimiento de la colonia. Se podrían hacer diferentes experimentos para probar esto, pero nos alejarían de los objetivos del trabajo.

Por otro lado, si el proceso de recombinación se lleva a cabo sólo con una de las mitades, al acoplar genotipo y fenotipo, no se debería observar crecimiento de colonia alguna en medio de selección, ya que por más que haya ocurrido la recombinación, el resultado final dará el mismo péptido inicial inactivo (figura 25). Estos experimentos fueron utilizados como controles negativos. Inesperadamente, después de bastantes horas de incubación se obtuvieron algunas colonias. Sin embargo, al analizar directamente el vector de estas células y retransformarlo en la cepa JMB9F', se determinó que eran falsas positivas. Otra observación es que están sucediendo arreglos inesperados en los que la región de recombinación se pierde. Este tipo de deleciones son resultado de la inestabilidad del sistema de recombinación *in vivo*. Una explicación de los falsos positivos podría ser que los vectores con una sola mitad de la PRAI en la cepa de selección JMB9F' complementaran de una u otra manera la función. Como se ve en las figuras 21 y 26 lo anterior no es una explicación, ya que estas construcciones en ninguna de las cepas de selección complementan la función de la PRAI. En realidad no se logró determinar qué es lo que causa con exactitud estas colonias falsos positivos, aunque se asociaron dos factores a su aparición. El primero es la diferencia que hay entre las dos cepas de selección, la F' y la F' original, ya que nunca se observaron falsos positivos con la cepa JMB9 original. Las diferencias que los explican podrían no tener que ver directamente con la presencia de los componentes de F, sino más bien con las consecuencias del proceso de conjugación bacteriana por el que se construyó la cepa F'. En este proceso no sólo se transfieren los genes relacionados con F, sino que se puede transferir más material genético de diversa naturaleza, el que podría estar contribuyendo a la aparición de falsos positivos. Para solucionar lo anterior se podría hacer de nuevo la cepa F' pero sin conjugación, con la transformación directa del plásmido F. De esta manera se evitaría la ganancia de material

genético adicional. El segundo factor relacionado con su aparición es la densidad celular. Se observó que los falsos positivos sólo ocurren cuando el número de células que se platean es muy alto (cuando en medio rico no es posible distinguir colonias aisladas). Al platear los controles negativos, si se hace con un número no tan alto de células, las colonias falsas positivas no aparecen. Otro punto que es importante mencionar es que nunca se encontraron en una misma caja colonias positivas y falsas positivas mezcladas. Probablemente por la poca representación de las falsas positivas o tal vez por la competencia con las positivas.

Independientemente de lo anterior, si se tuviera el tiempo, sería interesante explorar más a fondo las razones de estos falsos positivos, ya que existe la posibilidad de que ciertos rearrreglos peptídicos complementen la función de PRAI.

Proporciones de los genotipos de los fagémidos recombinados. Además de probar el hecho de la recombinación *in vivo* de PRAI, es importante analizar la eficiencia con la que ocurre este proceso. Para esto se analizaron las proporciones de los genotipos de los fagémidos recombinantes de dos maneras diferentes. La primera fue a través de la proporción que representan las colonias crecidas en medio de selección de las colonias crecidas en medio rico, ya que se esperaría que las colonias crecidas en medio de selección fueran el 25% de las crecidas en medio rico, al representar uno de los cuatro posibles genotipos. Se hicieron procedimientos con diferentes tiempos de recombinación para observar el efecto de este factor. Además, fue necesario ajustar la proporción de colonias esperadas con la proporción de colonias que normalmente crecen en el medio de selección cuando todas las células están transformadas con el mismo vector portador de la proteína silvestre. Como se observa en la figura 27, el sistema de selección *per se* representa un cuello de botella para el crecimiento de las colonias, sobretodo cuando existen varios factores restrictivos (los antibióticos y la falta de triptofano). Este factor intrínseco del sistema de selección disminuyó la proporción esperada considerablemente, alrededor de 12 veces, y le confirió una variabilidad también de magnitud considerable. Hay que resaltar que el cuello de botella que los sistemas de selección pueden representar por sí solos es una restricción importante en cualquier proceso de evolución dirigida; en este caso, por cada positiva que se obtenga se pueden perder once. En cada caso es recomendable entonces

caracterizar el sistema de selección usado para poder jugar con los diferentes factores restrictivos, de tal manera que la selección se optimice. En el presente estudio, por ejemplo, si no se utiliza la tetraciclina en el medio de mínimo el número de positivas perdidas por el método de selección disminuye considerablemente.

Con la proporción de colonias esperadas ajustada se pudo comparar los resultados obtenidos de los procedimientos con los diferentes tiempos de recombinación, como se muestra en la figura 28. Aquí se puede observar que las medias de las proporciones observadas se encuentran en el intervalo de la proporción esperada, además de que las proporciones observadas en los diferentes tiempos, no son diferentes entre sí debido a la gran variabilidad. La eficiencia del proceso es comparable con la esperada y el tiempo de recombinación a esta escala no es un factor que esté influyendo la eficiencia del mismo. Sin embargo, es notable la variabilidad obtenida de los resultados de los procesos de recombinación. Existen varios factores que sumados podrían explicar esta gran variabilidad. El primero es la variabilidad intrínseca del sistema de selección usado. Como se puede ver en la figura 28, la proporción esperada tiene una cierta variación, menor que la de los observados, pero que puede explicar cierta parte de la variación de éstos. Otra explicación que comprende varios fenómenos es la inestabilidad del sistema. Esta inestabilidad también fue observada por el grupo que desarrolló el sistema de recombinación *in vivo* de un sólo vector para anticuerpos (Sblattero *et al.*, 2000). Como se mencionó anteriormente, en éste trabajo también se detectaron de manera directa, al secuenciar algunos de los plásmidos de los falsos positivos, deleciones y rearrreglos causantes de la pérdida del gen completo de la proteína o de sus mitades, lo que al final tendría efecto en las proporciones de las colonias. Tres fenómenos podrían estar causando esta inestabilidad. Primero estarían los posibles problemas causados por la incompatibilidad de plásmidos. A pesar de que en este sistema se ha observado que la idea de que dos plásmidos diferentes con un mismo origen de replicación no pueden existir en una misma célula no se cumple, es posible que, más que no cumplirse, su efecto sea mínimo y esto esté haciendo que cada evento de recombinación tenga resultados variables. Otro problema de inestabilidad que tiene el sistema es que los sitios heterólogos, en este caso *loxP511* y *loxPwt*, recombinan entre sí, pero con una frecuencia muy baja (Siegel R. W. et al., 2001). Este tipo de fenómenos causarían la deleción de la región que se recombina y, por lo tanto,

una alteración en las proporciones de los genotipos finales. El último problema de inestabilidad serían los rearrreglos cromosomales que se han observado en cepas que tienen altos niveles de expresión de la Cre recombinasa. La recombinación entre los fágemidos y el cromosoma celular podrían resultar en alteraciones de los fágemidos. Además, la cepa Cre⁺ que se utiliza para la recombinación en estos experimentos no es Δ TrpF, por lo que podrían ocurrir recombinaciones remotas en las que el gen silvestre fuera adquirido por el fágemido proveniente del cromosoma celular.

Otros dos aspectos experimentales que podrían tener que ver con la variabilidad encontrada son, por un lado, las diferencias en multiplicidad de infección de fagos portadores de una mitad respecto a los fagos portadores de la otra mitad y, por el otro, la posibilidad de que al acoplar el genotipo al fenotipo ocurra una infección múltiple. En el primer caso al existir, por error experimental, mayor cantidad de fagos portadores de una mitad que portadores de la otra, muy posiblemente el equilibrio de la recombinación se vea desplazado en alguna dirección. Estos errores de infección no equivalente dependen, hasta cierto punto, de lo exactos que sean los métodos de titulación. En una ocasión, donde se sobreinfectó por descuido con alguna de las mitades, el resultado fue una notoria disminución en las colonias crecidas en medio mínimo. En el segundo caso, se afectaría la proporción final si suponemos que a pesar de infectar las células en una proporción menor de 1:1 con los fagos resultado de la recombinación, ocurre que dos fagos infectan una misma célula.

Finalmente, la suma de todos estos fenómenos podría explicar esta gran variabilidad. Algunos de estos factores se pueden controlar afinando los métodos experimentales, pero otros son intrínsecos de los sistemas utilizados. A pesar de la gran variabilidad, incluso cuando la recombinación parece haber ocurrido de manera muy alejada a lo esperado, los niveles de recombinación son suficientes para que el sistema sea una buena herramienta en la generación de variabilidad. Suponiendo que la eficiencia de recombinación es de un décimo de lo esperado y esperamos bibliotecas de entre 10^{11} y 10^{12} variantes, las bibliotecas generadas serían de entre 10^{10} y 10^{11} variantes. Este es un tamaño varios órdenes de magnitud mayor que los tamaños logrados con los métodos convencionales, en los que se emplea transformación por electroporación.

La otra forma en que se analizaron las proporciones de los genotipos de los vectores resultantes de la recombinación fue por medio de los patrones de restricción de éstos, utilizando dos enzimas diferentes. En este caso, al ser vectores obtenidos de colonias crecidas en medio rico, se espera encontrar proporciones cercanas al 25% de cada uno de los genotipos. Dos de los genotipos, el de la proteína completa y el que sólo tiene la primera mitad, resultaron en una proporción muy cercana a lo esperado, sin embargo los otros dos genotipos estuvieron muy por debajo. El resto són genotipos que parecen ser resultado de rearreglos. El problema de este análisis es que, debido al trabajo que implica metodológicamente, sólo se realizó para el resultado de una recombinación por 12 horas, lo que le resta valor estadístico. Además, sería necesario asegurarse que el porcentaje de genotipos desconocidos fuera mucho menor y que los supuestos rearreglos son en realidad rearreglos, y no cambios puntuales que modifican los patrones de restricción. Estos problemas metodológicos se pueden estar sumando a las causas de la variabilidad antes mencionadas, dando los resultados que se obtienen. Independientemente, como ya se mencionó, si la eficiencia de recombinación se da en un 50%, como es este caso, el sistema sigue siendo una herramienta atractiva en la generación de variabilidad, ya que se podrían lograr bibliotecas varios ordenes de magnitud mayores que con los métodos convencionales.

Utilización del sistema de recombinación *in vivo* en librerías.

Una vez implementado el sistema con la PRAI es importante conocer los alcances del mismo. Una forma es evaluar su capacidad de combinar variantes benéficas poco representadas en una biblioteca. Para esto se utilizó el sistema de dos mutantes deletéreas, situadas cada una en una de las mitades y su combinación, anteriormente explicado. Al hacer librerías de la doble mutante con diferentes concentraciones de las mutantes sencillas, recombinarlas y seleccionar la variante activa sin mutaciones, se puede analizar la capacidad del sistema de recombinación *in vivo* de combinar variantes poco representadas en una biblioteca. En la primera biblioteca creada, cuando las mutantes sencillas representan un milésimo de la librería, se obtuvieron fácilmente las variantes activas de PRAI y el proceso de selección donde se enriqueció la variante activa fue tal como se

esperaba. En la segunda librería, donde las mutantes sencillas están diluidas 1/100,000, también fue posible obtener variantes activas; sin embargo, su selección fue más difícil por la presencia de falsos positivos. El proceso de selección en este caso no fue como se esperaba, ya que se observó saturación del medio tiempo antes de la saturación observada con la primera librería. El análisis de estas células reveló que la mayoría eran falsas positivas. A pesar de esto, entre las falsas positivas fue posible rescatar positivas, transformando plásmidos obtenidos de un cultivo de medio mínimo líquido preenriquecido de la cepa de selección F'. La cepa original permite discernir, debido a que en ella no se dan falsos positivos. Las explicaciones más plausibles de que sólo en la segunda librería se hayan enriquecido con tal magnitud los falsos positivos son un lavado deficiente de las células antes de colocarlas en el medio de selección o alguna diferencia en la composición del medio de selección usado. Sin embargo, otra posibilidad sería una contaminación, la cual se eliminaría fácilmente rescatando los fagémidos y retransformando en células frescas. Es de esperar que en la primera librería las variantes activas fueran mucho más que en la segunda, esto representaría una competencia mucho mayor lo que resultaría en el enriquecimiento mayoritario de las positivas.

Por otro lado, existe la posibilidad de que la recuperación de las variantes activas se deba a cambios puntuales que reviertan las mutaciones deletéreas y no al proceso de recombinación de las mutantes sencillas. Esto es más difícil, ya que para revertir las mutaciones deletéreas no sólo es necesario un cambio en la posición indicada, sino un cambio que dé como resultado el aminoácido catalítico. Lo anterior no es posible con la tasa de mutación puntual observada en el sistema, que es nula o al menos imperceptible con el número de variantes analizadas. Además, en el caso de la primera mitad, existen otros marcadores silenciosos que nos permiten saber la procedencia de la mitad. La doble mutante tiene una mutación puntual silenciosa que elimina al sitio XhoI de la primera mitad de PRAI y la mutante sencilla tiene una mutación silenciosa (G por A). De esta manera, si la reversión de la mutación deletérea de la primera mitad se dió por mutación puntual, existiría la mutación que elimina el sitio XhoI de PRAI y no existiría la mutación silenciosa de G por A de la mutante sencilla. Al no observarse lo anterior, sabemos que al menos la primera mitad proviene de la mutante sencilla y no de la reversión de la mutación deletérea a través de cambios puntuales.

Para comparar el sistema con la metodología más usada para combinar variantes benéficas, el *gene shuffling*, se probó esta herramienta con la biblioteca que tiene las mutantes sencillas más diluidas (10^{-5}). Antes de hacer esto fue necesario hacer un control positivo que indicara que el *shuffling* se estaba efectuando de manera correcta. Esto se hizo con una biblioteca que tiene las mutantes sencillas poco diluidas (1/100). Al encontrar variantes activas después de barajar la librería anterior, a pesar de que el banco generado no fue muy grande ($\sim 1 \times 10^5$), se prosiguió a barajar la biblioteca con las mutantes sencillas en proporción 10^{-5} . Contrariamente a lo esperado, sí se obtuvieron colonias, aunque en mucho menor número que con la otra librería. Los resultados de crecimiento anteriores son más confiables, ya que se utilizó directamente la cepa JMB9F^r que no da falsos positivos - independientemente de esto se secuenciaron los plásmidos obtenidos de las colonias. Estos resultados indican que los límites del *gene shuffling* están por debajo de los probados, aunque a diferencia de lo que se realizó con la recombinación *in vivo* no se hicieron controles negativos que descartaran contaminación. Estimando la probabilidad de que ocurran los últimos resultados, es de esperarse que estos se deban a fenómenos de contaminación.

Con ambas metodologías fue posible combinar variantes que están presentes en un cienmilésimo de la población, sin embargo, los verdaderos límites no fueron encontrados. En ambos casos sería necesario recombinar y barajar librerías con variantes en menores proporciones. Además, es esencial repetir estos experimentos y realizar, en ambos casos, controles negativos rigurosos.

Independientemente, también es importante resaltar la sencillez del método de recombinación en comparación con el de *gene shuffling*. En el primero, una vez que se clona la librería primaria en el vector de recombinación no es necesario realizar ninguna otra clonación, se pueden realizar eventos de recombinación sucesivos sin clonaciones intermedias. Además una vez que la biblioteca se ha rescatado en fagos las transformaciones siguientes se llevan a cabo por infección, que es un proceso más simple y mucho más eficiente que la electroporación. Por el contrario, cada vez que se realiza un evento de *gene shuffling* es necesario reclonar la librería y retransformarla por

electroporación. Esto último, además de hacer el método más complejo, disminuye considerablemente el tamaño de la librería, hasta alrededor de 10^7 cuando el plásmido con que se transforma no está superenrollado.

ESTA TESIS NO SALE
DE LA BIBLIOTECA

Conclusión

Hasta ahora, los sistemas de recombinación *in vivo* dirigidos a crear variabilidad sólo habían sido aplicados a librerías de anticuerpos. En este trabajo, se implementa con éxito uno de estos sistemas de recombinación *in vivo* para la proteína fosforribosil antranilato isomerasa de *E. coli*, que es una enzima citosólica, representante de la amplia familia de los barriles α/β . La inserción del sitio necesario para la recombinación en la proteína no tiene un efecto importante sobre su actividad y permite la combinación de diferentes variantes preexistentes hacia la creación de nuevas variantes y, por lo tanto, de librerías más diversas. Sin embargo, al intentar implementar el sistema en dos variantes de otro miembro de los barriles α/β , la trifosfato isomerasa de *T. brucei*, se observa que la inserción del sitio de recombinación inactiva ambas variantes de la enzima. Aunque no se descarta la capacidad de recombinar, ni la utilidad del sistema, lo anterior representa una inconveniente sobrecarga de inicio, poniendo en duda los beneficios netos del sistema. Aún es difícil extrapolar estos resultados a otros barriles α/β , pero hay que tener en cuenta la mencionada desventaja. Es más probable que el sistema pueda ser implementado con mayor facilidad con otras proteínas barriles α/β similares a la PRAI o posiblemente, de manera más general, barriles α/β que presenten características ancestrales de duplicación y fusión de un medio barril original como HisA y HisF. Además, el encontrar nuevas secuencias que sirvan como péptidos conectores, tener más información acerca de los sitios óptimos de corte en las diferentes enzimas y la utilización de sistemas de autoedición (*autosplicing*) facilitará la generalización del método.

Por otro lado, la eficiencia del sistema, a pesar de ser variable, demuestra que la recombinación *in vivo* puede ser una herramienta en la generación de grandes bibliotecas, de más de 10^{10} variantes, y por lo tanto, un método importante en la generación de variabilidad para la evolución dirigida de proteínas. Su capacidad de combinar variantes benéficas poco representadas en una biblioteca no es totalmente conocida, pero se sabe que es posible al menos combinar variables que se encuentran tan diluidas como 1/100,000. Este trabajo es el primer acercamiento hacia el desarrollo de un método que promete ser útil en la búsqueda de nuevas propiedades enzimáticas.

Por último, y en un ámbito diferente, vale la pena mencionar que el sistema puede ser atractivo para trabajos encaminados a explicar la evolución del grupo de los barriles α/β . Con el fin de probar experimentalmente la hipótesis del surgimiento de los barriles α/β a partir de la duplicación y fusión de un medio barril ancestral, se podrían generar librerías de variantes de

diferentes barriles relacionados con características ancestrales y con el sistema de recombinación *in vivo* combinarlas en busca de barriles mixtos estructuralmente estables.

Bibliografía

- Alberts B., Bray D., Lewis J., Raff M., Roberts K. and Watson J. 1983. *Molecular Biology of the Cell*. Garland Publishing, Inc. EUA. Pags. 1146
- Altamirano M. M., Blackburn J. M., Aguayo C. and Fersht A. R. 2000. Directed evolution of new catalytic activity using the $\alpha\beta$ -barrel scaffold. *Nature* 403: 617-622
- Banner D. W., Bloomer A. C., Petsko G. A., Philips D. C., Pogson C. I., Wilson, I. A., Corran P. H., Furth A. J., Milman J. D., Offord R. E., Priddle J. D. and Waley S. G. 1975. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature* 255: 609-614
- Bertolaet B. L. and Knowles J. R. 1995. Complementation of fragments of triosephosphate isomerase defined by exon boundaries. *Biochemistry* 34: 5736-5743
- Borchert T. V., Abagyan R., Jaenicke R. and Wierenga R. K. 1994. Design, creation, and characterization of stable, monomeric triosephosphate isomerase. *Proc. Nat. Acad. Sci. USA* 91: 1515-1518
- Borchert T. V., Abagyan R., Radha Kishan K. V., Zeelen J. Ph. and Wierenga R. K. 1993. The crystal structure of an engineered monomeric triosephosphate isomerase, monoTIM: the correct modelling of an eighth-residue loop. *Structure* 1: 205-213
- Bork P., Gellerich J., Groth H., Hooft R. and Martin F. 1995. Divergent evolution of a β/α -barrel subclass: detection of numerous phosphate-binding sites by motif search. *Protein Science* 4: 268-274
- Brilli M., Liò P., Lazcano A. and Fani R. 2002. Evolution of TIM barrel: Multiple Gene elongation events in HisA. Memories of the 10th ISSOL Meeting, Oaxaca, México, Jun 2002
- Buchholz F. and Stewart A. F. 2001. Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nature Biotechnology* 19: 1047-1052
- Chen K., and Arnold F. H. 1993. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc. Natl. Acad. Sci. USA* 90: 5618-22

- Coco W. M., Levinson W. E., Crist M. J., Héctor H. J., Darzins A., Pienkos P. T., Squires C. H. and Monticello D. J. 2001. DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nature Biotechnology* 19: 354-9
- Cohen N., Abramov S., Dror Y. and Freeman A. 2001. *In vitro* enzyme evolution: the screening challenge of isolating the one in a million. *Trends in Biotechnology* 19: 507 – 510
- Copley R. and Bork P. 2000. Homology among α/β barrels: implications for the evolution of metabolic pathway. *Journal Molecular Biology* 303: 627-440
- Cramer A., Raillard S. A., Bermudez E. and Stemmer W. P. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 391: 288-91
- Creighton T. E. 1993. Proteins, Structures and Molecular Properties. 2ª edición. W. H. Freeman and Company. EUA. Pags. 507
- Davies J. and Riechmann L. 1995. An antibody VH domain with a *lox-Cre* site integrated into its coding region: bacterial recombination within a single polypeptide chain. *FEBS Letters* 377: 92-96
- Eberhard M., Tsai-Pflugfelder M., Bolewska K., Hommel U. and Kirschner K. 1995. Indoeleglycerol phosphate sythase - phosphoribosyl anthranilate isomerase: comparison of the bifunctional enzyme from *Escherichia coli* with engineered monofunctional domains. *Biochemistry* 34: 5419-5428
- Eder J. and Kirschner K. 1992. Stable substructures of eightfold α/β -barrel proteins: Fragment complementation of phosphoribosylanthranilate isomerase. *Biochemistry* 31: 3617-3625
- Eigen M. 1984. The origin and evolution of life at the molecular level. *Adv. Chem. Phys.* 55: 119-137
- Fani R. Lio P., Chiarelli I. and Bazzicupo M. 1994. The evolution of the histidine biosynthetic genes in prokaryotes: a common ancestor for hisA and hisF genes. *J Molecular Evolution* 38:489-495
- Farber G. and Petsko G. 1990. The evolution of α/β barrel enzymes. *Trends Biochemistry Science* 15: 228-234

- Fersht A. 1985. Enzyme Structure and Mechanism. 2ª edición. W. H. Freeman and Company. EUA. Capítulo 14, pags 369 – 388
- Fisch I., Kontermann R. E., Finer R., Hartley O., Soler-Gonzalez A. S., Griffiths A. D. and Winter G. 1996. A strategy of exon shuffling for making large peptide repertoires displayed on filamentous bacteriophage. *Proc. Natl. Acad. Sci. USA* 93: 7761-7766
- Fuentes G. P. 2001. Fragmentación del gene de la enzima fosforribosil antranilato isomerasa (PRAI) de E. coli y regeneración de su actividad mediante complementación de fragmentos. Tesis para obtener el grado de maestría en ciencia bioquímicas. Instituto de biotecnología, UNAM
- Geoffroy E., Sodoyer R. and Aujame L. 1994. A new phage display system to construct multicombinatorial libraries of very large antibody repertoires. *Gene* 151: 109-113
- Gerlt J.A. and Babbitt P. 2001. Baçrels in pieces? *Nature structural biology* 8: 5-7
- Gopaul D. N., Guo F. and Van Duynne G. D. 1998. Structure of the Holliday junction intermediate in Cre-loxP site-specific recombination. *EMBO Journal* 17: 4175-4187
- Griffiths A. D. Williams S. C., Hartley O., Tomlinson I. M., Waterhouse P., Crosby W. L., Kontermann R. E., Jones P. T., Low N. M., Allison T. J., et al. 1994. Isolation of high affinity human antibodies directly from large synthetic repertoires. *The EMBO Journal* 13: 3245-3260
- Guo F., Gopaul D. N. and Van Duynne G. D. 1997. Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature* 389: 40-46
- Henn-Sax M., Thoma R., Schmidt S., Hennig M., Kirschner K. and Sterner R. 2002. Two ($\beta\alpha$)₈-barrel enzymes of histidine and tryptophan biosynthesis have similar reaction mechanisms and common strategies for protecting their labile substrates. *Biochemistry* 41: 12032-42
- Höcker B., Beisman-Driemeyer S., Hettwer S. Lustig A. and Sterner R. 2001^a. Dissection of α/β 8-barrel enzyme into two folded halves. *Nature Structural Biology* 8: 32-36
- Höcker B., Jürgens C., Wilimanns M. and Sterner R. 2001^b. Stability, catalytic versatility and evolution of the ($\beta\alpha$)₈-barrel fold. *Current Opinion in Biotechnology* 12: 376-381

- Höcker B., Schmidt S. and Sterner R. 2002. A common evolutionary origin of two elementary enzyme folds. *FEBS Letters* 510: 133-135
- Hommel U., Eberhard M. and Kirschner K. 1995. Phosphoribosyl anthranilate isomerase catalyzes a reversible Amadori reaction. *Biochemistry* 34: 5429-5439.
- Jasanoff A., Davis B. y Fersht A. R. 1994. Detection of an intermediate in the folding of the $(\beta\alpha)_8$ -barrel N-(5'-Phosphoribosyl) anthranilate isomerase from *Escherichia coli*. *Biochemistry* 33: 6350-6355
- Jermutus L. and Pelletier J. 2001. Creating and evaluating protein diversity. *Current Opinion in Biotechnology* 12: 331 – 333
- Jürgens C., Strom A., Wegener D., Hettwer S., Wilmanns M. and Sterner R. 2000. Directed evolution of a $(\beta\alpha)_8$ -barrel enzyme to catalyze related reactions in two different metabolic pathways. *Proc. Nat. Acad. Sci. USA* 97: 9925-9930
- Kauffman S. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press, NY
- Knowles J. R. 1991. Enzyme catalysis: not different, just better. *Nature* 350: 121-124
- Landy A. 1989. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annual Review Biochemistry* 58: 913-949
- Lang D., Thoma R., Henn-Sax M., Sterner R. and Wilmanns M. 2000. Structural evidence for evolution of the α/β barrel scaffold by gene duplication and fusion. *Science* 289: 1546-1550
- Lodish H., Berk A., Zipursky S. L., Matsudaira P., Baltimore D. and Darnell J. 2000. *Molecular Cell Biology*. 4ª edición. W. H. Freeman and Company. EUA. Pags. 1084
- Luger K., Hommel U., Herold M., Hofsteenge J. and Kirschner K. 1989. Correct folding of circularly permuted variants of a $\beta\alpha$ barrel enzyme in vivo. *Science* 243: 206-210
- Luger K., Szadkowski H. and Kirschner K. 1990. An 8-fold $\beta\alpha$ barrel protein with redundant folding possibilities. *Protein Engineering* 3(4): 249-258
- Mainfroid V., Goraj K., Rentier-Delrue F., Houbrechts A., Loiseau A., Gohimont A. C., Noble M. E., Borchert T. V., Wierenga R. K. and Martial J. A. 1993. Replacing the $(\beta\alpha)$ -unit of E. coli TIM with its chicken homologue leads to a stable and active hybrid enzyme. *Protein Engineering* 6(8): 893-900

- Miles E.W., Yutani K. y Ogasahara K. 1982. Guanidine hydrochloride induced unfolding of the α subunit of tryptophan synthase and of the two α proteolytic fragments: evidence for stepwise unfolding of the two α domains. *Biochemistry* 21: 2586-2592
- Mills D. R., Peterson R. L. and Spiegelman S. 1967. An extracellular darwinian experiment with a self-duplicating nucleic acid molecule. *Proc Natl Acad Sci USA* 58: 217-24
- Nagano N., Orengo C. A. and Thornton J. M. 2002. One fold with many functions: the evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *Journal of Molecular Biology* 321: 741-765
- Ostermeier M., Shim J. H. and Benkovic S. J. 1999. A combinatorial approach to hybrid enzymes independent of DNA homology. *Nature Biotechnology* 17: 1205-9
- Priestle J. P., Grüter M. G., White J. L., Vincent M. G., Kania M., Wilson E., Jardetzky T. S., Kirschner K. and Jansonius J. N. 1987. Three-dimensional structure of the bifunctional enzyme N-(5'phosphorribosyl)-anthranilate isomerase-indole-3-glycerol-phosphate synthase from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 84: 5690-5694
- Reardon D. and Farber G. 1995. The structure and evolution of α/β barrel proteins. *FASEB* 9: 497-503
- Riechmann L. and Winter G. 2000. Novel folded protein domains generated by combinatorial shuffling of polypeptide segments. *Proc. Nat. Acad. Sci. USA* 97: 10068-10073
- Riechmann L. and Winter G. 2000. Novel folded protein domains generated by combinatorial shuffling of polypeptide segments. *Proc. Nat. Acad. Sci. USA* 97: 10068-10073
- Saab-Rincón G., Rivelino Juárez V., Osuna J., Sánchez F. and Soberón X. 2001. Different strategies to recover the activity of a monomeric triosephosphate isomerase by directed evolution. *Protein Engineering* 14: 149-155
- Sblattero D. and Bradbury A. 2000. Exploiting recombination in single bacteria to make large phage antibody libraries. *Nature Biotechnology* 18: 75-80

- Sblattero D., Lou J., Marzari R. and Bradbury A. 2001. In vivo recombination as a tool to generate molecular diversity in phage antibody libraries. *Reviews in Molecular Biotechnology* 74: 303-315
- Schliebs W., Thanki N., Eritja R. and Wierenga R. K. 1996. Active site properties of monomeric triosephosphate isomerase (monoTIM) as deduced from mutational and structural studies. *Protein Science* 5: 229-239
- Schmidt-Dannert C. 2001. Directed evolution of single proteins, metabolic pathways, and viruses. *Biochemistry* 40:13125 – 13136
- Sieber V., Martinez C. A. and Arnold F. H. 2001. Libraries of hybrid proteins from distantly related sequences. *Nature Biotechnology* 19: 456-60
- Siegel R. W., Jain R. and Bradbury A. 2001. Using an in vivo phagemid system to identify non-compatible loxP sequences. *FEBS Letters* 499: 147-153
- Stemmer W. P. 1994^a. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* 370: 389-91
- Stemmer W. P. C. 1994. DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* 91: 10747-10751
- Taylor S. V., Kast P. and Hilvert D. 2001. Investigating and engineering enzymes by genetic selection. *Angew. Chem. Int. Ed.* 40: 3310 – 3335
- Thoma R., Schwander M., Liebel., Kirschner K. and Sterner R. 1998. A histidine gene cluster of the hyperthermophile *Thermotoga maritima*: sequence analysis and evolutionary significance. *Extremophiles* 2: 379-389
- Tsurishita N., Fu H. and Warren C. 1996. Phage display vectors for in vivo recombination of immunoglobulin heavy and light chain genes to make large combinatorial libraries. *Gene* 172: 59-63
- Urfer R. and Kirschner K. 1992. The importance of surface loops for stabilizing an eightfold $\beta\alpha$ barrel protein. *Protein Science* 1: 31-45
- Waterhouse P., Griffiths A. D., Johnson K. S. and Winter G. 1993. Combinatorial infection and *in vivo* recombination: a strategy for making large phage antibody repertoires. *Nucleic Acid Research* 21: 2265-2266

- Wierenga R. K. 2001. The TIM-barrel fold: a versatile framework for efficient enzymes. *FEBS Letters* 492: 193-198
- Wierenga R. K., Noble M. E. M. and Davenport R. C. 1992. Comparison of the refined crystal structures of liganded and unliganded chicken, yeast and trypanosomal triosephosphate isomerase. *Journal Molecular Biology* 224: 1115-1126
- Wilmanns M., Priestle J. P., Niermann T. and Jansonius J. N. 1992. Three-dimensional structure of the phosphoribosylanthranilate isomerase: indoleglycerolphosphate synthase from *Escherichia coli* refined at 2.0 Å resolution. *Journal Molecular Biology* 223: 477-507
- Wilson Miles E. and Davies D. R. 2000. On the ancestry of barrels. *Science* 289: 1490
- Zahra D. G., Vancov T., Dunn J. M., Hawkins N. J. y Ward R. L. 1999. Selectable in-vivo recombination to increase antibody library size – an improved phage display vector system. *Gene* 227: 49-54
- Zitzewitz J. A., Gualfetti P.J., Perkons I.A., Wasta S.A. and Matthews C. R. 1999. Identifying the structural boundaries of independent folding domains in the α subunit of tryptophan synthase, a β/α barrel protein. *Protein Science* 8: 1200-1209