



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

CARACTERIZACIÓN DE PROTEÍNAS QUIMÉRICAS:
COMPLEMENTACIONES DE UN BARRIL TIM

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

B I Ó L O G O

P R E S E N T A:

HAVEN ARTURO LÓPEZ SÁNCHEZ



TUTOR: DR. LORENZO PATRICK SEGOVIA FORCELLA

2008



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.

INDICE

I.- INTRODUCCIÓN	5
1.1 LOS AMINOÁCIDOS	5
1.2 NIVELES DE ESTRUCTURA DE LAS PROTEÍNAS	7
1.3 ESTRUCTURA SECUNDARIA	8
1.4 DOMINIOS	12
1.5 PLEGAMIENTO DE LAS PROTEÍNAS	13
1.6 EVOLUCION DE LOS DOMINIOS	16
1.7 FUERZAS EVOLUTIVAS	18
1.8 TÉCNICAS ESPECTROSCÓPICAS PARA EL ESTUDIO DE PROTEÍNAS	20
II.- ANTECEDENTES	24
III.- OBJETIVOS	26
IV.- MATERIALES Y MÉTODOS	27
4.1 ANÁLISIS BIOINFORMÁTICO	27
4.2 CMI DEL BANCO H	28
4.3 CURVAS DE CRECIMIENTO	28
4.4 EXPRESIÓN DE LAS QUIMERAS FUSIONADAS A CAT	29
4.5 ENSAYO DE BRADFORD	29
4.6 SDS-PAGE	29
4.7 WESTERN BLOT Anti-CAT	29
4.8 RESTRICCIÓN DE CAT	30
4.9 EXPRESIÓN DE LAS QUIMERAS SIN CAT Y EXTRACCIÓN DE CUERPOS DE INCLUSIÓN	31
4.10 PURIFICACIÓN DE LAS QUIMERAS	31
4.11 RENATURALIZACIÓN (replegamiento) Y DIÁLISIS	31

4.12 DICROÍSMO CIRCULAR	32
4.13 FLUORESCENCIA	33
4.14 ANS	33
V.- RESULTADOS	33
5.1 ANÁLISIS BIOINFORMÁTICO	33
5.2 CONCENTRACIONES MÍNIMAS INHIBITORIAS (CMI _s)	34
5.3 CURVAS DE CRECIMIENTO	35
5.4 EXPRESIÓN DE QUIMERAS FUSIONADAS A CAT	36
5.5 SDS-PAGE	36
5.6 WESTERN BLOT Anti-CAT	38
5.7 RESTRICCIÓN DE CAT	40
5.8 EXPRESIÓN DE LAS QUIMERAS SIN CAT Y EXTRACCIÓN DE CUERPOS DE INCLUSIÓN	40
5.9 PURIFICACIÓN EN COLUMNA DE AFINIDAD	42
5.10 RENATURALIZACIÓN Y DIÁLISIS	43
5.11 DICROÍSMO CIRCULAR	45
5.12 FLUORESCENCIA	46
5.13 ANS	49
VI.- DISCUSIÓN Y CONCLUSIONES	50
APÉNDICE 1	55
VII.- LITERATURA CITADA	58

ABREVIATURAS

aa: aminoácidos.

C α : carbono α .

D. C.: dicroísmo circular.

PRAI: fosforibosil antranilato isomerasa.

IGPS: indolglicerolfosfato sintasa.

ePRAI: *E. coli* fosforibosil antranilato isomerasa.

ADN: ácido desoxirribonucleico.

CAT: cloramfenicol acetil transferasa.

CMI: concentración mínima inhibitoria.

O/N: *overnight*; toda la noche.

LB: medio Luria-Bertani.

OD: *optical density*; densidad óptica.

rpm: revoluciones por minuto.

PBS: *phosphate buffered saline*; *buffer* de solución salina y fosfatos.

PMSF: *phenylmethanesulphonylfluoride*, (Inhibidor de proteasas).

SDS: *sodium dodecyl sulfate*; duodecil sulfato sódico.

BSA: *bovine serum albumin*; albumina de suero bovino.

TBST: *Tris-Buffered Saline Tween-20*; *buffer* de solución salina, tris y tween 20.

TBS: *Tris-Buffered Saline*; *buffer* de solución salina y tris.

BCIP: *5-Bromo-4-chloro-3-indolyl phosphate*; 5-bromo-4-cloro-3-indol fosfato.

NBT: *nitro blue tetrazolium*.

GuHCl; GdnCl: cloruro de guanidine.

ANS: anilino-8-naftaleno sulfonato.

Cm: cloramfenicol.

IPTG: *Isopropyl-thio-2-D-galactopyranoside*; isopropil-tio-2-D-galactopiranosido.

I.- INTRODUCCIÓN

El nombre de las proteínas proviene del griego *prōtos* que significa “primero” y deriva de razones históricas en el estudio de la biología y la química. La abundancia de estas moléculas y la gran gama de actividades que llevan a cabo dentro de una célula es tal que se llegó a creer que eran las responsables directas de casi todos los fenómenos biológicos relevantes, incluso de la herencia, pero desde mediados del siglo pasado con el descubrimiento de las funciones biológicas del ADN se supo que no era así, aunque sin ellas ningún proceso celular sería posible. En términos generales, la importancia de las proteínas radica en que todo lo que sucede en una célula requiere de la intervención de éstas. Tal es la importancia que de hecho son los productos finales de la información genética en la gran mayoría de los casos. En una célula típica existen miles de proteínas diferentes, cada una codificada por un gen y cuya función es específica [1].

1.1 LOS AMINOÁCIDOS

Todas las proteínas en todas las especies de seres vivos están constituidas por los mismos aminoácidos (*aa*), un conjunto de 20 de ellos, que son llamados unidades o residuos. A pesar de contar con un conjunto limitado de aminoácidos la gran diversidad de proteínas encontrada en la naturaleza se debe a que éstas difieren en el número y secuencia de aminoácidos y pueden ordenarse en un número casi infinito de secuencias lo que al final determina su estructura y función. La secuencia de aminoácidos de una proteína, es decir la unión covalente de éstos a través de enlaces peptídicos, es llamada estructura primaria y es la propiedad que distingue más claramente a una proteína de otra [1].

Los 20 aminoácidos encontrados en las proteínas tienen todos un grupo carboxilo y un grupo amino unidos al mismo átomo de carbono (carbono α). Difieren

entre ellos en sus cadenas laterales (grupos R) que varían en tamaño, estructura y carga eléctrica lo que les confiere distintas características (Figura 1.1) que a su vez determinan su solubilidad, hidrofobicidad, polaridad (entre otras características) y el tipo de interacciones que pueden llevar a cabo con otras sustancias y entre ellos mismos. Por convención se denomina a los 20 aminoácidos de las proteínas como aminoácidos estándar, primarios o normales para distinguiriros de aquellos que sufren modificaciones después de la síntesis de las proteínas y de la misma manera para diferenciarlos de aminoácidos presentes en organismos vivos pero no en proteínas [1], [2].

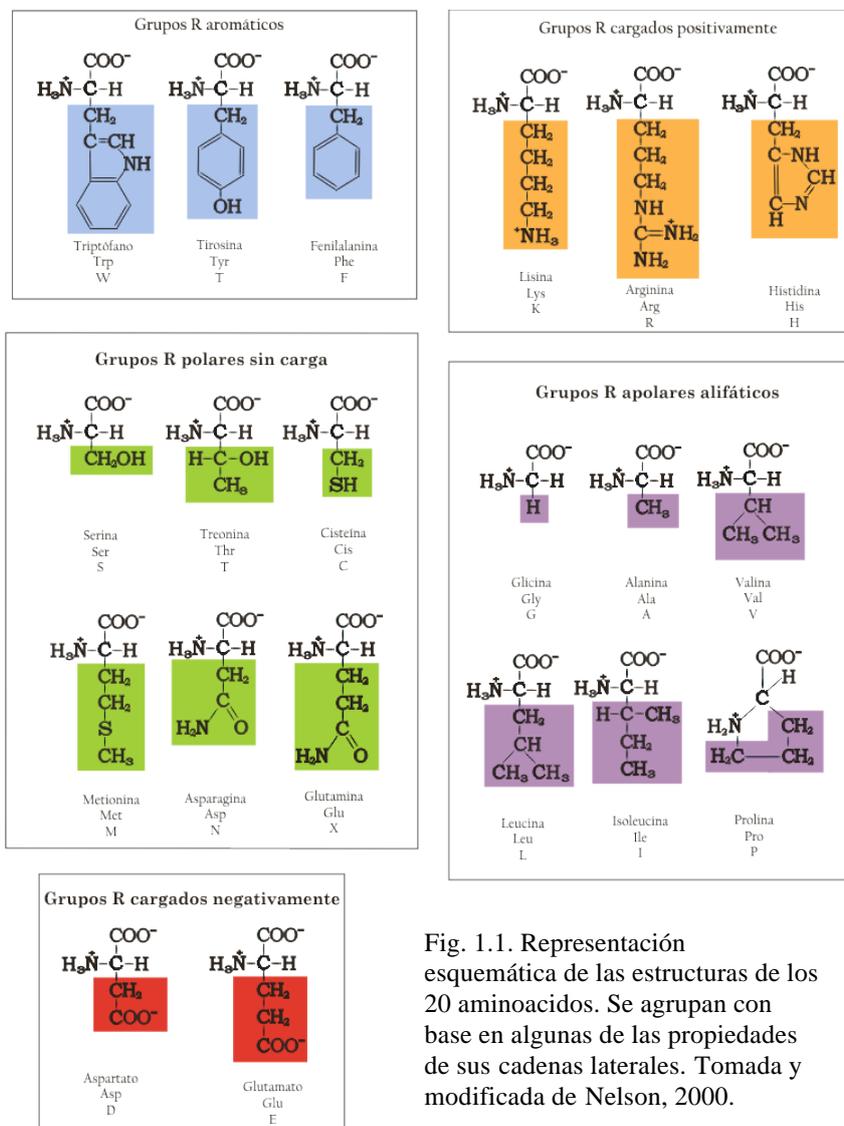


Fig. 1.1. Representación esquemática de las estructuras de los 20 aminoácidos. Se agrupan con base en algunas de las propiedades de sus cadenas laterales. Tomada y modificada de Nelson, 2000.

En todos los aminoácidos estándar, excepto en la glicina, el carbono α es asimétrico, esto quiere decir que está unido a cuatro grupos sustituyentes diferentes: un grupo carboxilo, un grupo amino, un hidrógeno y un grupo R que como se estableció anteriormente marca la diferencia. Gracias a esta característica se dice que el carbono α es un centro quiral, y esto permite que la configuración de los aminoácidos se dé en dos maneras distintas en el espacio, las cuales son imágenes especulares no superponibles. Estas dos formas o configuraciones se conocen como enantiómeros o esteroisómeros. Al ser moléculas quirales son ópticamente activas, es decir, pueden hacer girar el plano de luz polarizada con una dirección de rotación diferente para cada uno de los esteroisómeros [1], [2].

1.2 NIVELES DE ESTRUCTURA DE LAS PROTEÍNAS

De vuelta a las proteínas, se puede considerar que hay cuatro niveles de estructuración: a) Estructura primaria: es la secuencia de aminoácidos formada por enlaces covalentes entre ellos, llamados enlaces peptídicos. A este nivel no se especifica la disposición espacial de los aminoácidos [1], [2].

b) Estructura secundaria: los aminoácidos cercanos se asocian unos con otros para formar elementos de estructura secundaria. Éstos elementos son formados usualmente a través de interacciones entre los átomos de los aminoácidos que corresponden al esqueleto y no por interacciones de las cadenas laterales. Existen pocas clases básicas de estructuras secundarias de las cuales la hélice α y la hebra β son las más importantes. Al asociarse los elementos de estructura secundaria de una manera definida generan la estructura terciaria [3].

c) Estructura terciaria: puede definirse como la descripción de las relaciones espaciales entre todos los aminoácidos de un polipéptido o como su estructura tridimensional completa [2].

d) Estructura cuaternaria: este nivel ocurre en las proteínas que poseen varias cadenas polipeptídicas y se refiere a la relación espacial de los polipéptidos o subunidades en la proteína global [1], [2].

1.3 ESTRUCTURA SECUNDARIA

Al considerar la estructura secundaria resulta ser de utilidad clasificar a las proteínas en dos grandes grupos: fibrosas y globulares. En las proteínas fibrosas las cadenas polipeptídicas se ordenan en largos filamentos u hojas y suelen estar formadas por un solo tipo de estructura secundaria; las globulares se pliegan en forma esférica o globular siendo éstas más complejas al contener varios tipos de estructura secundaria (Figura 1.2).

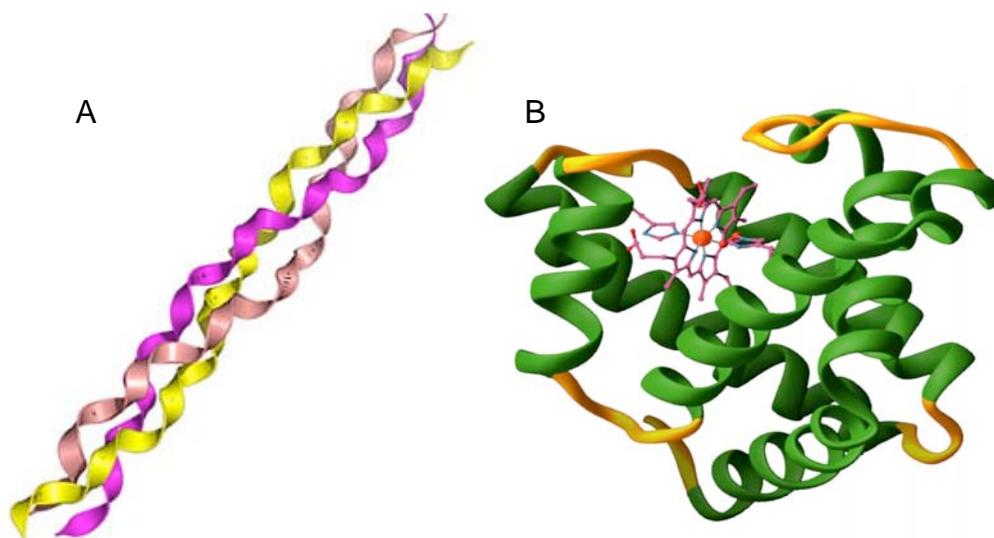


Fig. 1.2. Representaciones tipo caricatura de la estructura de: A) colágeno y B) mioglobina de esperma de ballena. Estas dos estructuras representan las formas fibrosa y globular de las proteínas respectivamente. A) Tomada de www.3dchem.com/imagesofmolecules; B) Tomada de Voet, 2004.

El primer tipo de elemento de estructura secundaria es una de las estructuras más simples que una cadena polipeptídica podría asumir tomando en cuenta la rigidez del enlace peptídico, es llamada la hélice α . En esta estructura el esqueleto polipeptídico, es decir los átomos unidos covalentemente en la secuencia de enlaces

peptídicos de la cadena ($C'-N-C\alpha-C'$), se encuentra enrollado compactamente a lo largo de un eje longitudinal y los grupos R de los aminoácidos sobresalen del esqueleto helicoidal. La unidad repetitiva es el giro de hélice que ocupa 0.56nm del eje longitudinal y cada giro incluye 3.6 aminoácidos. Cada vuelta de la hélice se mantiene unida a las vueltas adyacentes a través de enlaces hidrógeno. Los enlaces hidrógeno se forman entre el átomo de oxígeno de un residuo y un átomo de hidrógeno del residuo que se encuentra cuatro posiciones adelante en la cadena, los enlaces totales en una cadena sumados proporcionan una estabilidad considerable (Figura 1.3). La secuencia de aminoácidos es importante porque afecta la estabilidad de la hélice ya que las interacciones adicionales que se producen entre las cadenas laterales de los aminoácidos pueden estabilizar o desestabilizar esta estructura [1], [2], [3], [4], [5].

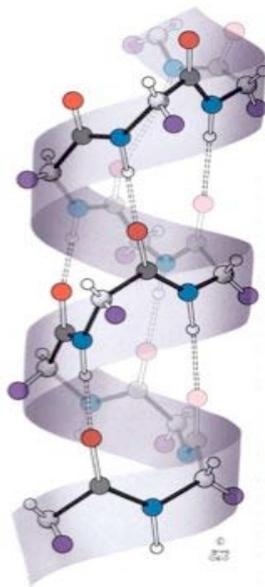


Fig. 1.3. Hélice α dextrógira. Los enlaces de hidrógeno entre los grupos N-H y los grupos C=O, que se hallan cuatro residuos atrás, están indicados por las líneas discontinuas. Tomada de Voet, 2004.

La otra estructura secundaria muy común en proteínas es la conformación β , en ella el esqueleto de la proteína se encuentra extendido en zig-zag. A menudo se encuentran

diferentes secciones del esqueleto dispuestas de manera adyacente formando una estructura que asemeja una serie de pliegues, a esta estructura se le llama hoja β plegada. En la conformación β los enlaces hidrógeno que la estabilizan pueden ser intracatenarios o intercatenarios, y se forman entre el grupo carbonilo (C=O) de una hebra y el grupo amino (NH) de las hebras adyacentes. Los grupos R de aminoácidos adyacentes se relacionan por una rotación de 180° y por ello sobresalen de la estructura en zig-zag en direcciones opuestas (Figura 1.4). Las cadenas polipeptídicas de una hoja β plegada pueden ser paralelas (misma orientación amino-carboxilo) o antiparalelas (orientaciones opuestas). Al igual que en la hélice α la composición de aminoácidos puede afectar a la estructura, cuando dos o más hojas plegadas se encuentran empaquetadas densamente en una proteína los grupos R de los aminoácidos de la superficie de contacto deben ser relativamente pequeños [1], [2] [3], [4], [5].

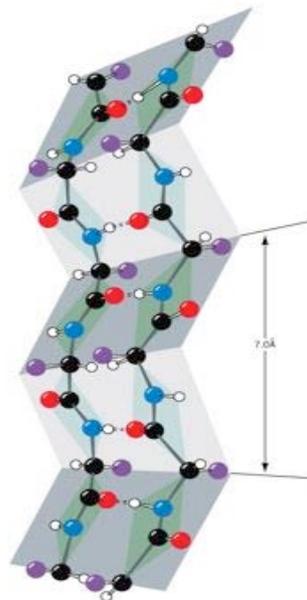


Fig. 1.4. Hoja β antiparalela de dos hebras, resalta su apariencia plegada. las líneas discontinuas indican los enlaces de hidrógeno. Los grupos R (morado) se distribuyen alternativamente en los lados opuestos de la hoja. Tomada de Voet, 2004.

Las estructuras secundarias regulares, hélices α y hebras β , representan poco menos de la mitad de las proteínas globulares. Los segmentos restantes son estructuras que no se repiten pero esto no quiere decir que sean menos ordenadas, simplemente son más difíciles de describir. Las proteínas globulares están constituidas por tramos de estructura secundaria, aproximadamente rectos, unidos por hebras de polipéptido que cambian abruptamente de dirección llamados giros de torsión β , nombrados así por que con frecuencia conectan conformaciones β antiparalelas. Una definición estricta para éstos giros es: el acercamiento cerrado (<7 Å) de dos átomos carbon α ($C\alpha$) de residuos separados por tres enlaces peptídicos. Estos giros son estabilizados por puentes de hidrógeno en los cuales el residuo aceptor y donador están separados por tres residuos. Aparecen casi siempre en la superficie definiéndola en gran medida [41]. Hay que diferenciar a los giros β de los *loops* o asas, éstos últimos son regiones más grandes, más de cinco residuos, que no están estabilizadas por puentes de hidrógeno y que no necesariamente causan giros y cambios de dirección de la cadena principal tan abruptos como los giros β , incluso éstas estructuras se observan conectando regiones secuenciales de la proteína pero que no se empaican contiguamente [41]. Por otro lado muchas proteínas poseen regiones verdaderamente desordenadas que con frecuencia ondean alrededor cuando se encuentran en disolución debido a que existen pocas fuerzas para mantenerlos en su sitio [2].

Además de los cuatro niveles estructurales antes mencionados se han tenido que definir dos más entre la estructura secundaria y la estructura terciaria, debido a que los límites no son claros. Se dice que un agrupamiento estable de varios elementos de estructura secundaria son una estructura supersecundaria o motivo (*motif*). Generalmente se utiliza este término para describir estructuraciones particularmente estables en muchas proteínas, a veces repetidas en una misma

proteína [6]. La forma más ocurrente es la unidad $\beta\alpha\beta$, o motivo $\beta\alpha\beta$, en la que la conexión transversal entre dos hebras paralelas de conformación β está constituida por una hélice α (Figura 1.5-a). Otro motivo es el β *hairpin* (pasador) que consiste de una hoja β antiparalela formada por segmentos secuenciales de polipéptido conectados por vueltas estrechas (Figura 1.5-b). Por su parte, la unidad $\alpha\alpha$ está formada por dos hélices α antiparalelas empaquetadas una contra otra con sus ejes inclinados de manera que permita una interacción favorable de las cadenas laterales de los residuos en contacto (Figura 1.5-c) [2], [6].

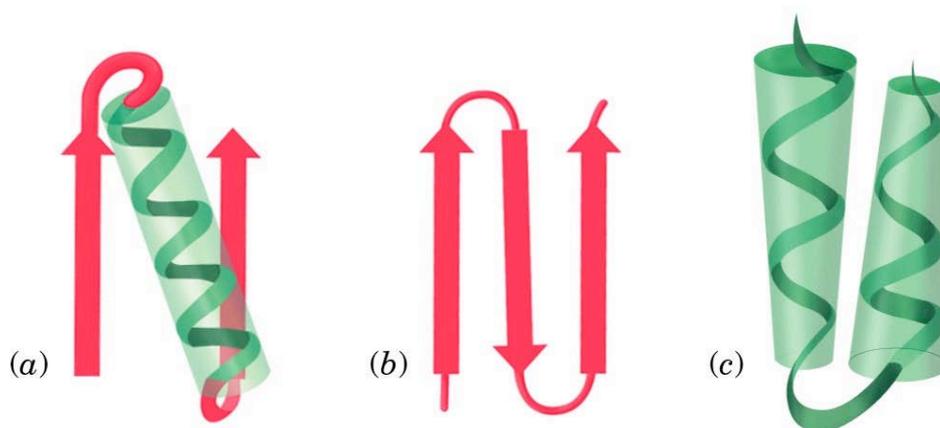


Fig. 1.5. Esquemas de tres tipos de motivos: a) Unidad $\beta\alpha\beta$; b) β *hairpin*; c) Unidad $\alpha\alpha$. Tomada de Voet, 2004.

1.4 DOMINIOS

Un dominio puede ser entendido como una región compacta de entre 40 a 400 aminoácidos (en promedio entre 100 y 200) y constituye una unidad estructural particular en una cadena polipeptídica de gran tamaño. Muchos dominios se pliegan independientemente del resto de la cadena dando lugar a estructuras termodinámicamente estables y así una sola cadena puede contener varios dominios dando como resultado una apariencia multilobular. En algunos casos los dominios tienen funciones específicas y separadas. Los dominios son unidades independientes

desde el punto de vista estructural, cada una con las características de una pequeña proteína [1], [2], [7].

Si bien existen muchas maneras en que los dominios pueden ser categorizados, la manera más simple de hacerlo es clasificándolos como dominios α , dominios β y dominios α/β [2], [8].

Es necesario decir que la definición de dominio puede variar y puede ser considerado por sus capacidades catalíticas como una región proteica la cual no necesariamente cumple con la definición de dominio estructural aquí mencionada.

1.5 PLEGAMIENTO DE LAS PROTEÍNAS

Para llegar a la estructura terciaria la cadena polipeptídica se ordena espacialmente al establecer interacciones residuo-residuo y entre sus residuos y el medio, esto es conocido como el plegamiento de las proteínas. La ruta del plegamiento es muy compleja y aún no se conocen con detalle los principios que dirigen el proceso.

En algunas proteínas se ha podido comprobar que el plegamiento tiene lugar a través de varios intermediarios discretos (bien definidos) y que en algunos de los primeros pasos se produce el plegamiento local de regiones de estructura secundaria. En el “modelo del armazón” (*framework model*) se ha propuesto que el plegamiento se da de manera jerárquica, en este supuesto las estructuras secundarias locales se formarían primero, seguidas por las interacciones de largo alcance y el proceso continuaría hasta completar el plegamiento [9]. En un modelo alternativo, el modelo del colapso hidrofóbico o del glóbulo fundido, el plegamiento se iniciaría con un colapso espontáneo de la cadena polipeptídica que daría lugar a un estado compacto estabilizado por las interacciones hidrofóbicas entre residuos no-polares [9] [10], [11], [15].

Fuera de cualquiera de estos modelos y teorías, es bien conocido que un cierto número de condicionamientos estructurales ayudan a dirigir la interacción entre regiones de estructura secundaria y en consecuencia el plegamiento. Las formas resultantes de estas interacciones son las estructuras supersecundarias ya mencionadas [1], [2], [6], [7]. En fechas recientes se ha acuñado el término *foldon* para describir a diferentes partes de la proteína que pueden plegarse en distintos tiempos casi independientemente, las cuales guían a la proteína a su estructura plegada en vez de a configuraciones desplegadas. Estas unidades se pliegan en un solo paso de manera cooperativa [12].

Por su parte, las interacciones débiles representan la última restricción termodinámica a la interacción de distintas regiones de estructura secundaria. Los grupos R de los aminoácidos se proyectan hacia el exterior de las estructuras α y β pero se ha observado una disposición de las cadenas laterales de los aminoácidos de acuerdo a su polaridad, los residuos no-polares como Val, Met, Ile y Phe se encuentran casi siempre en el interior de una proteína lejos del alcance del disolvente acuoso y esto se debe al establecimiento de interacciones hidrofóbicas responsables de tal distribución. Los residuos polares con carga como Arg, His, Lys, Asp y Glu se hallan situados casi invariablemente en la superficie de la proteína en contacto con el disolvente acuoso. Los grupos polares sin carga Ser, Thr, Asn, Gln, Tyr y Trp aparecen habitualmente en la superficie de la proteína pero con frecuencia aparecen en el interior de la molécula casi siempre unidos hidrofóbicamente a otros residuos, todos ellos ocultos. Es habitual que las proteínas pequeñas mantengan su conformación nativa mediante un cierto número de enlaces covalentes, principalmente puentes disulfuro, pues es más difícil enterrar residuos hidrofóbicos en estructuras pequeñas y en ellas el número potencial de interacciones débiles que puedan colaborar en la estabilización disminuye [2], [5], [13].

No todas las proteínas se pliegan espontáneamente, en una amplia variedad de células se han encontrado proteínas que facilitan el plegamiento de otras proteínas, se denominan acompañantes moleculares (“chaperonas”). Algunas de estas proteínas pueden unirse a cadenas polipeptídicas evitando la agregación inespecífica de cadenas laterales. La mayoría de las ocasiones actúan como guía en el plegamiento de algunos polipéptidos [1], [2], [10], [14], [15].

En este panorama general sobre el plegamiento, que permanece como un fenómeno incierto en la actualidad, debemos agregar que la conformación terciaria o nativa es la forma termodinámicamente más estable que puede asumir la cadena polipeptídica de una proteína globular pero no es rígida. Las proteínas globulares poseen cierto grado de flexibilidad en sus cadenas principales y presentan fluctuaciones internas. Muchas proteínas sufren pequeños cambios conformacionales mientras realizan su función biológica, en muchos casos los cambios están asociados a la unión de un ligando (molécula específica que se une a la proteína) y en otros casos a la unión al sustrato. La importancia del plegamiento y la estructura terciaria de una proteína radica en que su función biológica depende de las dos características anteriores, y como ya se había establecido tanto el plegamiento como la estructura dependen de la secuencia de aminoácidos [2], [13].

A mediados del siglo XX Christian Anfinsen *et al.*, [16], demostraron que la secuencia de aminoácidos determina la estructura terciaria debido a que la desnaturalización de algunas proteínas es reversible. La desnaturalización es una alteración extrema que anula la estructura tridimensional totalmente, puede llevarse a cabo por medio de calor, extremos de pH, disolventes orgánicos y mediante la exposición de la proteína a ciertos agentes como urea o cloruro de guanidina. Algunas proteínas globulares desnaturalizadas por estos métodos son capaces de recuperar su estructura nativa o terciaria en un proceso conocido como renaturalización, lo cual

sucede si son devueltas a condiciones en las que su conformación nativa es estable [16], [17].

Una prueba más de que la secuencia determina la estructura se ha obtenido del estudio de proteínas homólogas (relacionadas filogenéticamente) como las mioglobinas de vertebrados terrestres y marinos. Con ellas se ha podido observar una similitud de estructuras terciarias y de secuencias a pesar de que provienen de distintos orígenes llevando a la conclusión de que en la secuencia está determinada la estructura [7], [18]. A pesar de esto, los elementos estructurales y funcionales esenciales, en vez de los residuos aminoácidos, son los que se han conservado durante el cambio evolutivo [18].

1.6 EVOLUCION DE LOS DOMINIOS

El estudio de las estructuras tridimensionales de las proteínas ha revelado que la estructura proteica está más conservada evolutivamente que la secuencia, existen muchos ejemplos de dominios que adoptan estructuras en tercera dimensión altamente similares a pesar de la poca similitud aparente en las secuencias [19]. En la actualidad se acepta que proteínas que comparten un mismo plegamiento y muestran señales de ancestría común son homólogos remotos o distantes [19].

A través del análisis de secuencias se ha encontrado que algunos dominios tienen un origen muy antiguo debido a que se encuentran altamente esparcidos en cada uno de los tres dominios de la vida: arquea, bacteria y eukaria. La persistencia de dichos dominios implica que son hiperadaptables y/o esenciales dentro de los procesos celulares. Muchos dominios enzimáticos del metabolismo central (como los barriles TIM (β/α)₈ o proteínas con plegamiento tipo Rossmann) parecen deber su herencia a ancestros que precedieron al último ancestro común de los tres dominios de la vida [19].

Todas estas hipótesis sobre la emergencia de los dominios nos remontan a la pregunta: ¿a partir de qué evolucionaron estos dominios? Existe evidencia de que no sólo los dominios si no también sus plegamientos han cambiado, lo cual sugiere que la mayoría de los plegamientos modernos podrían haber surgido, a lo mucho, de unos cuantos ancestros [19]. Para ello, Söding *et al.*, sugieren que los dominios pueden ser descendientes de conglomerados de polipéptidos pequeños que fueron capaces de plegarse, un motivo, y brindar una función benéfica [20].

La idea de que los primeros dominios evolucionaron por recombinación de un conjunto limitado de unidades estructurales es atractiva por distintas razones: 1) La complejidad combinatoria de evolucionar un dominio completo en una sola pieza sobrepasa los límites; para un dominio de 100 residuos existen 20^{100} posibles secuencias y sólo una fracción ínfima de éstas sería capaz de plegarse [20]. En comparación, la optimización de una secuencia de 20 residuos requerida para fabricar una unidad de estructura supersecundaria está dentro del alcance de los sistemas biológicos; 2) No existen mecanismos no biológicos que puedan producir cadenas polipeptídicas del tamaño suficiente para formar un dominio y mucho menos para producir las cantidades necesarias para explorar el espacio de secuencia disponible; 3) Aún si la evolución *de novo* de dominios completos fuera posible sería demasiado ineficiente en relación al ensamblaje a partir de módulos, el cual permite la utilización de características optimizadas individualmente en un nivel de complejidad más alto. Además, los sistemas biológicos utilizan la modularidad a muchos niveles, desde el nivel molecular en donde se ha visto que para las proteínas el barajeo de módulos de tamaño menor a los dominios puede acelerar en mucho la evolución de estructuras plegadas, hasta el nivel multicelular; 4) Se ha encontrado que un creciente número de fragmentos de proteínas no-homólogas son similares en secuencia, estructura y a veces función, apuntando hacia un origen común [20].

Una vez que estos dominios y sus plegamientos se establecieron en el transcurso de la evolución y se encontraron de manera aislada se pudo dar el origen de las proteínas multidominio. Se dice que los dominios encontrados en estas proteínas multidominio son evolutivamente móviles y se han expandido durante el transcurso de la evolución encontrándose en proteínas que no están relacionadas en la actualidad [21].

Una de las hipótesis del surgimiento de las proteínas multidominio es que la duplicación y el barajeo por recombinación (*shuffling*) de dominios existentes son las dos fuerzas evolutivas más importantes que dirigen la evolución del proteoma [22].

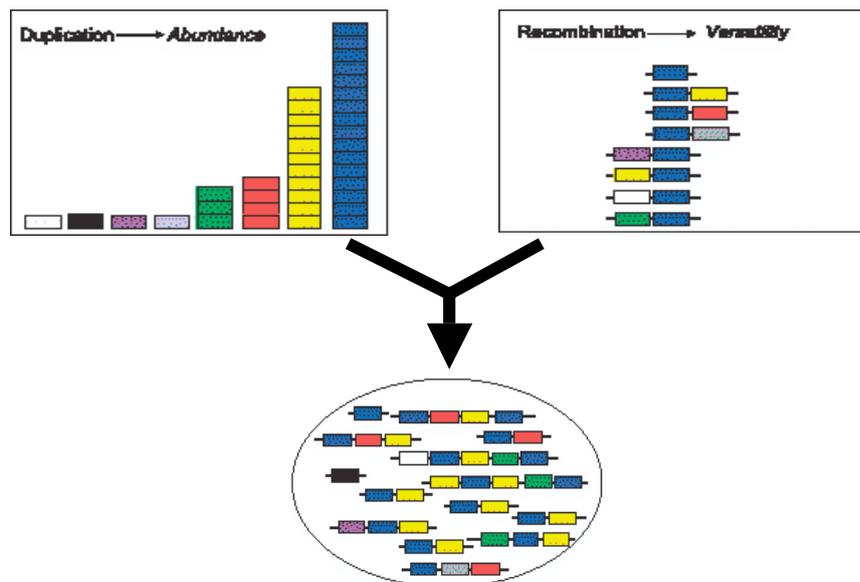


Fig. 1.6. Esquema en donde se representa el efecto de la duplicación y recombinación en la formación del proteoma. Tomado de Sánchez, L. 2006.

1.7 FUERZAS EVOLUTIVAS

A través del estudio en el campo de la evolución de proteínas, principalmente en estudios de recombinación homóloga (DNA *shuffling*) y de sustitución de bases, se ha reconocido que la diversidad proteica es producto de toda una jerarquía de eventos naturales que producen mutaciones; eventos como reordenamientos, deleciones,

transferencias horizontales, transposiciones y yuxtaposiciones no-homólogas en conjunto con la recombinación homóloga y la sustitución de bases [23].

La duplicación es una fuerza evolutiva central para la diversificación de dominios [20]. El efecto de esta fuerza resulta en la expansión de un dominio en términos de su abundancia y esto ha sido asociado al incremento de complejidad de un organismo (Figura 1.6) [22]. La evidencia experimental que sustenta el importante papel de la duplicación fue obtenida por Sterner *et al.*, en un trabajo en el cual se provee de bases para pensar que el plegamiento barril $(\beta\alpha)_8$ ha evolucionado a partir de la duplicación de un subdominio que comprendía cuatro unidades $(\beta\alpha)$. Ellos encontraron que la superposición de las mitades amino (-N) y carboxilo (-C) de las enzimas HisA y HisF (HisA-N y HisF-N; HisA-C y HisF-C) de *Thermotoga maritima*, ambas con un plegamiento $(\beta\alpha)_8$, expone altas similitudes entre las estructuras y las secuencias. Posteriormente al expresar por separado en *E. coli* los segmentos génicos codificantes de la mitad amino y la mitad carboxilo de HisF encontraron que se trataba de proteínas inactivas con estructura secundaria y terciaria bien definidas y al coexpresarlas *in vivo* encontraron mezclas de las dos mitades ensambladas que formaban un complejo catalíticamente activo. Con esto ellos sugieren la posibilidad de que este plegamiento haya evolucionado a partir de un medio barril ancestral a través de una serie de duplicaciones génicas, fusiones y diversificaciones [24].

Por su parte el barajeo o recombinación de dominios es el principal mecanismo para la rápida generación de nuevas combinaciones de dominios. En el presente se ha reconocido que este mecanismo es utilizado en diferentes sistemas como el sistema inmunológico de vertebrados el cual utiliza un puñado de dominios y los arregla en combinaciones casi infinitas para satisfacer los requerimientos del sistema de reconocimiento inmune en vertebrados [20]. Vogel *et al.*, [22] consideran

que la recombinación de dominios es el principal mecanismo que otorga versatilidad al proteoma (Figura 1.6).

Deem *et al.*, plantean que durante cualquier proceso evolutivo las proteínas, sus dominios y sus plegamientos, quedan atrapadas en mínimos de energía local dado lo cual son necesarios movimientos dramáticos como intercambios/ensamblajes (*swaps*) y yuxtaposiciones para salir de estas regiones de energía mínima y encontrar así nuevos plegamientos. Basados en simulaciones computacionales con algoritmos genéticos en las que buscan cuantificar el potencial evolutivo del espacio de secuencia de las proteínas, estos mismos autores proponen que la recombinación no-homóloga de DNA (*nonhomologous “swapping”*) es el paso limitante en la producción de nuevos plegamientos proteicos. Junto a los reordenamientos e inserciones, la recombinación no-homóloga permite la creación de nuevos plegamientos por medio de yuxtaposiciones novedosas de estructuras codificadas de manera conveniente [23].

Estos mismos autores han observado que a pesar de que la recombinación no-homóloga es menos tolerada que una mutación puntual, su naturaleza es más dramática y tiende a elevar la tasa evolutiva promedio [23].

Winter y Riechmann intentan hacer una mímica del proceso de recombinación no-homóloga a través del barajeo de pequeños segmentos polipeptídicos *in vitro* para probar éstas hipótesis. En su trabajo recombinan el DNA que codifica la región N-terminal de la proteína de choque frío A (CspA) con fragmentos al azar de DNA genómico de *Escherichia coli*. El resultado de estas recombinaciones fue clonado y desplegado en fagos seleccionando a las proteínas quiméricas resistentes a tripsina, es decir quimeras que lograron formar dominios plegados resistentes a proteólisis. Estos resultados son consistentes con las hipótesis planteadas y sugieren que cualquier secuencia en el genoma tiene el potencial de producir dominios plegados en combinación con otros segmentos del mismo genoma [25].

1.8 TÉCNICAS ESPECTROSCÓPICAS PARA EL ESTUDIO DE PROTEÍNAS

La espectroscopía fue originalmente el estudio de la interacción entre radiación y materia como una función de la longitud de onda (λ). De hecho, históricamente, la espectroscopía se refirió al uso de la luz visible dispersada de acuerdo a su longitud de onda (como se hace con un prisma). Después el concepto fue expandido en gran medida hacia la comprensión de cualquier medición de una cantidad como función de ya sea la longitud de onda o la frecuencia. La espectrometría es una técnica espectroscópica utilizada para encontrar la concentración o cantidad de una especie química en una muestra dada. El instrumento que se utiliza para realizar tales mediciones es un espectrofotómetro.

La mayoría de los métodos son diferenciados entre atómicos y moleculares de acuerdo a si son o no aplicados a átomos o moléculas. A la par de ésta diferenciación, pueden ser clasificados según la naturaleza de su interacción:

Espectroscopía de absorción: utiliza el rango del espectro electromagnético en el cual una sustancia absorbe. Este tipo incluye tanto espectroscopía de absorción atómica como varios tipos de técnicas moleculares, como la espectroscopía infrarroja y la espectroscopía de resonancia magnética nuclear.

Espectroscopía de emisión: utiliza el rango del espectro electromagnético en el cual una sustancia irradia (emite). La sustancia primero debe absorber energía, que puede provenir de una variedad de fuentes, lo cual después determina el tipo de la emisión subsecuente. Dentro de este tipo se encuentra la luminiscencia y una de las técnicas más usadas es la espectrofluorimetría.

Espectroscopía de dispersión: mide la cantidad de luz que una sustancia dispersa a ciertas longitudes de onda, ángulos de incidencia y ángulos de polarización. El proceso de dispersión es mucho más rápido que los de absorción y emisión.

En las últimas décadas se ha dado un crecimiento explosivo de la biología estructural derivado del desarrollo de la tecnología del ADN recombinante y los avances bioinformáticos. Así mismo, el Dicroísmo Circular (D. C.) ha incrementado su reconocimiento como una técnica espectroscópica valiosa para resolver estudios estructurales en condiciones bajo las cuales las proteínas operan (generalmente en solución), así como en otras condiciones para aportar mediciones de las tasas de cambio estructural de las proteínas [26].

El dicroísmo circular es un fenómeno óptico más que una técnica. Consiste en considerar que la luz polarizada en un plano está formada por dos componentes circularmente polarizados de igual magnitud, uno con rotación hacia la derecha (R) y el otro con rotación hacia la izquierda (L). El dicroísmo circular se refiere a la absorción diferencial de éstos dos componentes. Si después de pasar a través de la muestra examinada los componentes L y R son absorbidos en diferentes magnitudes, se dice que la radiación resultante posee una polarización elíptica (Fig. 1.7) [26].

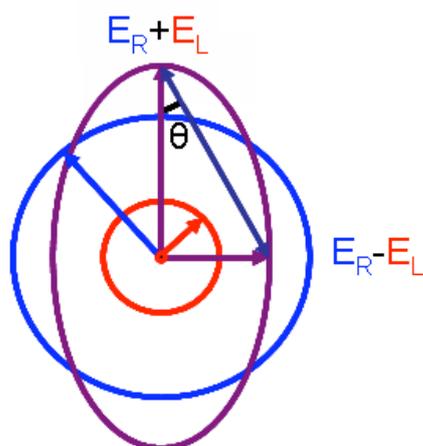


Fig. 1.7. Los componentes circularmente polarizados L (rojo) y R (azul) de la luz polarizada en un plano. Los componentes poseen distinta magnitud y la resultante es elípticamente polarizada (morada). Tomada de http://en.wikipedia.org/wiki/Circular_dichroism

Una señal de D. C. se observa cuando un cromóforo es quiral (ópticamente activo) y surge solo cuando ocurre absorción de la radiación. Las distintas bandas espectrales son asignadas a distintas características estructurales de la molécula por lo

que en estudios con proteínas se puede obtener información estructural complementaria a partir de ciertas regiones espectrales [26].

En proteínas los cromóforos de interés incluyen al enlace peptídico (absorción debajo de los 240 nm), cadenas laterales de aminoácidos aromáticos (absorción en el rango de 260 a 320 nm) y los enlaces disulfuro (bandas de absorción débiles y amplias alrededor de los 260 nm). Si un número de cromóforos del mismo tipo están en proximidad, ellos pueden comportarse como una sola unidad absorbente (exciton) el cual da origen a características espectrales únicas [26].

Los distintos tipos de información que pueden ser obtenidos de estudios de dicroísmo circular en proteínas son: a) composición de estructura secundaria; b) firmas de estructura terciaria; c) integridad de los sitios de unión de cofactores; d) información sobre las características estructurales globales de proteínas. Para este trabajo el interés particular se enfoca en la composición de estructura secundaria, y se sabe que los diferentes tipos de estructuras secundarias regulares encontradas en proteínas dan origen a espectros de D. C. característicos en el UV-lejano (Fig. 1.8) [26].

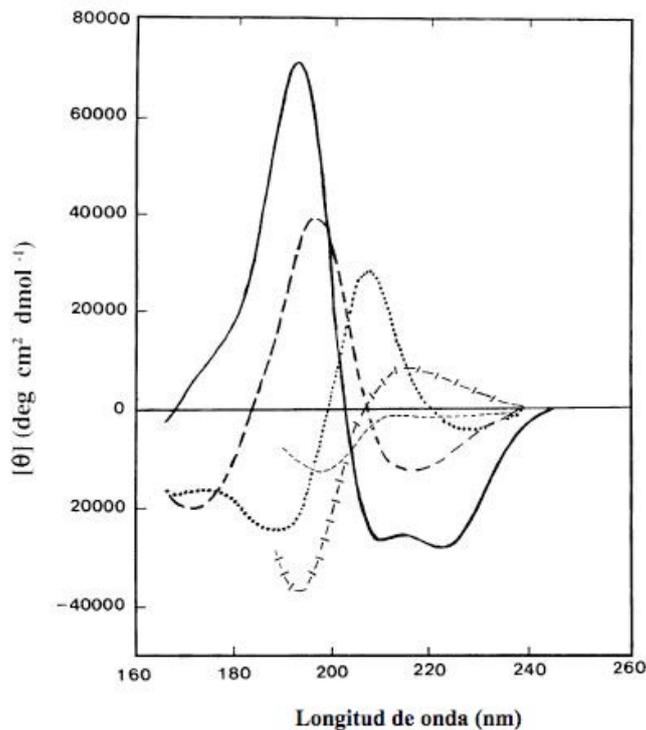


Fig. 1.8. Espectro en el UV-lejano asociado a varios tipos de estructura secundaria. Línea sólida, hélice α ; línea de punteado largo, hoja β -antiparalela; línea punteada, giro- β tipo I; línea cruzada, hélice 3_1 -extendida o hélice poli-(Pro) II; línea de punto corto, estructura irregular (Tomada y modificada de Kelly, 2005).

Por su parte, la espectroscopía de fluorescencia utiliza fotones de alta energía para excitar una muestra, la cual después emitirá fotones de energía más baja. Esta técnica se ha vuelto popular en las aplicaciones bioquímicas y médicas. El fenómeno de fluorescencia muestra varias características generales dentro de las cuales destaca el cambio de Stoke, que se entiende porque la fluorescencia ocurre típicamente en energías más bajas o longitudes de onda más largas, es decir, que la energía de emisión es típicamente menor que la de absorción. Los cambios de Stoke son más dramáticos para fluoróforos polares en solventes polares debido a la interacción del fluoróforo con su medio [27].

El grupo indol de los residuos de triptófano en proteínas es uno de tales fluoróforos solvente-sensitivos y el espectro de emisión del indol puede revelar la localización de los residuos de triptófano en proteínas. La emisión de un residuo expuesto en la superficie ocurre en longitudes de onda más largas que aquellas de un residuo de triptófano en el interior de la proteína. Las proteínas presentan el fenómeno

óptico de fluorescencia intrínseca, debido principalmente a los residuos aromáticos (Trp, Thr y Phe). El fluoróforo dominante en las proteínas es el grupo indol del triptófano, absorbe cerca de los 280 nm y emite alrededor de los 340 nm [27].

II.- ANTECEDENTES

Como antecedente directo a este proyecto se encuentra la tesis de licenciatura llamada: Obtención de proteínas plegadas por complementación con secuencias genómicas; realizada por Lorena P. Sánchez Sánchez, en 2006 [28]. El objetivo del trabajo fue obtener proteínas plegadas *de novo* por recombinación no-homóloga utilizando como templado un miembro de la familia de los barriles ($\beta\alpha$)₈, la PRAI de *E. coli*.

El barril TIM utilizado es el de la enzima Fosforibosil Antranilato Isomerasa (PRAI, por sus siglas en inglés) (Fig 2.1) codificada por el gen *trpF* de *E. coli*. Esta enzima participa en el tercer paso en la vía de síntesis del triptófano. La PRAI de *E. coli* se encuentra formando, de manera natural, parte de una enzima bifuncional asociada a la IGPS (indolglicerolfosfato sintasa), en donde IGPS cataliza la reacción siguiente en la vía. La proteína utilizada es una variante que ha sido separada para funcionar como monómero, posee una extensión de 200 aminoácidos y no tiene ningún elemento de estructura secundaria adicional [28]. Además esta PRAI no es un barril TIM perfecto ya que presenta un giro helicoidal adicional en el asa que conecta la hebra β 8 con la hélice α 8 y una degeneración de la hélice α 5 (ePRAI) [28].

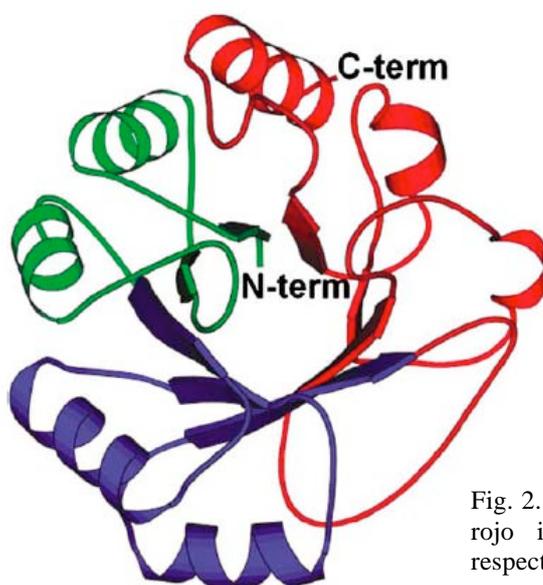


Fig. 2.1. PRAI de *E. coli* (ePRAI). Los colores verde, azul y rojo indican los segmentos 1-50, 51-105 y 106-197 respectivamente. Tomada de Akanuma, 2005.

En estudios previos de ingeniería de proteínas con PRAI se han generado variantes enzimáticamente activas con más de las 8 repeticiones ($\beta\alpha$) originales [28]. En un segundo caso, se construyó una variante ePRAI que contiene una duplicación en serie del quinto módulo ($\beta\alpha$) [29] demostrando así la alta tolerancia de PRAI a las modificaciones.

Para lograr el objetivo de obtener proteínas plegadas *de novo* por recombinación no-homóloga, L. Sánchez utilizó el programa SCHEMA, un algoritmo computacional que identifica los fragmentos de una proteína que pueden ser recombinados sin alterar la integridad de la estructura tridimensional [30]. Este programa calcula el número de interacciones entre aminoácidos y determina el número de interacciones que serían perturbadas al haber un evento de recombinación [28]. Al analizar el perfil obtenido para la PRAI se encontraron 8 picos bien definidos que caracterizan al plegamiento ($\beta\alpha$)₈. Este resultado fue complementado con el resultado de un alineamiento estructural de 12 proteínas homólogas a la PRAI en el cual se encontró un fragmento estructural conservado y se eligió a éste con la finalidad de que el fragmento a sustituir fuese un elemento común en los barriles ($\beta\alpha$)₈ [28]. Dicho fragmento corresponde a la región comprendida entre los aminoácidos 51 al 83 (96 pares de bases) en la PRAI de *E. coli*, es decir las estructuras ($\beta\alpha$)₃ β ₄ (Fig. 2.1). Este fragmento se encuentra dentro de una región sugerida como un importante sitio de nucleación para los eventos tempranos del plegamiento sin el cual la PRAI no podría plegarse [31]. Al haber seleccionado la región, procedió a hacer una restricción de la misma y a complementar a la proteína con fragmentos de ADN cortados al azar del genoma de *E. coli* [28]. Las proteínas quiméricas resultantes fueron seleccionadas por resistencia a cloramfenicol ya que el sistema de selección utiliza a la cloramfenicol acetil transferasa (CAT) como reportero de plegamiento. La CAT

fusionada a la proteína de interés soluble y/o plegada es capaz de hidrolizar al cloramfenicol, en el caso que la proteína de interés se encuentre en un estado insoluble (desplegada) entonces la CAT precipitaría junto con la fusión resultando en baja concentración intracelular y baja resistencia a cloramfenicol. Una proteína soluble produciría un incremento en la resistencia a cloramfenicol [32].

De estos procedimientos se obtuvieron un total de 28 clonas con insertos distintos que fueron identificados dentro del genoma de *E. coli*. El análisis de éstas secuencias reveló una alta diversidad con respecto al origen de las secuencias complementantes y con respecto al tamaño de las mismas. Otra característica a destacar es que aproximadamente la mitad de las proteínas resultantes contenían insertos en una fase distinta a la que codifican originalmente, es decir que el marco de lectura de la proteína original de la que provenían no se conservó [28].

III.- OBJETIVOS

General: Caracterizar a las proteínas quiméricas a nivel de expresión y estructura para obtener datos sobre el estado de plegamiento de las mismas y con ello poder aportar información sobre el origen de los plegamientos.

Particulares:

- 1.- Generar una estrategia de expresión adecuada para obtener cantidades suficientes de proteínas solubles.
- 2.- Realizar pruebas espectroscópicas, en particular de Dicroísmo Circular, para obtener información sobre la estructura de las proteínas quiméricas.

IV.- MATERIALES Y MÉTODOS

En este trabajo se utilizó el plásmido pT4-PRAI-CAT: promotor *ptrc*, resistencia a kanamicina [33]. El gen de la PRAI silvestre (en rojo en la Figura 4.1) se encuentra en dirección río abajo del promotor *trc*, la región seleccionada y retirada está marcada en negro dentro de la PRAI y se muestran los sitios de restricción correspondientes. En el extremo carboxilo de la PRAI se encuentra la CAT (azul en la figura 4.1). Ambos genes están separados entre sí por una región *linker* y al final de el gen de CAT se encuentra una secuencia de seis histidinas para una posterior purificación por cromatografía de afinidad (no se muestran).

Las cepas de *E. coli* utilizadas son: XL1-Blue: $F'::Tn10\ proA_+B_+lac_\Delta(lacZ)M15/recA1\ gyrA96\ (Nal,)\ thi\ hsdR17\ (r_k-m_{k+})\ glnV44\ relA1\ lac$; para medir la concentración mínima inhibitoria de cloramfenicol y para electroporar los vectores religados (ver más adelante).

La cepa JM101: $F'\ traD36\ proA^+B^+ lac^q\ \Delta(lacZ)M15/\Delta(lac-proAB)\ glnV\ thi$. fue utilizada para las curvas de crecimiento y la expresión de las fusiones así como para la expresión de las quimeras (pasos 4.3, 4.4, y 4.9 de la sección Métodos).

4.1 ANÁLISIS BIOINFORMÁTICO

Los alineamientos de los insertos se realizaron utilizando la herramienta BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) y se buscó una identidad del mayor al 98%.

Para las predicciones de estructura secundaria se utilizó la herramienta PSIPRED (V2.4) (<http://bioinf.cs.ucl.ac.uk/psipred/>).

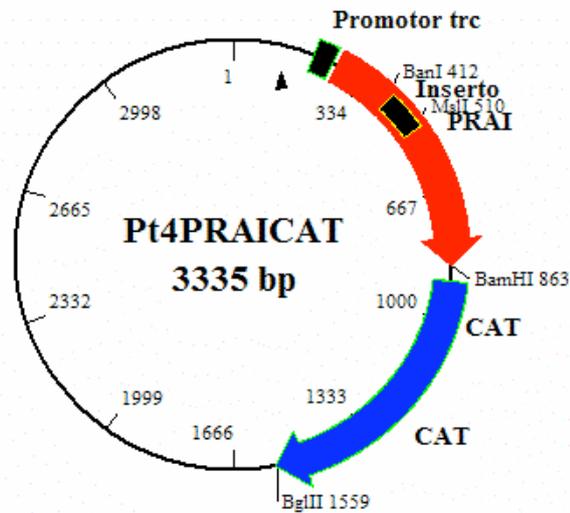


Figura 4.1. Plásmido Pt4-PRAI-CAT. En rojo se denota al gen correspondiente a la PRAI y dentro de él en negro se marca la región separada para la clonación del ADN cortado al azar de *E. coli*. En azul se muestra al gen de la CAT. También se muestran varios sitios de restricción utilizados.

4.2 CMI DEL BANCO H

El banco H es el último banco de quimeras fabricado por L. Sánchez, al cual no le determinó la CMI (Concentración Mínima Inhibitoria). Para determinarla se utilizaron células recién electroporadas XL1-Blue. Se inocularon gotas de 5µL de los cultivos O/N sin diluir y de las diluciones 10^{-2} , 10^{-4} y 10^{-6} en cajas de LB sólido con diferentes concentraciones de cloramfenicol: 0, 20, 40, 60, 80, 100, 120 y 150 µg/mL. Se incubaron las cajas a 30°C por 24 horas.

4.3 CURVAS DE CRECIMIENTO

A partir de cultivos O/N de 3mL se inoculó medio LB líquido con kanamicina para cada clona. Se incubaron a 30°C y se tomaron lecturas cada, media hora para observar el crecimiento, hasta 1 OD.

4.4 EXPRESIÓN DE LAS QUIMERAS FUSIONADAS A CAT

Se utilizó la cepa JM101. A partir de cultivos O/N de 3 mL se inoculó medio LB líquido con kanamicina para cada clona. Se incubaron a 30°C desde 0.1 OD hasta 1 OD (A_{600}) ó durante 8 horas. Posteriormente se centrifugó en frío a 4000 rpm durante 5 min. y se resuspendió el pellet en 500 μ L de PBS. Se sónico cada una durante 10 segundos en dos ocasiones y se agregó PMSF (100 μ M). A continuación se centrifugó en frío a 4000 rpm durante 10 min. y se separaron las fracciones soluble (sobrenadante) e insoluble (pellet); esta última fracción se resuspendió en 100 μ L de PBS con 1% de SDS. Una muestra de cada fracción fue hervida durante 5 min y estas muestras se corrieron en un gel desnaturizante en acrilamida al 12.5%.

4.5 ENSAYO DE BRADFORD

La concentración de las proteínas en ambas fracciones fue determinada a través del ensayo de Bradford (Bio-Rad), con una curva estándar de BSA y por espectroscopia a 595 nm.

4.6 SDS-PAGE

Se realizaron electroforesis (SDS-PAGE) [32] en geles de acrilamida al 12.5%, y se corrieron a 8 mA. Posteriormente se hicieron tinciones con azul de Coomasie 0.1% (*Coomasie Brilliant Blue*) y se lavó el exceso en una solución de ácido acético 10% y metanol 10%.

4.7 WESTERN BLOT Anti-CAT

Se realizó una transferencia a una membrana de nitrocelulosa que se dejó bloqueando O/N con una solución de TBST y leche en polvo al 1% a 4°C. Posteriormente se lavó

3 veces con TBST por 10 min. y se incubó con el anticuerpo anti-CAT 1:3000 en solución de TBS con leche Svelty al 0.1% durante 1 hora. Se lavó tres veces durante 10 min. con TBST. Se incubó con el anticuerpo anti-DIG 1:5000 en TBS con leche en polvo al 0.1% durante 1 hora. Se hicieron dos lavados de 10 min. con TBST. Por último se reveló la membrana con una solución de BCIP 10% (Zymed®) y NTB 10% (Zymed®) en agua mQ.

4.8 RESTRICCIÓN DE CAT

Para remover a la CAT de cada una de las clonas quiméricas se digirió el vector pT4-PRAI-CAT con las enzimas de restricción BamH1 y BglIII (New England Biolabs®) que cortan la principio y al final de la CAT respectivamente. Se incubó a 37°C durante 16 horas en buffer 3 (1X: 100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl₂, 1 mM Ditiotretitol, pH 7.9 a 25°C) (New England Biolabs®).

Después se llevó a cabo una electroforesis en gel de agarosa grado ADN al 1% (Bio-Rad®). Se cortaron las bandas correspondientes a un peso molecular aproximado a 2900 pb. La obtención del DNA se realizó con el kit de Qiagen® (*Qiaquick Gel Extraction Kit*). Los vectores se religaron utilizando la enzima T4-DNA ligasa (New England Biolabs®) durante 16 horas en buffer de ligasa T4 (50 mM Tris-HCl, 10 mM MgCl₂, 10 mM Ditiotretitol, 1 mM ATP, pH 7.5 a 25°C) (New England Biolabs®) y se hizo una limpieza de sales con butanol como sigue: se agregó butanol en relación 5:1 (butanol/ligación) y se mezcló en un vortex, después se centrifugó 5 min. a 13000 rpm, en seguida se extrajo el butanol y se lavó con un volumen de etanol al 70% frío, después se centrifugó 5 min. a 13000 rpm y se extrajo el etanol. Se dejó secar en Savant durante 5 min. a calor medio y se resuspendió en agua destilada [34].

El vector ligado se electroporó en células XL1-Blue que fueron plaqueadas en kanamicina (35µg/mL) y en cloramfenicol (20 µg/mL) como control.

4.9 EXPRESIÓN DE LAS QUIMERAS SIN CAT Y EXTRACCIÓN DE CUERPOS DE INCLUSIÓN

A partir de cultivos O/D de 3 mL se inocularon 200 mL de medio LB líquido con 20 µL kanamicina (35 µg/mL) para cada clona. Se incubaron a 37°C O/N. Para la extracción a partir de cuerpos de inclusión se centrifugó en frío a 4000 rpm durante 20 min., se desecho el sobrenadante y se resuspendió el pellet en 15 mL de *buffer* de lisis (*buffer* B: 8M urea, 10mM Tris-Cl, 100mM NaH₂PO₄, pH 8.0). Se sónico cada una durante un minuto a pulso constante en tres ocasiones para romper sin calentar. A continuación se centrifugó en frío a 8000 rpm durante 20 min. y se separaron el sobrenadante y el pellet (residuos); se filtró el sobrenadante a través de una membrana con poros de 0.45 µm.

4.10 PURIFICACIÓN DE LAS QUIMERAS

Se realizó una purificación por cromatografía de afinidad Ni-NTA para proteínas con cola de histidinas, en condiciones desnaturalizantes. La muestra filtrada fue aplicada lentamente en una columna de Ni-NTA (2 mL) y se realizó un lavado con 15 mL del *buffer* B; después se hizo otro lavado con 7 mL del *buffer* de lavado (*buffer* C: 8M urea, 10mM Tris-Cl, 100mM NaH₂PO₄, pH 6.3); se prosiguió con dos eluciones con 7 mL del *buffer* D (8M urea, 10mM Tris-Cl, 100mM NaH₂PO₄, pH 5.9) y 7 mL del *buffer* E (8M urea, 10mM Tris-Cl, 100mM NaH₂PO₄, pH 4.5) (*buffers* de elusión). Las distintas fracciones se recolectaron cada mL. Se determinó la pureza mediante un gel de acrilamida al 12.5%.

4.11 RENATURALIZACIÓN (replegamiento) Y DIÁLISIS

Se realizó una titulación a partir de pH 4.5 hasta pH 9.0 con NaOH 0.1M, debido a que las pruebas para dializar indicaron que el pH 9.0 y la presencia de glicerol eran favorables para la solubilización de las proteínas. Se dializó en pasos para replegar a la proteína. En el caso de G1SC partimos de las siguientes concentraciones en el *buffer* de la muestra: urea 4M, fosfatos 50mM, Tris-Cl 5mM, glicerol 10% y pH 9.0. Se dializó contra: urea 0.8M, fosfatos 50mM, Tris-Cl 0M, glicerol 10% y pH 9.0 a temperatura ambiente. En el siguiente paso llevamos a: fosfatos 50 mM y pH9 en frío.

En el caso de H17SC también titulamos de un pH 4.5 a un pH 9.0 con NaOH 0.1M. Se siguió la misma estrategia de diálisis en pasos para replegar a la proteína. Se utilizó β -mercaptoetanol 0.5 M para romper puentes disulfuro intermoleculares y evitar así la agregación. Se partió de las condiciones de la muestra: urea 4M, fosfatos 50mM, Tris-Cl 5mM, glicerol 10%, β -mercaptoetanol 0.5M, pH 9.0 y se dializó contra: urea 0.8M, fosfatos 50mM, tris 0mM, glicerol 10%, β -mercaptoetanol 50mM y pH 9.0 a temperatura ambiente. Posteriormente se dializó contra: fosfatos 50mM, glicerol 10%, β -mercaptoetanol 20mM y pH 9.0 en frío. En el último paso se dializó contra: fosfatos 50mM, pH 9.0 en frío. La concentración de las proteínas fue determinada a través del ensayo de Bradford, con una curva estándar de BSA y por espectroscopia a 595 nm.

4.12 DICROÍSMO CIRCULAR

El espectro del dicroísmo circular (D. C.) fue registrado para ambas proteínas a 25°C utilizando un espectropolarímetro modelo Jasco J-710. Se utilizó una celda plana de cuarzo de 1 cm. El espectro fue colectado desde 260 nm hasta 200 nm con una resolución de 1 nm y un tiempo de respuesta (integración) de 2 seg. Los espectros de

referencia (blancos) fueron obtenidos a partir de muestras de *buffer* fosfatos 50mM. Se hicieron dos lecturas, antes y después de calentar la muestra.

Se hizo una curva de desplegamiento con temperatura, desde 25°C hasta 90°C con resolución de 1°C con control de temperatura Peltier. Se utilizó una celda de 1 cm y se tomó la lectura a 222 nm.

4.13 FLUORESCENCIA

El cambio en la fluorescencia fue monitoreado para ambas proteínas. Se prepararon muestras de proteína con una concentración de 50µg/µL y con distintas concentraciones de GuHCl: 0M, 0.5M, 1M, 1.5M, 2M, 2.5M, 3M, 4M y 5M para desnaturalizar a las proteínas. Los espectros de emisión de fluorescencia fueron registrados en un espectrofluorómetro modelo ISS PC-1 a una longitud de onda de excitación de 280 nm y colectados desde 300 nm hasta 400 nm. Los espectros de las muestras fueron corregidos por sustracción del espectro de referencia del *buffer* en que estaban contenidas las muestras (fosfatos 50mM, GuHCl).

Se midió el espectro de emisión fluorescencia completo (de 300 nm a 400 nm) para cada grado centígrado, con una longitud de onda de excitación de 280 nm. Para G1SC en un intervalo de temperatura de 30°C hasta 70°C. Para H17SC en un intervalo de temperatura de 25°C hasta 80°C.

4.14 ANS

Se incubaron las muestras con GuHCl (como está descrito arriba) agregando ANS 1mM (Anilino-8-naftaleno sulfonato). Los espectros de fluorescencia fueron registrados utilizando un espectro modelo ISS PC-1. Se utilizó una longitud de onda de excitación de 360 nm y se colectaron los espectros de emisión entre 400 nm y 550

nm. Los espectros fueron corregidos por sustracción de del espectro de referencia de las muestras de proteína [50 μ g/ μ L] en *buffer* fosfatos 50 mM.

V. RESULTADOS y DISCUSIÓN

5.1 ANÁLISIS BIOINFORMÁTICO

Para identificar la región genómica de procedencia de los distintos fragmentos de ADN, cortado al azar del genoma de *E. coli* (insertos), se realizaron alineamientos utilizando la herramienta BLAST. De estos alineamientos obtuvimos también la proteína de origen a la que pertenecen los insertos. Los resultados de estas búsquedas (Apéndice 1, Tablas 1 y 2) nos indican que existe una gran variedad en cuanto a los orígenes de las secuencias complementantes o insertos, lo cual es consistente con la sugerencia de Riechmann *et al.* [25] sobre el potencial que tiene cualquier segmento del genoma para producir dominios plegados al combinarse con otros segmentos del mismo genoma. Dentro del banco H encontramos 3 clonas (H1, H3 y H4) que presentaron insertos formados por dos segmentos de distinta procedencia, es decir, dos segmentos provenientes de distintas regiones en el genoma que se unieron y se ligaron en la construcción comportándose como un solo inserto.

Posteriormente utilizando el programa PSIPRED (V2.4) se llevó a cabo una predicción de estructura secundaria para darnos idea de las estructuras que los insertos tendrían a formar. Los resultados sugieren que las proteínas quiméricas se pliegan utilizando combinaciones de hélices α y conformaciones β en la mayoría de los casos, sólo 6 de las quimeras (H18, H6, H11, H12, H17, H2) presentan un solo tipo de estructura secundaria en la predicción (ver tablas 1 y 2).

5.2 CONCENTRACIONES MÍNIMAS INHIBITORIAS (CMI_s)

Se determinó la Concentración Mínima Inhibitoria de cloramfenicol (CMI) a las clonas del banco H en concentraciones crecientes de cloramfenicol para darnos una idea del grado de plegamiento en que se encontraban las proteínas quiméricas

(Apéndice 1, Tablas 1 y 2). La finalidad de este experimento era obtener alguna clona con una gran resistencia a cloramfenicol, lo cual reflejaría un alto grado de plegamiento; sin embargo, ninguna de las 10 clonas del banco H y ninguna de las 28 clonas en total presento una CMI mayor a la de la fusión PRAI silvestre-CAT (PRAI-CAT, >150 $\mu\text{g/mL Cm}$). Dados estos resultados decidimos caracterizar a todas las clonas.

5.3 CURVAS DE CRECIMIENTO

Se realizaron curvas de crecimiento de todas las clonas a 30°C. A esta temperatura mostraron un comportamiento parecido a una curva de crecimiento típica para una población bacteriana [14]. En dichas curvas se observa una fase lag o de adaptación en los tiempos cercanos a cero con una pendiente cercanas a cero, seguida por una fase exponencial o de crecimiento, que se caracteriza por tener un pendiente muy pronunciada, y hacia el final se observa una fase estacionaria en la que el crecimiento se detiene [14]. En este caso todas las clonas presentan fases lag muy prolongadas y todas las clonas a excepción de 7C presentan fases de latencia o de adaptación cercanas a los 200 min., algunas de ellas mantienen ésta fase por mucho más tiempo (ej: F5, G9, G19, G26, G28, H1, H18, F4, H4, H17, H7, H11 Y H12) (Fig. 5.1).

En este experimento se utilizaron a la cepa JM101 transformada con PRAI consenso (SP) (Fig. 5.1-A) [33] como control negativo y a la misma cepa transformada con PRAI silvestre fusionada a CAT (P+C) como control positivo (Fig. 5.1-B). Ambas clonas presentan un comportamiento muy similar al de las demás clonas y varias de éstas últimas muestran un crecimiento más rápido que el de los controles por lo que asumimos que la presencia de las proteínas quiméricas recombinantes no está alterando el patrón de crecimiento de la cepa.

Se esperaba que conjuntando los datos obtenidos en las CMI y las dinámicas de crecimiento se pudieran elegir algunas clonas que presentaran características de crecimiento rápido y alta resistencia a cloramfenicol, al no encontrar este patrón decidimos continuar trabajando con todas las clonas. Los resultados mostrados en las dinámicas de crecimiento no representan repeticiones ni promedios por lo que estadísticamente no son significativos.

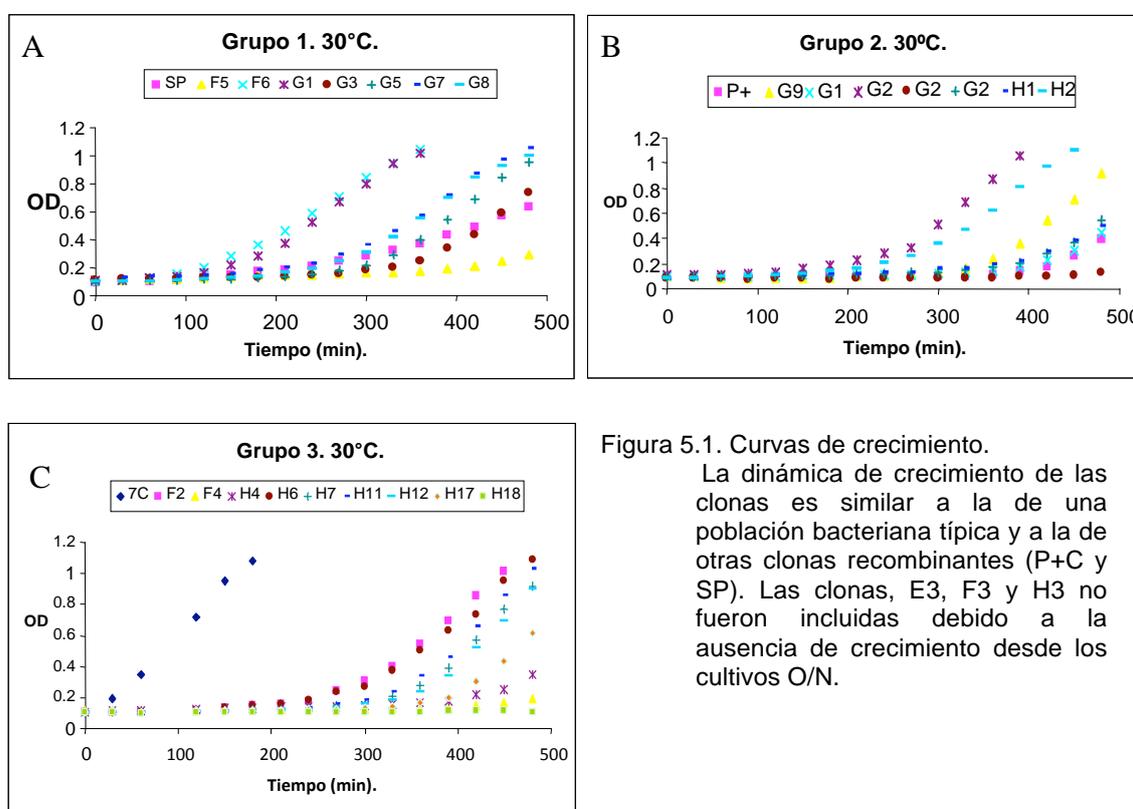


Figura 5.1. Curvas de crecimiento. La dinámica de crecimiento de las clonas es similar a la de una población bacteriana típica y a la de otras clonas recombinantes (P+C y SP). Las clonas, E3, F3 y H3 no fueron incluidas debido a la ausencia de crecimiento desde los cultivos O/N.

5.4 EXPRESIÓN DE QUIMERAS FUSIONADAS A CAT

El siguiente paso de la caracterización consistió en observar los niveles de expresión de las proteínas de interés en cada clona. Para ello se realizaron expresiones de las fusiones a CAT cuyo resultado fue observado en un gel de acrilamida. De nuevo la condición de temperatura a 30°C presentó los mejores resultados y no fue necesaria la

presencia de IPTG para mejorar la expresión, esto debido a que el promotor trc utilizado en nuestras fusiones es un promotor muy fuerte [35].

5.5 SDS-PAGE

Las células se rompieron por sonicación y se centrifugaron. Se separaron el sobrenadante (fracción soluble) y el *pellet* (fracción insoluble). Las dos fracciones resultantes de la expresión de cada quimera, es decir la fracción soluble y la fracción insoluble, fueron analizadas por SDS-PAGE. Los geles muestran para todas las quimeras un perfil de sobre expresión del producto recombinante en la fracción insoluble (Fig. 5.3), este patrón corresponde a los encontrados durante la formación de cuerpos de inclusión durante la expresión de las proteínas [36]. En la fracción soluble no fue posible encontrar la sobre expresión deseada bajo ninguna condición de expresión, los parámetros de temperatura, presencia/ausencia de inductor, tiempo de inducción y tamaño del cultivo fueron variados y no se encontró la sobre expresión en esta fracción.

Como se muestra en la imagen 5.2 fue difícil establecer la presencia de las quimeras en la fracción soluble. Las bandas esperadas entre un rango de 40 kD-60 kD no se presentan con la sobre expresión deseada. La importancia de observar a las proteínas recombinantes en dicha fracción radica en que las proteínas solubles están correctamente plegadas.

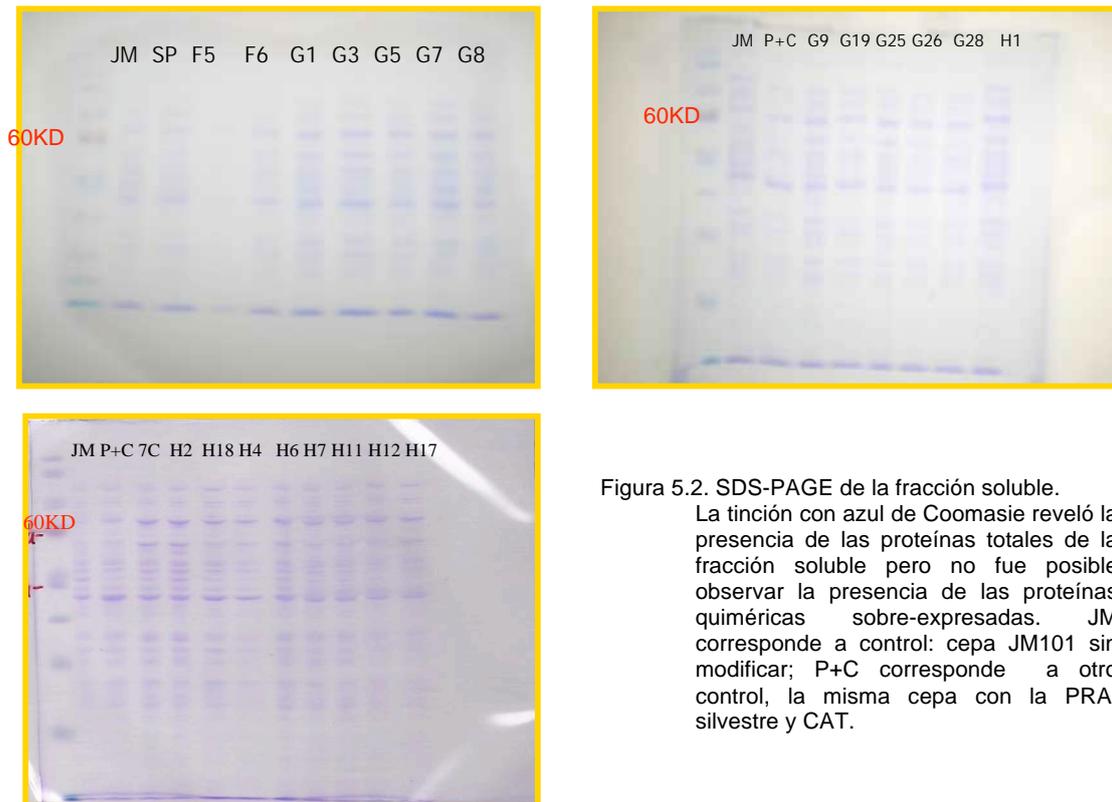


Figura 5.2. SDS-PAGE de la fracción soluble. La tinción con azul de Coomassie reveló la presencia de las proteínas totales de la fracción soluble pero no fue posible observar la presencia de las proteínas quiméricas sobre-expresadas. JM corresponde a control: cepa JM101 sin modificar; P+C corresponde a otro control, la misma cepa con la PRAI silvestre y CAT.

La sobreexpresión en la fracción insoluble (Fig. 5.3) nos indica que el plegamiento de las proteínas es incorrecto, al menos indica que no son solubles. A menudo la expresión de proteínas recombinantes constituye un factor de estrés para las bacterias debido a que compiten por energía y sustratos con las proteínas nativas así que la posibilidad de guardar o acumular el exceso de proteínas mal plegadas podría ser una manera de controlar el estrés ocasionado por la sobre expresión [37]. En la actualidad se sabe que existe un mecanismo dinámico intracelular basado en asistentes de plegamiento que recupera a las proteínas mal plegadas de los cuerpos de inclusión y las solubiliza al facilitar su correcto plegamiento, revisado a detalle en [38]. Debido a que las proteínas no se encontraron en formas solubles decidimos llevar a cabo un Western blot anti-CAT de la fracción soluble en búsqueda de las fusiones.

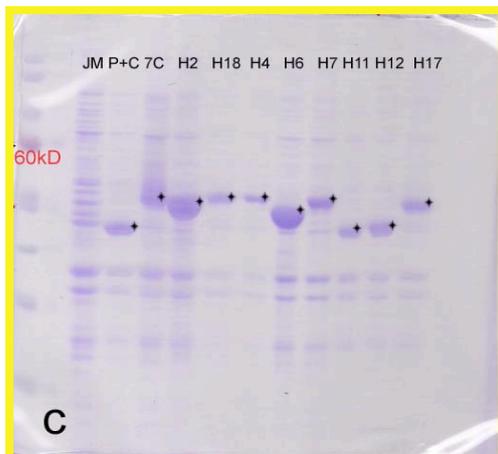
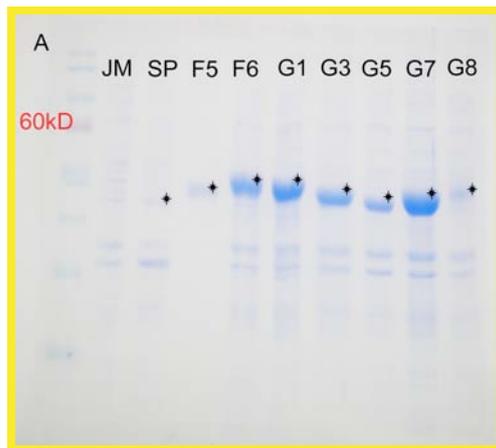


Figura 5.3. SDS-PAGE de la fracción insoluble. La sobre expresión de las quimeras es evidente. Durante el proceso de expresión se dio la formación de agregados intermoleculares (de las quimeras) llamados cuerpos de inclusión lo que llevó a su precipitación y asociación con otros componentes celulares como membranas. Este patrón de agregación corresponde a proteínas incorrectamente plegadas que forman asociaciones a través de regiones hidrofóbicas expuestas [36]. SP es un control y corresponde a la PRAI consenso [33].

5.6 WESTERN BLOT anti-CAT

El resultado de esta técnica confirmó la presencia de todas las proteínas quiméricas fusionadas a CAT en la fracción soluble, es decir proteínas quiméricas con un plegamiento correcto (Fig 5.4).

El Western blot muestra dos resultados principales. El primero de ellos es que todas las clonas presentan una banda por debajo de las quimeras correspondiente a CAT (Fig. 5.4), esto es debido probablemente a una proteólisis. Este fenómeno dentro del sistema revela que las mediciones de las CMI's son poco confiables ya que el grado de resistencia a cloramfenicol presentado por cualquiera de las clonas puede ser debido únicamente a la resistencia conferida por la CAT separada, o es posible que sea una resistencia compuesta entre la CAT separada y la fusionada a la quimera en cuestión. Cualquiera que sea el caso, el valor de la CMI de una quimera dada no está

reflejando el grado de plegamiento de la quimera puesto que todas las clonas presentan a la CAT sola que es capaz de hidrolizar al antibiótico.

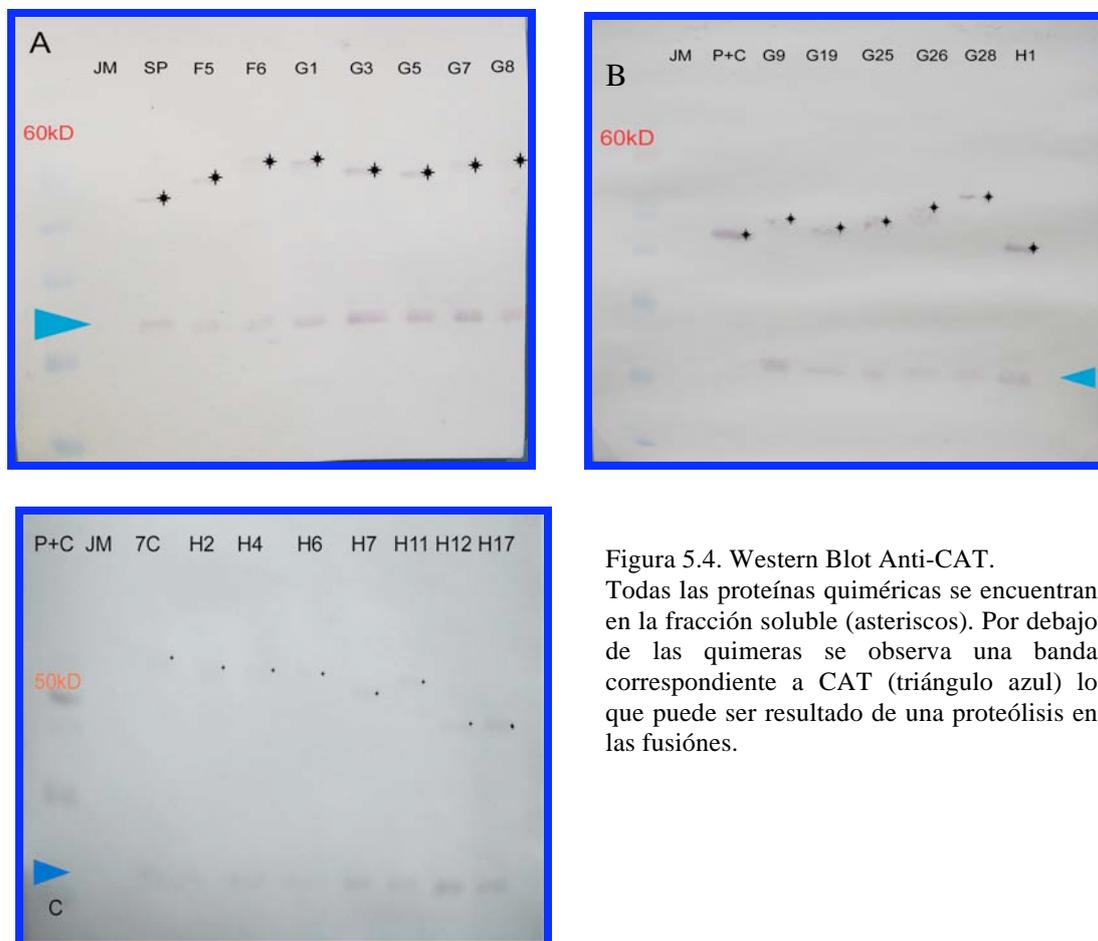


Figura 5.4. Western Blot Anti-CAT. Todas las proteínas quiméricas se encuentran en la fracción soluble (asteriscos). Por debajo de las quimeras se observa una banda correspondiente a CAT (triángulo azul) lo que puede ser resultado de una proteólisis en las fusiones.

El segundo resultado obtenido es una diferencia en los niveles de expresión de las quimeras. Como se observa en la figura 5.4 hay bandas que se distinguen mejor, dentro de la misma membrana y en relación a las otras membranas, ej: G1, G3, G5, H1, H11 y H17. Se decidió tomar este resultado como criterio para proseguir con sólo 8 de las clonas: G1, G3, G5, H2, H11, H12, H17 y H18.

5.7 RESTRICCIÓN DE CAT

Se purificó el plásmido de cada clona. A los 8 vectores de las clonas elegidas se les hicieron digestiones dobles para separar a las quimeras de la CAT y así obtener a las

quimeras con una cola de histidinas. Las enzimas utilizadas fueron BamHI y BglII. La quimera H12 fue descartada ya que se encontró un sitio de restricción de BamHI dentro del inserto (Fig. 5.5). Después se ligaron las digestiones con la T4 DNA ligasa para que una cola de histidinas previamente colocada en el plásmido quedara fusionada al extremo carboxilo de las quimeras. El resultado de las ligaciones fue transformado en la cepa XL1-Blue y posteriormente sembrado en medio LB sólido con kanamicina (resistencia del vector) y en medio LB sólido con cloramfenicol. La selección se llevó a cabo recuperando a las colonias resistentes a kanamicina y sensibles a cloramfenicol para cada clona. Todas las clonas excepto H12 fueron elegidas para realizar ensayos de expresión sin CAT. H12 presentó un sitio de restricción BamHI por lo que se descarto.

Al observar que bajo ninguna condición se aumentaba la expresión en la fracción soluble se decidió trabajar con la fracción insoluble y obtener así a las quimeras a partir de los cuerpos de inclusión para posteriormente ser plegadas in vitro.

5.8 EXPRESIÓN DE LAS QUIMERAS SIN CAT Y EXTRACCIÓN DE CUERPOS DE INCLUSIÓN

Los vectores fueron retransformados en la cepa JM101 para los ensayos de expresión

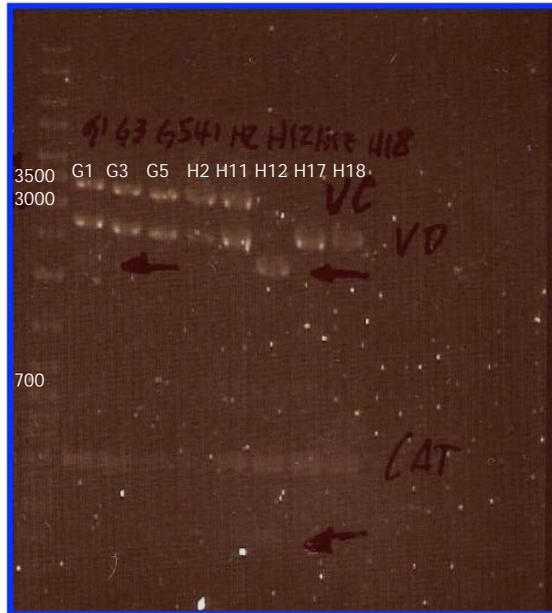


Figura 5.5. Digestiones dobles. Se separó a las quimeras de CAT utilizando las enzimas de restricción BamHI y BglIII. Los productos de la digestión son dos: vector digerido ~3000pb (VD) y CAT 600pb.

sin CAT. En estos ensayos se encontró que la mejor expresión se dio a una temperatura de 37°C O/N. El patrón observado corresponde al mismo observado para las fusiones a CAT, se encontró sobreexpresión en la fracción insoluble y no se encontró dicha sobre expresión en la fracción soluble bajo ninguna condición (Fig 5.6).

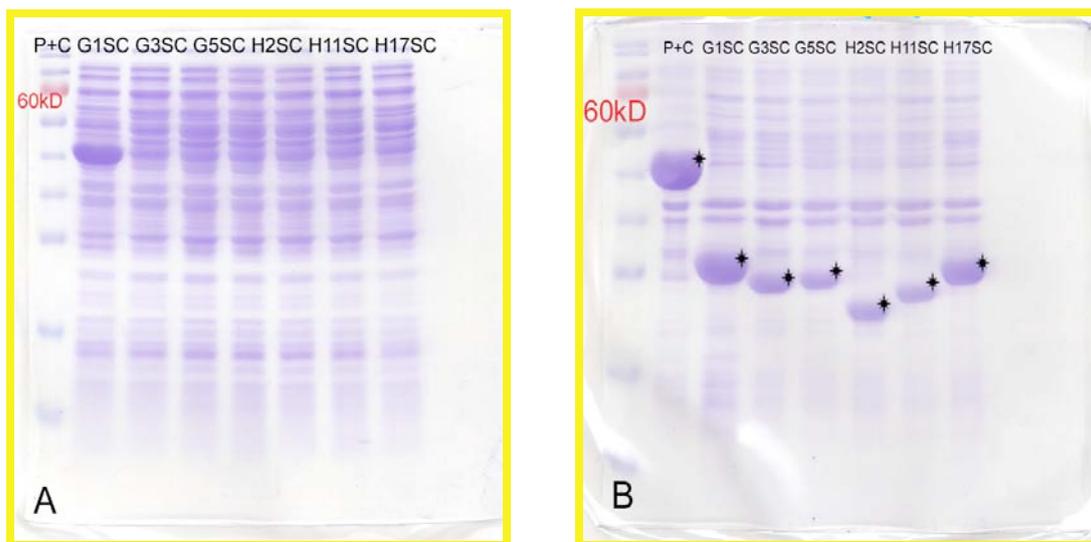


Figura 5.6. Expresión de las clonas sin CAT. En (A) la fracción soluble, en (B) la fracción insoluble. Los gels nos muestran una sobre expresión en la fracción insoluble al igual que sucedía con las fusiones a CAT. La clona H18 no se muestra debido a falta de crecimiento en el O/N. Las clonas mantienen el mismo nombre con la terminología SC (sin CAT) agregada al final.

5.9 PURIFICACIÓN EN COLUMNA DE AFINIDAD

La purificación fue realizada como se describe en la sección de métodos. De las 7 quimeras que fueron probadas solo G1SC y H17SC mostraron un patrón de pureza como el esperado, es decir, una banda del peso esperado, en gran concentración y sin restos de las demás proteínas hacia los últimos pasos de la purificación (últimos carriles del gel) (Fig. 5.7). Las demás quimeras presentaron bandas dobles cercanas al peso esperado, por lo que se decidió no seguir trabajando con ellas.

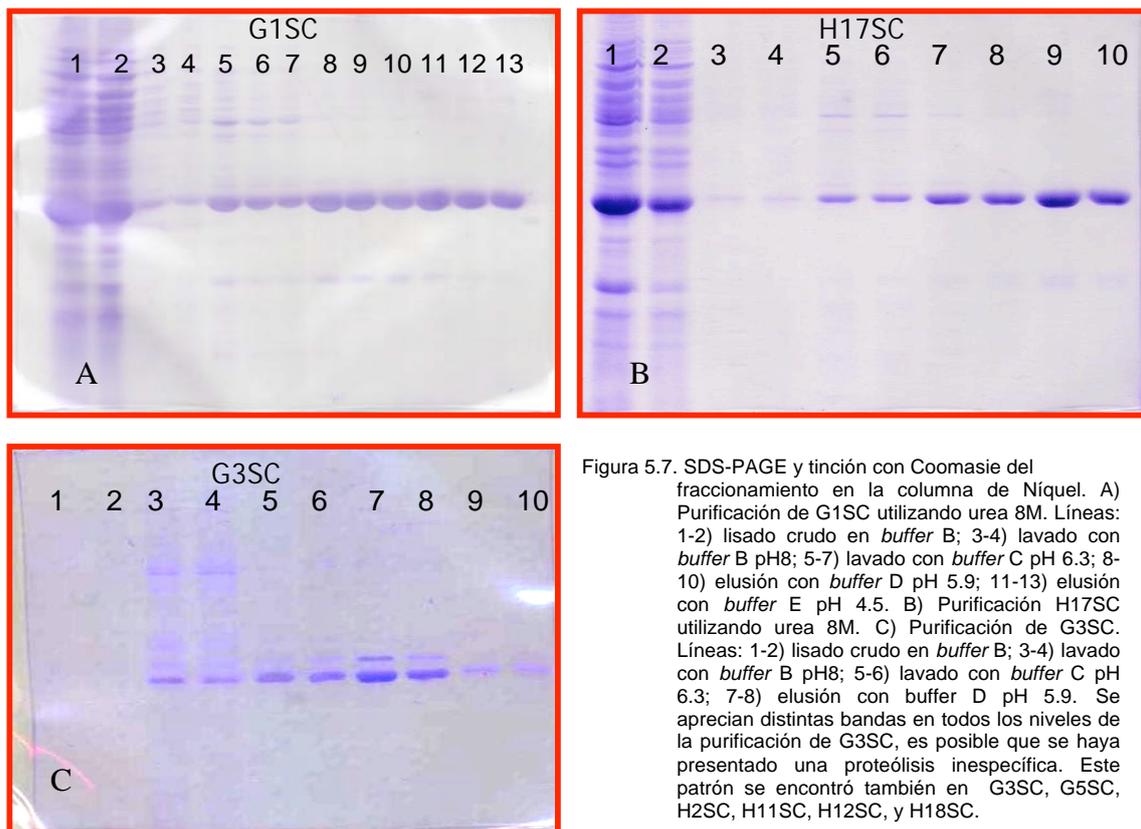


Figura 5.7. SDS-PAGE y tinción con Coomassie del fraccionamiento en la columna de Níquel. A) Purificación de G1SC utilizando urea 8M. Líneas: 1-2) lisado crudo en *buffer* B; 3-4) lavado con *buffer* B pH8; 5-7) lavado con *buffer* C pH 6.3; 8-10) elusión con *buffer* D pH 5.9; 11-13) elusión con *buffer* E pH 4.5. B) Purificación H17SC utilizando urea 8M. C) Purificación de G3SC. Líneas: 1-2) lisado crudo en *buffer* B; 3-4) lavado con *buffer* B pH8; 5-6) lavado con *buffer* C pH 6.3; 7-8) elusión con *buffer* D pH 5.9. Se aprecian distintas bandas en todos los niveles de la purificación de G3SC, es posible que se haya presentado una proteólisis inespecífica. Este patrón se encontró también en G3SC, G5SC, H2SC, H11SC, H12SC, y H18SC.

La presencia de varias bandas en la purificación es una prueba de una posible proteólisis inespecífica que se sugería desde los resultados del Western Blot Anti-CAT. Dado que la cola de histidinas está presente en el extremo carboxilo de las quimeras y éstas se están uniendo a la resina Ni-NTA se puede deducir que la

proteólisis se da en el extremo amino. Es factible pensar que ésta fuese una región desordenada y desprotegida en las fusiones con CAT la cual estaba disponible para proteasas inespecíficas de manera que el plegamiento estaba comprometido desde un inicio.

5.10 RENATURALIZACIÓN Y DIÁLISIS

Se intentaron varios acercamientos para determinar las condiciones de renaturalización (replegamiento) óptimas. Encontramos que para G1SC las condiciones de pH 9.0 y glicerol 10% fueron favorables. Debido a que durante la última elusión de la purificación se utilizó un buffer pH 4.5 se tuvo que realizar una titulación con NaOH 0.1M para llevar a la proteína a un pH 9.0. Este paso fue necesario para evitar una transición lenta entre ambos pHs durante el replegamiento y así evitar agregación y precipitación al acercarse por mucho tiempo al punto isoeléctrico (pI=5.37; predicción utilizando la herramienta *Compute pI/Mw Tool*) de esta proteína quimérica. La presencia del glicerol como agente estabilizador fue necesaria (ver tabla 3).

Tabla 3. Condiciones probadas para la solubilización de G1SC.

	NaPO ₄ 50mM + 300mM NaCl	NaPO ₄ 50mM	NaPO ₄ 50mM + glicerol 10%
pH 6.0	Precipitado	Precipitado	Precipitado
pH 7.0	Precipitado	Precipitado	Precipitado
pH 7.5	Precipitado	Precipitado	Soluble
pH 8.0	Precipitado	Soluble	Soluble
pH 9.0	Precipitado	Soluble	Soluble

El glicerol está reconocido como un osmolito que estabiliza la conformación nativa de las proteínas y las protege de varios tipos de estrés, en condiciones ambientales desfavorables las protege de la desnaturalización y pérdida de función [42]. La estabilidad de las proteínas también es afectada por las sales, las sales monovalentes como el cloruro de sodio a menudo afectan la estabilidad al modificar la fuerza iónica de la solución, la cual a su vez puede estabilizar o desestabilizar a las proteínas dependiendo de la distribución de cargas dentro de una proteína específica [43]. En el caso de G1SC la presencia de sales fue desestabilizadora causando agregación y precipitación (tabla 3). Las diálisis subsecuentes para obtener a la proteína sin urea ni glicerol fueron realizadas como se indica en la sección Métodos.

En el caso de H17SC las condiciones de solubilización fueron similares (ver tabla 4). Se tituló con NaOH 0.1M para llevar a la proteína a pH 9.0 (pI=5.97). H17SC presenta 4 cisteínas dentro de su secuencia por lo que además del glicerol como estabilizador fue necesaria la presencia de β -mercaptoetanol (ver métodos) para romper los puentes disulfuro intermoleculares que estaban causando agregación y precipitación. La presencia de sales no favoreció la solubilidad de esta quimera (tabla 4), parece ser un agente desestabilizador. Al término de las diálisis obtuvimos a ambas proteínas solubles en *buffer* fosfatos 50 mM y a través del ensayo de Bradford y espectroscopia a 280 nm se determinó que la concentración resultante para G1SC fue de 229 $\mu\text{g}/\mu\text{L}$ y para H17SC fue de 167 $\mu\text{g}/\mu\text{L}$.

Tabla 4. Condiciones probadas para la solubilización de H17SC.

	NaPO ₄ 50mM + 300mM NaCl	NaPO ₄ 50mM	NaPO ₄ 50mM + glicerol 10%
pH 6	Precipitado	Precipitado	Soluble
pH 7	Precipitado	Precipitado	Soluble
pH 7.5	Precipitado	Precipitado	Soluble
pH 8	Precipitado	Precipitado	Soluble
pH 9	Precipitado	Soluble	Soluble

5.11 DICROÍSMO CIRCULAR

La figura 5.8-a muestra que la proteína quimérica G1SC replegada posee estructura secundaria. La absorción por debajo de los 240nm es debida principalmente al enlace peptídico y la curva obtenida corresponde a la curva asociada a las hélices α [26]. En la misma figura se aprecian dos curvas, la primera (verde) corresponde a la lectura tomada antes de calentar la proteína hasta 90°C, la segunda (naranja) corresponde a la lectura tomada después del calentamiento. Como se observa las diferencias entre ambas curvas son mínimas. La finalidad de calentar a la proteína era saber a través del dicroísmo circular si la estructura secundaria sufría cambios al aplicar altas temperaturas, experimento comúnmente llamado “*melting*”. El comportamiento de la curva (Fig. 5.8-b) nos indica que no existe un desplegamiento de la proteína a causa de altas temperaturas, por ello las curvas del D. C. no cambiaron. Estos resultados indican que la estructura secundaria de la quimera G1SC posee una alta estabilidad térmica.

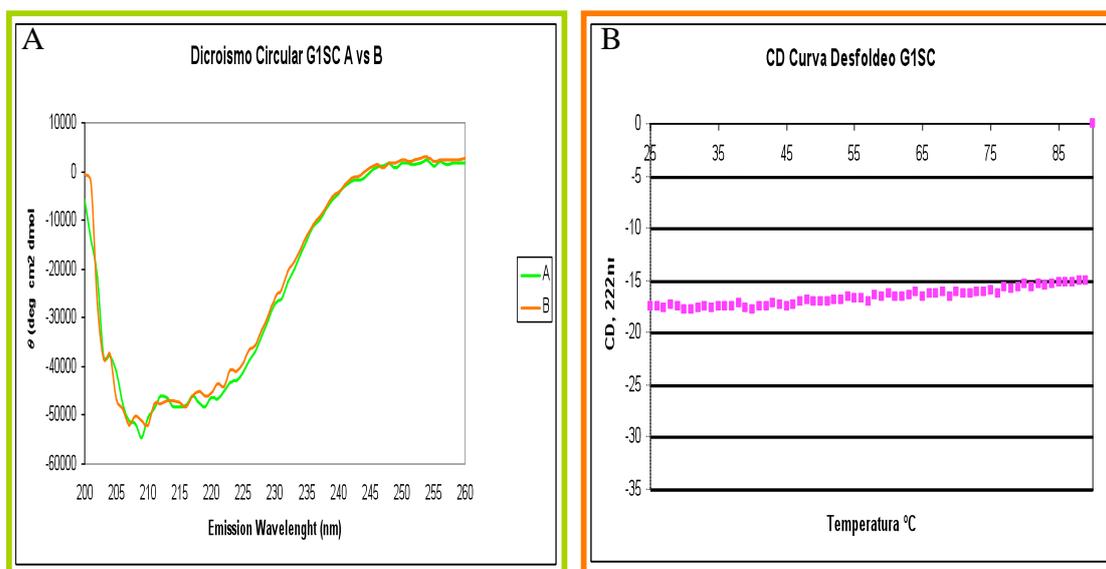


Figura 5.8. Caracterización de la estructura secundaria de G1SC. A) Espectro CD de G1SC. Mediciones del CD UV-lejano a 25°C. La curva corresponde a la curva de hélices α . B) *Melting* de G1SC, CD a 222nm. La señal del CD no cambió significativamente al aumentar la temperatura.

La figura 5.9-a muestra la señal del dicroísmo circular de la proteína H17SC replegada. Se muestra señal de estructura secundaria correspondiente a hélices α . El calentamiento de la muestra no afectó la señal obtenida en una segunda lectura (A vs B en la fig. 5.9-A). Ambas curvas mantienen el mismo comportamiento. La curva del “*melting*” (fig. 5.9-b) muestra muy poco cambio provocado por la temperatura. H17SC se comporta de la misma manera que G1SC, posee una estructura secundaria muy estable que no sufre desplegamiento a causa del aumento de la temperatura.

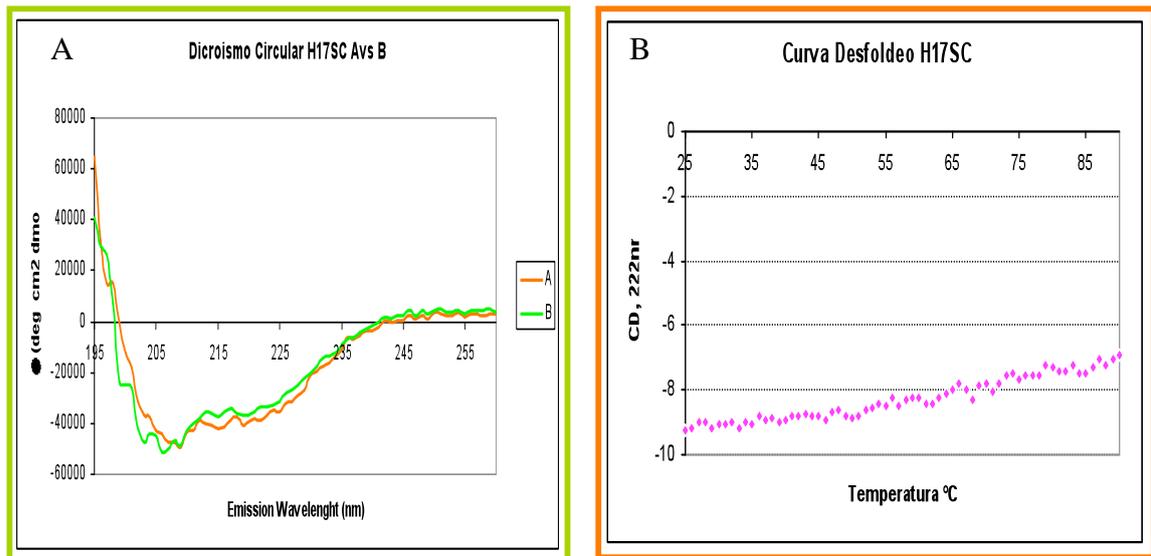


Figura 5.9. Caracterización de la estructura secundaria de H17SC. A) Espectro CD de H17SC. Mediciones del CD UV-lejano a 25°C. B) Melting de H17SC, CD a 22nm. El término desfoldeo en el título de la figura es un error de traducción, mejor usar “despliegamiento”.

5.11 FLUORESCENCIA

Para obtener información sobre la estructura terciaria de las quimeras se realizó el análisis del UV-cercano por fluorescencia. Se añadieron concentraciones crecientes de cloruro de guanidina para desplegar a las proteínas y se monitoreo el efecto sobre la estructura a través de la fluorescencia. En la figura 5.10-A se muestra el espectro de emisión de fluorescencia de la quimera G1SC a distintas concentraciones de cloruro de guanidina, se aprecia un desplazamiento hacia el rojo conforme aumenta la concentración del caotrópico. La emisión intrínseca de las proteínas es dominada principalmente por el triptofano (W; alrededor de los 340 nm); de los triptófanos que contiene una proteína los que están escondidos dentro de la estructura nativa emiten hacia el azul y su emisión cambia hacia el rojo cuando son expuestos a solventes, es decir cuando la proteína sufre despliegamiento. Este es el caso de la quimera G1SC, se aprecia el cambio hacia el rojo en la figura 5.10-A.

En el panel B de la figura 5.10 se observa el cambio en la emisión de fluorescencia en relación al aumento de temperatura, el comportamiento de la curva no corresponde al observado en muchas proteínas pequeñas unidominio (Fig. 5.11), como lo es G1SC.

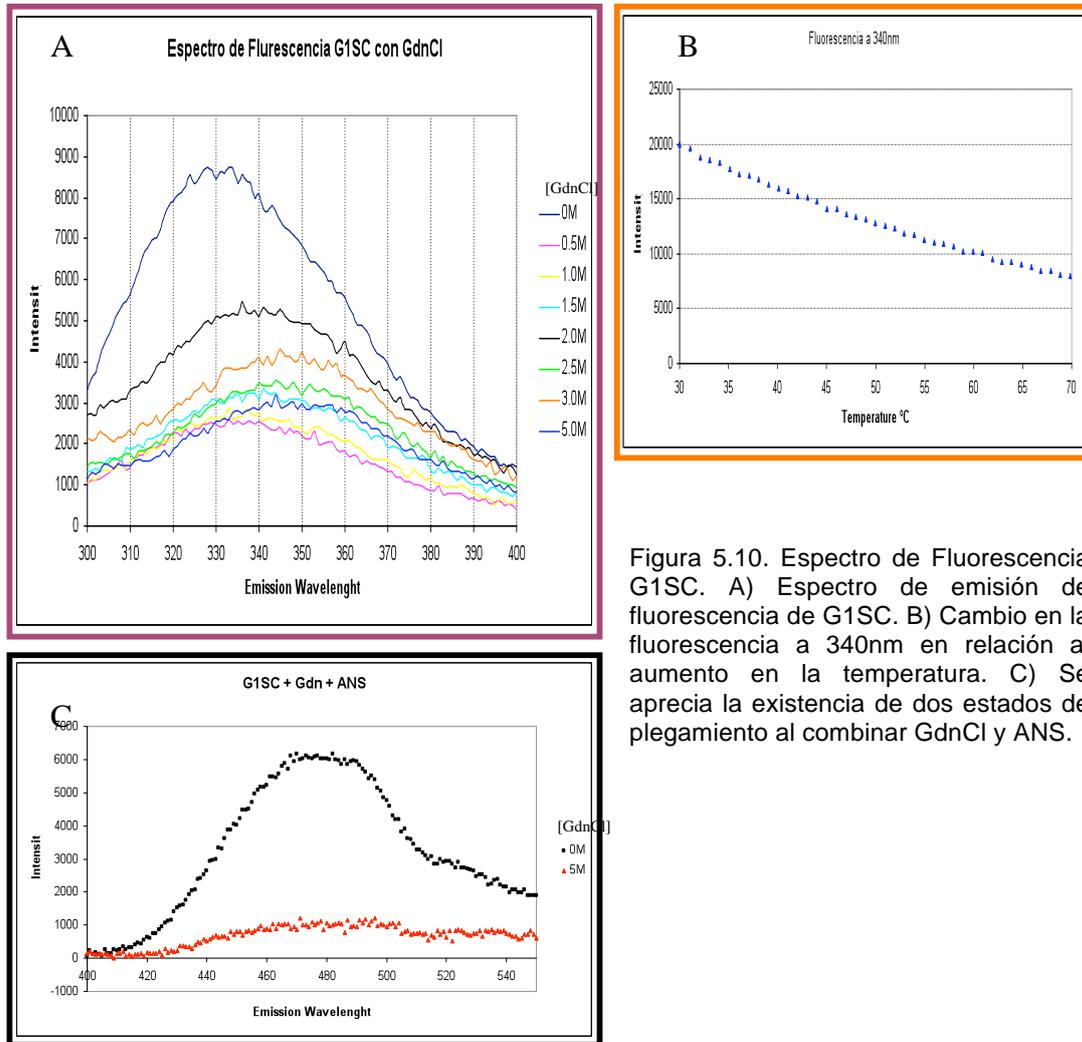


Figura 5.10. Espectro de Fluorescencia G1SC. A) Espectro de emisión de fluorescencia de G1SC. B) Cambio en la fluorescencia a 340nm en relación al aumento en la temperatura. C) Se aprecia la existencia de dos estados de plegamiento al combinar GdnCl y ANS.

La figura 5.11 muestra 3 zonas dentro de la curva, las dos mesetas o regiones con poca pendiente correspondientes a dos estados termodinámicamente distintos y una zona de la curva con pendiente pronunciada correspondiente a una zona de transición entre los dos estados termodinámicos [44]. Este comportamiento de desnaturalización al aumentar la temperatura ha sido reportado para muchas proteínas pequeñas monoméricas, dentro de las cuales hay barriles TIM [45].

La proteína H17SC muestra comportamientos semejantes a G1SC. Presenta un corrimiento al rojo apreciado en el espectro de emisión de fluorescencia (Fig. 5.12-A y C) que indica un desplegamiento a causa del agente caotrópico. La emisión de fluorescencia a 342 nm al aumentar la temperatura (Fig. 5.12-B) no muestra el comportamiento esperado para una proteína pequeña (Fig. 5.11) [13], [39].

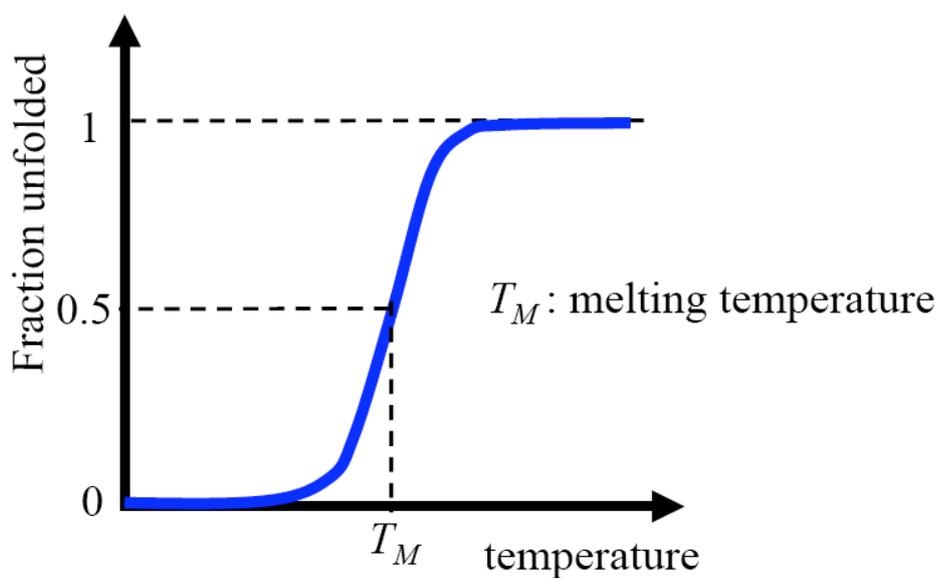


Figura 5.11. Esquema que representa la fracción (proporción) de moléculas de proteína en una solución que están desplegadas en función de la temperatura. En el eje de las ordenadas está la temperatura, en el eje de las coordenadas está la fracción de moléculas desplegadas. El punto en el que la mitad de las moléculas se encuentran desplegadas debido a la temperatura se llama *melting temperature* (T_M). Tomada de [44].

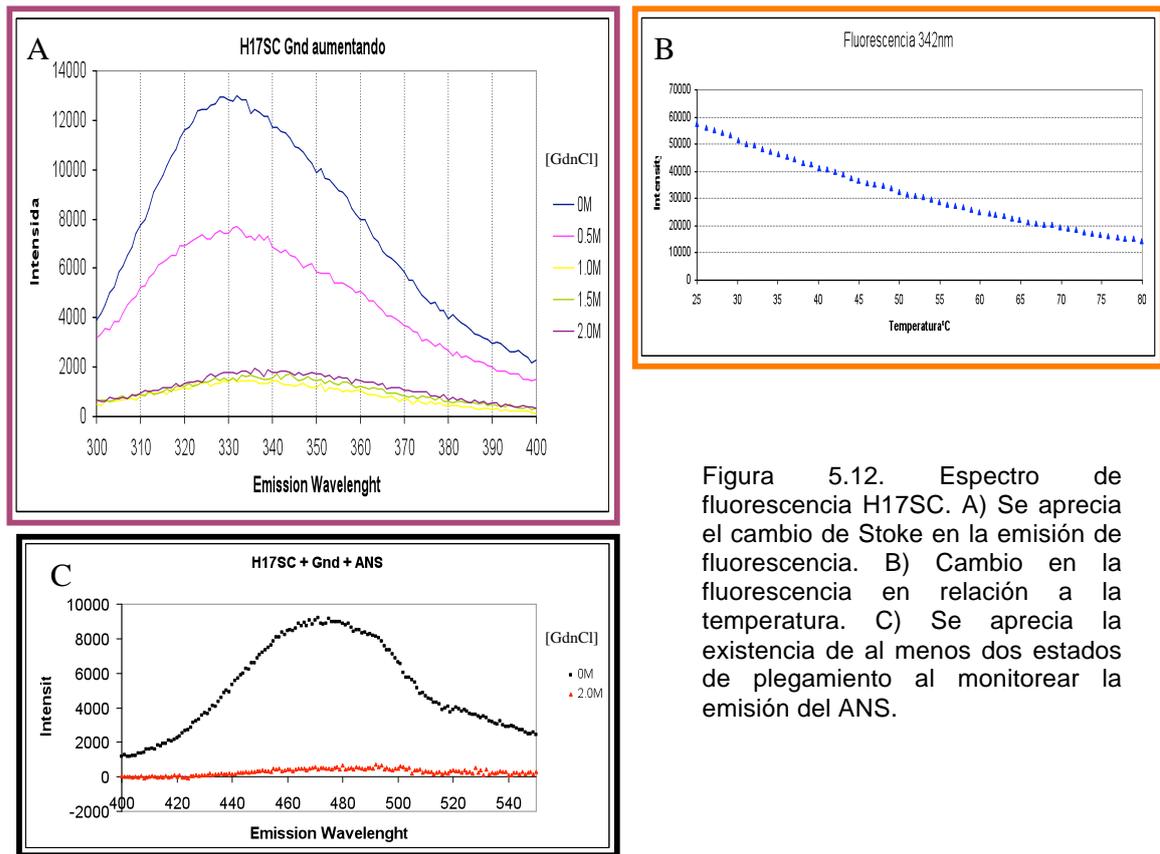


Figura 5.12. Espectro de fluorescencia H17SC. A) Se aprecia el cambio de Stoke en la emisión de fluorescencia. B) Cambio en la fluorescencia en relación a la temperatura. C) Se aprecia la existencia de al menos dos estados de plegamiento al monitorear la emisión del ANS.

5.12 ANS

Se realizó una prueba con ANS (1-anilinaftaleno 8-sulfonato) para saber más sobre la estructura terciaria de las quimeras. El ANS es una sonda química que se adhiere de manera no covalente en las regiones con ambientes no-polares o hidrofóbicos de las proteínas y emite, en el rango del UV en el espectro electromagnético, de manera preferencial en estas regiones [27].

En ambas figuras (panel C, figuras 5.10 y 5.12) se aprecian claramente las diferencias para el estado “nativo” (0M GdnCl) de la quimera y para el estado desplegado (5M GdnCl para G1SC y 2M GdnCl para H17SC). El estado nativo presenta una alta intensidad de fluorescencia en ambos casos lo que indica una gran unión de ANS a las quimeras, el estado desplegado muestra una intensidad mucho más baja con respecto al estado “nativo”.

Este comportamiento sugiere dos fenómenos principales: 1) los altos niveles de intensidad de fluorescencia presentados en el estado nativo de las dos quimeras son probablemente producto de una gran cantidad de regiones hidrofóbicas superficiales en la topología de las moléculas que proporcionan los ambientes no-polares en los que se une y emite el ANS; 2) en el estado no-nativo o desplegado de las quimeras los ambientes no-polares necesarios para la unión y emisión de ANS ya no están presentes debido al efecto desnaturalizante del GdnCl, provocando muy poca unión de ANS y muy poca emisión.

VI. CONCLUSIONES

De las 28 clonas con las que se empezó el trabajo solo G1, G3, G5, H2, H11, H12, H17 y H18 fueron seleccionadas para continuar con los experimentos de purificación, no se siguió con las demás clonas debido a los bajos niveles de expresión mostrados durante el Western blot o debido a la ausencia de crecimiento durante los preparativos de ese experimento y en etapas previas de expresión.

Durante los preparativos de los experimentos de purificación se encontró que H12 poseía un sitio de restricción que no permitió continuar con esta clona (ver resultados sección 5.7) y H18 no creció previo a los ensayos de expresión.

Para G3, G5, H2, H11 y H12 es claro que su caracterización no puede continuar debido a la proteólisis evidenciada durante la purificación. Esto mismo podría ser un problema presentado en las otras quimeras que fueron descartadas.

En el caso de G1SC y H17SC los resultados obtenidos en esta caracterización son: a) se obtuvo evidencia clara de la presencia de estructuras secundarias en ambas quimeras a través del dicroísmo circular; b) los experimentos con GdnCl muestran que si existe un desplegamiento (corrimiento al rojo en fluorescencia) en ambas quimeras, esto es que si existe estructura terciaria; c) los experimentos con temperatura (*melts*) indican que las estructuras secundarias son muy estables, al menos que son muy resistentes a altas temperaturas; sin embargo, no se obtiene el patrón de comportamiento esperado para una proteína globular pequeña (una transición acorde al modelo de dos estados) [13] ni a través del D. C. ni por fluorescencia; d) las pruebas con ANS y GdnCl confirman la existencia de distintos estados de plegamiento, no solo dos estados, por lo que el mecanismo parece ser mucho más complejo que una reacción de dos pasos [13]. Estas mismas pruebas sugieren un alto contenido de superficies hidrofóbicas en las estructuras “nativas”

(altos niveles de fluorescencia de ANS en bajas concentraciones de GdnCl) lo cual podría ser la causa de precipitación, formación de cuerpos de inclusión y otros fenómenos de agregación durante la expresión y purificación, en vez de una estrategia de expresión errónea.

Esta evidencia sugiere que las quimeras presentan el estado de glóbulos fundidos (molten globules), intermediarios cinéticos de plegamiento que se caracterizan por: 1) poseer estructuras secundarias intactas, como las aquí mostradas por el D. C.; 2) cambio en el ambiente de las cadenas laterales observables espectroscópicamente, como las aquí observadas con los agentes caotrópicos; 3) son casi tan compactos como las proteínas en estado nativo, los núcleos de plegamiento son densos debido a las interacciones van der Waals que se mantienen por la presencia de las estructuras secundarias y las interacciones de largo alcance producidas por la estructura terciaria. Los residuos no-polares en el núcleo permanecen en contacto aunque no tan estrechamente como en el estado nativo [13].

Este último punto concuerda con la evidencia mostrada en este trabajo en los experimentos con ANS: existen señales altas de ANS en bajas concentraciones de GdnCl debido al bajo nivel de empaquetamiento, es decir, en una estructura con menor compactación, pero que mantiene las interacciones de un núcleo, se permitiría la formación de cavidades hidrofóbicas, por tal motivo se pueden observar altas lecturas de hidrofobicidad en los estados “nativos” (bajas concentraciones de GdnCl) y además se pueden observar desplegamientos al incrementar las concentraciones de GdnCl. Estos desplegamientos son posibles debido a que un glóbulo fundido preserva muchos elementos de estructura que pueden ser desnaturalizados con los caotrópicos [13].

El estado de glóbulo fundido está propuesto como un “intermediario cinético de la ruta de plegamiento el cual es diferente de la estructura nativa, cuya formación es determinada principalmente por interacciones no-específicas de los aminoácidos con su entorno (agua y el núcleo hidrofóbico del glóbulo), es decir por interacciones hidrofóbicas” [13]. En el caso de estas quimeras, el resultado obtenido es el estado nativo de ambas aunque en el contexto de las proteínas naturales se comportan como un intermediario de plegamiento. Esto lleva a pensar que a través de recombinación no-homóloga se logró obtener péptidos plegados con varias características de plegamiento favorables, pero que no están optimizadas. En términos del plegamiento, se puede decir que el plegamiento de ambas quimeras se encuentra atrapado en un mínimo energético que no es el mínimo nativo, esta condición es poco usual para las proteínas pequeñas las cuales generalmente muestran transiciones altamente cooperativas entre ambos estados y los estados no-nativos no están significativamente poblados [40].

El fenómeno de recombinación no-homóloga es una posibilidad probada en este y otros trabajos, sus aportaciones a la diversidad proteómica al acelerar las tasas evolutivas tienen gran efecto en la evolución de las proteínas [23]. Durante este trabajo en su totalidad se prueba que es experimentalmente posible el fenómeno de recombinación no-homóloga de estructuras del tamaño de subdominios, o del tamaño de estructuras supersecundarias. Sin embargo, para poder hacer un discernimiento de los verdaderos alcances de este fenómeno es necesaria la generación de bancos muy grandes y diversos así como la implementación de un sistema de selección adecuado.

Así mismo es necesaria una determinación más precisa de las estructuras de estas quimeras así como una serie de experimentos en los cuales se puedan dilucidar los mecanismos de plegamiento y se puedan determinar las constantes físico-químicas

necesarias para poder hablar con mayor certeza sobre las quimeras. Las interacciones hidrofóbicas son no-específicas y gracias a ello podrían existir diferentes empaquetamientos del núcleo hidrofóbico, menos estables que el estado nativo pero más estables que el estado desplegado [40]. Con el establecimiento de tales parámetros sería posible hacer precisiones sobre el grado de plegamiento de las quimeras y como resultado sería posible hacer inferencias sobre efectos reales de la recombinación no-homóloga como herramienta de innovación estructural.

Nuestros resultados concuerdan con las aseveraciones de Bogarad *et al.*, y Lupas *et al.*, en donde plantean que la recombinación no-homóloga es una fuerza que acelera la evolución pero es menos tolerada que otros eventos que dirigen la evolución. La presencia de glóbulos fundidos o estructuras parcialmente plegadas e incluso de estructuras incorrectamente plegadas constituye un gasto energético para las células el cual desencadena un estado de estrés y dispara una serie de procesos metabólicos que intentan establecer un equilibrio dentro de la célula [36]. En términos selectivos esta condición no es favorable. Inclusive la formación de agregados intracelulares como los cuerpos de inclusión son un fenómeno derivado de la sobre-expresión, no se presenta como un fenómeno natural a causa del estrés que ello implica [36]. La existencia de glóbulos fundidos está restringida a pruebas *in vitro* y su existencia en sistemas biológicos es tema de debate, existen teorías sobre una posible función de este estado dentro de las células. Otro tema de debate es el planteamiento del estado de glóbulo fundido como precursor de nuevos plegamientos [13].

En varios sentidos, evolutivos, fisico-químicos y moleculares, sería de interés investigar si el efecto de la recombinación no-homóloga en conjunto con el efecto de otras fuerzas evolutivas, como la mutación, son capaces de semejar o hacer una

mímica del camino evolutivo por el cual se logren establecer las interacciones específicas necesarias y determinantes de una estructura terciaria “rígida”, a partir de estas quimeras y su estado de glóbulo fundido.

Se sabe que el esquema evolutivo del proteoma es multifactorial por lo proponemos la implementación de estrategias de mutagénesis dirigidas y mutagénesis al azar, no como herramientas biotecnológicas, sino como un intento de introducir la presencia de otras fuerzas evolutivas en el sistema, un intento de hacer una mímica no sólo de un evento aislado si no de las relaciones entre recombinación no-homóloga y otras fuerzas que dirigen la evolución del proteoma. Creemos que la combinación de dichas fuerzas es lo que lleva a la optimización de los plegamientos, con lo que se ampliaría el esquema evolutivo y se buscarían variantes de éstas quimeras con mejores características termodinámicas.

APÉNDICE 1.

Tabla 1. Características de 18 clonas.

<i>CLONA</i> (<i>tamaño</i> <i>del inserto</i>)	<i>GEN de</i> <i>ORIGEN</i>	<i>ORIGEN PROTÉICO</i>	<i>EN FASE</i> <i>CON EL</i> <i>GEN DE</i> <i>ORIGEN</i>	<i>PRED.</i> <i>ESTRUC.</i> <i>SEC.</i>	<i>CMI</i> <i>Cm</i> ($\mu\text{g/mL}$)
5C (107 aa)	FepA	Receptor de la membrana externa para el transporte del complejo hierro-sideróforo.	Si	$\beta\beta\beta\alpha\beta$	60–80
7C (73 aa)	Fiu	Posible receptor de la membrana externa para el transporte de hierro.	No	$\beta\alpha\alpha\alpha\beta$	40
E3 (68 aa)	YagR	Posible oxidoreductasa con un dominio de unión a molibdeno.	Si	$\alpha\beta\alpha\beta$	40
F2 (71 aa)	yegE +	Posible diguanilato ciclasa (“GGDEF domain signalling protein”).	Si	$\beta\beta\alpha\beta$	40
	YgbJ	Posible deshidrogenasa con un dominio Rossmann de unión a NAD(P).	No		
F3 (52 aa)	YgcP	Posible proteína reguladora (anti-terminador).	No	$\beta\alpha\beta\alpha\beta$	100
F4 (50)	YdcS	Posible transportador de espermidina/putrescina, miembro de la superfamilia ABC de transportadores.	No	$\beta\alpha\beta\alpha\beta$	40–60
F5 (65 aa)	Rnt	Ribonucleasa T.	No	$\beta\alpha\alpha\beta\alpha$	40
F6 (49 aa)	YdiJ	Posible oxidoreductasa ligada a FAD.	Si	$\beta\alpha\beta\alpha$	100–120
G1 (68 aa)	Cbl	Activador transcripcional de unión a DNA de la biosíntesis de cisteína.	Si	$\beta\beta\alpha\beta$	80–100
G3 (50 aa)	FecA	Transportador de citrato férrico de la membrana externa.	Si	$\beta\alpha\alpha\alpha$	40
G5 (44 aa)	YkgB	Posible proteína de la membrana interna.	No	$\beta\alpha\alpha\alpha$	80
G7 (59 aa)	frvR	Posible regulador del catabolismo de polisacáridos.	Si	$\beta\alpha\beta\alpha\beta$	100
G8 (70 aa)	YdjZ	Posible proteína de la membrana interna.	No	$\alpha\beta\alpha\beta$	60
G9 (51 aa)		Posible S-transferasa.	No	$\beta\alpha\beta\alpha$	60
G19 (34 aa)	AcnB	Aconitato hidratasa.	No	$\alpha\beta\alpha$	60

G25 (60 aa)	Pide	Posible aminopeptidasa.		No	$\beta\alpha\beta\alpha$	60
G26 (68 aa)	YbgP	Posible chaperona ("assembly protein").		No	$\beta\beta\beta\alpha\beta\alpha\beta$	80-100
G28 (75 aa)	Zwf	Glucosa-6-fosfato deshidrogenasa.	1-	No	$\beta\alpha\beta\alpha$	60-80
PRAI +CAT	-	-		-	-	>150
XL1-Blue	-	-		-	-	<20

Tabla 1. Se muestran: nombre y tamaño de los fragmentos en cada clona; el gen de origen de cada fragmento; la proteína de origen (o el origen hipotético); la fase con respecto al gen de origen; la predicción de estructura secundaria (PSIPRED V2.4) y la CMI de cloramfenicol. Tomada de L. Sánchez, 2006.

Tabla 2. Características de las clonas del banco H.

<i>CLONA</i>	<i>ORIGEN GEN</i>	<i>ORIGEN PROTÉICO</i>		<i>EN FASE CON EL GEN DE ORIGEN</i>	<i>PRED. ESTRUC. SEC.</i>	<i>CMI Cm (μg/mL)</i>
H1 (136aa)	thrL (58aa)	Subunidades del transportador AI2.	del	No	$\alpha\alpha\alpha$	120
	ygeY (77aa)	Peptidasa (predicción).			$\beta\alpha\beta\alpha\beta\alpha$	
H2 (29aa)	DinJ	Antitoxina Dinj (predicción).		Si	$\alpha\alpha$	150
H3 (73aa)	nanT (39aa)	Transportador de ácido siálico.		Si	-	80
	yafY (34aa)	Regulador transcripcional de unión a DNA (predicción).			-	
H4 (90aa)	traM (56aa)	Región del factor F de transferencia.		Si	-	100
	yhfK (34aa)	Proteína hipotética.				
H6 (30aa)	thrL	Proteína de la membrana interna (predicción).		Si	α	120
H7 (76aa)	trpB	Triptofano sintasa, subunidad β .		Si	$\beta\alpha\beta\alpha$	80
H11 (36aa)	YdjA	Proteína hipotética.		Si	α	110-120
H12 (36aa)	gabT	Aminobutirato aminotransferasa.		Si	β	120

H17 (58aa)	thrL	Proteína hipotética.	Si	$\alpha\alpha$	120
H18 (45aa)	nadB	Aspartato oxidasa.	Si	$\alpha\alpha$	150
PRAI +CAT	-	-	-	-	>150
XL1-Blue	-	-	-	-	<20

Tabla 1. Se muestran: nombre y tamaño de los fragmentos en cada clona; el gen de origen de cada fragmento; la proteína de origen (o el origen hipotético); la fase con respecto al gen de origen; la predicción de estructura secundaria (PSIPRED V2.4) y la CMI de cloramfenicol. (-) indica no definido.

LITERATURA CITADA

1. Nelson, D. L., Cox, M. M. *Lehninger Principles of Biochemistry*. 3 ed. 2000, New York, USA: Worth Publishers.
2. Voet, D., J. Voet. *Biochemistry*. 2004: Wiley and Sons. 1591.
- 3.- Watson, J.D., T. A. Baker, S. P. Bell, A. Gann, M. Levine, R. Losick, *Molecular biology of the gene*. 5 ed. 2004, New York: Benjamin Cummings / Cold Spring Harbor Laboratory Press.
4. Pauling, L., R. B. Corey. *Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets*. Proceedings of Natural Academy of Sciences, 1951. **37**: p. 729-40.
5. Chothia, J., M. Levitt, D. Richardson. *Structure of proteins: packing of α -helices and pleated sheets*. Procedures of Natural Academy of Sciences, 1977. **74**: p. 4130-34.
6. Salem, G. M., E. G. Hutchinson, C. A. Orengo, J. M. Thornton. *Correlation of observed fold frequency with the occurrence of local structural motifs*. Journal of Molecular Biology, 1999. **287**: p. 969-81.
7. Lupas, A. N., C. P. Ponting, R. B. Russell. *On the evolution of protein folds: are similar motifs in different protein folds the result of convergence, insertion, or relics of an ancient peptide world?* Journal of Structural Biology, 2001. **134**: p. 191-203.
8. Chothia, C., T. Hubbard, S. Brenner, H. Barns, A. Murzin. *Protein folds in the all- β and all- α classes*. Annual Review of Biophysics and Biomolecular Structure, 1997. **26**: p. 597-627.
9. Baldwin, R. L., *How does protein folding get started?* Trends in Biochemical Sciences, 1989. **14**(7): p. 291-4.
10. Ruddon, R. W., E. Bedows. *Assisted protein folding*. The Journal of Biological Chemistry, 1997. **272**: p. 3125-28.
11. Dill, K., S. Chan. *From Levinthal to pathways to funnels*. Nature Structural Biology, 1997. **4**: p. 10-19.
12. Panchenko, A., Z. Luthey-Schulten, P. Wolynes. *Foldons, protein structural modules, and exons*. Proceedings of Natural Academy of Sciences, 1996. **93**: p. 2008-13.
13. Creighton, T., ed. *Protein folding*. 1992, W. H. Freeman and Company: New York. 547.
14. Madigan, M., Martinko, J., Parker, J. *Brock: Biología de los microorganismos*. 2001, Madrid: Prentice Hall. 1064.
15. Netzer, W. J., F. U. Hartl. *Protein folding in the cytosol: chaperonin-dependent and -independent mechanisms*. Trends in Biochemical Sciences, 1998. **23**: p. 68-73.
16. Sela, M., F. H. White Jr, C. B. Anfinsen. *Reductive Cleavage of Disulfide Bridges in Ribonuclease*. Science, 1957. **125**: p. 691-2.
17. Anfinsen, C. B. *Principles that govern the folding of protein chains*. Science, 1973. **181**: p. 223-30.
18. Grishin, N. V. *Fold change in evolution of protein structures*. Journal of Structural Biology, 2001. **134**: p. 167-85.
19. Ponting, C. P., R. R. Russell. *The natural history of protein domains*. Annual Review of Biophysics and Biomolecular Structure, 2002. **31**: p. 45-71.
20. Söding, J., A. N. Lupas. *More than the sum of their parts: on the evolution of proteins from peptides*. BioEssays, 2003. **25**: p. 837-46.
21. Kolkman, J. A., W. P. C. Stemmer. *Directed evolution of proteins by exon shuffling*. Nature Biotechnology, 2001. **19**: p. 423-28.

22. Vogel, C., S. A. Teichmann, J. Pereira-Leal. *The relationship between domain duplication and recombination*. Journal of Molecular Biology, 2005. **346**: p. 355-65.
23. Deem, M., L. Bogarad. *A hierarchical approach to protein molecular evolution*. Proceedings of Natural Academy Sciences, 1999. **96**: p. 2591-5.
24. Höcker, B., C. Jürgens, M. Wilmanns, R. Sterner. *Stability, catalytic versatility and evolution of the $(\beta\alpha)_8$ -barrel fold*. Current opinion in biotechnology, 2001. **12**: p. 376-81.
25. Winter, G., L. Riechmann. *Novel folded protein domains generated by combinatorial shuffling of polypeptide segments*. Proceedings of Natural Academy Sciences, 2000. **97**: p. 10068 – 73.
26. Kelly, S., T. Jess, N. Price. *How to study proteins by circular dichroism*. Biochimica et Biophysica Acta, 2005. **1751**: p. 119-139.
27. Lacowicz, J., *Principles of fluorescence spectroscopy*. 1999, New York: Kluwer Academia/Plenum Publishers. 678.
28. Sánchez, L. P. *Obtención de proteínas plegadas por complementación con secuencias genómicas*. 2006. Tesis Profesional. Facultad de Ciencias, UNAM, México D. F.
29. Urfer, R., K. Kirschner. *The importance of surface loops for stabilizing an eightfold $\beta\alpha$ barrel protein*. Protein Science, 1992. **1**: p. 31-45.
30. Voight, C., C. Martinez, Z. Wang, S. Mayo, F. Arnold. *Protein building blocks preserved by recombination*. Nature Structural Biology, 2002. **9**: 553-58.
31. Akanuma, S., A. Yamagishi. *Identification and characterization of key substructures involved in the early folding events of a $(\alpha/\beta)_8$ -barrel protein as studied by experimental and computational methods*. Journal of molecular Biology, 2005. **353**: p. 1161-70.
32. Maxwell, K. L., et al., *A simple in vivo assay for increased protein solubility*. Protein Science, 1999. **8**: p. 1908-11.
33. Farías Rico, J. A. *Generación de un andamiaje estable, diseño de un barril TIM consenso*. 2005. Tesis Profesional. Facultad de Ciencias, UNAM, México D. F.
34. Sambrook, J., D. W. Russell. *Molecular Cloning: A Laboratory Manual*. 2001. New York: Cold Spring Harbor Laboratory. 2100.
35. Baneyx, F. *Recombinant protein expression in Escherichia coli*. Current Opinion in Biotechnology, 1999. **10**: p. 411-21.
36. Schrödel, A., A. de Marco. *Characterization of the aggregates formed during recombinant protein expression in bacteria*. BMC Biochemistry, 2005. **6**: p. 10.
37. Mortensen, K., H. Sorensen. *Advanced genetic strategies for recombinant protein expression in Escherichia coli*. Journal of Biotechnology, 2005. **115**: p. 113-28.
38. Ben-Zvi, A. P., P. Goloubinoff, *Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones*. Journal of Structural Biology, 2001. **135**: p. 84-93.
39. Zitzewitz, J. A., P. J. Gualfetti, I. A. Perkons, S. A. Wasta, C. R. Matthews. *Identifying the structural boundaries of independent folding domains in the α subunit of tryptophan synthase, a (β/α) barrel protein*. Protein Science, 1999. **8**: p. 1200-09.
40. Watters, A., D. Baker. *The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection*. Cell, 2007. **128**: p. 613-24.
41. Chou, K. C., J. R. Blinn, *Classification and prediction of beta-turn types*. Journal of Protein Chemistry, 1997. **16**: p. 575-95.
42. Zancan, P., M. Sola-Penna. *Trehalose and glycerol stabilize and renature yeast*

- inorganic pyrophosphatase inactivated by very high temperatures*. Archives of Biochemistry Biophysics, 2005. **444**: p. 52-60.
43. Dominy, B.N., D. Perl, F. Schmid, C. Brooks III. *The effects of ionic strength on protein stability: the cold shock protein family*. Journal of Molecular Biology, 2002. **319**: p. 541-54.
44. Fersht. A. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. 1999. New York: Freeman. 631.
- 45.- Benitez-Cardoza, C. G., A. Rojo-Dominguez, A. Hernandez-Arana, *Temperature-induced denaturation and renaturation of triosephosphate isomerase from Saccharomyces cerevisiae: evidence of dimerization coupled to refolding of the thermally unfolded protein*. Biochemistry, 2001. **40**: p. 9049-58.