

00381

22.  
2ej



UNIVERSIDAD NACIONAL AUTONOMA  
DE MEXICO

FACULTAD DE CIENCIAS  
DIVISION DE ESTUDIOS DE POSGRADO

AISLAMIENTO Y CARACTERIZACION DE UN GEN DE  
CHAPERONINA 60 (CPN60), DURANTE LA MORFOGENESIS  
DEL NODULO SIMBIONTE EN EL FRIJOL COMUN  
(*Phaseolus vulgaris*).

T E S I S

QUE PARA OBTENER EL GRADO ACADEMICO DE

DOCTOR EN CIENCIAS  
(BIOLOGIA)

P R E S E N T A:

M.en C. VICTOR MANUEL VALDES LOPEZ

DIRECTOR DE TESIS: DR. FEDERICO ESTEBAN SANCHEZ RODRIGUEZ

MEXICO, D. F.

1999

1999



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.

**d e d i c a t o r i a**

**A TODOS AQUELLOS CON QUIENES HE IDO RECORRIENDO EL CAMINO DE LA VIDA**

**GRACIAS FEDERICO**



**A CECILIA, MI MUJER DE TODA LA VIDA**

**A JULIAN**

**A EMILIO**

• TAMIZADO DE LA BIBLIOTECA DE HOJA Y COMPARACION DE SECUENCIAS DE LAS CLONAS AISLADAS	73
• TAMIZADO DE LA BIBLIOTECA DE NODULO	74
• SECUENCIACION Y ANALISIS DE LA CLONA DE NODULO	74
• DETERMINACION DEL PATRON DE EXPRESION (NORTHERN BLOT)	75
• SOUTHERN BLOT GENOMICO. CARACTERIZACION DEL GEN DE CHAPERONINA	75
• MATERIAL Y METODOS	77
• AISLAMIENTO DEL DNA DEL PLASMIDO	77
• TAMIZADO DE BIBLIOTECAS DE cDNA	77
• INDUCCION DE NODULOS	78
• AISLAMIENTO DE DNA DE PLANTA	78
• SOUTHERN BLOTS GENOMICOS	79
• NORTHERN BLOT	79
• OBTENCION Y ANALISIS DE LAS SECUENCIAS DE DNA	79
• OTROS METODOS DE DNA RECOMBINANTE	80
• RESULTADOS Y DISCUSION	81
• DETERMINACION DE LAS CONDICIONES DE HIBRIDACION. SOUTHERN GENOMICO DE <i>Phaseolus vulgaris</i> CON LA SONDA cpn60- $\beta$ DE <i>Arabidopsis thaliana</i>	81
• TAMIZADO DE LA LIBRERÍA DE HOJA DE FRIJOL CON LA SONDA cpn60- $\beta$ DE <i>Arabidopsis thaliana</i>	81
• SUBCLONACION DEL INSERTO DE LA CLONA DE HOJA EN EL pBLUESCRIPT SK	82
• SECUENCIA Y COMPARACION DE LA CLONA CHAPPHAS-6.1 DE HOJA DE FRIJOL	85
• TAMIZADO DE LA BIBLIOTECA DE NODULO UTILIZANDO COMO SONDA LA CLONA CHAPPHAS-6.1	87
• SECUENCIA Y COMPARACION DE LA CLONA CHAPNOD-4.2 DE NODULO DE FRIJOL	88
• TAMIZADO SECUNDARIO DE LA BIBLIOTECA DE NODULO PARA AISLAR UNA CLONA COMPLETA	90
• RESTRICCION DE LAS CLONAS PARA VISUALIZAR EL TAMAÑO DE LOS INSERTOS	92
• SECUENCIA DE LA CLONA COMPLETA AISLADA DE LA BIBLIOTECA DE NODULO (CHAPNOD-5.1)	94
• CARACTERISTICAS GENERALES DE LA SECUENCIA:	
• RNY	96
• UTILIZACION DE CODONES	97

	PAGINA
• CARACTERISTICAS DEL PEPTIDO LIDER	101
• COMPOSICION DE AMINOACIDOS	103
• HIDROPATIA	105
• COMPARACION DE LA SECUENCIA DERIVADA DE AMINOACIDOS DEL PEPTIDO MADURO	107
• DETERMINACION DEL PATRON DE EXPRESION DE LA CLONA CHAPNOD-5.1 (NORTHERN BLOT)	109
• COMPARACION DE LA SECUENCIA DE LAS CLONAS DE HOJA Y NODULO DE FRIJOL	111
• CARACTERIZACION GENOMICA DE <i>cpn60-β</i> DE <i>Phaseolus vulgaris</i>	112
• CONCLUSIONES FINALES	116
• PERSPECTIVAS	116
• BIBLIOGRAFIA	117

---

## RESUMEN

Las proteínas denominadas de stress, tienen la función de ayudar a que los polipéptidos en las células adquieran o recuperen su conformación tridimensional nativa. Dentro de estas, las Hsp60s / Cpn60 (chaperoninas), se caracterizan por formar grandes oligómeros de 14 subunidades las cuales forman estructuras cilíndricas con una cavidad interior donde se ha propuesto que se lleva a cabo el plegamiento de los polipéptidos substratos. Si bien estas chaperonas moleculares pueden sobreexpresarse en respuesta a factores ambientales o biológicos, normalmente forman parte del repertorio constitutivo de proteínas celulares. Además, durante procesos normales de diferenciación celular, las Hsp60s también pueden aumentar su concentración intracelular, particularmente cuando concomitantemente con los procesos de desarrollo, existe una elevada tasa de síntesis de proteínas.

Durante la interacción simbiótica de las bacterias del género *Rhizobium* con las raíces de las leguminosas, ocurren procesos de diferenciación y morfogénesis, que culminan con la formación de estructuras denominadas nódulos. Eventualmente, la bacteria penetra al interior de las células de la raíz diferenciándose en los llamados bacteroides. Este proceso requiere que tanto la bacteria como la planta expresen genes específicos. Los genes que expresa la planta son denominados nodulinas. La culminación de esta interacción llega en el momento en que se inicia la fijación del nitrógeno.

Ha sido demostrado que en diferentes eventos de endosimbiosis, las chaperoninas juegan un papel importante. Por ejemplo, cuando *Salmonella* infecta macrófagos se sobreexpresa su Hsp60. Una bacteria no identificada que forma una endosimbiosis con *Amoeba*, manifiesta señales de stress y particularmente mantiene un alto nivel de expresión de Hsp60. Uno de los casos más dramáticos es el de las bacterias endosimbiontes del género *Buchnera*, que se localizan en células especializadas de insectos áfidos. Estas bacterias prácticamente solo sintetizan una proteína: la simbionina, la cual es un homólogo de Hsp60.

Dentro del género *Rhizobia*, también existen reportes similares. Cuando *Bradyrhizobium japonicum*, establece la interacción simbiótica, esta bacteria sobreexpresa sus Hsp60s. Además, en *Bradyrhizobium* su Hsp60 está codificada por una pequeña familia multigénica. Uno de los miembros de esta familia presenta un patrón de regulación dependiente de la activación de la fijación del nitrógeno. Por su lado, uno de los genes de Hsp60 en *Rhizobium meliloti*, está involucrado en la capacidad funcional del activador transcripcional NodD. Dicho de otro modo, existe bastante evidencia que relaciona a las chaperoninas bacterianas con eventos

endosimbióticos dentro de células eucariontes, tanto animales como vegetales. Sin embargo, no existen reportes desde el punto de vista del citoplasma huésped.

Así, nos hemos preguntado, si durante la morfogénesis del nódulo y la infección bacteriana, la planta también requiere de estas chaperonas moleculares.

Usando una sonda homóloga de *cpn60-β*, de *Arabidopsis thaliana*, se efectuó un tamizado de una librería de cDNA de hoja de *Phaseolus vulgaris*, de la cual se aislaron varias clonas homólogas de esta chaperonina. Una de éstas, fue utilizada a su vez como sonda para tamizar ahora una librería de cDNA de nódulos de esta misma planta. Nuevamente, se aislaron varias clonas. Se escogió la de mayor tamaño para secuenciarla, estableciéndose que corresponde a la *cpn60-β*. Por la similitud con diferentes secuencias de planta y de otras especies, el cDNA aislado corresponde a la chaperonina que se localiza en el plástido de la célula vegetal. De manera inesperada al efectuar los alineamientos, se observó que la secuencia de las clonas aisladas de hoja y de nódulo son idénticas. Asimismo, el patrón de restricción e hibridaciones de tipo Southern utilizando la región 3' no traducida de la clona de nódulo (región gen específica), indican que todas las clonas aisladas son idénticas. Para ampliar este resultado, se efectuaron Southern genómicos, los cuales son compatibles con la hipótesis de que en el frijol esta proteína está codificada por un solo gen. Este resultado contrasta con el reporte de que en *Arabidopsis thaliana* esta *cpn60-β*, está codificada por una familia multigénica de al menos cuatro miembros. Adicionalmente, experimentos de tipo northern, señalan que este gen se expresa a niveles basales en la raíz pero se sobrexpresas a medida que avanza la formación de los nódulos con un máximo en el día 12, decreciendo posteriormente. Durante estas etapas se sintetizan varias proteínas importantes para el metabolismo del nódulo, cuya localización es plástidica.

## INTRODUCCION

### EL STRESS CALORICO (HEAT SHOCK)

Desde que fuera descrito por el genetista napolitano Ferruccio Ritossa en 1962 (Ritossa, 1962), la respuesta al stress calórico (**heat shock**) ha fascinado a generaciones de biólogos (Hightower, 1991). Originalmente se describió como el cambio dramático que sufrían los cromosomas politénicos de las glándulas salivales de *Drosophila busckii* cuando éstas eran sometidas a un aumento de temperatura. Como respuesta a este tratamiento se podían observar en los cromosomas notables engrosamientos que se denominaron "puffs". Eventualmente, se demostró que estos puffs, eran RNA, lo cual claramente indicaba un incremento en la transcripción de estas regiones cromosómicas (Ashburner & Bonner, 1979). Dicho de otro modo, los puffs representan genes cuya expresión es inducida por un factor físico como el calor. Las posibilidades de estudiar la regulación de la expresión génica en estos puffs abrieron muchos caminos a la imaginación (Lindquist, 1986).

Necesariamente, la transcripción en los puffs, tendría que estar acompañada de la traducción de los mRNA's producidos. Así, en 1973 Alfred Tissières y Herschel K. Mitchell demostraron, que como resultado del stress, además de los puffs, se producían selectivamente solo ciertos polipéptidos suspendiéndose la síntesis de la mayor parte de los polipéptidos celulares (Tissières, et al., 1974). Estas proteínas se presentan en una amplia gama de pesos moleculares y fueron denominadas como Heat Shock Proteins (**Hsps**); (proteínas de choque térmico). Actualmente han sido clasificadas como de alto peso molecular (HMW-Hsps), (alrededor de los 100-90 kDa); Hsp70s (alrededor de 70 kDa); Hsp60s (alrededor de 60 kDa) y de bajo peso molecular (LMW-Hsps), (alrededor de 15-30 kDa) (Lindquist & Craig, 1988). Esta respuesta al stress calórico no es exclusiva de *Drosophila*, sino es universal. Ha sido observada en todos los organismos en los que ha sido buscada, de eubacterias a arqueobacterias y del ratón al frijol. Se presenta en prácticamente todos los tejidos y todos los tipos celulares de organismos pluricelulares, así como en explantes de tejidos y en células en cultivo (Lindquist, 1986). Sin embargo, vale la pena señalar dos aspectos. Primero, dependiendo del organismo, las Hsps que específicamente se inducen pueden variar (Lindquist 1981). Segundo, las Hsps son inducidas a temperaturas muy diferentes dependiendo del organismo, pero siempre la temperatura de inducción refleja las condiciones normales a las que vive el organismo (Huey & Bennett 1990). Por ejemplo, *Escherichia coli* induce sus Hsps al pasar de 37 °C a 50 °C; la levadura al pasar de 25 °C a 39 °C; *Drosophila* al pasar de 25 °C a 36.5 °C y en general las plantas adaptadas a temperaturas ambientales incluyendo cultivares de soya, chícharo, maíz y trigo comienzan a sintetizar Hsps cuando la temperatura excede 32-33 °C. Por su

lado, las bacterias hipertermofílicas inducen sus Hsps cuando pasan de 95 °C a 105 °C y los peces árticos cuando pasan de 0 °C a 5-10 °C (Parsell & Lindquist 1993). Dicho de otro modo, los organismos inducen sus Hsps cuando la temperatura se incrementa arriba de lo que es considerado normal para ese organismo particular, de manera que no existe un umbral universal de temperatura. Por ejemplo en vertebrados el desarrollo de hipertermia, fiebre y la inducción de la síntesis de Hsps, dependerá de si el organismo controla su temperatura o de si depende del medio ambiente para regularla (Kluger 1990). Otro aspecto importante, es que estas proteínas se sobreexpresan por efecto del stress; lo cual significa que normalmente existe un nivel basal de éstas. Dicho de otro modo, forman parte del repertorio cotidiano de genes que se expresan de manera constitutiva (house-keeping genes) (Ang et al., 1991).

---

**TABLA 1.**

**ALGUNAS CONDICIONES QUE INDUCEN LA EXPRESION DE PROTEINAS DE STRESS (Welch, 1993)**

---

**FACTORES AMBIENTALES**

- STRESS POR CALOR
- INHIBIDORES DEL METABOLISMO ENERGETICO
- METALES PESADOS DE TRANSICION
- ANALOGOS DE AMINOACIDOS
- AGENTES QUIMIOTERAPEUTICOS

**ENFERMEDADES**

- INFECCIONES VIRALES
- FIEBRE
- INFLAMACION
- ISQUEMIA
- HIPERTROFIA
- DAÑO OXIDATIVO
- MALIGNIDAD

**FACTORES CELULARES NORMALES**

- CICLOS DE DIVISION CELULAR
  - FACTORES DE CRECIMIENTO
  - DESARROLLO Y DIFERENCIACION
- 

Sin embargo, durante algún tiempo, la función específica de estas proteínas fue un misterio. No fue sino hasta que se demostró que estas Hsps se inducían por otros factores, tanto bióticos como físicos, además del calor, que se comenzó a esbozar cual podría ser el papel de estas proteínas (Welch, 1993). Hoy sabemos que además del calor, en diferentes organismos, la repuesta de stress

puede ser disparada posteriormente a la anoxia (hipoxia) , 2,4 dinitrofenol, azida de sodio, arsenito de sodio, etanol, peróxido de hidrógeno, intoxicación por metales pesados de transición (cadmio, zinc, mercurio), análogos de aminoácidos, neoplasia, fiebre, infecciones virales e inclusive por salicilatos y prostaglandinas (**Hightower, 1991; Lindquist, 1986**). En la **Tabla 1**, se listan de manera genérica algunos de estos factores y además se destaca que la síntesis de proteínas de stress puede ser disparada por fenómenos de diferenciación celular , morfogénesis y desarrollo (**Bond & Schlesinger, 1987**).

En 1980, Lawrence E. Hightower notó que lo que tienen en común algunos de los factores (listados en la **tabla 1**), es que en potencia, son factores que podrían actuar como agentes desnaturalizantes de proteínas (**Hightower, 1980**). Con una anticipación visionaria, Hightower en consecuencia propuso que la acumulación de proteínas anormales o desnaturalizadas en una célula sería el gatillo que dispararía la repuesta de stress. Algunos años después, esta propuesta de Hightower comenzó a ser confirmada cuando se demostró que la inyección de proteínas desnaturalizadas en células procariontes o eucariontes, efectivamente activaban a los genes de heat-shock (**Goff, & Goldberg, 1985; Ananthan, et al., 1986**). Por esas fechas Hugh R. B. Pelham de manera independiente, llegó a conclusiones similares a las de Hightower, proponiendo en particular que la Hsp70, la Hsp90 y las Grps (Glucose-regulated proteins) podrían de alguna manera, reconocer proteínas intracelulares desnaturalizadas y reconvertirlas a su estructura biológica correcta (**Pelham, 1986**). Mas recientemente, se ha sugerido que no es la desnaturalización per se lo que dispara la respuesta de stress, ya que no todas la proteínas desnaturalizadas inyectadas la estimulan. Mas bien, se ha propuesto que es la posibilidad de formar agregados lo que distingue a las proteínas desnaturalizadas que inducen la respuesta y se ha demostrado que la formación de grandes agregados se correlaciona con la magnitud de la respuesta (**Mifflin & Cohen 1994**). En este sentido, desde luego el calor ha sido uno de los factores mas utilizados para estudiar la desnaturalización y la renaturalización de proteínas. Asimismo, los procesos de diferenciación, habitualmente van acompañados por un incremento en la síntesis de nuevos polipéptidos, algunos de los cuales son sintetizados de manera masiva. Así entonces, se propuso que la función de las Hsps sería, por un lado ayudar a que si por alguna razón, los polipéptidos celulares perdieran su conformación nativa, estas Hsps participarían ayudando a que otras proteínas recuperen su conformación tridimensional funcional, protegiendo así las funciones celulares. Por otro lado, en etapas durante las cuales se incrementa la tasa de síntesis de proteínas, las Hsps estarían presentes ayudando a que las proteínas recién sintetizadas adquieran su conformación nativa correcta. Adicionalmente, hay ocasiones en las que algunos polipéptidos deben de atravesar membranas de uno a otro lado. Desde luego en este caso, el

polipéptido al cruzar la membrana pierde su conformación nativa y debe recuperarla al llegar a su destino (**Vierling, 1991**).

## **SECUENCIA, ESTRUCTURA Y FUNCION DE PROTEINAS**

Justificadamente, se puede decir que el plegamiento de una proteína es un evento de importancia central en biología (**Ellis, 1994a; Maddox, 1994a**). De hecho, la adquisición de la estructura tridimensional de una proteína, es una importante extensión implícita en el Dogma Central de la Biología Molecular (**Crick, 1958; Ellis, 1994a**). Formulada en su forma más simple, la traducción de la secuencia de nucleótidos del gen, en la secuencia lineal de residuos de aminoácidos en la cadena polipeptídica, aporta la información necesaria para que al adoptar su estructura tridimensional, el polipéptido adquiera su función biológica. El problema del plegamiento de un polipéptido fue reconocido hace más de medio siglo por Anson y Mirsky quienes en 1925 observaron que la desnaturalización es un proceso reversible (**Rose & Wolfenden, 1993**). Posteriormente, en 1936, Mirsky y Linus Pauling hipotetizaron que las proteínas nativas tenían una estructura característica, la cual se pierde con la desnaturalización (**Mirsky & Pauling, 1936**). Estas ideas culminaron finalmente con la demostración experimental de este paradigma central de la biología molecular por Christian Anfinsen en la década de los sesenta (**Anfinsen, 1973**). En sus experimentos clásicos, Anfinsen purificaba una proteína, la sometía a algún agente desnaturizante (con lo cual la proteína perdía su función), después removía el agente desnaturizante y si ponía a la proteína en condiciones favorables de temperatura, pH y fuerza iónica, después de un tiempo, la proteína recuperaba su función, indicando desde luego que había readquirido su conformación tridimensional nativa de manera espontánea. A partir de esos experimentos, decimos entonces, que la estructura tridimensional nativa funcional de una proteína es aquella que adopta en condiciones estándar de temperatura, pH y fuerza iónica, y que, en la secuencia de residuos de aminoácidos de esta proteína existe la información biológica para que este plegamiento se logre (**Rose & Wolfenden, 1993**). Sin embargo, aun desde las épocas de Anfinsen, había algunos problemas. Por ejemplo: estos experimentos en general funcionaban con polipéptidos pequeños - de alrededor de 100-150 residuos de aminoácidos como la ribonucleasa, las globinas o la insulina- (**Wang & Tsou, 1991**). Cuando se trataba de repetir estos experimentos con polipéptidos mayores, como la albúmina o la ovoalbumina (más de 600 residuos de aminoácidos), ésta nunca recuperaba ni su estructura, ni su función, sufriendo lo que se denominó una desnaturalización irreversible (**Pace, 1990**). Dicho de otro modo, la coagulación de la ovoalbumina, es una desnaturalización irreversible. Estos resultados planteaban un problema, ya que si se estaba proponiendo un paradigma, este debería de ser

de aplicación universal y no limitado a algunos polipéptidos. Sería como proponer que una proteína debería tener un "estado de gracia" para que la renaturalización ocurriera. La solución a esta aparente paradoja, está desde luego en la forma como un polipéptido es normalmente sintetizado en los ribosomas.

En la estructura tridimensional de una proteína, se dan contactos e interacciones de regiones que en la estructura primaria están alejadas, (a veces más y a veces menos). Estas son interacciones estereoespecíficas no covalentes que pueden ser iónicas, puentes de hidrógeno, interacciones de van der Waals y desde luego interacciones hidrofóbicas (**Rose & Wolfenden, 1993**). Estas desde luego, se dan entre aminoácidos particulares, los cuales se pueden encontrar en otras regiones del polipéptido, de tal manera, que en realidad existen un gran número de interacciones posibles entre los diferentes aminoácidos, y estas posibles asociaciones aumentan con el peso molecular del polipéptido. El plegamiento de un polipéptido puede ser representado de manera esquemática como una serie de estados intermediarios que la proteína va adoptando como si fuera entrando en un túnel rugoso (**Wolynes, et al., 1995**). Antes de plegarse correctamente, un polipéptido puede adoptar un gran número de estados intermedios, algunos de los cuales pueden actuar como trampas de arena disminuyendo la velocidad y las probabilidades de llegar a la conformación tridimensional correcta. Así, cuando se trabaja con un polipéptido no muy grande, las interacciones incorrectas pueden modificarse en un lapso razonable de tiempo, hasta llegar a las interacciones que definen el estado de menor energía, que es lo que llamamos la conformación nativa. George Lorimer, acostumbra referirse a estas interacciones, como posibilidades termodinámicas alternativas de plegamiento que puede adquirir un polipéptido en determinadas condiciones y destaca que algunas de estas conformaciones pueden ser difíciles de rescatar hasta llegar a la conformación de menor energía (**Lorimer, 1996a**). Al aumentar la longitud de la cadena polipeptídica entonces, el número de conformaciones incorrectas es tan grande que para fines prácticos se llega a la desnaturalización irreversible. Nuevamente, ¿Cómo entonces los polipéptidos de alto peso molecular adquieren en la célula su conformación tridimensional correcta?

La respuesta consiste en que las proteínas, grandes o pequeñas, van adquiriendo su conformación tridimensional al ir siendo sintetizadas en los ribosomas. De esta manera, las diferentes interacciones entre los aminoácidos se reducen, ya que por ejemplo, un dominio en la región amino terminal, quedará plegado aun antes de que los aminoácidos correspondientes a la región carboxilo terminal sean incorporados en la cadena (**Gething & Sambrook, 1992**). El mismo tipo de argumento se podría aplicar a una proteína que va a cruzar de un compartimento celular a otro (**Hendrick & Harli, 1993**). De igual manera, en este caso la cadena polipeptídica, al ir atravesando la membrana se va plegando paulatinamente.

De cualquier manera, es curioso, que en términos generales, tanto en sistemas procariontes como eucariontes, el peso molecular de los polipéptidos individuales, se mantenga en lo que consideraríamos un intervalo de bajo peso molecular. Por ejemplo los proyectos de secuenciación de los genomas de *Escherichia coli*, *Haemophilus influenzae*, *Mycoplasma genitalium*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Mus musculus*, *Arabidopsis thaliana* y *Homo sapiens*, muestran que el tamaño **promedio** de los correspondientes polipéptidos se ubica alrededor de los 30 kDa, i.e. alrededor de 300 residuos de aminoácidos (Koonin, et al., 1996; Friedrich, 1996). Desde luego que en estos sistemas se encuentran polipéptidos de mayor tamaño, algunos incluso de mucho más de 1.000 residuos de aminoácidos, pero la tendencia general es que si una célula va a usar una proteína de gran tamaño, lo usual es que se integre por el ensamble de varios polipéptidos, iguales o diferentes de menor tamaño. En este respecto, es notable que en *E. coli*, la distribución mayoritaria de tamaños de sus polipéptidos sea trimodal con máximos en los 15, 30 y 45 kDa. Tal vez estos resultados indican que el casete estructural-informacional mínimo, corresponde a alrededor de 150 residuos de aminoácidos (Savageau, 1986). Tal vez este sería el tamaño "ideal" para lograr el menor número de problemas durante el plegamiento de las cadenas polipeptídicas. Es notable, que al observar el plegamiento tridimensional de polipéptidos de "alto peso molecular", estos se pliegan en dominios y subdominios estructurales que inclusive pueden tener independencia funcional (por ejemplo, vide *infra* el caso particular de la estructura tridimensional de GroEL, una Hsp60).

Una hipótesis relevante en este sentido tiene que ver con la selección evolutiva de los polipéptidos utilizados por los sistemas biológicos desde su origen hasta el presente (un período aproximado de  $3.5 \times 10^9$  años). Si consideramos **300 residuos de aminoácidos** como el tamaño promedio de los polipéptidos actuales y si asumimos que estas 300 posiciones van a ser ocupadas por los 20 aminoácidos habituales en cualquier orden posible, tendríamos que el número total de secuencias diferentes de aminoácidos sería  $20^{300}$  o sea  $2.03 \times 10^{390}$ , que desde luego es un número muy grande. Este número ha sido frecuentemente discutido en la literatura, para ejemplificar la enorme potencialidad informacional de los sistemas genéticos. Sin embargo, otro enfoque podría consistir en preguntarnos cuantas de todas esas secuencias podrían tener un significado biológico. Esto es, para que una secuencia tenga un valor de selección debe de ser capaz de tener un papel funcional o estructural que sea objeto de selección natural. Inmediatamente, nos podemos imaginar una serie de secuencias que no tendrían esta capacidad. Por ejemplo 300 veces leucina, no parecería tener mucho atractivo; 300 veces aspártico-glutámico tampoco. ¿Dónde comenzarían las secuencias que pudieran tener un papel biológico? La respuesta la podemos obtener de los propios sistemas biológicos. Un procarionte

típico como *E. coli* tiene alrededor de **5,000 genes**. Un eucarionte típico como *H. sapiens*, tiene alrededor de **80,000**. Suponemos que hay variaciones en el número de genes tanto en procariontes como en eucariontes y desde luego, estos dos organismos son típicos desde nuestro punto de vista. Así, podríamos decir que en **procariontes** el número de genes promedio podría ser de **2,500** y en **eucariontes** de **20,000** (Bird, 1995). La siguiente pregunta es cuántas especies procariontes y eucariontes existen en la actualidad. Desde luego, nadie tiene una respuesta definitiva. Sin embargo, podemos hacer una aproximación (May & Nee, 1995). Primero, tendríamos que suponer que existen mas especies procariontes que eucariontes, digamos **5 a 1**. Segundo, solo de **insectos** se asume que hay aproximadamente **5 x 10<sup>6</sup>** especies. Estos organismos son los animales mas abundantes y tal vez representan al menos a la mitad de las especies animales. Entonces, tendríamos que suponer que debe de haber el doble de especies animales (May, 1992). Si asumimos que al menos debe de existir un número similar de plantas, finalmente tendríamos **20 x 10<sup>6</sup> especies eucariontes**. Si multiplicamos este número por cinco obtenemos **100 x 10<sup>6</sup> especies de procariontes**, para un total de **120 x 10<sup>6</sup> especies** presentes hoy en día en este planeta. Si ahora multiplicamos estos valores por el número de genes promedio que se calculo antes, nos resulta que en **eucariontes** debe haber:

$(20 \times 10^3) \times (20 \times 10^6) = 400 \times 10^9$  genes y en **procariontes**  $(2.5 \times 10^3) \times (100 \times 10^6) = 250 \times 10^9$  genes. La **suma total** sería de **650 x 10<sup>9</sup>**, que desde luego es un número muy pequeño (por grande que parezca). Ahora, habría que considerar cuantas especies ha habido desde el origen de la vida hasta el presente, lo cual desde luego nos lleva a calcular otro número imaginario. Sin embargo, en general se asume que el **99.9 %** de las especies están extintas (Crichton, 1995). Desde luego nadie sabe si el número de especies presentes en un momento dado se mantiene constante. Es probable que no. Pero podemos asumir lo contrario para fines de este cálculo. De cualquier manera nadie sabe cual es la tasa de extinción de las especies, si bien se ha dicho que cada día se extingue una especie, y que en general el tiempo promedio de permanencia de una especie en el planeta es de solo 4 millones de años (Crichton, 1995). Así, si multiplicamos cualquiera de estos números para calcular el valor total de genes (polipéptidos), tendríamos

**a)** Considerando 99.9 % de especies extintas  $(650 \times 10^9) \times 1,000 = 6.5 \times 10^{14}$ . **b)** Considerando una especie extinguida por día, asumiendo desde luego la aparición de una nueva especie que substituye a la extinta, de tal manera que la población total se mantiene constante tendríamos: (número promedio de genes x número de días totales en toda la historia de la vida en este planeta) =  $(5.416 \times 10^3) \times (1.2775 \times 10^{12}) = 6.92 \times 10^{15}$ . **c)** Considerando 4 millones de años para que se recambien totalmente las especies del planeta  $(650 \times 10^9) \times (3.5 \times 10^9 / 4 \times 10^6) = (650 \times 10^9) \times (875) = 5.6875 \times 10^{14}$ . Como se ve, estos valores (a, b, y c), coinciden aproximadamente en el orden de magnitud. Pero aunque asumiéramos que tenemos un error de seis ordenes de

magnitud, que es poco probable, e hiciéramos una predicción de que en toda la vida del planeta, en los sistemas biológicos ha habido del orden de  $1 \times 10^{20}$  diferentes genes; cuando contrastamos este valor con  $2.03 \times 10^{370}$ , que asumimos como el número total de polipéptidos de 300 residuos de aminoácidos, vemos que en el mejor de los casos los sistemas biológicos apenas y han utilizado una pequeñísima fracción ( $10^{-370}$ ), del supuesto potencial genético.

Podríamos en este momento repetir la pregunta, ¿que características tienen estas secuencias que las hacen tan especiales, de tal manera que han sido seleccionadas para ser usadas por los sistemas biológicos? Nuevamente, parte de la respuesta la dan los mismos sistemas biológicos.

En los cálculos anteriores, hemos obtenido números totales, haciendo la suposición de que cada una de estas secuencias es diferente de las demás. Dicho de otro modo, al tomar una pequeña muestra ( $10^{20}$ ) del gran conjunto de secuencias posibles ( $2.03 \times 10^{370}$ ), la probabilidad de que dos secuencias se parezcan es extraordinariamente baja ( $2.03 \times 10^{-370}$ ), de manera que al comparar estas secuencias, que supuestamente representarían una muestra al azar, no deberíamos de encontrar siquiera dos que se parecieran entre sí. Esto en realidad no es tan cierto. Cuarenta años de biología molecular que incluyen dos décadas de secuencia intensiva de genes a lo largo de la escala filogenética, que de hecho hoy en día incluye muestras de organismos (especies) extintas, nos han mostrado que los sistemas biológicos operamos bajo un fuerte principio de conservación.

A nivel de genes y sus polipéptidos codificados, lo que vemos es que evolutivamente son pocos los genes radicalmente nuevos que aparecen. Si consideramos los genes del repertorio básico de estructura y metabolismo, estos se encuentran desde los mycoplasmas, hasta el hombre (por tomar dos extremos), (Fraser, et al., 1995; Adams, et al., 1995). Desde luego existen genes específicos de grupos taxonómicos, pero lo más común es que encontremos genes equivalentes (homólogos) a todo lo largo de la escala filogenética. Desde luego, sabemos que ha habido momentos en la evolución en que ocurren incrementos notables en el número de genes. Los dos eventos más notables ocurrieron poco después de la aparición de los eucariontes y después de la aparición de los vertebrados (Bird, 1995). Sin embargo, aun en estos casos sabemos que la evolución a nivel molecular tiene mas que ver con la duplicación de genes que con el surgimiento de secuencias totalmente nuevas.

Se ha dicho que la mutación es la materia prima de la evolución. Sin embargo, ésta no actúa sobre secuencias vírgenes, sino sobre secuencias ya probadas de los mismos genes. Cuando un gen se duplica y la mutación actúa sobre una de las copias, es posible que aparezca un cambio que pueda modificar levemente la estructura y la función, aunque desde luego el resto de la secuencia se mantiene muy parecida. Hablamos entonces de genes parálogos. Este proceso ha derivado en la aparición en los genomas de lo que se denominan familias multigénicas de las

cuales hay muchísimos casos bien evidenciados como el de las globinas (mioglobina y hemoglobina), en vertebrados. Pero aun en organismos más sencillos aparecen estas reiteraciones. Por ejemplo en el nemátodo *Caenorhabditis elegans*, que tiene un pequeño genoma de 100 mega bases (100 millones de pares de bases) y donde se calcula que debe haber no mas de 13-14 mil genes el número de familias multigénicas es elevado (Hodgkin, et al., 1995). Aun en el caso de la arquea *Methanococcus jannaschii*, cuyo genoma de 1.66 mega bases acaba de ser secuenciado completamente, se han encontrado 100 familias multigénicas de las cuales la mayor contiene 16 miembros (Bull, et al., 1996). Este panorama de genes o secuencias que son utilizadas y vueltas a utilizar con pequeñas variaciones ha sido conceptualizado por Susumu Ohno bajo lo que el ha llamado el **prevalente principio de la recurrencia repetitiva**, (Ohno & Ohno, 1986), que no es otra cosa mas que puntualizar que posiblemente el número total de genes (polipéptidos) que se han utilizado en la biosfera en este planeta durante mas de 3,500 millones de años, no importa en cuantas especies, es pavorosamente pequeño, cuando se compara con los números presentados anteriormente. Es difícil decir cual es este número. Pero estos resultados nos sugieren entonces que este valor sería mucho menor que  $10^{14}$  o  $10^{20}$ . Algunos cálculos de Wally Gilbert, sugieren que este número podría ser tan pequeño como  $10^4$  o  $10^5$  (Dortt, et al., 1990).

Esto nos regresa de nuevo entonces a la pregunta inicial respecto a qué características tienen las secuencias de los polipéptidos que las han hecho tan especiales. Desde luego no lo sabemos. Sin embargo, hay algunas sugerencias, que por cierto tendrían que ver precisamente con las potencialidades de plegamiento tridimensional de las proteínas.

Cuando se sintetizan oligopéptidos con secuencias similares, o derivadas de genes reales, estos oligos en general tienden a adoptar estructuras termodinámicamente estables y a veces inclusive con algo de estructura secundaria. Por el contrario, cuando estos oligopéptidos son sintetizados usando secuencias lo mas al azar posible, que no tengan similitud con ninguna secuencia real conocida, estos péptidos presentan una gran inestabilidad termodinámica y desde luego, no forman ninguna clase de estructura secundaria (Li, et al., 1996; Kardar, 1996). Recientemente entonces, ha sido nuevamente propuesto, que la característica mas básica que llevó a la selección de los polipéptidos actuales (y por extensión de los pasados), es simple y sencillamente, su capacidad de plegamiento, que resulta en estructuras compactas termodinámicamente estables (Sali, et al., 1994; Baldwin, 1994). Estos postulados desde luego enfatizan el hecho de que el plegamiento de las proteínas es un evento central en la biología.

Ahora bien, regresando a los experimentos de Anfinsen, estos tienen un segundo problema además de la limitación del tamaño. Este problema, tiene que ver con una complicación que ha plagado a la bioquímica desde sus inicios (Ellis, 1996). Para descubrir características particulares

del comportamiento molecular, se acostumbra llevar a cabo la purificación de los componentes del sistema y montar lo que llamamos un ensayo *in vitro*, que era exactamente lo que hacía Anfinsen, montar un ensayo para demostrar el replegamiento de las proteínas, bajo condiciones experimentalmente controladas. Este es el problema. Las condiciones de tubo de ensayo, están notablemente alejadas de lo que ocurre en realidad en el interior de la célula. En particular, sabemos que en el citoplasma la concentración de proteína es de alrededor de 5 mM (**Hendrick & Hartl, 1993**). Esta es una altísima concentración de proteína con la que ningún bioquímico se atrevería a ensayar. De hecho los experimentos de Anfinsen se llevaban a cabo, con proteínas purificadas y notoriamente diluidas (en el rango nM o  $\mu$ M a lo más). Cuando Anfinsen aumentaba la concentración de proteína, aunque ésta fuera pequeña, después de desnaturalizar (esto sí se podía), la renaturalización se hacía imposible ya que se formaban grumos y coágulos (**Anfinsen, 1973**). Esto se ha interpretado como que al intentar replegarse, en concentraciones elevadas la proteína, parcialmente plegada puede tener un gran número de contactos e interacciones incorrectas ya no intramoleculares sino intermoleculares con un gran número de proteínas vecinas. En este segundo problema, es más difícil invocar las características de plegamiento paulatino del polipéptido en el ribosoma. Hasta donde sabemos, durante la traducción el ribosoma, los polipéptidos que están siendo sintetizados tienen contactos y están abiertos con el medio, donde desde luego además del solvente y iones hay gran cantidad de proteínas. En realidad, durante la síntesis del polipéptido, cuando aun no se ha sintetizado completamente y no se ha plegado, habría una gran cantidad de posibilidades de asociaciones equivocadas.

Se llega así, a la justificación de la función de las proteínas de Heat Shock como componentes esenciales del repertorio cotidiano de polipéptidos celulares. Si bien es cierto que la secuencia de aminoácidos, posee en potencia la información para determinar la estructura nativa de la proteína, en la célula el plegamiento de los polipéptidos requiere de la ayuda adicional de proteínas cuyo papel es colaborar a que esta estructura se logre (**Hendrick & Hartl, 1993**). Desde luego, además, están las circunstancias en las que diversos factores pueden inducir una pérdida o inestabilidad de la conformación tridimensional. Pero cabe destacar de manera central el papel cotidiano de las Hsps, las cuales justificadamente entonces son denominadas chaperonas moleculares (*vide infra*).

¿Cómo logran desempeñar su función las diferentes clases de chaperonas moleculares? Este es uno de los problemas centrales que aún no han sido resueltos. Además, cada clase de Hsps difiere de la otra y en realidad hoy no sabemos cuantas clases en total podría haber. Pero cada una desempeña su papel de manera diferente. Tal vez, esta diversidad de mecanismos y estrategias para mantener la conformación nativa de los polipéptidos tiene que ver con lo que quedó señalado anteriormente. La conformación tridimensional nativa es una de tantas

conformaciones que puede adquirir la proteína. Además, ésta es notablemente sensible a pequeños cambios de temperatura pH y fuerza iónica. Si por alguna razón la proteína altera su conformación requiere ayuda para recuperarla.

## **PATOLOGIAS MOLECULARES RESULTANTES DE PLEGAMIENTOS ERRÓNEOS**

El Dogma Central de la Biología Molecular, es en su esencia un postulado de tipo negativo que establece que una vez que la información ha llegado a la proteína ya no puede salir de ahí (Crick, 1958). En mas detalle, los flujos de información ácido nucleico-ácido nucleico y ácido nucleico-proteína son posibles, mientras que los flujos de información proteína-ácido nucleico y proteína-proteína son imposibles.

Estos principios fueron enunciados por primera vez por Francis Crick en 1958 y se convirtieron en una piedra angular de la biología molecular desde entonces. Sin embargo, el Dogma Central, frecuentemente ha sido mal interpretado en el sentido de DNA hace RNA, RNA, hace proteína. En 1970, con motivo del descubrimiento simultáneo en los laboratorios de Howard Temin y David Baltimore, de la transcriptasa reversa, se dijo que el Dogma central había sido violado y que en vista del descubrimiento de que el flujo de información RNA a DNA era posible, el Dogma se podría considerar como una sobre-simplificación. Desde luego, que el Dogma no es violado por dicho flujo de información utilizado por los retrovirus ya que cae dentro de los flujos ácido nucleico-ácido nucleico. Sin embargo, en aquel entonces, el descubrimiento de ese flujo de información mediado por la transcriptasa reversa, motivó, que Francis Crick, escribiera un nuevo artículo en Nature donde expuso de manera más extensa las características y postulados del Dogma (Crick, 1970). En aquel artículo, como en el primero, Crick establece que el Dogma es un postulado de tipo negativo. Adicionalmente, en el nuevo artículo, clasifica a los flujos de información posibles en **generales** (replicación, transcripción y traducción) y **especiales** (retrotranscripción, replicación de RNA y el flujo no descrito hasta la fecha DNA a proteína). Desde luego enfatiza que si alguno de los flujos **prohibidos** por el Dogma llegara a descubrirse, cuestionaría hasta la raíz los conceptos derivados de éste. Es curioso, que en aquel artículo de 1970, Crick mencione como un problema el hecho de que se desconociera aún la naturaleza física del agente infeccioso de la enfermedad denominada "scrapie" (que se podría traducir como "rasquiña"). Esta, es una enfermedad neurodegenerativa, que afecta a las ovejas. Tiene una evolución muy lenta, (el período de incubación desde la infección hasta que aparecen los síntomas puede ser de años); por lo que se decía que era producida por un virus lento, si bien los intentos de determinar el tamaño de la partícula infecciosa, indicaban que tenía un peso molecular menor que el de cualquier otro virus que hubiera sido descrito (Mestel, 1996). La

enfermedad es incurable y una vez que se manifiestan los síntomas de la enfermedad ésta progresa con relativa rapidez, causando la muerte del animal en pocos meses. Un problema asociado, y que es de donde la enfermedad toma su nombre, es que las ovejas desarrollan una gran comezón que las impulsa a rascarse compulsivamente. Esta conducta anormal trae como consecuencia el daño irreparable a los vellones. Desde el punto de vista comercial esto es una tragedia, ya que los criadores de ovejas para producir lana, saben que cuando aparecen en los rebaños algunas ovejas infectadas, se tiene que sacrificar al resto, ya que tarde o temprano todo el rebaño va a presentar la enfermedad. Otra característica de esta enfermedad es que se transmite principalmente por vía oral. En el caso de las ovejas, cuando éstas están infectadas y paren en el campo, los restos de la placenta quedan dispersos en las hierbas que les sirven como alimento, lo cual lleva a que la infección se propague (**prion home page**). Mas recientemente, a partir de mediados de los años ochenta, ha aparecido, principalmente en Inglaterra y Europa otra enfermedad, el llamado "mad cow disease" (síndrome de las vacas locas). Esta encefalitis espongiiforme de los bovinos, también ha alcanzado niveles epidémicos causando enormes pérdidas económica (**Prusiner, 1995**). Para evitar la propagación de la infección, enormes cantidades de vacas infectadas son cremadas. En este caso, la prensa, ha contribuido a que se desate el pánico, por la posibilidad de que la enfermedad se propague al humano. Estas ideas han aparecido en parte porque aunque no se sabe a ciencia cierta, aparentemente en su inicio esta enfermedad en los bovinos se generó porque el ganado fue alimentado con un complemento alimenticio que en parte consistía de carne y hueso provenientes de óvidos infectados de scrapie (**Nota:** Bóvido: Miembro de la familia *Bovidae*, incluye al ganado vacuno, borregos y cabras. Bovino: por ejemplo el buey y la vaca. Óvido: mamífero bóvido del género *Ovis*; *Ovis aries* el borrego). Esto significa, que las partículas infecciosas, pueden atravesar las barreras específicas. Sin embargo, a la fecha no se ha demostrado ninguna asociación entre la aparición y propagación de la enfermedad en los bovinos y una mayor incidencia de este tipo de enfermedades en los humanos. Tal vez la mejor razón para suponer que esta enfermedad no pasara fácilmente al hombre, es que éste, los óvidos y los bóvidos compartimos el último ancestro común hace aproximadamente 70 millones de años, mientras que las ovejas y el ganado vacuno se separaron apenas hace menos de 20 millones de años. Esta clase de enfermedades neurodegenerativas, parece ser particular de los vertebrados, y si bien son poco comunes, también en el humano han sido descritas. Dentro de las patologías de este tipo en el humano, se conocen el Kuru, el síndrome de Creutzfeldt-Jakob (**CJD**), el síndrome de Gerstmann-Sträussler-Scheinker (**GSS**) y el Insomnio Familiar Fatal (**FFI**) (**Bessen, 1996**). El Kuru, llamó mucho la atención de investigadores como Vincent Zigas y D. Carleton Gajdusek. El Kuru solo se ha descrito entre la tribu Fore en Papua, Nueva Guinea, donde se le llama "la muerte sonriente". Esta extraña

enfermedad se caracteriza por una falta de coordinación (ataxia), que a menudo es seguida por demencia y siempre es fatal. En este caso, la infección ocurría después del canibalismo ritual de despojos de miembros de la tribu que habían fallecido de esta enfermedad neurodegenerativa. La incidencia de ésta era mayor entre las mujeres y los niños, ya que estos consumían el cerebro, mientras que los hombres, que consumían el músculo, presentaban una menor incidencia de este síndrome (**Mestel, 1996**). Actualmente, gracias a medidas profilácticas obvias, implantadas en parte por Gajdusek y los que le siguieron, esta enfermedad prácticamente ha desaparecido. En el caso del **CJD**, **GSS** y otras enfermedades similares en humanos, parece que el principal componente es genético y solo ocasionalmente se puede desarrollar la infección en el humano por contagio. Una de las pocas posibles fuentes de infección es el transplante de cornea, implantes de dura madre o electrodos en el cerebro, uso de instrumental quirúrgico contaminado o inyección de hormona del crecimiento derivada de la glándula pituitaria de humanos enfermos. Este último problema se ha reducido por el uso de esta hormona recombinante (**Prusiner, 1995**). Como sea, el principal interés en el aislamiento y caracterización del agente infeccioso, no se derivó de las enfermedades en el humano, ya que como problema de salud son enfermedades menores, circunscritas, por lo general, a pequeños grupos, poblaciones y familias. Por el contrario, como quedó indicado más arriba, en los países como Nueva Zelanda, Australia y el Reino Unido, donde la cría de ovejas para la producción de lana es una de las principales empresas nacionales, la investigación de la etiología, la caracterización del agente infeccioso, la prevención y la posible cura del scrapie adquieren carácter de prioridad nacional. Es por esta razón que los grupos de investigación en esta área han podido tener un gran apoyo económico. El primer avance importante ocurrió a inicios de la década de los ochenta, cuando Stanley Prusiner logró aislar al elusivo agente infeccioso (**Prusiner, 1982**). La purificación del agente infeccioso, recuerda mucho los experimentos realizados en 1944 cuando Oswald Avery, Colin McLeod y Mclin McCarty, reportaron la purificación del agente transformante. Los experimentos clásicos de Avery *et al.*, llevaron a la demostración de que la molécula de DNA, era la portadora de las características hereditarias (**Avery, MacLeod & McCarty, 1944**). Utilizando una serie de métodos y estrategias experimentales, en 1982, Prusiner logró purificar al agente infeccioso. El resultado más inesperado fue que dicho agente consistía exclusivamente de una proteína de un peso molecular aproximado de alrededor de 27-30 kDa. Por las características de ser proteináceo e infeccioso, Prusiner bautizó a esta proteína como **prión** (**Prusiner, 1982**). Desde luego el problema más crítico que planteaba el descubrimiento de los priones, era el de su multiplicación. Se postularon varias hipótesis para explicarla. De estas la que destacaba por sus implicaciones, era la herética posibilidad de que la proteína del prión tuviera la capacidad para codificar ya fuera a otra proteína igual o a un segmento de ácido

nucleico (DNA o RNA), que fuera posteriormente descodificado por una de las vías tradicionales (Prusiner, 1982). Dicho de otro modo estas dos posibilidades de ser demostradas en definitiva violarían el Dogma Central de la Biología Molecular. Esta posibilidad desencadenó una gran cantidad de especulación y desde luego hizo que los priones (y Prusiner), saltaran a la fama internacional. Sin embargo, por mucha publicidad que tuviera el asunto, juiciosamente, Prusiner decidió explorar posibilidades experimentales más concretas. Así, en 1985, Prusiner y su grupo, en asociación con Leroy Hood y con Charly Weissmann, publicó el aislamiento del gen del príon (Oesch, et al., 1985). Previamente, Prusiner había logrado reproducir un modelo experimental del scrapie en el hámster, que tenía ventajas en particular en cuanto al tiempo que tarda en desarrollarse la enfermedad. A partir de cerebros de hámsters infectados, Prusiner y sus colaboradores, construyeron una librería de cDNA. Esta fue tamizada con oligonucleótidos derivados de las secuencias parciales que habían sido obtenidas de la proteína del príon. La clona positiva, aunque truncada, de 1918 nucleótidos de largo presentaba un marco de lectura abierto suficiente para codificar un polipéptido de 240 residuos de aminoácidos. La secuencia derivada de aminoácidos correspondía perfectamente con varios péptidos derivados directamente del príon. Hibridaciones de tipo Southern, mostraron que éste era un gen celular presente tanto en el genoma de los hámsters infectados como en los hámsters sanos. Esto es, el gen del príon no llega con la infección sino está presente en el genoma del huésped potencial. Adicionalmente estas hibridaciones realizadas con DNA genómico de ratón y de humano, demostraron la presencia de secuencias homólogas al príon en estos genomas. Posteriormente, ha sido demostrado que este gen se encuentra presente en prácticamente todos los vertebrados (Basler, et al., 1986). El siguiente paso, fue analizar el patrón de expresión de este gen utilizando hibridaciones de tipo northern. Estos experimentos dieron resultados sorprendentes. Si se sigue la presencia de priones (proteína) desde animales recién infectados a lo largo de la incubación y del desarrollo de la enfermedad, el título de priones aumenta entre tres y cuatro ordenes de magnitud. Era de esperarse que la cantidad de mRNA que codifica al príon tuviera un aumento similar. Cuando se determinó la cantidad de este mRNA, se encontraron con que este transcrito ya tiene una expresión basal antes de que los animales sean inoculados. Mas extrañamente, los niveles del mRNA no aumentan al progresar la infección sino se mantienen constantes, invariantes, a pesar de que la proteína se está acumulando. Aún más; northerns realizados con otros tejidos, demostraron que el gen del príon básicamente es un gen constitutivo ya que se detecta su expresión no solo en el cerebro (donde se manifiesta la enfermedad), sino que se expresa en corazón, pulmón, páncreas, bazo, testículos y riñón. Solo en el hígado, parece no expresarse este gen. Estos resultados paradójicos, comenzaron a ser explicados cuando se analizó la proteína sintetizada por este mRNA tanto en animales sanos como en animales

infectados. Usando anticuerpos poli-clonales anti-prión, pudieron detectar proteínas equivalentes en ambos tipos de animales. Sin embargo, después de un tratamiento moderado con proteinasa K, los anticuerpos detectaron la proteína solo en los extractos de cerebros de animales infectados. La interpretación de estos resultados es que existen dos formas de esta proteína (Oesch, *et al.*, 1985). La primera, la normal, que se expresa de manera constitutiva y que no es infecciosa, la denominaron proteína scrapie celular (**PrP<sup>C</sup>**). La segunda, la que aparece y se acumula en los cerebros infectados la denominaron proteína scrapie-scrapie (**PrP<sup>Sc</sup>**). Esta última, es capaz de formar agregados que precipitan, son resistentes a proteasas y van formando las placas amiloides que son características de este grupo de enfermedades neurodegenerativas y que en última instancia son las directamente responsables de la vacuolización del cerebro y eventualmente de la muerte. ¿Cuál es la relación de estas dos formas de la proteína? Se ha demostrado, que ambas tienen la misma estructura primaria. En realidad el gen codifica normalmente a la proteína PrP<sup>C</sup>. Sin embargo, de acuerdo a las reglas del plegamiento tridimensional de las proteínas, la PrP<sup>C</sup>, puede sufrir un plegamiento ligeramente erróneo (Kocisko, *et al.*, 1994; Beyreuther & Masters, 1994). Esta forma equivocada de la proteína, tiende entonces a asociarse con otros polipéptidos similares erróneamente plegados, modificando la cinética normal de síntesis y degradación cotidiana de esta proteína. De hecho ha sido demostrado, que durante la infección la tasa de síntesis de la proteína normal no se altera (lo cual es consistente con el hecho de que la transcripción del gen no se modifica (Prusiner, 1991). Una predicción de este modelo, sería que la aparición de priones se lleva a cabo de manera constante durante toda la vida de un individuo. Sin embargo, normalmente la acumulación de estos polipéptidos erróneamente plegados no alcanza el umbral crítico para que se manifieste la enfermedad. En caso de que el individuo fuera expuesto a cantidades considerables de priones, estos rebasarían la concentración crítica y su acumulación se aceleraría. Dicho de otro modo, cuando se define la conformación nativa de un polipéptido de acuerdo a los experimentos de Anfinsen, se considera que para adquirir esta conformación un polipéptido, pasa por varios estados intermediarios y que a veces desde el punto de vista termodinámico, algunos de estos intermediarios son más estables que otros. En el caso de los priones esta conformación alternativa es notablemente estable y resistente ya que tiene la propiedad de formar agregados macromoleculares. Es notable que estas conformaciones alternativas de la proteína del prión, aunque tienen estructuras primarias idénticas, difieren en sus estructuras secundarias. La PrP<sup>C</sup> tiene un alto contenido de  $\alpha$  hélice (42 %) y muy poca  $\beta$  conformación. Por el contrario, la PrP<sup>Sc</sup> tiene 43% de  $\beta$  conformación y solo 30 % de  $\alpha$  hélice (Pan, *et al.*, 1993). Recientemente, se ha reportado la estructura tridimensional de la proteína del prión (Riek, *et al.*, 1996). En el caso de las enfermedades hereditarias en humanos (CJD y GSS), se han identificado mutaciones en el gen

del prión que favorecen o predisponen a que la proteína adquiera la conformación errónea (Hsiao, *et al.*, 1989). Así, en estos casos el umbral de concentración crítica se alcanza mucho antes que en la versión silvestre de este gen. Experimentos con ratones transgénicos que expresan una versión mutante del prión (tomando como modelo la mutación del humano), llevan a una neurodegeneración espontánea en estos animales, lo cual apoya la hipótesis anterior (Hsiao, *et al.*, 1990). Notablemente, aún no se ha logrado establecer con certeza cual es la función normal de esta proteína ni en el cerebro, ni en los otros órganos donde se expresa. Por ejemplo, ratones en los que se ha introducido una disrupción en el gen del prión y que por lo tanto no expresan la proteína, no parecen presentar un fenotipo particularmente alterado (Büeler, *et al.*, 1992). Es de interés mencionar que de acuerdo a las hipótesis de multiplicación de los priones, estos ratones sin el gen del prión no desarrollan la enfermedad cuando son inoculados con priones infecciosos (Prusiner, *et al.*, 1993). Estos resultados son explicados por la razón de que si no hay una síntesis celular continua de Pr<sup>Pc</sup>, no existe el substrato para que se acumule Pr<sup>Psc</sup>. Tal vez solo hasta que estos ratones envejecen, manifiestan algunos problemas motores y la pérdida de células de Purkinje en el cerebelo (desde luego en ausencia de infección), (Sakaguchi, *et al.*, 1996). También, se ha reportado que la proteína del prión podría estar implicada en la transmisión sináptica, la regulación de los ritmos circadianos y el sueño, pero aún no está claro su papel específico (Collinge, *et al.*, 1994; Tober, *et al.*, 1996). A nivel de secuencia de aminoácidos no tiene ninguna similitud con ninguna otra secuencia reportada, lo cual hace difícil inferir su función. A nivel de secuencia, este gen solo se ha ubicado en vertebrados. Se ha descrito que en levadura ocurren unas patologías moleculares transmisibles, muy semejantes a la que ocurre con los priones. En el primer ejemplo, una proteína de la levadura llamada **Ure2p**, la cual está involucrada en el metabolismo de compuestos nitrogenados, puede ocasionalmente convertirse a una forma inactiva que promueve la conversión de toda la proteína Ure2p a la forma inactiva (Wickner, 1994). Sorprendentemente, este fenotipo (**URE3**), puede ser "curado" temporalmente, cultivando a las levaduras en presencia de 1-5 mM de hidrocloreuro de Guanidina (GndHCl, un conocido agente desnaturizante). También, como se podría esperar, la sobre-producción de Ure2p aumenta la frecuencia de la conversión a la forma "patológica" (Masison & Wickner, 1995). El otro ejemplo en levadura es el del fenotipo denominado (**PSI\***). En este caso, la proteína identificada, **Sup35p**, está involucrada en el término de la traducción. Su región carboxilo terminal tiene homología con el factor de elongación EF-1 $\alpha$ . La pérdida de su función normal por conversión a la forma erróneamente plegada, causa una gama de problemas en la célula ya que afecta la actividad de varios tRNAs supresores que actúan a nivel de los codones de término de todos los mRNAs. De hecho se ha propuesto que Sup35p es un factor de término de la traducción (Cox, 1994).

Estas proteínas han sido consideradas como análogos de la proteína del prión y comienzan a ser conocidas como priones de levadura, en el sentido de que también son proteínas infecciosas. Estos descubrimientos sugieren que el fenómeno de los priones podría ser un fenómeno más general, que no estaría restringido a proteínas particulares que pueden causar problemas neurológicos en mamíferos.

Desde luego, Ure2p y Sup35p, no tienen homología a nivel de secuencia ni con la PrP ni entre ellas. De hecho, ni siquiera tienen el mismo peso molecular. PrP tiene 254 residuos de aminoácido, Sup35p, 685 y Ure2p, 354. PrP y Sup35, poseen ambas en la región amino terminal, una repetición en tandem (5 veces) de un octapéptido o un nonapéptido de secuencia similar (consenso PHGGGWGQ / PQGGYQQYN) respectivamente. Por el otro lado Ure2p, no posee estas secuencias repetidas (Cox, 1994).

No obstante algunas similitudes de estos fenómenos son interesantes. Al igual que en el caso del prión, la sobreexpresión de cualquiera de estas proteínas en la levadura, aumenta la frecuencia con la que se manifiesta la patología. (PSI<sup>+</sup>), puede ser "curada" por concentraciones relativamente altas de hidrócloruro de guanidina, alta concentración de sal, alcoholes, azúcares ácidos grasos, dimetil sulfoxido y dietil-etano-sulfonato (Tuite, 1994). Una diferencia importante es que ni de Ure2p ni de Sup35p se han descrito mutaciones que aumenten la propensión al cambio conformacional, análogas a CJD o GSS. De hecho, la delección de Sup35p o la mutación PNM (psi no more, que cambia glicina a aspártico en el codón 58), hace que el fenotipo (PSI<sup>+</sup>) ya no se pueda manifestar. Asimismo, no se ha logrado la infección con las proteínas de levadura puras. La infección solo se logra si la célula receptora, recibe a la proteína infecciosa con citoplasma de la célula infectante (citoducción) (Weissmann, 1994). Por otro lado, recientemente el grupo de Susan Lindquist ha mostrado que cuando Sup35p manifiesta el fenotipo (PSI<sup>+</sup>) forma agregados e inclusiones celulares, mientras que esta proteína tiene una distribución uniforme en las levaduras no infectadas (Patino, et al., 1996). Esta característica es típica de PrP que también forma inclusiones que derivan en las placas amiloides del cerebro con la patología. Es interesante el que este fenómeno estudiado por Lindquist, pudiera tener implicaciones más amplias, ya que este grupo demostró que el estado de agregación de Sup35p depende de Hsp104, una chaperona molecular inducida por stress medioambiental (Parsell, et al., 1994). Curiosamente, el fenotipo (PSI<sup>+</sup>) solo se manifiesta a concentraciones moderadas, digamos fisiológicas de Hsp104. De manera notable, cuando Hsp104 se sobreexpresa o cuando está ausente, el fenotipo (PSI<sup>+</sup>) no se manifiesta (Chernoff, et al., 1995). Este fenómeno podría ser la base de la herencia de una característica adquirida. Análogamente a la participación de Hsp104, se ha demostrado que la sobreexpresión de otra "heat shock protein", la Hsp70, temporalmente también elimina al fenotipo (PSI<sup>+</sup>). De hecho, se ha propuesto que la razón por la cual diversos agentes desnaturalizantes,

mencionados mas arriba, "curan" a los fenotipos (PSI<sup>+</sup>) y (URE3) de la levadura, es precisamente porque inducen la sobreexpresión de proteínas de stress. Hsp104 se incrementa de tres a cinco veces en presencia de 5-10 mM de GndHCl (**Chernoff, et al., 1995**). En este sentido, es curioso que la eficiencia de la conversión *in vitro* de PrP<sup>C</sup> a PrP<sup>Sc</sup>, se aumente en presencia de 3 M GndHCl, pero no ocurra cuando GndHCl se aumenta a 6 M (**Kocisko, et al., 1994**). Recientemente se ha reportado que la proteína del prión PrP<sup>C</sup>, interactúa con la proteína de stress Hsp60 (**Edenhofer et al., 1996**). En conjunto, estos resultados sugieren que tanto en levadura como en el scrapie, el asunto de los priones está íntimamente relacionado con el problema central del plegamiento tridimensional de los polipéptidos (**Patino, et al., 1996**). De hecho, retomando las ideas básicas de los experimentos de Anfinsen, uno puede fácilmente imaginarse una serie de circunstancias donde el plegamiento de un polipéptido a su estructura nativa puede salir mal y casi se podría predecir que con alguna frecuencia algo puede salir mal (**Taubes, 1996**). En este sentido, se ha comenzado a considerar que varias enfermedades podrían estar causadas por plegamiento incorrecto de proteínas específicas. Cada caso podría implicar procesos y fenómenos no idénticos y no en todos los casos se producen proteínas infecciosas. No obstante, parece claro que la enfermedad de Alzheimer se produce por agregación de proteínas de tipo  $\beta$ -amiloide. La fibrosis cística ocurre por plegamiento erróneo del regulador de conductancia transmembranal (CFTR), (en este caso, además se dan interacciones inadecuadas con la Hsp70). El síndrome de Marfan es provocado por plegamiento erróneo de la fibrilina. En las cataratas hay agregación de cristalinas. Inclusive, en el escorbuto, la falta de vitamina C evita la formación de hidroxiprolinas y la colágena se pliega incorrectamente (**Thomas, et al., 1995**).

Así, de lo anterior nuevamente se recupera la idea de que el plegamiento tridimensional, para que una proteína adquiera su conformación nativa, es un evento central en los sistemas biológicos. Este proceso es tan delicado que puede requerir de pequeños ajustes para que ocurra de manera satisfactoria en las células, lo cual remarca la importancia de las proteínas que han sido denominadas de manera general chaperonas moleculares (*vide infra*).

## **HEAT-SHOCK, FOLDING AND THE CHAPERONIN CONNECTION**

Como quedó señalado anteriormente, una de las características desconcertantes del fenómeno de stress calórico es la gran variedad de proteínas de heat-shock, que son sintetizadas (**Morimoto et al., 1990**). Además de las proteínas clásicas de stress *i.e.* (las familias de Hsp90, Hsp70, Hsp60 y las de bajo peso molecular sHsps), existen otras como las Hsp104, la Hsp47, la ubiquitina y otras proteínas reguladas por glucosa (Grp), que pueden inducirse por factores particulares como anoxia o azida de sodio y que pueden o no estar presentes en determinado organismo (**Welch,**

1990). En algunos casos, la posible función de algunas de éstas es muy específica. Por ejemplo, se ha sugerido que la Hsp47 está asociada a la síntesis, ensamblaje y exporte de la colágena (Sauk *et al.*, 1994; Nagata, 1996). Actualmente se sabe que varias de estas proteínas pueden actuar como chaperonas moleculares, - si bien no todas las Hsps son chaperonas moleculares, ni todas las chaperonas moleculares se comportan como Hsps- (Kelley & Georgopoulos, 1992). De hecho, otros factores medioambientales como el frío o la salinidad estimulan la expresión de proteínas no relacionadas con el stress calórico y que no se comportan como chaperonas moleculares (VanBogelen & Neldhardt, 1990).

**El concepto general de chaperonas moleculares**, fue planteado inicialmente de manera formal por John Ellis quien integrando una serie de ideas y datos dispersos, las definió como un grupo de proteínas no relacionadas que son capaces de promover el plegamiento y el ensamblaje correctos de otras proteínas, pero que no forman parte ellas mismas de las estructuras funcionales finales (Ellis, 1987). Esta idea ya había sido esbozada por Hugh Pelham, quien un año antes, en un artículo especulativo, había argumentado que las Hsp70, las Hsp90 y las Grps, podrían estar involucradas en el ensamble y desensamble de proteínas (Pelham, 1986). Si bien nunca manejó el término de chaperona, en ese artículo fundamental, Pelham proponía que durante el heat shock las proteínas se desnaturalizaban parcialmente, exponiendo regiones hidrofóbicas que inducían la formación de agregados insolubles. Postulaba entonces, que las Hsp70, Hsp90 así como las proteínas reguladas por glucosa (Grps), se unían a las regiones hidrofóbicas expuestas limitando las interacciones entre éstas y promoviendo la resolubilización de agregados. Asimismo, especulaba con la posibilidad de que las Hsps reconocieran igualmente a los polipéptidos nacientes los cuales, en su opinión, por definición están "desnaturalizados", de tal manera que se comprendía el porque las Hsps se requieren no solo en el stress calórico, sino durante todas las etapas de la vida de una célula (Pelham, 1986).

Las chaperonas moleculares se encuentran a todo lo largo de la escala filogenética y muchas de ellas son clasificadas como proteínas de stress (Lindquist, 1986), si bien como quedó señalado anteriormente, también se ha demostrado que tienen una función esencial bajo condiciones normales (Ang *et al.*, 1991).

Más recientemente las chaperonas moleculares han sido definidas como proteínas que se unen y estabilizan conformeros inestables de otros polipéptidos, de tal manera que facilitan *in vivo* ya sea su plegamiento, su ensamblaje en oligómeros, su transporte a compartimentos subcelulares específicos e inclusive su degradación (Hendrick & Hartl, 1993). Las chaperonas moleculares no contienen información estérica para especificar el plegamiento correcto. Más bien, ha sido propuesto que su mecanismo de acción consiste en prevenir interacciones incorrectas dentro y

entre polipéptidos en estado no-nativo, de manera que se incrementa el rendimiento (cantidad final de polipéptidos plegados), pero no la tasa (velocidad) de la reacción de plegamiento (Hartl, 1996). Esta característica distingue a las chaperonas moleculares de otra clase de proteínas llamadas catalizadoras de plegamiento como las proteína disulfuro isomerasas y las peptidil-prolil isomerasas, que aceleran pasos lentos en la ruta de plegamiento de algunas proteínas i.e. el rearreglo de los puentes disulfuro de proteínas de secreción y la isomerización *cis-trans* de los enlaces peptídicos que preceden a los residuos de prolina respectivamente (Gething & Sambrook, 1992).

### La nucleoplasmina.

Curiosamente, el término de chaperona molecular había sido acuñado originalmente muchos años antes, para describir la función especializada de la proteína nuclear llamada nucleoplasmina, la cual permite el ensamblaje de los nucleosomas en la cromatina de extractos de oocitos de *Xenopus laevis* (Laskey, et al., 1978). Se puede lograr la reconstitución *in vitro* de nucleosomas a partir de histonas y DNA sometiendo a estos a una diálisis prolongada con una concentración arriba de 1M de NaCl y alrededor de 5M de urea. Sin embargo si el experimento se realiza usando condiciones fisiológicas de fuerza iónica, lo que resulta es una rápida aparición de agregados inespecíficos en vez de nucleosomas. Sin embargo, en los experimentos de Laskey, si las histonas eran primero incubadas con una pequeña cantidad de homogenado de huevos de *Xenopus* antes de la adición del DNA, el resultado era la formación de nucleosomas, lo cual sugería la participación de un factor adicional requerido para el ensamblaje de nucleosomas *in vivo*. El grupo de Laskey purificó y caracterizó a este factor que resulto ser una proteína ácida de aproximadamente 29 kDa, la cual denominaron originalmente proteína ensambladora de nucleosomas y posteriormente se le dio el nombre de nucleoplasmina. La nucleoplasmina es requerida exclusivamente para el ensamble de los nucleosomas pero no forma parte de él. Esto es, la información estereoespecífica reside en las histonas y no en la nucleoplasmina. Por lo tanto el papel de la nucleoplasmina no consiste en aportar información para el ensamble de nucleosomas, sino reducir las cargas positivas de los monómeros de las histonas de manera que predominen interacciones correctas entre estas y el DNA. En las palabras de Laskey y sus colaboradores, "Sugerimos que el papel de la proteína que hemos purificado es el de una 'chaperona molecular' la cual previene interacciones iónicas incorrectas entre las histonas y el DNA" (Laskey, et al., 1978). Es notable la similitud de estas ideas originales con el concepto actual de las chaperonas moleculares. Resulta sin embargo paradójico, que actualmente la nucleoplasmina no sea considerada frecuentemente como una chaperona molecular. Esto es debido a las notables diferencias entre el mecanismo de acción relativamente sencillo de la

nucleoplasmina y las otras proteínas que son consideradas verdaderas asistentes del plegamiento.

### **Rubisco Subunit Binding Protein (RuSBP).**

La mayor parte del conocimiento que tenemos de las chaperonas moleculares, ha provenido principalmente de la investigación de las Hsp70 y las Hsp60 (*vide infra*). Sin embargo, en este punto vale la pena una aclaración de nomenclatura. Así como el término de chaperona molecular fue originalmente aplicado a la nucleoplasmina, y posteriormente ampliado a otras proteínas por John Ellis; el término **chaperonina**, fue originalmente aplicado para designar a la "Rubisco Subunit Binding Protein" (RuSBP, actualmente designada cpn60), y a las proteínas GroEL/GroES de *E. coli* (**Hemmingsen, et al., 1988**). En ese trabajo presentado por varios de los autores mas eminentes del campo, se menciona que basados en la similitud de sus secuencias derivadas de aminoácidos el producto polipeptídico del gen *groEL* y la RuSBP, son homólogos. Un poco antes en ese mismo año, se había descrito que a su vez GroEL estaba relacionada inmunológica y estructuralmente con una abundante proteína mitocondrial que se inducía por stress calórico, de peso molecular de aproximadamente 58 kDa aislada del protozooario *Tetrahymena thermophila* (**McMullin & Hallberg, 1988**). En el mismo trabajo, se describía que los anticuerpos dirigidos contra esta proteína, cruzaban contra proteínas mitocondriales similares de *Saccharomyces*, *Xenopus*, *Zea mays* y humanos y se sugería que esta clase de proteínas podrían tener algún papel de protección del daño inducido por el stress calórico, ayudando a estabilizar los procesos de ensamblaje macromolecular (**McMullin & Hallberg, 1988**). Al siguiente año se clonaron y caracterizaron los genes de la Hsp60 mitocondrial de levadura y humano, demostrándose que eran homólogos de GroEL y RuSBP (**Reading et al., 1989; Jindal, et al., 1989**). Es claro que el escenario estaba puesto para la aparición formal de las chaperonas moleculares.

### **LA CONEXIÓN**

El locus *groE* de *E. coli*, fue originalmente identificado porque mutaciones en este sitio bloqueaban la morfogénesis del ensamblaje de la cabeza de los fagos  $\lambda$  y T4 (**Georgopoulos, et al. 1972, 1973; Takano & Kakefuda, 1972; Coppo et al., 1973; Sternberg 1973a, 1973b**). Posteriormente, se demostró que el locus *groE* consiste de dos genes que codifican productos polipeptídicos de pesos moleculares de aproximadamente 65 y 15 kDa, en el orden: promotor-*groEL-groES* (**Tilly et al., 1981**).

### **La Rubisco.**

Resulta curioso que fue en particular el análisis del ensamblaje de la ribulosa bifsosfato carboxilasa-oxigenasa (Rubisco) lo que aclaró el significado de los descubrimientos anteriores

(Ellis, 1990a). La Rubisco, directa o indirectamente juega un papel vital en el metabolismo de la célula vegetal, ya que es el principal catalizador que incorpora al carbón dentro de compuestos orgánicos a partir del bióxido de carbono atmosférico durante la fotosíntesis (Gatenby & Ellis, 1990). La biogénesis de la Rubisco es inusualmente compleja e involucra a la luz como el factor que dispara a dos sistemas genéticos, uno localizado en el cloroplasto y el otro en el núcleo (Kuhlemeler, et al., 1987). A pesar de su papel fundamental, la Rubisco es un mal catalizador ya que tiene tanto una baja afinidad por el bióxido de carbono, así como un número de recambio bajo; de tal manera que los organismos autotróficos dedican una considerable cantidad de energía biosintética a producir muchas moléculas de esta enzima (Ellis, 1990a). De hecho se ha dicho que este accidente metabólico estructural, ha llevado a que la Rubisco sea la proteína más abundante en la biosfera (Ellis, 1979). Además, como si fuera poco; aparentemente como una consecuencia inevitable de su mecanismo de acción como carboxilasa, esta enzima cataliza una reacción de oxigenación (Gatenby & Ellis, 1990). Esta reacción de oxigenación es el primer paso de la fotorespiración, un proceso específico de las plantas, que salva los productos de la reacción de oxigenación. Esta vía de salvamento no es totalmente eficiente; algo de carbón se pierde como bióxido de carbono durante su operación, de manera que esta pérdida es el principal factor limitante de la productividad fotosintética en las plantas (Ellis, 1990a). Así, este proceso "calamitoso", que parece ser una consecuencia de la imperfección de la Rubisco, ha llevado a profundizar el estudio de esta enzima y a considerar la posibilidad de diseñar una carboxilasa vegetal con menos actividad de oxigenasa utilizando métodos de ingeniería genética (Mizioroko & Lorimer, 1983). Utilizando métodos de clonación molecular se ha logrado la expresión en *E. coli* de enzima (Rubisco), activa a partir de los genes clonados de diversos microorganismos procariontes (Gatenby & Ellis, 1990). Sin embargo, todos los métodos para ensamblar una Rubisco activa de plantas en *E. coli* han fracasado de manera que la mutagénesis sitio dirigida no ha podido ser aplicada (Ellis, 1990a). Así, esta problemática ha impulsado el estudio del mecanismo de la síntesis y el ensamble de la Rubisco (Ellis, 1990a).

En la bacteria purpúrea, no-sulfurosa, fotosintética *Rhodospirillum rubrum*, la Rubisco es un dímero ( $L_2$ ) de dos subunidades idénticas de peso molecular de 50.5 kDa (Schneider, et al., 1986). Por el otro lado, en la mayoría de los organismos fotosintéticos ya sean procariontes y desde luego en los cloroplastos, la Rubisco es una proteína oligomérica ( $L_8S_8$ ), que contiene ocho subunidades grandes (52 kDa) y ocho subunidades pequeñas (-14 kDa), (Mizioroko & Lorimer, 1983). Cada subunidad mayor contiene un sitio activo, pero cada uno de éstos incluye residuos de las subunidad adyacente de manera que un requisito para su actividad es el ensamblaje mínimo de dímeros. En este sentido entonces el oligómero  $L_8$  se consideraría un tetrámero de estos dímeros (Schneider, et al., 1986). En organismos procariontes los genes que codifican a las subunidades

mayor y menor se localizan en el mismo operón y se ensamblan en oligómeros funcionales en el citosol bacteriano (Tabita, 1988). Por el otro lado, en organismos eucariontes, la Rubisco se localiza exclusivamente dentro de los plástidos de las células vegetales (Gatenby & Ellis, 1990). La subunidad menor está codificada en el genoma nuclear y es importada al plástido después de su síntesis en los ribosomas citoplasmáticos, mientras que la subunidad mayor, está codificada por el genoma del plástido y es sintetizada por los ribosomas del mismo plástido, (Dean *et al.*, 1989).

Fue justamente el estudio del ensamblaje de las subunidades mayores de la Rubisco usando cloroplastos aislados, lo que llevó al descubrimiento de la que sería conocida como la chaperonina plástica (Barraclough & Ellis, 1980). El análisis del producto de la síntesis de proteínas del cloroplasto del chícharo, en geles de poliacrilamida en condiciones no-desnaturalizantes, reveló que la mayoría de las subunidades mayores no migran con los oligómeros preexistentes de la Rubisco. Experimentalmente se demostró que la causa de este retraso electroforético, se debe a que se encuentran asociadas no-covalentemente a otra abundante proteína de alto peso molecular compuesta de monómeros de 60 kDa. Esta proteína fue bautizada como la proteína unidora de la subunidad mayor y está codificada por genes nucleares (Ellis, 1990a). Estas observaciones llevaron a la propuesta de que la unión transitoria de la subunidad mayor a esta proteína podría ser un paso obligatorio en ensamblaje de la Rubisco (Barraclough & Ellis, 1980). Dicho de otro modo, los resultados anteriores implicaban que las subunidades mayores de la Rubisco, tendían a tener interacciones incorrectas entre si, y que estas asociaciones producían estructuras que no eran funcionales (Ellis, 1990a). Es interesante hacer notar que después de ser sometidas a un agente desnaturalizante, las subunidades mayores de Rubisco vegetal, tienden a formar agregados insolubles, aun en presencia de subunidades menores nativas (Gatenby & Ellis, 1990). Como fuera, la proposición de que el ensamblaje de una proteína oligomérica requiriera de la ayuda de otra proteína, carecía de antecedentes en 1980, de manera esta propuesta atrajo mas escepticismo que entusiasmo, y así, estas ideas fueron abordadas en lo referente a la Rubisco, casi exclusivamente por solo dos laboratorios durante la siguiente década, e.g. Roy, *et al.*, 1982; Hemmingsen & Ellis, 1986. Durante la mayor parte de aquel período la otra proteína fue conocida como la proteína unidora de la subunidad mayor de la Rubisco, hasta el descubrimiento en 1988 de que las subunidades pequeñas, cuando son importadas al cloroplasto, también se unen a esta proteína (Gatenby, *et al.*, 1988). El nombre fue recortado entonces, retirando el sustantivo "grande" ; pero el nuevo apelativo no duraría más de un año cuando, como se señaló antes, nació el nombre de chaperonina (Hemmingsen, *et al.*, 1988).

## CLASES Y CARACTERISTICAS DE LAS CHAPERONAS MOLECULARES

En la **tabla 2**, se listan a las familias de chaperonas moleculares.

---

**TABLA 2**  
**FAMILIAS DE CHAPERONAS MOLECULARES**

---

- Hsp 100
  - Hsp 90
  - Hsp 70
  - Hsp 60
  - LMW Hsp's
- 

La familia de las **Hsp100**, parecen ser exclusivas de eucariontes, si bien se ha reportado cierta identidad con ClpB, una proteasa de *E. coli* (**Wülfing & Plücker, 1994**). Como fue discutido antes, Hsp104 de levadura es la proteína más estudiada de esta familia, está implicada en fenómenos de termotolerancia promoviendo la solubilización de agregados de proteínas y forma oligómeros de seis subunidades (**Parsell, et al., 1996**).

Las **Hsp90**, se encuentran tanto en procariontes como en eucariontes donde se localizan tanto en el citoplasma como en el retículo endoplasmático, pero no han sido descritas en organelos (**Becker & Craig, 1994**). Se ha descrito que su estructura nativa es dimérica (**Lindquist & Craig, 1988**). Hsp90 es una de las proteínas más abundantes en condiciones fisiológicas y su síntesis se incrementa notablemente por heat-shock (**Buchner, 1996**). En la levadura existen dos copias de este gen y se ha demostrado que mutaciones individuales en cualquiera de ellos (HSP83/HSC83), son viables (aunque presentan alteraciones de crecimiento y no crecen arriba de 37 °C, sin embargo las dobles mutantes son letales (**Lindquist & Craig, 1988**). En contraste, la delección del gen *htpG*, que codifica al único miembro de esta familia en *E. coli*, no es letal (**Bardwell & Craig, 1988**). Existen dos explicaciones para estos resultados. Podría ser que existiera otra proteína en *E. coli* que cubriera la pérdida de su Hsp90. Por el otro lado podría ser que la función de esta proteína fuera esencial en levadura pero no en *E. coli* (**Lindquist & Craig, 1988**). Funcionalmente Hsp90 ha sido más estudiada en eucariontes donde se le ha involucrado en la activación de

receptores esteroides y la activación y maduración de ciertas kinasas. En este contexto, se ha demostrado que la Hsp90 se asocia a varios otros factores proteicos y a proteínas que han perdido su conformación nativa (Buchner, 1996). La sobreexpresión de Hsp90 de eucariontes bajo condiciones de heat-shock sugiere una función general de protección a la desnaturalización, como ha sido demostrado para la Hsp90 de *E. coli* (Bardwell & Craig, 1987). Recientemente, utilizando sistemas de plegamiento *in vitro*, ha sido demostrado que la Hsp90 purificada, reconoce y une proteínas no nativas y en consecuencia previene agregación no específica durante el plegamiento (Wiech *et al.*, 1992; Jakob & Buchner, 1994). Sin embargo, a diferencia de otras chaperonas moleculares, la función de Hsp90 parece ser independiente de ATP (Buchner, 1996). Durante su función de plegamiento, Hsp90 parece interactuar con intermediarios conformacionales que ya poseen una cantidad significativa de estructura secundaria, sin embargo, se desconocen las características conformacionales específicas reconocidas por Hsp90 (Jakob *et al.*, 1995).

Las Hsps de bajo peso molecular [small heat shock proteins (sHsp) / Low Molecular Weight Heat Shock Proteins / (LMW Hsp's), MW 15-30 kDa], son particularmente conspicuas en plantas donde se les puede localizar tanto en el citoplasma como en organelos (mitocondrias y cloroplastos), (Lindquist, 1986). Por el otro lado, también se encuentran en el citoplasma de otros organismos como *Drosophila*, mamíferos y levadura, así como en *E. coli* (Buchner, 1996). Las sHsps, son un grupo heterogéneo de proteínas y diferentes organismos poseen un número diverso de éstas que van desde una en levadura hasta más de treinta en plantas superiores (Lindquist & Craig, 1988). A pesar de esta heterogeneidad, se les puede considerar un grupo de proteínas homólogas con base a una similitud relativa a nivel de estructura primaria, así como similitud estructural y funcional (Lindquist, 1986). En plantas se han descrito seis familias de estas proteínas que constituyen la superfamilia de las small Hsps. Estas Hsps, pueden formar oligómeros de 200 a 350 kDa (Lee *et al.*, 1996). Se ha demostrado que Hsp25 y Hsp27 de ratón y humano pueden influir en el proceso de plegamiento de sustratos modelo y que se unen a proteínas desnaturalizadas por condiciones de stress calórico (Jakob *et al.*, 1993). Asimismo, las sHsps de mamífero confieren termotolerancia y los homólogos de *E. coli* se asocian específicamente con cuerpos de inclusión que contienen proteínas sobreexpresadas, sugiriendo que estas Hsps podrían tener funciones de chaperonas moleculares (Buchner, 1996). Evidencia directa de esta función ha sido obtenida de experimentos *in vitro*, que demuestran que las sHsps se unen selectivamente a proteínas desnaturalizadas (Jakob, *et al.*, 1994). En el ojo, la  $\alpha$ -cristalina (MW 20 kDa), constituye hasta el 30% de la proteína del cristalino del ojo. Esta proteína está formada por dos subunidades (A y B) que tienen una similitud de aproximadamente 58% entre sí, pero además, tienen una

considerable similitud (40%), con sHsps de otros organismos (Rao *et al.*, 1994). Habitualmente se sugiere que la función de esta proteína es contribuir a las propiedades refractivas de este órgano. Sin embargo, se ha demostrado que para mantener la transparencia del ojo, las  $\alpha$ -cristalinas también tienen funciones de chaperonas moleculares previniendo el agregamiento irreversible e inespecífico de otras proteínas del cristalino del ojo (Horwitz, 1992; Rao *et al.*, 1994). Con la notable excepción de las  $\alpha$ -cristalinas del ojo, todos los demás miembros de la familia de las Hsps pequeñas no son indispensables. Asimismo, parece ser que la función de las sHsps no es regulada por ATP (Buchner, 1996).

Las **Hsp70s** han sido de las proteínas de stress más caracterizadas (Pelham, 1990). Curiosamente, el gen de *E. coli*, *dnaK*, fue identificado originalmente como un gen del huésped que se requería para la replicación del bacteriófago  $\lambda$  (Georgopoulos, 1977; Sunshine *et al.*, 1977). Posteriormente se ha demostrado que homólogos de DnaK se localizan a todo lo largo de la escala filogenética incluyendo procariontes, el citoplasma y el retículo endoplásmico de los eucariontes, mitocondrias y cloroplastos (Boorstein, *et al.*, 1994; Rensing & Maier, 1994). Sus funciones incluyen plegamiento de polipéptidos, su translocación a través de membranas e inclusive juegan un papel en la degradación de proteínas (Pelham, 1990). Esta versatilidad de las Hsp70 proviene de su función básica que es la unión reversible dependiente de ATP, de segmentos hidrofóbicos expuestos en una proteína desnaturalizada. Mientras que la unión resulta en la estabilización de la conformación desplegada, la liberación controlada podría permitir a la proteína entrar en la vía de plegamiento (Flynn, *et al.*, 1989). Estos ciclos de unión y liberación están modulados por otras proteínas asociadas que a su vez son consideradas Hsps, **DnaJ (Hsp40)** y **GrpE** (MW 21.7 kDa), (Georgopoulos *et al.*, 1990). A nivel estructural, las Hsp70 podrían no tener estructura cuaternaria, *i. e.* funcionan como monómeros (Buchner, 1996). A nivel de estructura tridimensional, estas proteínas se pliegan en dos grandes dominios. El dominio en el extremo amino terminal (44 kDa) es responsable de una actividad de ATPasa asociada con la función de estas proteínas. Este dominio es seguido de una región (18 kDa) que contiene el sitio de unión del péptido y un segmento más variable de alrededor de 10 kDa (Wang *et al.*, 1993). El dominio amino terminal ha sido cristalizado y se ha reportado su estructura tridimensional (Flaherty *et al.*, 1990). Curiosamente al hacerse un alineamiento estructural se observó que su plegamiento tridimensional es notablemente parecido a la estructura tridimensional la actina G y a un dominio equivalente de la hexoquinasa (Flaherty *et al.*, 1991). Esta similitud estructural y funcional (hidrólisis de ATP), ha llevado inclusive a sugerir homología entre estas proteínas, postulándose que éstas, además de las ciclinas bacterianas, comparten un origen evolutivo común (Bork *et al.*, 1992). Sin embargo, hay que hacer notar que a nivel de estructura primaria, estas proteínas prácticamente

no tienen similitud, lo cual indica un fenómeno de convergencia evolutiva más que homología. De hecho, si se considera la tasa de cambio evolutivo a nivel de aminoácidos (reloj molecular) de la actina (1.55%/100 millones de años), (McDowell *et al.*, 1996) y de las Hsp70 en el dominio de hidrólisis de ATP, (2.5%/100 millones de años), (Hughes, 1993; Boorstein *et al.*, 1994), se llega a la conclusión de que estos dos genes habrían compartido el último ancestro común hace más de 4 mil millones de años (antes del origen de la vida), lo cual es desde luego absurdo. Por otro lado, se llegó a reportar que el dominio en el extremo carboxilo tenía una ligera similitud con moléculas del complejo mayor de histocompatibilidad (MHC), por lo cual se propuso que este dominio estaba implicado en el reconocimiento de los polipéptidos sustrato (Rippman *et al.*, 1991). Si bien esto último ha sido confirmado *i. e.* que el dominio carboxilo interactúa con la proteína sustrato; recientemente utilizando un análisis multidimensional de NMR, se ha demostrado que la topología de este dominio no se parece a la de las proteínas del MHC (Morshauer *et al.*, 1995). El ciclo de unión y liberación, dependiente de ATP ha sido elucidado de manera general para el homólogo bacteriano DnaK y sus cofactores DnaJ y GrpE (Palleros *et al.*, 1993). Cuando DnaK se une a ATP, la unión y liberación del péptido ocurre rápidamente. En contraste cuando tiene ADP el ciclo se lleva a cabo lentamente. DnaJ y GrpE, aceleran la interconversión entre estos dos estados (Flynn *et al.*, 1989). En cada evento de liberación una fracción del polipéptido alcanza su estructura nativa. El restante, que se encuentra como un intermediario inestable de plegamiento, se une nuevamente y se repite el ciclo (Schmid *et al.*, 1994). De manera significativa, se ha demostrado que DnaK/DnaJ, liberan a su sustrato proteico en una conformación desplegada (Harfi, 1996). Respecto a las características que reconocen las Hsp70 en sus sustratos, se ha demostrado que estas chaperonas moleculares se unen a péptidos en una conformación extendida (Landry *et al.*, 1992). En particular, se ha comprobado que Bip (un miembro de las Hsp70), se une preferencialmente a péptidos de 4-7 residuos de longitud, que tienen un relativamente alto contenido de residuos alifáticos, hidrofóbicos (Flynn *et al.*, 1991).

### **THE LORD OF THE RINGS (HSP60)**

Dentro de las chaperonas moleculares, las Hsp60 (GroEL/chaperoninas60/cpn60/TRiC) constituyen el grupo más complejo. Han sido descritas a todo lo largo de la escala filogenética localizándose en eubacterias, cianobacterias, arquea y eucariontes. En estos últimos, inicialmente fueron reportadas en organelos celulares (plástidos y mitocondrias), pero hoy sabemos que también se encuentran en el citosol (Gupta, 1995). Si bien todas estas proteínas comparten notables similitudes estructurales y funcionales; la similitud a nivel de secuencia entre las chaperoninas citoplasmáticas y las de arquea por un lado contra las chaperoninas bacterianas, mitocondriales

y plásticas por el otro, es tan baja que resulta difícil decidir si constituyen un grupo homólogo o no (Gupta, 1990a; North, 1991; Ellis, 1992; Gupta, 1995). Por este motivo y como estas últimas (las clásicas), son las mejor caracterizadas; en lo que sigue me referiré principalmente a éstas y posteriormente se discutirá la familia TRiC.

En *E. coli*, *groEL*, la chaperonina mejor caracterizada, forma un operón con *groES*, localizado a los 93.5 minutos del genoma de esta bacteria (Tilly *et al.*, 1981; Guest & Nice, 1978). GroEL y sus homólogos poseen una capacidad débil de hidrólisis de ATP, cuya actividad es modulada por GroEL (Chandrasekhar *et al.*, 1986; Bochkareva *et al.*, 1988; Ostermann *et al.*, 1989; Goloubinoff, 1989). Más recientemente, se ha reportado que tanto GroEL como GroES son capaces de unir ATP; aunque aparentemente solo GroEL posee capacidad de hidrólisis de este nucleótido (Martin *et al.*, 1993b). Se ha propuesto que el significado de la capacidad de unión de ATP de GroES es el de facilitar la unión de ATP por GroEL (Ellis 1993). Homólogos de *groEL* han sido localizados en todos los organismos en los que se ha buscado (Gupta, 1995). Inclusive, los nuevos proyectos de macro-secuenciación de genomas invariablemente detectan este gen (Adams *et al.*, 1995; Fraser *et al.*, 1995; Bult *et al.*, 1996). Curiosamente en estos genomas no siempre se detectan homólogos de las otras Hsps. Por ejemplo, en la arquea metanogénica *Methanococcus jannaschii*, los genes de DnaK, no aparecen, pero de chaperonina si (Bult *et al.*, 1996).

Como se indicó antes, las chaperoninas tienen la función de promover el plegamiento dependiente de ATP de proteínas tanto en condiciones de stress como bajo condiciones normales. En general, estas proteínas son inducibles por stress, aunque habitualmente existen genes duplicados homólogos ("cognates"/ cognados, parientes ó vinculados), que tienen un patrón de expresión constitutivo (Hemmingsen, *et al.*, 1988). Paradójicamente, se ha reportado que el homólogo del cloroplasto, no es inducible por calor (Vierling, 1991; Zabaleta, *et al.*, 1994a). Recientemente se ha reportado que GroEL podría tener un papel adicional, regulando la vida media del RNA mensajero (Georgellis, *et al.*, 1995). Se ha especulado que esto, podría tener implicaciones que vincularan a las chaperoninas con el mundo ancestral de RNA (Hartl, 1996).

### **Estructura.**

Las chaperoninas de tipo GroEL, están constituidas de dos anillos, cada uno de los cuales está formado por siete subunidades, cada una de alrededor de 60 kDa. Estos, cooperan con un cofactor, GroES en *E. coli*, que está constituido por un solo anillo heptamérico de subunidades de alrededor de 10 kDa (Saibil & Wood, 1993). La estructura de GroEL ha sido ampliamente estudiada por microscopía electrónica (McMullin & Hallberg, 1988; Langer, *et al.*, 1992a; Saibil & Wood, 1993) y recientemente se ha reportado la estructura tridimensional de este gran complejo (Braig, *et al.*, 1994). El complejo cilíndrico tiene 150 Å de altura y 140 Å de ancho, que encierra

una cavidad central donde se ubica el sitio de unión del polipéptido. Cada subunidad de 57 kDa, está conformada por tres dominios: el apical, el intermedio y el ecuatorial. El extremo amino terminal se localiza en el dominio ecuatorial. De ahí, la cadena polipeptídica pasa al dominio intermedio, luego al apical; regresa al intermedio y finalmente retorna al dominio ecuatorial (Braig, *et al.*, 1994). El dominio ecuatorial, contiene el sitio de unión al ATP y aporta la mayor parte de los residuos que determinan la interacción entre las subunidades (Braig, *et al.*, 1994; Fenton, *et al.*, 1994). El dominio más pequeño, el intermedio, tiene regiones potenciales que podrían actuar como bisagras en los sitios de unión con los otros dominios y que permitirían el movimiento de estos (Hartl 1994). Por su lado, la estructura tridimensional del heptámero de GroES revela una forma similar a un domo ó cúpula, con un orificio central, que se acoplaría fácilmente sobre el orificio central de GroEL (Hunt, *et al.*, 1996; Mande, *et al.*, 1996).

### **Unión del sustrato.**

En ausencia de GroES, GroEL, puede unir hasta dos moléculas de polipéptidos en el centro de cada uno de los anillos a nivel del dominio apical (Hartl, 1996). Sin embargo durante el ciclo normal de plegamiento, solo una de las cavidades es ocupada, lo cual sugiere cooperatividad entre los dos heptámeros (Langer, *et al.*, 1992a; Chen, *et al.*, 1994). No obstante, el requisito de que la chaperonina funcione como dos anillos toroidales, podría no ser absoluto, ya que se ha descrito que en mitocondrias humanas la cpn60 funciona como un solo anillo heptamérico (Virtanen, *et al.*, 1992a). Como sea, la formación de un complejo estable, probablemente requiere que el polipéptido haga contacto con varias de las subunidades en el heptámero. Se ha demostrado que la proteína unida se encuentra en una conformación colapsada similar a lo que se ha descrito como el "glóbulo fundido" (Marfin, *et al.*, 1991; Landry, *et al.*, 1992; Robinson, *et al.*, 1994; Zahn, *et al.*, 1996). Este estado de glóbulo fundido fue propuesto desde los inicios de los setentas como un intermediario de plegamiento que ya contiene estructura secundaria similar a la proteína nativa, pero que aún no adopta su estructura tridimensional definitiva (Pititsyn 1996). De hecho se ha propuesto, que las proteínas solo son capaces de existir en tres estados termodinámicos: el nativo, el glóbulo fundido y el desnaturalizado, que recordarían a los tres estados de la materia: sólido, líquido y gaseoso (Pititsyn 1995). Sin embargo, otro reporte señala que la chaperonina reconoce a su sustrato en un estado mas desnaturalizado que el glóbulo fundido clásico (Okazaki, *et al.*, 1994). Estos resultados contradictorios se podrían deber a discrepancias en el significado preciso del término glóbulo fundido empleado por diferentes grupos aunque se apliquen a la misma proteína (Kuwajima 1996). Sin embargo, es claro que las chaperoninas reconocen a su sustrato cuando ya tiene bastantes características nativas ya sea estructura  $\alpha$  o  $\beta$  (Landry & Gierasch, 1991; Schmidt & Buchner, 1992), pero aún expone algunas

regiones hidrofóbicas hacia el solvente; dicho de otro modo la interacción de GroEL con su sustrato es de naturaleza hidrofóbica (Martin *et al.*, 1991; Lin, *et al.*, 1995; Horowitz, *et al.*, 1995; Gibbons & Horowitz, 1995). De hecho, en la estructura cristalina de GroEL, se pueden observar en el dominio apical, un cierto número de residuos hidrofóbicos expuestos hacia la cavidad central. Mutaciones en estos residuos bloquean la capacidad de unión de polipéptidos (Fenton, *et al.*, 1994). Inclusive se sabe que estos dominios apicales pueden sufrir un movimiento en *bloc*, con respecto al dominio intermedio, lo que permitiría que en la cavidad central se acomodaran sustratos polipeptídicos de tamaños diversos en el rango de 15-60 kDa; los cuales tendrían oportunidad de interactuar con los mencionados residuos hidrofóbicos de la chaperonina (Martin, *et al.*, 1992; Saibil & Wood, 1993). La idea de que cpn60 interactúa y protege regiones hidrofóbicas expuestas en sus polipéptidos sustrato, es reforzada por el reporte de que micelas de lípidos y detergentes anfipáticos pueden imitar la acción de la chaperonina en el plegamiento de polipéptidos desnaturalizados (Zardeneta & Horowitz, 1992). Así, fuera de estas características de las proteínas sustrato: estar parcialmente desnaturalizadas en la forma de "glóbulo fundido", exponiendo algunas regiones hidrofóbicas; las cpn60 no muestran ninguna predilección por regiones específicas o secuencias de residuos de aminoácidos particulares. Es por esta razón que se dice que las chaperoninas son paradójicamente promiscuas (Landry & Gierasch 1994).

### **El ciclo de plegamiento.**

En términos generales hay tres niveles funcionales que contribuyen a la alta eficiencia del plegamiento de polipéptidos por las chaperoninas. 1) Prevención de la formación de agregados debido a la unión de polipéptidos parcialmente desnaturalizados. 2) Localización de polipéptidos desnaturalizados en un compartimento de plegamiento aislado. 3) Re-unión y rearrreglo estructural de los polipéptidos que no han alcanzado su estructura nativa (Hartl & Martin, 1992). Si bien la unión del polipéptido no-nativo es independiente de cualquier factor, la capacidad de GroEL para proveer un compartimento de plegamiento depende generalmente de cambios conformacionales que son inducidos por la llamada co-chaperonina GroES/cpn10 (Chandrasekhar, *et al.*, 1986; Martin, *et al.*, 1993a; Chen *et al.*, 1994). La liberación del péptido o su re-unión requiere de la alternancia de ciclos de unión-desunión GroES-GroEL que son regulados por la capacidad de hidrólisis de ATP que posee GroEL (Bloom, *et al.*, 1983; Hemmingsen & Ellis, 1986; Bochkareva *et al.*, 1988; Goloubinoff, *et al.*, 1989; Martin *et al.*, 1993a). Estudios de microscopía electrónica han revelado que GroES se une solo a uno de los extremos de GroEL, en el dominio apical (Langer *et al.*, 1992a; Chen *et al.*, 1994). Otros autores sin embargo, han reportado la formación de complejos simétricos donde GroES es visualizada en los dos extremos

de GroEL, argumentando además la importancia funcional de este complejo (**Azem et al., 1994; Schmidt et al., 1994; Todd, et al., 1994**). Más recientemente se ha cuestionado la relevancia de dichos complejos y se ha sugerido que en dichos reportes se utilizaron concentraciones elevadas no fisiológicas de magnesio así como un pH alto (**Engel et al., 1995**). En apoyo a los complejos asimétricos, cabe considerar que desde hace tiempo que las cinéticas de reacción bioquímicas han indicado que la capacidad óptima de plegamiento e inhibición de la actividad de ATPasa se logra con una estequiometría 1:1 de tetradecámeros de GroEL a heptámeros de GroES (**Chadrsekhar et al., 1986; Goloubinoff et al., 1989; Langer et al., 1992a; Martin et al., 1993a; Hayer-Hartl et al., 1995**). Estos puntos de discrepancia sirven para enfatizar que aún son muchos los aspectos oscuros en la función y mecanismo de acción de las chaperoninas (**Buchner, 1996**). Como sea, en complejos simétricos o asimétricos, se ha reportado que la unión de GroES conduce al desplazamiento de la proteína unida en la región del dominio apical de GroEL, forzándola a que penetre más profundamente a una región hidrofílica permisiva del plegamiento (**Martin et al., 1993a; Hartl 1994; Mayhew et al., 1996; Weissman et al., 1996**). Además, la asociación de GroES bloqueando el orificio central de GroEL, también tendría la función de prevenir la difusión del sustrato al medio externo, ya que el paso de un anillo al otro del deca-tetra-oligómero no está libre (**Braig et al., 1994; Mayhew & Hartl, 1996**). Todos estos pasos de unión del sustrato, unión de GroES, hidrólisis de ATP, liberación del sustrato, re-unión del mismo, involucran una compleja red de interacciones alostéricas caracterizadas por una cooperatividad positiva en la unión e hidrólisis del nucleótido en cada anillo de GroEL y una cooperatividad negativa para la unión del sustrato entre los anillos (**Martin et al., 1993a; Todd et al., 1994; Hayer-Hartl et al., 1995; Hartl 1996**). Una descripción detallada de cada paso, queda fuera de la intención de esta introducción. Sin embargo, hay algunos puntos que sería interesante mencionar. El costo energético de cada ciclo de plegamiento es alto. Se ha calculado que por cada polipéptido plegado se gastan alrededor de 100-130 moléculas de ATP (**Martin et al., 1991**). Este gasto probablemente se debe a que cada monómero del tetra-deca-oligómero hidroliza varias moléculas de ATP durante múltiples ciclos de unión de GroES antes de que el sustrato adquiera su conformación nativa (**Martin et al., 1993a**). Aunque los detalles varían, en general se acepta que durante el ciclo de plegamiento los dos anillos de GroEL no son equivalentes. En el anillo denominado *cis* se asocian el sustrato y posteriormente GroES. En este momento podrían ocurrir cambios conformacionales asociados a una rápida hidrólisis de ATP, pero no necesariamente plegamiento del sustrato. Casi simultáneamente el otro anillo denominado *trans* une ATP, pero su hidrólisis no es inmediata (este anillo puede unir a GroES, pero no se asocia al sustrato). Sorprendentemente esta unión de ATP dispara un "reloj" de  $t_{1/2} = 15-30$  segundos. Después de este tiempo se hidroliza el ATP con la subsecuente liberación de GroES dando la

oportunidad al sustrato de salir al medio ya sea plegado o para unirse nuevamente al anillo *cis* y repetir el ciclo. Se ha propuesto que el significado funcional de este "timer" posiblemente está relacionado con un compromiso entre la velocidad de plegamiento y la facilidad de plegarse de los diferentes sustratos de GroEL. Un timer más rápido tal vez sería inadecuado para polipéptidos lentos (some misfolders; slow folders; no folders) y un timer más lento sería desventajoso para proteínas que requieren menos asistencia de la chaperonina (fast folders; some misfolders) (Weissman *et al.*, 1996). Así, conceptualmente el plegamiento del sustrato ocurre en una serie de reacciones iterativas que corresponden a los ciclos de hidrólisis de ATP (Todd *et al.*, 1996).

### **La Jaula de Anfinsen.**

Como quedó señalado anteriormente, desde las épocas de Anfinsen, se ha planteado una especie de paradoja entre el plegamiento polipeptídico *in vivo* contra *in vitro* (Anfinsen 1973; Ellis 1996). Actualmente se piensa que una de las diferencias fundamentales entre estos procesos es lo que se ha llamado conglomeración macromolecular (Hartl 1996). Este concepto se refiere al hecho de que una fracción considerable del volumen celular está ocupado por proteínas y otras macromoléculas que alcanzan una concentración de alrededor de 340g por litro (Zimmerman & Milton 1993). Una consecuencia directa de esta conglomeración macromolecular es la de incrementar las constantes de asociación por arriba de las prevalentes en soluciones diluidas por varios ordenes de magnitud (Ellis & Hartl 1996). Por esta razón se entiende que el plegamiento exitoso *in vitro* sea favorecido al usarse bajas concentraciones de proteína y bajas temperaturas (Anfinsen, 1973). En este sentido es de notar que inclusive los experimentos de plegamiento *in vitro* que utilizan chaperonas moleculares, ni siquiera se acercan a las condiciones prevalentes en el interior de las células (Ellis & Hartl 1996). Se ha propuesto que se podrían simular condiciones de alto nivel de congestionamiento incluyendo en los experimentos polímeros de alto peso molecular como el Ficoll 70 y el Dextran 70 (Zimmerman & Milton, 1993). Sin embargo, es posible que ni así acabaríamos de tener una idea clara del nivel de aglomeración, ni de las múltiples posibilidades de interacciones equivocadas que pueden ocurrir dentro de una célula (Goodsell 1991).

Así, estas consideraciones sugieren que el simple hecho de confinar al polipéptido sustrato dentro de la cavidad de GroEL, sería equivalente a efectuar un procedimiento similar a los protocolos de Anfinsen. Al entrar la proteína a este medio aislado, la proteína alcanza una dilución infinita. Dicho de otro modo, en una primera aproximación, el plegamiento dentro de la cavidad de GroEL, sería equivalente al plegamiento en el tubo de ensayo, resolviéndose el problema de agregación. Estas consideraciones fueron expuestas por primera vez por Nilsson y

Anderson (**Nilsson & Anderson 1991**). Posteriormente, honrando los experimentos pioneros de Christian Anfinsen en este campo, este modelo de plegamiento en aislamiento ha sido llamado "la jaula de Anfinsen" (**Salbil et al., 1993**). La localización de la proteína en confinamiento en la jaula de GroEL, podría tener consecuencias físicas adicionales. A la manera de un agujijón, al pasar la proteína de interacciones iniciales de reconocimiento hidrofóbicas con la chaperonina, a una región más hidrofílica de la jaula (*vide supra*) (**Martin et al., 1993a; Hartl 1994; Mayhew et al., 1996; Weissman et al., 1996**), se podría promover la internalización de los residuos hidrofóbicos expuestos mas efectivamente que en solución (**Hartl 1996**). Llevando la analogía de la jaula un poco mas allá, se ha mencionado que el papel de GroES es equivalente al del carcelero (**Ellis, 1994b**). Es curioso como el uso de palabras del lenguaje común, en el terreno científico produce, a veces no intencionalmente, conceptos extraños. La definición de diccionario de chaperón, se refiere a una persona, habitualmente una mujer madura, que acompaña a una dama, habitualmente no casada, para guiarla, protegerla y evitar que se den relaciones incorrectas ("*dangereux liasons*"/*dangerous liaisons*) (**Ellis 1990b**). Implícito en esta definición está el hecho de que habitualmente la chaperona tiene un comportamiento gentil, bondadoso, e inclusive ocasionalmente puede actuar en el sentido opuesto a aquel que se supone debe seguir (La Celestina era una chaperona). Ahora con el concepto de la jaula, esta gentileza se pierde. Las proteínas "mal portadas" son encarceladas, e inclusive hay un celador (GroES), que no las deja salir hasta que han regresado al "buen camino". Este enfoque, implícitamente contiene un elemento de castigo. Las proteínas se portan mal y son enviadas a la cárcel. No será sorprendente que próximamente se mencione en la literatura del campo, que el costo energético del plegamiento es equivalente a pagar una fianza en moléculas de ATP.

### **Plegar o desplegar.**

Además de que el modelo de la caja de Anfinsen tiene atractivos conceptuales y mecánicos muy obvios, existe una buena cantidad de evidencia experimental a favor de este modelo (e. g. **Ostermann et al., 1989; Martin et al., 1991; Martin et al., 1992; Langer et al., 1992a; Martin et al., 1993a; Salbil et al., 1993; Chen et al., 1994; Hayer-Hartl et al., 1995; Mayhew et al., 1996; Weissman et al., 1996**). Sin embargo, por otro lado existen también numerosos trabajos que sugieren que las chaperoninas más bien tienen un papel de desnaturalizar al polipéptido (**Bochkareva et al., 1988; Zahn et al., 1994; Todd et al., 1994; Weissman et al., 1994; Ranson et al., 1995; Taguchi & Yoshida 1995; Zahn et al., 1996**). En general en dichos reportes si bien el producto final es la proteína nativa, durante la fase cíclica de interacción GroEL - substrato, se pueden detectar intermediarios de plegamiento en solución después de que han entrado y salido de la jaula. La conclusión razonada de estos resultados es que una proteína trataría de adquirir su conformación nativa en

el citosol, objetivo que en ocasiones no lograría, quedando "atrapada" en una conformación inadecuada. La participación de la chaperonina sería entonces desnaturalizar a ese intermediario y regresarlo al citosol, para que la molécula intentara nuevamente plegarse "espontáneamente" (Weissman *et al.*, 1996; Todd *et al.*, 1996). Así, estos puntos de vista opuestos acerca de lo que realmente hace una chaperonina en su interior -(de lo que no hay duda es de que el sustrato tiene que entrar al interior del anillo)- por el momento no están resueltos (Ellis & Hartl, 1996). A veces se argumenta que unos u otros protocolos experimentales no son correctos (Hartl 1996). Otras veces, se concede que los resultados discrepantes se podrían deber a la naturaleza del sustrato *i. e.* algunas proteínas como la dihidrofolato reductasa se pliegan dentro del anillo, mientras que otros como la rodanasa salen de la jaula (¿se escapan?), desnaturalizadas ("como locas"), pero eso sí, en un estado de compromiso de plegamiento ("a la próxima si la logro") (Mayhew *et al.*, 1996; Weissman *et al.*, 1996). Recientemente un reporte, propone que los sustratos podrían salir a veces plegados y a veces desnaturalizados (Burstin *et al.*, 1996). Una solución a estos conflictos y contradicciones ha sido esbozada recientemente. Al analizar la "desnaturalización" de Rnasa T1 por GroEL se determinó que la chaperonina tiene una alta afinidad por este estado (desnaturalizado) de la proteína, de tal manera que el equilibrio global de conformeros se inclina hacia este estado, pero no cataliza la desnaturalización (Walter *et al.*, 1996). Dicho de otro modo, desde el punto de vista conceptual, la distinción entre si la chaperonina pliega o desnaturaliza a su sustrato, podría ser trivial o semántica. Se podría decir que GroEL puede desnaturalizar a un intermediario mal plegado por medio de un mecanismo termodinámico acoplado, cuando la energía libre de Gibbs de la unión a GroEL es mayor que la estabilidad conformacional del intermediario y cuando su tasa de desnaturalización es alta (Walter *et al.*, 1996). Ideas similares habían sido propuestas por John Ellis desde 1990. En particular Ellis decía que era posible representar gráficamente el modelo de acción de una chaperonina, visualizando la energía libre para el plegamiento polipeptídico en la célula contra dos posibles resultados: la proteína adquiere rápida y espontáneamente su conformación nativa o se pliega incorrectamente y queda atrapada en un intermediario de plegamiento. Ellis especulaba que para algunas proteínas la adquisición de la estructura correcta podría estar favorecida tanto cinética como termodinámicamente y por lo tanto no se requerirían chaperonas. En otros casos las chaperonas podrían ser esenciales para favorecer la adquisición de la estructura nativa, inhibiendo ya fueran las vías de plegamiento incorrecto o ya fuera disminuyendo la energía de activación de la vía de plegamiento correcto (Ellis 1990b). Aparentemente en el caso de la chaperonina citoplasmática TRIC, existe evidencia de que las proteínas sustrato mantienen y posiblemente adquieren su conformación nativa en el interior de la chaperonina (Tan *et al.*, 1995a).

## Consideraciones adicionales

Si la función de las chaperoninas es promover el plegamiento y ensamble de otras proteínas ¿cómo se pliega y ensambla una chaperonina? La respuesta es ella misma. Ha sido demostrado que GroEL se auto-chaperonea. Esto es, cuando se hacen experimentos de replegamiento *in vitro* de GroEL, se requiere la presencia de GroEL nativo para que se formen nuevos oligómeros (Lissin *et al.*, 1990). Aún más, la Hsp60 de mitocondria es sintetizada en el citoplasma y debe de atravesar la membrana de la misma para desempeñar su función (Hartl & Neupert 1990). En este sentido la Hsp60 se comporta como cualquier otra proteína que sea importada al organelo y requiere de la presencia de Hsp60 pre-ensamblada funcional para poder formar oligómeros (Cheng *et al.*, 1990). Sin embargo, es posible que los oligómeros se ensamblen a partir de monómeros nativos (Lissin 1995). También, se ha reportado un protocolo que requiere sulfato de amonio y magnesio/ATP, mediante el cual es posible renaturalizar una pequeña cantidad de GroEL *in vitro* (Ybarra & Horowitz 1995). Sin embargo, parecería que *in vivo*, cpn60 es de las proteínas que obligadamente requieren cpn60. Por el otro lado cpn10 parece ser capaz de plegarse y ensamblarse sin ayuda ni de otra chaperonina, ni de protocolos rebuscados de plegamiento *in vitro* (Mascagni *et al.*, 1991). Recientemente estos resultados han sido confirmados y además se ha demostrado que GroES es capaz de asistir al plegamiento asistido de GroEL por ella misma (Searle *et al.*, 1996). Dicho de otro modo GroES se puede plegar espontáneamente, pero GroEL requiere GroEL y mejora su plegamiento si GroES asiste en la reacción.

Si bien *in vitro* las diferentes Hsps pueden actuar solas, *in vivo* en general, cooperan unas con otras en una compleja red de asociaciones (Buchner, 1996). En particular, en la mitocondria la Hsp70 y la Hsp60 interaccionan secuencialmente con los mismos substratos (Manning-Krieg *et al.*, 1991). De igual manera en *E. coli*, las proteínas substrato primero se unen al complejo DnaK y DnaJ, para posteriormente con la participación de GrpE, ser transferidas a GroEL/GroES (Langer *et al.*, 1992b). Adicionalmente a lo anterior, se ha demostrado que en sistemas eucariontes a medida que la proteína va saliendo del ribosoma, interactúa secuencialmente con Hsp70, Hsp40 (el homólogo de DnaK) y finalmente con la chaperonina del citoplasma eucarionte TRiC (Frydman *et al.*, 1994; Frydman & Hartl 1996).

En el espacio interior de la chaperonina, parece que se puede acomodar una proteína globular de hasta 90 kDa (Ellis, 1993). A este respecto, un problema no resuelto se refiere a la cuestión de cómo una proteína de mayor peso molecular podría entrar en esta cavidad central (Hartl 1996).

Recientemente un estudio de inmuno-microscopía electrónica indica que si bien el 80-85% de la chaperonina mitocondrial de mamíferos se localiza justamente en las mitocondrias; el 15-20 % restante se localiza en sitios discretos extra-mitocondriales. Estos sitios incluyen algunos próximos a

la membrana externa de las mitocondrias, pero también se localizan cerca del retículo endoplásmico, la superficie celular, vesículas no identificadas y en peroxisomas. Experimentos de inmunoprecipitación muestran que en condiciones normales estos *foci*, no corresponden al precursor sin procesar de Hsp60. De hecho cuando se bloquea el importe a las mitocondrias con el ionóforo nonactina, solo se puede detectar el precursor, sugiriendo que Hsp60 tendría que ser importada a la mitocondria, ser procesada y después regresar a estos otros sitios. Estos estudios plantean interesantes preguntas acerca del posible papel de Hsp60 en sitios extra-mitocondriales (Soltys & Gupta 1996).

### **¿Cuáles son los substratos fisiológicos de las chaperoninas?**

Si bien el intervalo de substratos que han sido ensayados *in vitro* es relativamente limitado, por el otro lado hay mucha evidencia experimental que indica que el rango de proteínas que interaccionan o requieren de chaperoninas es muy amplio. En primer lugar, se sabe que los productos polipeptídicos codificados por *groEL* y *groES*, son esenciales para la viabilidad celular y son requeridos para el crecimiento a todas las temperaturas. Mutantes en los que se han suprimido estos genes no son viables (Fayet *et al.*, 1989). Asimismo se ha demostrado que la amplificación y/o sobreexpresión de *groEL/groES*, suprime ya sea el efecto de mutantes termosensibles particulares (*dnaA46* / implicado en el inicio de la replicación en *oriC*), así como de múltiples productos polipeptídicos, lo cual sugiere un papel general de estas chaperoninas en las vías de plegamiento (Fayet *et al.*, 1986; Van Dyk *et al.*, 1989). Por su lado, se ha demostrado que la Hsp60 mitocondrial es esencial para el ensamblaje de proteínas que son importadas a las mitocondrias de levadura (Cheng *et al.*, 1989). También hay trabajos de que las proteínas que son importadas al cloroplasto requieren *cpn60* (Lubben *et al.*, 1989). Este tipo de evidencia ha llevado a considerar que las chaperoninas son promiscuas y que *in vivo* obligatoriamente muchas proteínas las requieren aun en circunstancias donde no hay stress ni sobreexpresión (Viitanen *et al.*, 1992b). No obstante lo anterior, no es totalmente claro qué significa "muchas proteínas", ni que tanto se deben generalizar este tipo de conceptos. Dicho de otro modo no está totalmente establecido qué proporción de proteínas interactúan necesariamente con las chaperoninas (Hartl 1996). Aún más, la rodanasa es uno de los substratos tradicionales en los ensayos *in vitro* de plegamiento asistido por chaperoninas, sin embargo se han reportado protocolos de plegamiento con una recuperación mayor del 80% en ausencia de chaperonina (Mendoza *et al.*, 1991).

En la siguiente sección que sigue quisiera acotar y delinear algunas ideas que me parecen relevantes al respecto del plegamiento asistido por las chaperoninas.

A veces en los experimentos de plegamiento *in vitro*, una chaperonina particular puede mostrar diferencias con respecto a substratos diferentes (**Martin et al., 1993a; Mayhew et al., 1996**); inclusive a veces no se logra plegar a una proteína desnaturalizada. Por ejemplo, ni GroEL, ni cpn60 mitocondrial son capaces de asistir en el plegamiento de  $\beta$ -actina ni  $\alpha$ -tubulina (ambas proteínas se localizan en el citoplasma de células eucariontes donde interactúan con la chaperonina citoplasmática, *vide infra*). Ambas proteínas son unidas y liberadas en ciclos por GroEL/mcpn60, pero no se produce proteína nativa como producto final (**Tan et al., 1995b**). Una conclusión obvia de estos resultados es que la estructura de las proteínas y la maquinaria de plegamiento han co-evolucionado de tal manera que las chaperoninas muestran preferencias y particularidades dependiendo del organismo o del compartimento celular (**Harl 1996**).

Por otro lado, un estudio reciente del grupo de Arthur L. Horwich, quien usó una mutación termosensible para apagar la producción de GroEL *in vivo*, concluye que un mínimo de 30% de las proteínas solubles en *E. coli* requieren GroEL para plegarse correctamente en condiciones de heat-shock. En general, en estos experimentos al pasar las células a la temperatura no-permisiva, la tasa global de síntesis de proteínas disminuye (como ya había sido descrito), pero en particular usando geles bidimensionales, se observa que algunas proteínas si bien son traducidas no alcanzan su conformación nativa y forman agregados insolubles (**Horwich et al., 1993**). Se puede enfatizar entonces que un importante grupo de proteínas necesitan de GroEL. En este sentido, los resultados del grupo de Horwich a *grosso modo*, concuerdan con otras observaciones *in vitro* que indican que aproximadamente el 50% de las proteínas solubles de *E. coli* se unen a GroEL después de diluir el agente desnaturalizante (**Vitnanen et al., 1992**). Pero la otra cara de la moneda está en que ambos resultados también indican, que entre el 50 y el 70% de las proteínas de *E. coli*, no requieren la ayuda de GroEL. Un caso particular es el que la  $\beta$ -galactosidasa no se ve afectada ni en su actividad específica, ni en su solubilidad (**Horwich et al., 1993**). Así una interpretación colateral sería que muchos más polipéptidos o bien utilizan asistencia de otras chaperonas moleculares como Hsp90, Hsp70, sHsps; o bien logran adquirir su conformación nativa de una manera espontánea. Respecto al primer punto, efectivamente se sabe que las chaperonas moleculares llegan a ser hasta cierto punto intercambiables e inclusive dispensables (**Buchner 1996; Bardwell & Craig 1988**). Respecto al segundo punto, desde luego la posibilidad de que algunas proteínas no requieran de asistencia *in vivo* para plegarse ha sido considerada (**Ellis 1990b; Todd et al., 1996; Walter et al., 1996; Weissman et al., 1996**). Al respecto se ha reportado que no todas las proteínas desnaturalizadas cuando son micro-inyectadas, son capaces de disparar la respuesta de stress (**Mifflin & Cohen 1994**). Pero aun así la pregunta subsiste ¿en concreto cuántas proteínas realmente requieren chaperoninas para plegarse *in vivo*? Recientemente George H. Lorimer ha sugerido una respuesta concreta (**Lorimer 1996b**). De

manera independiente, supongo, John Ellis y Franz Ulrich Hartl han llegado a conclusiones similares (Ellis & Hartl 1996). El razonamiento es como sigue. En lo general tenemos una idea razonable de la complejidad interior de una célula bacteriana como *Escherichia coli* (Goodsnell 1991). Considerando  $1.56 \times 10^{-13}$  g de proteína por célula, y suponiendo un peso molecular promedio de 40,000 daltones; entonces podemos considerar que en promedio en el interior de *E. coli* debe de haber alrededor de  $2.35 \times 10^6$  cadenas polipeptídicas. Sabiendo el tiempo de duplicación (40 minutos), se puede calcular la tasa mínima de síntesis de proteínas que correspondería a alrededor de 60,000 cadenas polipeptídicas por minuto por célula. Inclusive podríamos calcular con estas cifras que en promedio cada uno de los 35,000 ribosomas de *E. coli*, sintetiza un polipéptido cada 35 segundos, que corresponde a aquellos reportes de Howard Dintzis referentes a que en cada ribosoma se incorporan en la cadena polipeptídica 10 aminoácidos por segundo. Pero el punto realmente relevante es que cada 35 segundos salen de los ribosomas 35,000 nuevas cadenas polipeptídicas que deben plegarse correctamente, y este proceso continúa durante la fase de crecimiento exponencial, para que al final de cada ciclo celular deberán haber sido sintetizadas y plegadas mas de dos millones de cadenas polipeptídicas. La pregunta que se hace Lorimer es ¿existe suficiente GroEL/GroES en la célula para contender con esta profusión de síntesis de proteínas? La respuesta concreta, directa es: no, no hay chaperonina que alcance para esta tasa de plegamiento. Pero creo que vale la pena abordar un poco sobre el enfoque de Lorimer.

Para tener una respuesta a su pregunta; de hecho para que el plantear la pregunta sea posible: Lorimer tuvo que reunir una cantidad de datos dispersos en la literatura, que han sido obtenidos pacientemente por una gran cantidad de investigadores. Estos no son datos espectaculares (como calcular el volumen celular o la concentración de macromoléculas y agua), pero son resultados específicos que al irse juntando pueden dar una visión mucho mas amplia. Un punto de vista que a menudo pasa desapercibido, es que actualmente la biología molecular continúa siendo una disciplina principalmente descriptiva (Maddox 1994b). Por ejemplo es claro que los varios proyectos de secuenciación de genomas completos pretenden en esencia tener la lista de genes y su secuencia de nucleótidos (Osion, 1995). Claramente, por si mismo este catálogo no nos va a decir nada acerca de la lógica molecular de funcionamiento de un sistema biológico. Cada vez que se reporta la secuencia de un nuevo genoma la impresión que le queda a uno es de atiborramiento. Es claro que hoy por hoy, cientos de miles o millones de pares de bases son simplemente indigeribles (Short, 1995). Para tener una idea de la magnitud de los proyectos de secuenciación, el catálogo de genes se ha comparado con la tabla periódica de los elementos i. e. 100, 000 genes en el humano vs. 100 elementos en la tabla (Lander, 1996). No obstante el aparente estado de confusión que actualmente nos pueda parecer que se deriva de estos

proyectos, es claro que si queremos movernos hacia la siguiente etapa del conocimiento, vamos a tener que intentar asimilar los datos y encontrar nuevas formas de visualizar los genomas. El momento actual, es el momento previo a una posible síntesis e integración del conocimiento que tenemos de los sistemas biológicos a nivel molecular, que va a cambiar radicalmente nuestra forma de ver la vida. Desde luego el saber totalmente la secuencia de los 4,000 genes de *E. coli* o de los 100, 000 genes del humano no nos va a decir qué es un hombre, ni siquiera qué es una bacteria. Pero ya se anticipa que este conocimiento va a tener un impacto notable no solo en la ciencia sino en nuestra vida cotidiana. Actualmente estamos apenas comenzando a olfatear en las listas de genes, ya se anticipa que no pasará mucho tiempo antes de que comencemos a extraer conclusiones mas generales (**Maddox, 1994b**). Por ejemplo, aunque en realidad no esté secuenciado ni el 5% de genoma humano, ya se anuncia que "alegremente" nos dirigimos a la era postgenómica (**Nowak, 1995**). Desde luego, no ha pasado desapercibido el hecho de que las proteínas son los efectores teleonómicos de la economía celular (**Monod, 1971**) y que si queremos entender a los genomas necesitamos entender al "proteoma"; queriendo este término significar las relaciones estructurales, funcionales y regulatorias de las proteínas en las células (**Kahn, 1995; Lander, 1996**). Es curioso a este respecto, que si bien entendemos en lo general el código genético, aún no comprendamos lo que se ha llamado **la segunda mitad del código genético** que son precisamente las reglas de plegamiento de los polipéptidos, esto es, cómo es que la información adimensional del gen produce estructuras tridimensionales funcionales que dan a las proteínas sus propiedades biológicas (**Ellis & van der Vies, 1991**). En concreto, antes de la gran síntesis, viene la fase descriptiva y de recopilación de datos. Una analogía podría ser ilustrativa. Antes de que Darwin pudiera hacer esa gran síntesis conceptual que conocemos como la Teoría General de Evolución por Selección Natural; Lineo tuvo que producir un sistema de clasificación coherente de los organismos. Dicho de otro modo. El sistema de clasificación de Lineo de acuerdo a las características morfológicas de los seres vivos, constituye un catálogo ordenado de la complejidad biológica que fue fundamental para que Darwin pudiera extraer conclusiones mucho más amplias y generales (**Darwin 1872**) . Lineo demarcó las características de los individuos a nivel específico. Darwin trascendió este nivel y percibió, que lo importante en la evolución es la población (**Jacob 1970**). En este sentido, Darwin se anticipó inclusive a Boltzman, Maxwell y Gibbs al incorporar conceptos que abarcan grandes números. Cuando pensamos que el biólogo molecular moderno no puede dejar de considerar la ley de acción de masas, aunque a veces no sea sencillo hacerlo (**Maddox 1994b**), implícitamente se está rindiendo un tributo al raciocinio de Darwin. Si aislamos y caracterizamos moléculas sin considerar su entorno físico real estamos limitando nuestra perspectiva. Tanto a nivel evolutivo como a nivel molecular, cantidades específicas de individuos o moléculas logran que eventos altamente

improbables ocurran con certeza, o como elegantemente lo dijera el gran Jacques Monod: del azar surge la necesidad (**Monod, 1971**). Regresando al planteamiento inicial, Sydney Brenner ha enfatizado que actualmente la biología molecular ha regresado a la vieja tradición científica de observación, medición y cuantificación; una ganancia en su opinión. Brenner opina que aquellos que quieren saberlo todo acerca de una sola cosa, deben ser comprensivos con aquellos que quieren aprender un poco de fenómenos más generales y ciertamente antes de los experimentos cruciales, de las grandes integraciones, están los catálogos, los datos concretos, las medidas y cuantificaciones (**Brenner, 1992**). No es necesario ser demasiado rebuscado, un manejo sencillo de los datos puede llegar a revelar mucho más de lo que se piensa. Ese es justamente el enfoque de Lorimer.

Habiendo calculado la tasa de síntesis de proteínas, Lorimer utilizando datos de la literatura, calcula la cantidad de GroEL en la célula bacteriana. Se sabe a *grosso modo* que *E. coli*, contiene 1.35 y 0.24% de GroEL y GroES respectivamente respecto a la proteína total. Durante el stress calórico GroEL puede llegar a constituir hasta el 15% de la proteína total (**Lindquist, 1986**). Sabiendo el peso molecular de los monómeros y su ensamblaje en oligómeros, se puede calcular que en *E. coli* durante la fase de crecimiento exponencial existen **1,580 partículas de GroEL<sub>14</sub> y 3,072 de GroES<sub>7</sub>**. La razón de la presencia de aproximadamente el doble de oligómeros de GroES es posiblemente una consecuencia de que ambos genes se encuentran en un operón policistrónico (**Tilly et al., 1981**). Por otro lado, se ha sugerido que esta asimetría de concentraciones podría dar apoyo a la propuesta de que el intermediario funcional de la chaperonina es la estructura simétrica con dos moléculas de GroES asociadas a los extremos de GroEL, *vide supra* (**Lorimer 1996**). Sin embargo, un punto de vista alternativo sería que, dado que GroES se asocia y disocia de GroEL durante los ciclos de plegamiento, justamente se esperaría que su concentración fuera mayor para facilitar el encuentro por difusión del heptámero y del decatetramero. Ahora, promediando las tasas de plegamiento asistido por chaperoninas, Lorimer calcula que a 37 °C éstas van de 0.07 min<sup>-1</sup>, a 1.05 min<sup>-1</sup>. Aun considerando la tasa más rápida y multiplicando por el número de moléculas de GroEL, se concluye que a lo sumo se podrían plegar 1,660 cadenas polipeptídicas por minuto, por célula, o sea 36 veces menos que la tasa mínima de producción de cadenas polipeptídicas. Dicho de otro modo, la chaperonina de *E. coli*, solo es responsable del plegamiento de una pequeña proporción (**2-5%**), del total de proteínas en la célula (**Lorimer, 1996b**). Como se indicó anteriormente Ellis y Hartl, utilizando cálculos similares concluyen que GroEL, alcanza para asistir al plegamiento del 7.5% del total de polipéptidos (**Ellis & Hartl 1996**). Estos resultados sugieren entonces que o bien muchas proteínas alcanzan su conformación nativa de manera espontánea en ausencia de stress o bien son asistidas por otras chaperonas moleculares (**Lorimer 1996b; Ellis & Hartl 1996; Buchner, 1996**).

Entonces el hecho de que las chaperoninas puedan asistir al plegamiento de substratos proteicos *in vitro*, no necesariamente prueba que tal o cual proteína es su substrato *in vivo*, y entonces aparece ahora un punto de vista algo contrastante con la supuesta promiscuidad de las chaperoninas: el 90-95% de las proteínas de *E. coli* normalmente no necesitan a GroEL (**Lorimer, 1996b**). Por supuesto en condiciones de stress, donde se incrementa la tasa de síntesis de GroEL y disminuye la tasa global de síntesis de proteínas (**Horwich et al., 1993**), el papel de la chaperonina debe ser más relevante y de hecho esto explicaría su promiscuidad. Sin embargo, es curioso que en *Drosophila* la Hsp60 no se comporte como una proteína clásica de stress calórico (**Lakhota & Singh 1996**). En apoyo a este rango de substratos más limitado de las chaperoninas, recientemente se ha reportado que el sistema de chaperonina no es esencial para el plegamiento de todas las proteínas que son importadas a las mitocondrias de la levadura (**Rospert et al., 1996**). Como sea, vemos que actualmente el péndulo está oscilando de la hipótesis de que las chaperoninas interactúan con una gran cantidad de substratos, a un enfoque que circunscribe mas a los polipéptidos que *in vivo* requieren chaperoninas. Pero aun así, hay dos aspectos que no pueden ser minimizados: **a)** la presencia universal de homólogos de GroEL a lo largo de la escala filogenética que habla de su papel fundamental e indispensable (**Zeilstra-Ryalls, et al., 1991; Gupta 1995**) y **b)** el hecho de que *groEL/groES* sean indispensables a todas las temperaturas y que mutaciones en estos genes no son viables (**Fayet et al., 1989**). A este respecto, Lorimer calcula que si *E. coli* contiene alrededor de 1,860 proteínas diferentes el 4% de éstas equivalen a aproximadamente 74 proteínas. Si todas éstas, son absolutamente dependientes de GroEL para alcanzar su conformación nativa, entonces la delección de *groEL/groES* sería el equivalente de simultáneamente suprimir los genes de esas 74 proteínas. Si sólo uno de estos codificara un producto polipeptídico esencial para la sobrevivencia de la célula, se entendería por que el efecto de deletar a *groEL/groES* es letal (**Lorimer 1996b**). A este respecto, utilizando una estrategia de mutagénesis por inserción al azar, se ha estimado que la frecuencia de productos génicos esenciales en *Bacillus subtilis* (cuyo genoma es de tamaño similar al de *E. coli*), es de aproximadamente 7-8% (6:79). Estos genes mapean a todo lo largo del genoma de esta bacteria y algunos como *dnaA*, *dnaB* y *rpoD*, han sido identificados (**Itaya 1995**). Desde luego esta coincidencia de cifras, no implica que justamente esos genes esenciales sean los mismos que requieren a GroEL para plegarse; pero refuerza la idea de que al menos algunos de los productos polipeptídicos que requieren a la chaperonina podrían ser esenciales (**Lorimer 1996b**). Una conclusión de los experimentos arriba descritos es que el tamaño mínimo para discriminar entre lo vivo y lo no vivo se ubica entre 318-562 kpb de tamaño de genoma (**Itaya 1995**). En este respecto, el mayor tamaño de genoma reportado para un virus es de 370 kbp (**Itaya 1995**) y el de *Mycoplasma genitalium*, reportado como el genoma más pequeño para

un organismo de vida libre es de 580 kbp (**Fraser et al., 1995**). *Mycoplasma* tiene homólogos de *dnaK*, *dnaJ*, *grpE* y *groEL/ES*.

### **Aspectos particulares de las chaperoninas en plantas (vegetales).**

Como se describió antes la chaperonina RuSBP (*cpn60*), fue la que permitió desarrollar a profundidad el concepto de chaperona molecular (**Gatenby & Ellis, 1990**). Cuando esta proteína plastídica es purificada a homogeneidad y analizada en geles desnaturizantes de poliacrilamida se pueden resolver dos polipéptidos de movilidad muy similar que han sido denominados *cpn60- $\alpha$*  y *cpn60- $\beta$*  (**Hemmingsen & Ellis, 1986; Musgrove et al., 1987**). No hay evidencia de que este tipo de "subunidades" se encuentren ni en eubacterias, ni en mitocondrias (**Gatenby & Ellis, 1990**). En los plástidos estas llamadas subunidades  $\alpha$  y  $\beta$  se encuentran en concentraciones equivalentes y tienen entre sí alrededor del 50% de identidad a nivel de aminoácidos; esto es no son mucho más similares entre sí, de lo que lo son a GroEL (**Musgrove et al., 1987; Martel, et al., 1990**). Sorprendentemente no se sabe si los polipéptidos  $\alpha$  y  $\beta$ , se localizan en el mismo o en diferentes tetradecámeros *i. e.* no se sabe si la chaperonina del plástido forma homo o hétero oligómeros con estas subunidades, ni si cada uno de los péptidos forma un heptámero, ni si existen oligómeros que tuvieran diferente cantidad de  $\alpha$  y  $\beta$  (**Viljanen et al., 1995**). Los genes de las *cpn60- $\alpha$*  y la *cpn60- $\beta$*  de *Brassica napus*, han sido clonadas y expresadas en *E. coli* tanto juntos como de manera individual. Desafortunadamente, los niveles de expresión en estos estudios fueron bajos además de que las proteínas recombinantes formaron partículas mixtas con la GroEL endógena. A pesar de lo anterior se pudo concluir que para que  $\alpha$  pueda incorporarse en tetradecámeros se requiere la expresión de las subunidades  $\beta$ , mientras que esta última se ensambla en oligómeros sin requerir de  $\alpha$ . Interesantemente en estos estudios, las bacterias mostraron un crecimiento anormal, además de que la sobre-expresión no regulada de la chaperonina de *B. napus* fue letal para *E. coli* (**Cloney, et al., 1992a; Cloney et al., 1992b; Wu et al., 1993**). Un efecto similar ha sido observado cuando se expresa la *cpn60- $\beta$*  de *Arabidopsis* en *E. coli* (**Valdés-López, observación no publicada**). Se sabe que ocasionalmente la expresión de genes eucariontes en *E. coli* puede tener efectos tóxicos para la bacteria (**Herskowitz 1987**). Sin embargo, dada la similitud estructural y funcional de las chaperoninas éste es un resultado inesperado. Como explicación se ha propuesto que la expresión de las chaperoninas vegetales perturban las funciones normales de GroEL, o que interfieren con algún papel regulatorio sobre el regulón de heat shock (**Wu et al., 1993**). Asimismo, en lo que considero un intento desesperado, se ha reportado la co-expresión de los genes de *cpn60  $\alpha$*  y/o  $\beta$ , la co-chaperonina del plástido (el homólogo de GroES) y las subunidades L y S de la Rubisco de trigo. El resultado: nada se

ensambla correctamente (Cloney *et al.*, 1993). Por el otro lado de la moneda, plantas transgénicas de tabaco que expresan e importan a GroEL recombinante al cloroplasto, no muestran diferencias ni de crecimiento ni de desarrollo respecto a los controles y aparentemente se ensambla algo de GroEL<sub>14</sub> correctamente, además de moléculas mixtas (Wu, *et al.*, 1993). Asimismo, se han reportado efectos pleiotrópicos en plantas transgénicas que expresan niveles altos o bajos de cpn60- $\beta$  (Zabaleta *et al.*, 1994b; Vitanen *et al.*, 1995). Curiosamente, se ha reportado que la expresión en *E. coli* de la Hsp60 mitocondrial de *Homo sapiens* resulta en el ensamblaje correcto del anillo homo-heptamérico funcional de esta chaperonina, sin que se reporte toxicidad en la célula bacteriana (Singh & Gupta, 1992). Así, posiblemente los resultados anteriores pudieran reflejar algo de la divergencia evolutiva de esta familia de proteínas. Hace tiempo que utilizando secuencias de genes y proteínas, se ha propuesto que tanto mitocondrias como cloroplastos se derivaron de procariontes particulares (Schwartz & Dayhoff, 1978). En este sentido, claramente, las Hsp60 de mitocondria forman un grupo taxonómico que las agrupa con las chaperoninas de las  $\alpha$ -proteobacterias (subdivisión  $\alpha$  de las bacterias púrpuras Gram-negativas), mientras que las cpn60 de cloroplastos de plantas y algas rojas se agrupan con aquellas de cianobacterias, lo que sugiere un origen evolutivo común en ambos casos (Viale & Arakaki, 1994; Gupta 1995).

Otra característica particular de las chaperoninas en plantas superiores, es que solo se inducen moderadamente por calor (Zabaleta *et al.*, 1994a; Vitanen *et al.*, 1995). Sin embargo, en el alga unicelular *Chlamydomonas reinhardtii*, tanto cpn60- $\alpha$ , como cpn60- $\beta$ , se sobreinducen entre 10 y 30 veces por efecto del calor (Thomson *et al.*, 1995).

A diferencia de bacterias y mitocondrias, en cloroplastos, la co-chaperonina cpn10 (homólogo de GroES), es del doble de tamaño y consiste de dos moléculas cpn10 que se han fusionado "cola con cabeza", formando un único polipéptido de 21 kDa (Berish, *et al.*, 1992). Se ha propuesto, que esta organización binaria podría reflejar alguna adaptación particular de las plantas superiores en respuesta a las subunidades  $\omega/\beta$  de las cpn60 (Vitanen *et al.*, 1995). Esta doble organización de la co-chaperonina está conservada filogenéticamente y se ha demostrado en hepáticas, musgos, helechos, gimnospermas, monocotiledóneas y dicotiledóneas (Baneyx, *et al.*, 1995). Esta co-chaperonina (rebautizada como cpn21) expresada ya sea completa o ambas "mitades" por separado, pueden funcionar correctamente *in vivo*, en mutantes de *E. coli* deficientes de GroES e interactúan con GroEL (Baneyx, *et al.*, 1995). Sorprendentemente tanto GroES, como el homólogo mitocondrial son a su vez compatibles con cpn60 de plástido *in vitro* (Vitanen, *et al.*, 1995). Complicando un poco el panorama, se ha descrito que GroEL funciona bien con cpn10 mitocondrial además de cpn21 y desde luego con GroES (Vitanen *et al.*, 1995). Sin embargo, otro reporte concluye que *in vitro*, usando a GroEL

como chaperonina los dominios aislados de cpn21 son incapaces de ayudar como co-chaperoninas (**Bertsch & Soll 1995**). Por su lado, la cpn60 de mitocondria de mamífero requiere estrictamente a la co-chaperonina de origen mitocondrial (**Viljanen et al., 1992a**), pero la cpn60 de mitocondria de levadura no puede funcionar con GroES, pero es parcialmente compatible con cpn21 de plantas. Se ha propuesto que este efecto posiblemente sería mediado por la mitad de cpn21 que se pareciera mas a la cpn10 mitocondrial (**Viljanen et al., 1995**). Recientemente, usando anticuerpos contra cpn60 se ha encontrado que a diferencia del estroma en los cloroplastos, en el lumen de los tilacoides se encuentra una sola isoforma de cpn60 y que cpn10 es de tamaño "normal" (**Schlicher & Soll 1996**).

Por último cabría mencionar que en las mitocondrias de plantas se localiza una chaperonina similar a las encontradas en las mitocondrias de animales (**Prasad & Hallberg 1989; Prasad & Steward 1992; Cloney et al., 1994; Cole et al., 1994; Gupta, 1995**)

### **TRIC (TCP-1 containing Ring Complex).**

Para 1990, las chaperoninas habían sido descritas en bacterias, cianobacterias, mitocondrias y cloroplastos y se había establecido su papel fundamental para el plegamiento correcto de otros polipéptidos (**Morimoto et al., 1990**). Sin embargo, existía una paradoja. Ninguna de las chaperoninas descritas se localizaba en el citoplasma de células eucariontes. Si la función de asistir al plegamiento de otros polipéptidos era universal, el porque no aparecían en este compartimento celular era un misterio. Considerando la hipótesis endosimbiótica para el origen de mitocondrias y cloroplastos (**Margulis, 1970**), parecía como si el fenómeno fuera exclusivamente de tipo procarionte. Dicho de otro modo las chaperoninas citoplasmáticas brillaban por su ausencia. Justamente en ese año, Radhey S. Gupta, condujo una búsqueda amplia de similitud usando los bancos de almacenamiento de secuencias y detectó que había algunas proteínas que mostraban alguna similitud con las chaperoninas. En particular reportó que la mayor similitud era mostrada por una oscura proteína citoplasmática compuesta de monómeros de 62 kDa, denominada "t complex polypeptide 1" (TCP-1) y propuso que ésta podría ser la versión citoplasmática de la chaperonina (**Gupta, 1990a**). En este estudio, se mostraba que si bien la identidad a nivel de aminoácidos era extraordinariamente baja (17-20%), considerando criterios de similitud entre las substituciones, era posible inferir homología. Utilizando un enfoque similar John Ellis llego a las mismas conclusiones (**Ellis 1990**). Si bien estos dos trabajos fueron los primeros que detectaron a las chaperoninas citoplasmáticas, todavía pasaría algún tiempo antes de la confirmación experimental.

La proteína TCP-1, había sido descrita como el producto de un gen codificado en el complejo *t* del ratón (**Willison et al., 1989**). Los cromosomas *t* son variantes del cromosoma 17 del ratón que

existen de manera natural en las poblaciones y que se caracterizan por tener dos grandes inversiones (**Herrmann et al., 1987**). Esta región es llamada complejo *t* y se ha calculado que incluye alrededor de 100 loci genéticos (**Lyon et al., 1988**). Además del ratón había sido demostrado que existían homólogos de este gen en humanos, *Drosophila*, y levadura, si bien la comparación de las secuencias derivadas de aminoácidos no habían permitido asignar una función a estas proteínas (**Willison et al., 1989**). Había sido demostrado que TCP-1 está asociada a cambios en la espermatogénesis, expresándose de manera notable en los testículos del ratón adulto y en menor grado en muchos otros tejidos como la médula ósea, el timo, el cerebro y el corazón (**Willison et al., 1989; Kubota et al., 1992**). Usando anticuerpos monoclonales contra TCP-1 se había demostrado la presencia de proteína en extractos solubles de chícharos pero no en cloroplastos o mitocondrias (**Ellis 1990a**). Curiosamente se reportó también que utilizando sueros anti-chaperonina (GroEL pero no RuSBP), se podían visualizar por microscopía electrónica polipéptidos citoplasmáticos en células de hojas y coleoptilos de avena (**Grimm et al., 1991**).

Pero el momento clave, vino cuando se clonó el gen de la proteína principal de heat-shock (factor termofílico 55 / TF55), de la arquea termofílica *Sulfolobus shibatae* (**Trent et al., 1991**). El razonamiento detrás de estos estudios fue que se ha sugerido que arquea está evolutivamente relacionada al citosol eucarionte (**Woese & Fox 1977; Woese, 1981; Woese et al., 1990**). En particular, *Sulfolobus shibatae* habita a 75 °C, pero cuando es cambiada a 85 °C adquiere termotolerancia y la síntesis de TF55 se sobreinduce notablemente (**North 1991**). TF55 fue entonces purificado, demostrándose por microscopía electrónica de transmisión de barrido que el complejo nativo forma dos anillos de nueve subunidades cada uno, notablemente similares a los formados por las chaperoninas clásicas. Asimismo TF55 hidroliza ATP y se asocia a polipéptidos desnaturalizados. Por otro lado, utilizando la secuencia parcial de aminoácidos de TF55, se logró aislar el gen correspondiente. Sorprendentemente, la secuencia derivada de aminoácidos mostró no estar claramente relacionada con las chaperoninas sino tener alta identidad 36-40% con TCP-1, lo cual sugería fuertemente que esta última efectivamente podría ser la elusiva chaperonina citoplasmática (**Trent et al., 1991**).

La comprobación definitiva vino cuando se demostró que efectivamente una proteína homóloga a TCP-1 organizada en dos anillos toroidales y presente en el citoplasma de células eucariontes, tenía la capacidad de asistir en el plegamiento de las proteínas citoplasmáticas actina, tubulina y contractina, todas relacionadas con el citoesqueleto (**Gao et al., 1992; Yaffe et al., 1992; Melki et al., 1993**). En este sentido, utilizando microscopía electrónica se ha demostrado que al igual que en GroEL, la proteína substrato se localiza en la concavidad central de TRiC (**Marco et al., 1994**). Interesantemente, se reportó que TCP-1, formaba un complejo oligomérico de 800-950 kDa que contenía además, de cuatro a seis polipéptidos adicionales (con pesos

moleculares individuales entre 53-72 kDa) y posiblemente dos moléculas de Hsp70, de tal manera que a pesar de las similitudes con cpn60, este complejo tenía características únicas (Lewis *et al.*, 1992). Este aspecto comenzó a aclararse cuando se determinó la secuencia parcial de aminoácidos de algunas de estas otras proteínas. Todas mostraron una alta identidad (40-60%) con la propia TCP-1, proponiéndose entonces que estaban codificadas por una nueva familia multigénica (Frydman *et al.*, 1992). Es en este trabajo cuando se acuña el término **TRiC (TCP-1 Ring Complex)** (Frydman *et al.*, 1992). Posteriormente también se ha usado el término "**Chaperonin Containing TCP-1 / CCT**" (Kubota *et al.*, 1994). Algunos autores para distinguir a los grupos de chaperoninas las han clasificado como chaperoninas de grupo I o clásicas (GroEL/Hsp60/cpn60) y chaperoninas del grupo II o citoplasmáticas (TCP-1/CCT/TRiC) (Hemmingsen 1992; Kubota *et al.*, 1995). Este mismo grupo (Kubota *et al.*, 1994) , logró la clonación y secuenciación de seis de los nueve genes del ratón. Las secuencias tienen entre si de 25-36% de identidad, incluyendo a TCP-1 y 27-39% de identidad con TF55. Se propuso entonces, en base a esta divergencia, que CCT llevaría a cabo funciones más complejas requeridas para el ensamblaje de distintas proteínas del citoplasma eucarionte (Kubota *et al.*, 1994). Este optimista punto de vista se veía reforzado por reportes de que TRiC permitía el plegamiento de luciferasa y tubulina (Frydman *et al.*, 1992); los reportes de que había variación tejida específica en las subunidades (Kubota *et al.*, 1994); la clonación de genes homólogos en *Arabidopsis* (Mori *et al.*, 1992) y la propuesta de que esta chaperonina estaba involucrada en la biogénesis del fitocromo en células vegetales (Mummer, *et al.*, 1993). Este último reporte fue posteriormente retirado (Mummer, *et al.*, 1994) al clonarse el gen de la supuesta chaperonina y demostrarse que en realidad correspondía a  $\beta$ -glucosidasa (Gus-Mayer *et al.*, 1994). Curiosamente se informó que GroEL si es capaz de promover la renaturalización y la formación de oligómeros funcionales del fitocromo *in vitro* (Grimm *et al.*, 1993).

Sin embargo, si bien algunos reportes sugieren que TRiC pudiera estar involucrada en el plegamiento de algunos polipéptidos adicionales, la mayor parte de la evidencia sugiere fuertemente que TRiC "chaperonea" particularmente a actina y tubulina (Craig *et al.*, 1994; Kubota *et al.*, 1995; Tian *et al.*, 1995b; Hartl 1996). Estos resultados plantean varios problemas. En primer lugar, si TRiC es una chaperonina específica, ¿por qué requiere de tantas subunidades? De hecho se ha especulado que estas moléculas hetero-oligoméricas son desventajosas desde el punto de vista que la probabilidad de que una mutación que afectara a alguno de los ocho o nueve genes es mayor que en el caso de moléculas homo-oligoméricas (Kubota *et al.*, 1995). Inclusive en el caso de arquea las chaperoninas homólogas, no parecen tener más de dos subunidades (Phipps *et al.*, 1993; Knapp *et al.*, 1994). Así, la estructura hetero-oligomérica de TRiC debe de tener importantes ventajas, aunque por el momento no está claro cuales serían éstas; si

bien la estructura de subunidades diferentes parece estar conservada desde ciliados y protozoarios hasta mamíferos (**Kubota et al., 1995**). Podría ser que actina y tubulina fueran proteínas especiales dado que constantemente se ensamblan y desensamblan dentro de la dinámica constante de cambio del citoesqueleto (**Solomon 1991**). Dado que mutantes termosensibles en los genes que codifican las diferentes subunidades de TRiC afectan fenotípicamente a actina y tubulina, se ha propuesto que evolutivamente TRiC se ha encargado de contender con substratos problemáticos como estas proteínas (**Kubota et al., 1995**). Una explicación alternativa podría ser que durante el ensamble y desensamble del citoesqueleto existen múltiples oportunidades de interacciones incorrectas (*dangerous liaisons*). De hecho se sabe que actina puede interactuar con más de 110 proteínas diferentes las cuales regulan y modulan su ensamblaje (**Sheterline & Sparrow, 1994**). De hecho utilizando las cantidades estimadas de TCP-1 (**Lewis et al., 1992**), se puede calcular que en células germinales o en células en cultivo hay aproximadamente  $2-3 \times 10^5$  partículas de TRiC por célula. Esta alta concentración se explicaría justamente en función de la dinámica del citoesqueleto. En este sentido una paradoja sería el aparentemente largo tiempo que TRiC utiliza en su ciclo de plegamiento. Por ejemplo el tiempo de plegamiento para la luciferasa es de aproximadamente 10 minutos (**Frydman et al., 1992**); por su lado actina requiere de 2-3 minutos y tubulina de 5-10 (**Sternlicht et al., 1993**). Así, la perspectiva general parece regresar a una paradoja: ¿Porqué en el citoplasma eucarionte las chaperoninas (TRiC), parecen mostrar esta selectividad hacia proteínas del citoesqueleto? Ahora bien, por un lado TRiC parece estar involucrada en el plegamiento inicial de actina y tubulina y posiblemente en el de algunas proteínas de los filamentos intermedios. Pero por el otro lado otras chaperonas moleculares (Hsp70, Hsp90 y Hsp27) están involucradas en el ensamble y desensamble y/o en la protección al heat shock de los centrosomas, microtúbulos, microfilamentos y filamentos intermedios (**Liang & MacRae, 1997**). Dicho de otro modo, los requisitos particulares de asistentes del plegamiento de las proteínas del citoesqueleto son notablemente complejos y apenas comienzan a aparecer algunos resultados tendientes a esclarecer estas complejas interacciones.

Curiosamente, la posibilidad de que las chaperoninas estuvieran involucradas en el ensamble de citoesqueleto ya había sido postulada. La diferencia, es que se había pensado en un mecanismo que implicaría a las chaperoninas mitocondriales y se proponía que tubulina y actina entrarían a la mitocondria interactuarían con la Hsp60 de este organelo y saldrían al citoplasma a formar los microtúbulos y los microfilamentos (**Gupta 1990b**). Esta propuesta no ha tenido apoyos experimentales posteriores.

Desde luego lo anterior, no excluye la posibilidad de que TRiC participe en el plegamiento de otros polipéptidos del citosol ya que ha sido demostrado que a diferencia de GroEL, TRiC

interactúa y asiste al plegamiento de polipéptidos a medida que estos van emergiendo del ribosoma *in vitro* (Frydman *et al.*, 1994). Es curioso que dentro de las pocas proteínas para las que se ha demostrado la participación de TRiC, se encuentra la proteína de la cápside del virus de la hepatitis B (Lingappa *et al.*, 1994). A este respecto es interesante que se ha reportado que el gen de TCP-1, no es inducible por calor (Kubota *et al.*, 1995), mientras que TF55 (termosoma) de arquea si lo es (Trent *et al.*, 1991; Phipps *et al.*, 1993). Asimismo, es posible que en arquea, el termosoma participe en el plegamiento de otros polipéptidos ya que se ha demostrado que el homólogo de *Sulfolobus solfataricus* previene la inactivación irreversible a 70-90 °C de varias enzimas termofílicas y mesofílicas *in vitro* (Guagliardi *et al.*, 1994). Sin embargo, es posible que en arquea TF55 también participe en el plegamiento de posibles proteínas del "citoesqueleto". En este sentido FtsZ, el supuesto homólogo procarionte de tubulina ha sido localizado en el genoma de *Methanococcus jannaschii* (Bull *et al.*, 1996). Aún más, el homólogo de FtsZ, de la arquea extremófila *Halobacterium salinarium*, muestra mayor similitud hacia las tubulinas eucariontes que con los homólogos eubacterianos (Margolin *et al.*, 1996). Por el otro lado, en arquea no se ha localizado el gen que codificaría para FtsA (Bull *et al.*, 1996). Esta proteína, aunque tiene un bajo nivel de expresión, es esencial para la división celular en *E. coli* y se ha hipotetizado que sería el equivalente de actina en procariontes (Sánchez *et al.*, 1994). Como fue discutido más arriba, la supuesta relación estructural entre esta proteína, Hsp70, hexoquinasas y actina (Bork *et al.*, 1992), en el mejor de los casos representaría un caso de convergencia, no de divergencia y además como se dijo no aparece en el genoma de *M. jannaschii*. Interesantemente, se ha propuesto que en arquea el "citoesqueleto" podría estar constituido por su propia chaperonina, ya que ha sido visualizado por microscopía electrónica que tanto *in vivo*, como *in vitro* se forman filamentos de chaperonina (Trent, *et al.*, 1997).

Por último, parece que ni TRiC, ni el termosoma de arquea requieran co-chaperoninas similares a GroES. A este respecto se ha reportado que dos factores (cofactores A y B), pueden estimular el plegamiento de tubulina, pero no de actina *in vitro* (Gao *et al.*, 1993). El cofactor A actúa como dímero, pero al igual que GroES, modula la actividad de ATPasa de la chaperonina citoplasmática. Sin embargo, la secuencia de aminoácidos derivada de la clona de cDNA, define un polipéptido de 13 kDa, sin ninguna homología discernible ni a GroES, ni a ninguna otra secuencia reportada (Gao *et al.*, 1994). Es interesante que la proteína Gp23 de la cápside del fago T4, que interactúa con GroEL, no requiere de GroES. En vez de ésta, su ensamblaje correcto requiere de la proteína Gp31, codificada por el genoma de T4. Esta proteína no tiene similitud con GroES, pero puede sustituirla en el ensamblaje de T5 y  $\lambda$  *in vivo*, así como participar en el ensamblaje asistido por GroEL, *in vitro* e *in vivo* de Rubisco de cianobacteria (van der Vlies *et al.*, 1994). En el caso del termosoma de *Pyrodictium occultum* visualizado por microscopía

electrónica, se ha observado que en el sitio donde debería aparecer la apertura en los anillos se observa una masa que parece ocultar dicha apertura. Se ha sugerido como una alternativa que en este termosoma, los equivalentes de GroEL/GroES, se encuentran integrados en un solo polipéptido (Phipps *et al.*, 1993). A este respecto, recientemente se ha reportado una cierta similitud entre los aminoácidos 314-431 de GroEL de *E. coli* y GroES de *Bacillus subtilis* (Gupta 1995). La región de solapamiento de 94 residuos muestra 31 identidades, más 19 sustituciones conservativas. Anteriormente, otro reporte demostraba similitud entre cpn10 y cpn60. Sin embargo en este caso la zona de solapamiento correspondía a los aminoácidos 114-234 de cpn60 (Martell *et al.*, 1990). Desde luego una posibilidad que explicaría estos resultados, sería que efectivamente hubiera dos regiones en la chaperonina 60 que tuvieran similitud con cpn10/GroES. Sin embargo, si esto fuera cierto, un alineamiento de cpn60/GroEL contra si misma, debería mostrar una duplicación interna, cosa que no se detecta (Valdés-López resultados no publicados). Además, si cualquiera de las dos regiones de GroEL que supuestamente muestran similitud de secuencia con GroES fueran realmente significativas, se esperaría poder detectar similitud estructural. Justamente, al comparar de manera sencilla las estructuras secundarias y tridimensionales de los monómeros respectivos no se observa ninguna similitud. GroES topológicamente corresponde a un barril de tipo  $\beta$  irregular i. e. contiene exclusivamente regiones en  $\beta$ -conformación (Hunt *et al.*, 1996). Por el otro lado, a nivel de estructura secundaria, las dos supuestas regiones de solapamiento de GroEL presentan:  $\alpha$ -hélice seguida en el primer caso (hacia el lado amino) y alternada en el segundo (hacia el lado del carboxilo) de  $\beta$ -conformación (Braig *et al.* 1994). Desde luego en ambos casos, es posible hacer un alineamiento que haga coincidir algunas regiones de  $\beta$ -conformación en los dos transectos e inclusive percibir alguna tendencia de similitud de tres o cuatro aminoácidos en un segmento particular. Pero en lo general, a nivel de estructura secundaria, la coincidencia parece ser mas bien casual. A nivel de estructura terciaria, la similitud es todavía menos probable. Curiosamente, en ambos casos, las regiones de supuesto solapamiento se localizan principalmente en el dominio intermedio de GroEL. En el primer caso (aminoácidos 114-234 de GroEL), la región de solapamiento se inicia al final del dominio ecuatorial, pasa al dominio intermedio y supuestamente terminaría al inicio del dominio apical. En el segundo caso (314-431), la región de solapamiento se iniciaría al final del dominio apical, pasaría al intermedio y terminaría al inicio del dominio ecuatorial. Dicho de otro modo, en el caso de GroEL, ambos transectos están localizados en dominios separados, mientras que la cadena polipeptídica en GroES se organiza en una sola estructura globular. Asimismo, una inspección visual de los pares estereoscópicos de ambas proteínas en relación a la dirección que sigue la cadena polipeptídica, no indica ningún parecido que no sea incidental (Valdés-López observaciones no publicadas). Aún mas, si dichas similitudes existieran al menos sería lógico

esperar que funcionalmente las regiones fueran equivalentes. Como quedó indicado más arriba, se ha reportado que GroES también tiene capacidad de unión de nucleótidos trifosfatados (**Martin et al., 1993b**). Sin embargo, esta pequeña secuencia de aminoácidos en GroES (61-74), no tiene correspondencia estructural de ningún tipo con la región reportada de unión a ATP de GroEL (aminoácidos 470-497, localizados en el dominio ecuatorial) (**Martin et al., 1993b**), ni con el bolsillo de unión a ATP visualizado en la estructura tridimensional que involucra principalmente a los residuos 22-90, en el dominio ecuatorial de GroEL (**Fenton et al., 1994**).

En conclusión, dados los problemas de relaciones estructurales, funcionales y evolutivos que se presentan al comparar diferentes polipéptidos, que no están claramente relacionados entre sí, a veces los criterios se hacen tan laxos que las conclusiones difícilmente se pueden validar con otros enfoques. Sin embargo, en el caso de la inferencia de homología entre la familia TRiC y la familia de GroEL, esta conclusión parece razonable por el momento. Sin embargo, realizando un alineamiento estructural entre miembros de las dos familias, se ha reportado que existe un alto nivel de identidad entre todas las chaperoninas, particularmente en las regiones que corresponden al sitio de unión de ATP de GroEL en el dominio ecuatorial. También se pueden detectar homologías tanto en la parte ascendente, como en la parte descendente del dominio intermedio, pero en el dominio apical, el cual está implicado en la unión del sustrato, la homología entre las dos familias brilla por su ausencia (**Kim et al., 1994**). Las implicaciones funcionales de estas diferencias no son claras, pero posiblemente tengan que ver con la clase de sustratos reconocidos por estos dos tipos de chaperonina, ya que por el otro lado son notablemente similares en estructura tridimensional y función.

## LA NODULACION

Todas las formas de vida dependen del nitrógeno. Sin embargo, aunque la atmósfera terrestre contiene grandes cantidades de este elemento, pocos organismos lo pueden asimilar directamente. De hecho este es el único mineral que no puede ser absorbido directamente del suelo por las raíces de las plantas. Para que pueda ser utilizado es necesario que el nitrógeno se reduzca a amonio ( $\text{NH}_3$ ) o a nitratos ( $\text{NO}_3$ ). Este tipo de compuestos se encuentran naturalmente en el suelo en pequeñas cantidades y se originaron principalmente en la atmósfera reductora de la tierra primitiva. En la actualidad sólo se pueden producir por algunos fenómenos naturales como incendios, actividad volcánica o relámpagos (**Lehninger 1982**). En su estado molecular ( $\text{N}_2$ ), el nitrógeno atmosférico es un gas muy poco reactivo y para ser incorporado al material orgánico los tres enlaces covalentes que mantienen unidos a los átomos en pares deben ser rotos mediante un proceso denominado fijación del nitrógeno que tiene un elevado costo energético

(Losick & Kaiser 1997). En la naturaleza, la fijación del nitrógeno es efectuada por unas cuantas especies de bacterias. Sin ellas, la vida habría desaparecido hace mucho tiempo y el nitrógeno que fijan eventualmente regresa a la atmósfera gracias a otros microorganismos que liberan este gas cuando degradan desechos orgánicos. Además de costoso, el proceso de fijación del nitrógeno es un proceso altamente sensible al oxígeno. Esto se debe a que la enzima que cataliza la fijación del nitrógeno, la nitrogenasa, se inactiva por el oxígeno razón por la cual los organismos fijadores de nitrógeno han tenido que evolucionar mecanismos de protección de este elemento. La nitrogenasa es una enzima compleja que cataliza la conversión del nitrógeno atmosférico en amoníaco que posteriormente se convierte en glutamina, glutamato y finalmente en todos los demás aminoácidos (Kim & Rees 1992).

Algunas bacterias pueden fijar nitrógeno en vida libre ya sea en el suelo o en el agua. Estos organismos no simbióticos fijan nitrógeno para su propio uso y son llamados diazotrofos. Sin embargo los fijadores mas eficientes establecen una simbiosis con plantas superiores en la cual la energía para la fijación del nitrógeno y en general la protección al oxígeno es aportada por la planta. En dos clases de interacciones simbióticas la parte procarionte son bacterias del suelo (rhizobia en las simbiosis con leguminosas y *Frankia* en las simbiosis con actinorhizas). En la simbiosis con *Gunnera* el nitrógeno es fijado por la cianobacteria *Nostoc* (Myolna et al., 1995). Con mucho, la simbiosis con leguminosas, es la fuente individual de nitrógeno orgánico mas grande en el ciclo global del nitrógeno y es el proceso que ha sido más estudiado. Desde la roma antigua, las cualidades enriquecedoras del suelo de las leguminosas fueron reconocidas y se desarrollaron sistemas de rotación de cultivos, usando leguminosas para abonar el suelo (Nap & Bisseling 1990). Asimismo, en mesoamérica nuestros antepasados usaban el co-cultivo del maíz y del frijol con el mismo propósito (Sánchez, F. comunicación personal)

Las leguminosas pueden formar interacciones simbióticas con bacterias de los géneros *Rhizobium*, *Bradyrhizobium* y *Azorhizobium*, denominadas colectivamente rhizobia. Cuando estas bacterias se encuentran en vida libre son incapaces de fijar nitrógeno. Una especie particular de rhizobia es capaz de interactuar con una o número limitado de plantas. Las bacterias se adhieren a las raíces de la planta provocando el curvamiento característico de los pelos radiculares (día 0). Posteriormente las bacterias invaden la planta por medio de un tubo denominado hilo de infección. Mientras tanto, las células interiores en el cortex de la raíz se comienzan a dividir y forman el primordio del nódulo (días 1-10). El hilo de infección entra a estas células y las bacterias se diferencian en las formas endosimbióticas llamadas bacteroides. En ese momento se inicia la fijación del nitrógeno, (día 10-14). Posteriormente en el caso de nódulos determinados comienza la senescencia (día 20). Cada nódulo maduro contiene aproximadamente la mitad de su peso en bacteroides intracelulares los cuales han perdido la

mayor parte de su propia pared celular. La membrana plasmática de las bacterias se conserva, y a su vez se rodea por una membrana formada por la célula huésped, la llamada membrana peribacteroidal (**PBM**) y al conjunto se le denomina simbiosoma. El nitrógeno fijado como amonio es transportado de los bacteroides al citoplasma de la célula vegetal donde es asimilado y transportado finalmente al resto de la planta (**Brewin 1991**). Todo este proceso requiere de un continuo intercambio de señales por parte de la bacteria y de la planta así como de la expresión específica de genes de ambos simbiosomas.

### **Nodulinas.**

Las proteínas expresadas por la planta durante la morfogénesis del nódulo son llamadas nodulinas (**van Kammen 1984**). Los genes de nodulinas han sido operacionalmente definidos como tempranos y tardíos de acuerdo al momento de su expresión durante el desarrollo del nódulo. Las nodulinas tempranas se expresan antes de que se inicie la fijación del nitrógeno durante la fase de infección y están involucradas en la formación del nódulo. La expresión de las nodulinas tardías comienza cerca del momento cuando se va a iniciar la fijación del nitrógeno, y están involucradas en la interacción con el endosimbionte y con la especialización metabólica del nódulo (**Mylona et al., 1995**). Se ha demostrado que algunas nodulinas se expresan en otros órganos de la planta además del nódulo *i. e.*, no son genes específicos del proceso de nodulación. De hecho es posible que las verdaderas nodulinas, en el sentido que solo se expresen en el nódulo, sean escasas. Sin embargo el concepto de nodulina provee el marco conceptual y molecular de desarrollo del proceso simbiótico (**Sánchez et al., 1991**).

La mayor parte de las **nodulinas tempranas** tienen un alto contenido de residuos de prolina y se ha propuesto que por lo tanto son componentes de la pared celular (**Nap & Bisseling 1990**). Dentro de las nodulinas tempranas, *ENOD40* y *ENOD12* son de las primeras en detectarse durante la etapa de la infección (**Kouchi & Hata 1993**). Sin embargo, la infección por la bacteria tarda uno o más días dependiendo del huésped y del hospedero. A esta fase se le llama la etapa de pre-infección y prácticamente no se han identificado genes específicos de este período. Sin embargo, las peculiaridades bioquímicas y genéticas de los pelos radiculares son determinantes para que se establezca la infección. Durante el proceso de invasión debido a la acción de cambios localizados en la pared de los pelos radiculares, se desarrolla una nueva pared celular alrededor del sitio de penetración, adelante de las bacterias. Se han identificado varias proteínas específicas de los pelos radiculares (RH proteins), pero no se ha establecido su función (**Sánchez et al., 1991**).

Recientemente sin embargo se ha reportado la inducción transitoria de un gen de peroxidasa (*Rip1*) en la etapa de pre-infección en alfalfa (*Medicago truncatula*). Se ha especulado que el

papel de esta peroxidasa es el de restringir la infección por *Rhizobium*, de una manera análoga a otros mecanismos que ocurren en las interacciones planta-patógeno, permitiendo que la planta se prepare para el momento de la infección (Cook *et al.*, 1995). En este sentido, a pesar de que eventualmente la simbiosis entre la bacteria y la planta será benéfica para esta última, la infección es altamente regulada por la planta. Un mecanismo de control de la infección es la llamada inhibición por retroalimentación que se evidencia como una susceptibilidad transitoria de las células de los pelos radiculares a la infección por las bacterias (Bhuvaneshwar *et al.*, 1981). Mutantes en esta función continúan produciendo nódulos. Un segundo posible mecanismo para controlar la infección involucra la supresión temprana de la infección en la zona de nodulación. De hecho, solo la minoría de las infecciones iniciales persisten y solo algunas llegan a colonizar el tejido del nódulo. En los sitios donde se detienen las infecciones que abortan se presentan características clásicas de inducción de mecanismos de defensa de la planta, denominadas reacción hipersensible (Vasse *et al.*, 1993). Se ha reportado que el fitoregulador etileno está implicado en la respuesta de las plantas a ataques por patógenos (Ecker 1995). Recientemente se ha presentado evidencia de que el mecanismo fisiológico de supresión de la infección por rhizobia involucra justamente al etileno (Penmetsa & Cook 1997). En particular, se ha reportado que la inhibición de la síntesis de etileno causa un notable incremento de infecciones bacterianas persistentes con un aumento notable en el número de nódulos. Por el otro lado, la aplicación exógena de este gas reduce la nodulación. Estos dos hechos se correlacionan con los resultados de que mutantes de *Medicago truncatula* insensibles a ese gas sean hiper-infectadas por la bacteria simbiote. Todavía no se ha determinado ni el mecanismo de acción del etileno para detener la infección, ni como es que algunas células no responden al fitoregulador de tal manera que la infección por rhizobia ocurra (Penmetsa & Cook 1997). Como sea, es claro que el proceso de simbiosis está íntimamente relacionado con los procesos más habituales de patogénia, de tal manera que tanto la planta como los rhizobia controlan sus respectivos arsenales moleculares de defensa y ataque para que finalmente se de la interacción mutuamente benéfica.

Por su parte, las **nodulinas tardías** constituyen un grupo mas amplio y mejor caracterizado de proteínas que se expresan alrededor del día 12 cuando comienza la fijación del nitrógeno. Las nodulinas tardías ayudan a crear las condiciones para la fijación y asimilación del nitrógeno, participan en el metabolismo de carbohidratos y en la biogénesis de amidas y ureidos (Sánchez *et al.*, 1991). Dentro de las enzimas de asimilación del amonio están la glutamino sintetasa (GS) y la glutamato sintasa (GOGAT), las cuales incrementan su actividad junto con la actividad de nitrogenasa (Sánchez *et al.*, 1991). De la GS existen dos isoformas de la enzima: la citoplasmática (GS<sub>1</sub>) y la del cloroplasto/plástido (GS<sub>2</sub>), codificadas en ambos casos, por una pequeña familia

multigénica. En los nódulos radicales la forma GS<sub>1</sub> predomina, aunque también se ha observado la expresión de GS<sub>2</sub> en nódulos de frijol (Bennet *et al.*, 1989). En algunas plantas se ha reportado la expresión nódulo específica de GS<sub>1</sub>, mientras que en otras plantas lo que ocurre es un significativo incremento transcripcional de los genes de GS. En este caso se habla de expresión nódulo incrementada (Temple *et al.*, 1995), lo cual significa que un gen que normalmente se expresa en otros tejidos es reclutado durante la simbiosis. Como se verá más adelante, ésta podría ser la regla, más que la excepción. Dentro de las enzimas que participan en el metabolismo del carbón se encuentran la fosfoenolpiruvato (PEP) carboxilasa y desde luego la sacarosa sintasa. En ambos casos se han descrito isoformas nódulo específicas (Sánchez *et al.*, 1991). Dependiendo del tipo de nódulos que forman (indeterminados o determinados), las leguminosas transportan el nitrógeno asimilado como amidas o como ureidos. Como ejemplo de enzimas involucradas en estas vías metabólicas están la aspartato aminotransferasa (AAT) de alfalfa y la uricasa de frijol. La AAT está codificada por una pequeña familia multigénica y se han documentado isoformas citosólicas, mitocondriales y plastídicas. En alfalfa la isoforma AAT-2 se localiza en los plástidos y su expresión en nódulos es varias veces mayor que en hojas y raíces (Gantt *et al.*, 1992). Se ha descrito una uricasa nódulo específica de soya y de frijol. Esta enzima se localiza en los peroxisomas de células no-infectadas (Sánchez *et al.*, 1987). En nódulos determinados se ha reportado una familia de nodulinas tardías en el rango de 30 kDa, cuya función no se ha esclarecido (Sánchez *et al.*, 1991). En el frijol el representante de esta familia, la nodulina 30, tiene un nivel de transcripción aun más alto que el de leghemoglobina. El transcrito correspondiente se acumula abundantemente en las células infectadas en la región central de nódulos maduros. Sorprendentemente, el producto polipeptídico correspondiente es apenas detectable (Campos *et al.*, 1995).

La nodulina tardía más abundante, es la leghemoglobina (Lbs). Esta proteína, fue detectada por primera vez en leguminosas en 1947 (Kellin & Smith 1947), y justamente es la responsable del color rojo que se ve al cortar un nódulo. Por su similitud estructural y funcional con las hemoglobinas animales se ha propuesto que la Lbs tiene en los nódulos la función de acarrear oxígeno y facilitar su difusión a los bacteroides (Appleby 1984). Ciertamente el análisis de la similitud de las hemoglobinas entre plantas y animales sugiere que los genes respectivos se derivan de un ancestro común (Lesk & Chothia 1980). A nivel genómico, sin embargo, en los genes de plantas se encuentran cuatro exones a diferencia de los genes de animales donde solo se localizan tres (Holland & Blake 1987). Los exones correspondientes a la región del amino y del carboxilo, son completamente equivalentes en plantas y animales. La diferencia se localiza en el exón central de animales que es justamente donde en plantas, este exón se divide en dos por la presencia de un intrón adicional. Es interesante que este intrón fue predicho antes de ser descubierto. Por las

características del péptido codificado por el exón central en las globinas animales (mioglobina y hemoglobina), se derivó la posibilidad de que ancestralmente los genes de globinas de vertebrados, hubieran tenido este intrón y que posteriormente se hubiera perdido, dando como resultado la fusión de dos exones en uno solo (Go 1981; Blake 1993). El descubrimiento de este intrón en los genes de leghemoglobina, apoyan aquella propuesta y sugieren que este intrón se conservo en la línea evolutiva que se derivó hacia las plantas (Østergaard Jensen et al., 1981). Es interesante que los genes homólogos de insectos no tengan ningún intrón (Antoine & Nlesing 1984), lo cual descarta cualquier posibilidad de que el gen de plantas hubiera sido adquirido por transferencia horizontal a partir del de animales (insectos) por un vector de tipo viral (Lewin 1984). Este no parece ser el único caso en que los genomas vegetales son más conservadores respecto a la pérdida de intrones. Por ejemplo, en el caso de los genes de actina se presenta una situación similar: en plantas la posición de los intrones está muy conservada y en animales es muy variable (Sheterline & Sparrow 1994). Sin embargo, dado que se han secuenciado comparativamente menos genes de plantas que de animales, es difícil por el momento generalizar esta característica de los genomas vegetales (Bird 1995). Se ha señalado que la leghemoglobina podría ser una "verdadera proteína simbiótica" ya que la síntesis del grupo heme es aportada por el bacteroide, mientras que la globina está codificada por el genoma de la planta (Nap & Bisseling 1990). Esta situación es de alguna manera similar a la síntesis de la hemoglobina en las células eritroides, donde la síntesis del grupo heme es llevada a cabo de manera parcial por las mitocondrias (Sano & Granick 1961). En este sentido es interesante notar que la simbiosis entre los rhizobia y las leguminosas, particularmente en el momento en que las bacterias penetran al citoplasma de la célula vegetal, podría ser un modelo de los eventos iniciales que derivaron en el establecimiento definitivo de la endosimbiosis que culminó con la aparición de mitocondrias y cloroplastos. En el caso de la fijación del nitrógeno parecería que el siguiente paso lógico evolutivo sería efectivamente la colonización definitiva del citoplasma vegetal por los rhizobia. Desde luego habría que considerar que existen muchos otros casos de parasitismos intracelulares y que hasta donde sabemos solo algunos pocos han derivado en una endosimbiosis permanente i. e. mitocondrias y plástidos. Sin embargo, recientemente se han descrito los llamados endosimbiontes secundarios, en particular un "organelo" , el apicoplasto, presente en protozoarios parásitos del phylum Apicomplexa como *Toxoplasma* y *Plasmodium*, el cual aparentemente se ha derivado de la endosimbiosis a partir de un alga verde completa (Köhler et al., 1997). Asimismo, se ha sugerido que los eventos que han derivado en endosimbiosis definitivas ocurrieron al menos media docena de veces durante las etapas relativamente tempranas de la evolución celular (Vogel 1997). En este sentido patogenia y simbiosis podrían estar íntimamente relacionadas. Por el otro lado, los eucariontes mas divergentes como *Giardia* y

*Trichomonas*, parecían no tener mitocondrias, y se había sugerido que tal vez las perdieron posteriormente a su adquisición (**Pace 1997**); sin embargo aparentemente la pérdida no fue total. En *Trichomonas vaginalis* se ha descrito un organelo, el hidrogenosoma, que es en realidad una mitocondria altamente modificada. Tan modificada que de hecho ha perdido su genoma y la función de respiración. La evidencia a favor de esta hipótesis proviene del análisis que muestra que en el núcleo de *Trichomonas* existen genes de origen mitocondrial. Curiosamente estos son los genes de las proteínas de heat shock Hsp10, Hsp60 y Hsp70 (**Palmer 1997**). El análisis filogenético de la secuencia de estas proteínas, muestra que se asocian con el grupo monofilético compuesto exclusivamente por homólogos mitocondriales. Estos datos establecen entonces que las mitocondrias y los hidrogenosomas tienen un ancestro eubacteriano común e implican que estos eucariontes (tricomonados), que parecen surgir muy temprano en la evolución celular, ya contenían al ancestro endosimbionte (**Bul et al., 1996**). Un enfoque alternativo podría ser que en realidad que los eventos de endosimbiosis son altamente improbables ya que solo ocurrieron algunas pocas veces en los orígenes de las células eucariontes y actualmente, aunque aparentemente existen oportunidades para su establecimiento, en realidad no se desarrollan mas allá del caso de la interacción simbiótica que permite la fijación del nitrógeno e. g. *rhizobia/leguminosa*; *Frankia/actinorhizas*; *Nostoc/Gunnera* (**Mylona et al., 1995**).

Regresando a la simbiosis y a la leghemoglobina; la función de esta proteína es fundamental. La Lb constituye hasta el 25% de la proteína total soluble en nódulos maduros, y controla la concentración de oxígeno en este órgano. Su presencia permite un flujo controlado de oxígeno a los bacteroides que debido a su alta actividad metabólica, requieren del oxígeno y al mismo tiempo proporciona protección al daño que podría sufrir la nitrogenasa (**Nap & Bisseling 1990**). En el chícharo, se ha demostrado que mutaciones que reduzcan la expresión ya sea de la globina o de la síntesis del grupo heme producen un fenotipo donde ocurre la infección pero los bacteroides carecen de actividad de nitrogenasa (**Suganuma et al., 1995**). Las Lbs son codificadas por una pequeña familia multigénica. En la soya se han identificado cuatro isoformas, cuya expresión se activa un poco antes de la fijación del nitrógeno (**Østergaard Jensen et al., 1981**). Las diferentes isoformas poseen ligeras diferencias en su afinidad por el oxígeno, y su expresión se induce de manera ligeramente desfasada durante la maduración del nódulo (**Sánchez et al., 1991**). Al respecto de si Lb es verdaderamente una autentica nodulina, se ha reportado que existen genes homólogos a ésta en otras plantas que no forman nódulos simbióticos, incluida la monocotiledónea centeno (**Anderson et al., 1996**). Por ejemplo, en *Trema tomentosa* (Ulmaceae), el patrón de expresión de su Lb es tejido específico, detectándose en la raíz pero no en las hojas de esta planta. En *Parasponia andersonii* (relacionada con *Trema*), que

si forma nódulos, se ha demostrado que Lb no se detecta en hojas, pero si en raíces no infectadas. Cuando *Parasponia* es nodulada, el nivel de expresión de su Lb se incrementa alrededor de mil veces (**Bogusz et al., 1988**). En soya, se ha descrito un gen que codifica para una hemoglobina adicional a la simbiótica que se expresa en cotiledones, tallos de plántulas, raíces, hojas jóvenes y en algunas células en nódulos, si bien en estos últimos el nivel de expresión muy bajo. Este gen tiene solo 58% de similitud con el gen de la leghemoglobina simbiótica de la propia soya, pero tiene 85-87% de similitud con las hemoglobinas de las plantas no-leguminosas *Parasponia*, *Casuarina* y cebada (**Anderson et al., 1996**). En *Casuarina glauca*, (que forma nódulos simbióticos con *Frankia*), también se han descrito hemoglobinas simbióticas y no-simbióticas. Además, ambos genes mantienen sus patrones de expresión específicos cuando son introducidos en la leguminosa *Lotus corniculatus*, sugiriendo que los mecanismos de control de la expresión de las leghemoglobinas se ha conservado evolutivamente (**Jacobsen-Lyon et al., 1995**). Otro aspecto interesante es la estructura de las globinas. Por ejemplo, en *Trema* al igual que en la leguminosa soya y la no-leguminosa *Casuarina* (*Frankia*), el producto nativo de Lb es monomérico; en *Parasponia*, la Lb nativa es dimérica y en el insecto *Chironomus*, existen isoformas monoméricas y homodiméricas (**Bogusz, et al., 1988; Østergaard Jensen et al., 1981; Antoine & Niessing 1984**). En este sentido, se podría decir que las formas monoméricas de Lb se parecen más a la mioglobina, mientras que las formas diméricas podrían representar un estado de transición hacia la forma tetramérica de la hemoglobina de los vertebrados. Una situación similar parece haber ocurrido en etapas posteriores a la separación de los genes de mioglobina y de globina en vertebrados, ya que se ha reportado que la hemoglobina de las lampreas (*Agnata*), es monomérica (**Lanfranchi et al., 1994**). Sin embargo, se ha reportado que la hemoglobina de la bacteria *Vitreoscilla*, es dimérica (**Wakabayashi et al., 1986**), mientras que en gusanos como la de la lombriz de tierra, o gusanos que vive en pozas hidrotermales, su hemoglobina forma agregados de cientos de miles a millones de Daltons (**Zhu et al., 1996; Zal et al., 1996**). Recientemente, utilizando fragmentos de proteína aislados de trabéculas óseas de *Tyrannosaurus rex*, se han inmunizado ratas y los anticuerpos producidos reaccionan positivamente con hemoglobinas purificadas de aves y mamíferos. La explicación más parsimoniosa, es que los dinosaurios tenían hemoglobina; si bien con dicha evidencia, no es posible asegurar si éstas eran tetraméricas (**Schweltzer et al., 1997**). Dicho de otra manera, el criterio de si una globina forma monómeros, dímeros o tetrameros no implica una clara relación evolutiva. Por el otro lado, la afinidad de la globina por el oxígeno tampoco parece ser un criterio, ya que por ejemplo, la constante de asociación por el O<sub>2</sub> de las leghemoglobinas (monoméricas) es de 120 mM<sup>-1</sup> s<sup>-1</sup>, mientras que la de la mioglobina de ballena es de solo 19 mM<sup>-1</sup> s<sup>-1</sup> (**Harutyunan et al., 1995**), sugiriendo que dicha afinidad se correlaciona con las condiciones de

aerobiosis o microaerobiosis y no con la distancia filogenética. A pesar de las deficiencias de los métodos de comparación de secuencias, ésta parece ser la alternativa más razonable para decidir que en efecto, la Lb debe ser considerada una hemoglobina y no una mioglobina (**Arredondo-Peter & Escamilla 1991**). En este sentido es interesante que la leghemoglobina de lupin (leguminosa), muestre una mayor similitud con la hemoglobina de la bacteria *Vitreoscilla* (34% de identidad), que con globinas de vertebrados (23% de identidad). En este sentido aún es prematuro especular si estas hemoglobinas comparten un ancestro o si son el producto de evolución convergente (**Wakabayashi et al., 1986**).

Finalmente, las nodulinas en general llevan a cabo tareas que individualmente no son exclusivas de los nódulos, sino que se encuentran en otras partes de la planta e inclusive en otras plantas que no forman nódulos. Así se ha propuesto que la formación de los nódulos evolucionó a partir de modificaciones relativamente menores en las vías de diferenciación radicular, y en la cual algunos genes y vías de diferenciación comunes a todas las plantas se adaptaron para contender con las demandas fisiológicas, regulatorias y morfológicas específicas de la simbiosis (**Nap & Bisseling 1990; Bisseling et al., 1996**). En consecuencia la comprensión de las alteraciones que han sufrido estos procesos y como podrían ser alteradas podría llevar a diseñar estrategias para conferir a plantas no-leguminosas la capacidad de establecer simbiosis con microbios fijadores de nitrógeno (**Mylona et al., 1995**).

### **Nodulación, día -3 / 0**

La interacción planta-bacteria se inicia con señales químicas que actúan a distancia en la rizosfera. En particular, la planta produce y exporta una señal química en la forma de un flavonoide (**Fisher & Long 1992**). Estas moléculas funcionan como quimioatrayentes y además, aparentemente penetran a la bacteria y estimulan la expresión de los genes bacterianos de nodulación (*nod*). Los flavonoides son productos normales en las plantas y en particular diferentes leguminosas sintetizan diferentes flavonoides, los cuales actúan de manera específica sobre diferentes especies de Rhizobia. Estos flavonoides actúan a concentraciones extremadamente bajas ( $10^{-7}$  -  $10^{-8}$  M) y estimulan la expresión de los genes *nod* a los pocos minutos de su exposición, lo cual hace que estas señales sean particularmente adecuadas para la localización mutua en la rizosfera (**Sánchez et al., 1991; Fisher & Long, 1992**). A su vez, los genes *nod* codifican enzimas involucradas en la síntesis del factor de nodulación. Este, es un lipo-quito-oligosacárido que causa cambios morfológicos en la raíz de la planta (**Vijn et al., 1993**). El gen *nodD* codifica para una proteína reguladora que constituye el control central de regulación del resto de los genes *nod*. En particular la región amino terminal de NodD está involucrada en la unión al DNA, mientras que la región, más variable del carboxilo terminal está involucrada en la interacción

con los flavonoides (Fisher & Long 1992). A excepción de *nodD*, los genes *nod* no se expresan en la bacteria de vida libre. Los genes *nodABC*, comunes a todos los rizobia son los responsables de la síntesis del núcleo básico del factor Nod, mientras que genes específicos controlan la decoración del factor con grupos laterales, los cuales en última instancia son los responsables de la especificidad de la interacción simbiótica. Los factores Nod purificados a concentraciones de  $10^{-12}$  M, son capaces de estimular la diferenciación de células epidérmicas en pelos radiculares, de provocar la deformación de estos pelos radiculares y formación del hilo de pre-infección (Mylona et al., 1992; Downie 1994). Al respecto de ambas señales, es interesante que las señales emitidas por la planta, los flavonoides pertenecen a una clase de compuestos que incluyen a las moléculas de defensa de tipo fitoalexinas. Isoflavonas de la soya no solo inducen la expresión de los genes *nod* en *B. japonicum*, sino que también disparan la resistencia bacteriana a la fitoalexina de la soya. A su vez la señal de la bacteria, el factor Nod, estructuralmente se asemeja a un elicitor, en el sentido de que la quitina es una señal que normalmente las plantas interpretan como evidencia de un patógeno. Tal vez sean las modificaciones del factor Nod, lo que determina que la leguminosa acepte a los rizobia como simbioses en vez de desencadenar una respuesta de defensa contra el invasor (Fisher & Long 1992). Una alternativa que no es excluyente de la anterior, es que los factores Nod imiten fitoreguladores del crecimiento vegetal. A este respecto, la morfogénesis del nódulo, incluyendo expresión de nodulinas tempranas puede ser simulada por citosinina y por inhibidores del transporte de auxinas (Fisher & Long 1992; Hirsch et al., 1989).

En las especies de *Bradyrhizobium*, de crecimiento lento, los genes simbióticos se localizan en el cromosoma principal, mientras que en las especies de *Rhizobium*, de crecimiento rápido, estos genes están contenidos en megaplásmidos (pSyms). Además, en estas especies se pueden llegar a presentar de uno a diez plásmidos diferentes llegando a representar hasta más del 40% del total del DNA. Normalmente tenemos la creencia de que las bacterias están limitadas respecto a la cantidad de genes que pueden contener. La presencia de estos megaplásmidos, que más parecen cromosomas, indica que lo anterior podría ser una sobre simplificación debida a nuestro conocimiento limitado de la vida a nivel de microorganismos (Pace 1997). Es interesante que en *Azotobacter vinelandii*, una bacteria gram-negativa del suelo que es capaz de fijar nitrógeno atmosférico, se haya reportado que cuando se cultiva en medio rico, llega a tener hasta más de 80 copias de su genoma al llegar a la fase estacionaria (el tamaño haploide de su genoma es de 4.7 mega pares de bases i. e. muy similar al del genoma de *E. coli*). Cuando *Azotobacter* es cultivada en medio mínimo, no alcanza estos altos niveles de ploidía, lo cual sugiere que habitualmente esta amplificación no ocurre fuera de condiciones de laboratorio. Se ha reportado que el número normal de copias de su genoma es cercano a cuatro, i. e. realmente

no es una bacteria poliploide (**Nagpal et al., 1989; Manna & Das 1993; Maldonado et al., 1994**). En *Rhizobium etli* (antes *R. leguminosarum* biovar *phaseoli*), se localizan seis megaplásmidos que tienen tamaños entre 150 a 600 kilobases. En el pSym, de 390 kilobases, se localizan la mayoría de los genes de la simbiosis, pero los genes contenidos en los otros cinco plásmidos también contribuyen al establecimiento exitoso de la simbiosis (**Girard et al., 1991**). En este plásmido, también se localizan los genes *nif* y *fix*, esenciales para la fijación del nitrógeno, dentro de los cuales se encuentra el gen estructural de la nitrogenasa (**Quinto et al., 1985**). A pesar de la importancia intrínseca de los genes *Nif* y *fix*, mutantes es éstos son capaces de inducir la formación de nódulos en los cuales se expresan nodulinas y poseen una morfología similar al tipo silvestre. Aparentemente entonces, la fijación del nitrógeno es la consecuencia y no un prerrequisito para el desarrollo del nódulo (**Nap & Bisseling 1990**). Además de los genes *NoD*, *nif* y *fix* se han descrito otros genes (*exo*, *lps* y *nodV*) que están involucrados en la estructura de la superficie externa de la bacteria (**Nap & Bisseling 1990**).

### **Nodulación, días 1 / 10**

El primer cambio, visible microscópicamente, de la planta en respuesta al contacto con *Rhizobium*, en particular con el factor Nod, es la deformación y el enrollamiento de los pelos radiculares. Estos cambios morfológicos están precedidos por despolarización de la membrana plasmática, cambios en el flujo de calcio, rearrreglos de los filamentos de actina y un incremento del reflujo citoplasmático (**Mylona et al., 1995**). Respecto al citoesqueleto, éste es una estructura esencial del citoplasma eucarionte y modificaciones en el citoesqueleto podrían explicar muchos de los cambios morfológicos que se encuentran en etapas tempranas y tardías de la interacción planta-*Rhizobium*. La posibilidad de que el factor Nod pudiera disparar cambios en los niveles de calcio, los cuales a su vez podrían alterar significativamente al citoesqueleto, había sido predicho (**Sánchez et al., 1991**); recientemente esta relación ha comenzado a ser estudiada (**Ehrhardt et al., 1996**). Una vez que la bacteria entra en contacto con el pelo radicular, los rhizobia quedan atrapados, resultando en una alta concentración de factores. Este contacto desencadena la infección propiamente dicha, que se inicia con una hidrólisis muy localizada de la pared celular de las células vegetales (**van Spronsen et al., 1994**). Lo anterior, resulta en la formación de una estructura tubular denominada hilo de infección, a través del cual la bacteria entra a la planta. La bacteria en el interior del hilo de infección se encuentra rodeada de una matriz que parece consistir de compuestos secretados tanto por la planta como por la bacteria. El hilo de infección atraviesa células corticales de la raíz por medio de puentes citoplasmáticos alineados radialmente que son denominados hilos de preinfección. Estas estructuras siempre preceden la formación del hilo de infección propiamente dicho y atraviesan la vacuola central de las células

corticales centrales. Aunque estas células corticales externas que forman el hilo de preinfección nunca se dividen, los cambios morfológicos recuerdan a aquellos que se observan en células que entran al ciclo celular. El factor Nod purificado induce la formación del hilo de preinfección pero los hilos de infección no se forman, lo cual sugiere que para que se formen, se requiere de la presencia de los rhizobia (**van Brussel et al., 1992**). Concomitantemente con la formación del hilo de infección, las células corticales se activan mitóticamente formando el primordio del nódulo. En leguminosas tropicales como la soya y el frijol, son las células corticales externas las que se activan, formando estructuras globosas denominadas nódulos determinados. En leguminosas de climas templados como el chícharo y la alfalfa se activan las células corticales internas, formando estructuras alargadas denominadas nódulos indeterminados (**Mylena et al., 1995**). En los nódulos indeterminados se presenta un gradiente morfogénico desde un meristemo apical persistente, hasta la región proximal de senescencia (**Vasse et al., 1990**). En nódulos determinados el meristemo deja de dividirse en una etapa temprana del desarrollo. Como resultado, todas las células del tejido central se encuentran en el mismo estado de diferenciación y senescen al mismo tiempo (**Brewin 1991**). Utilizando *Rhizobium meliloti* marcado con proteína verde fluorescente (GFP), se ha propuesto que en el interior del hilo de infección solo las bacterias que se localizan en la punta de éste son capaces de crecer y dividirse. Dicho de otro modo, debe de existir comunicación molecular entre la bacteria y la planta de tal manera que las bacterias distales no proliferan activamente. Una posibilidad podría ser que la planta controlara la cantidad de nutrientes a lo largo del hilo de infección. La explicación de este fenómeno podría ser que con esta estrategia se limitaría el crecimiento de otras especies de bacterias que pudieran ser acarreadas junto con el *Rhizobium* (**Gage et al., 1996**). Como se menciono anteriormente, en este sentido la observación de que frecuentemente, la planta termina el crecimiento del hilo de infección y se presenta una reacción de tipo hipersensible, podría ser un mecanismo lógico de prevenir la entrada de otras especies contaminantes de bacterias, a la vez que serviría para regular el número de nódulos que se desarrollan en la planta (**Vasse et al., 1993**).

#### **Nodulación, días 10 / 14**

Así, a medida que el hilo de infección penetra, las bacterias proliferan en la punta de éste, hasta que el hilo de infección llega al primordio del nódulo. Una vez que se establece el contacto, el hilo de infección se bifurca y penetra al interior de algunas de las células centrales del primordio (**Losick & Kaiser 1997**). En ese momento, las bacterias salen de las puntas del hilo de infección hacia el citoplasma de la célula vegetal. Esta liberación es un proceso endocítico, en el que la bacteria termina rodeada de la membrana peribacteroidal (**PBM**). Esta, forma la interfase entre el citoplasma vegetal y el endosimbionte a través de la cual se intercambian señales y metabolitos

y se evita una respuesta de la planta en contra de la bacteria intracelular (**Verma 1992**). Interdispersas entre las células infectadas, se localizan células más pequeñas que no son infectadas (las **células no infectadas**). Subsecuentemente los endosimbiontes pierden su pared celular, crecen, se multiplican y se diferencian en bacteroides, ocupando eventualmente la mayor parte del volumen de las células infectadas. Los bacteroides rodeados de la PBM, son denominados simbiosomas. Se ha propuesto que el simbiosoma tiene las propiedades de un compartimento lítico que se neutraliza constantemente por el amonio producido por los bacteroides. De acuerdo a esta hipótesis, se esperaría que la falta de fijación de nitrógeno llevara a la degradación del bacteroide. Efectivamente existe evidencia de degradación prematura de los bacteroides en mutantes de *Rhizobium* que no fijan nitrógeno (**Hirsch & Smith 1987**). Durante la formación del simbiosoma, el crecimiento de los rhizobia y la síntesis de la membrana que lo rodea están sincronizadas, de tal manera que se requiere de una síntesis masiva de esta membrana plasmática. En las leguminosas, ésta se incrementa 30 veces (**Verma 1992**). La PBM tiene una composición diferente a la de la membrana plasmática normal que le permite desempeñar funciones especiales. A través de la PBM se intercambian amonio, el grupo hemo, compuestos carbonados y posiblemente los productos de la asimilación del amonio (**Verma & Hong 1996**). En general se desconocen las proteínas que participan en estos procesos. De hecho, la PBM contiene varias nodulinas (como la nodulina-26), e inclusive proteínas bacterianas (**Mytona et al., 1995**). Los bacteroides expresan un sistema de asimilación de ácidos dicarboxílicos, lo cual sugiere que son éstos la fuente primaria de carbón aportada por la planta al bacteroide. Llegados a este estado, las bacterias finalmente expresan la nitrogenasa y el amonio, producto de la fijación del nitrógeno es exportado a la planta. La forma en que el nitrógeno fijado es transportado depende de la planta. En las plantas que forman **nódulos indeterminados**, el nitrógeno fijado se exporta como amidas, mientras que en las que forman **nódulos determinados**, se exporta como ureidos. En ambos casos, el amonio exportado es asimilado en el citoplasma por vía de la glutamino sintetasa vegetal (**GS**). En nódulos determinados algunos componentes de la asimilación del amonio y la biosíntesis de ureidos pueden estar espacialmente separados. Mientras que la GS es expresada tanto en células infectadas como no infectadas (**Miao et al., 1991**), la uricasa (nodulina-35), una enzima clave en la oxidación de purinas que cataliza la oxidación de ácido úrico a alantoina, solo ha sido detectada en los peroxisomas de las células no infectadas (**Hanks et al., 1981; Nguyen et al., 1985; Sánchez et al., 1987**). La alantoinasa, que cataliza el siguiente paso de la oxidación de purinas, también ha sido localizada en células no infectadas (**Hanks et al., 1981**). También, en nódulos determinados las células no infectadas parecen estar involucradas en el transporte del nitrógeno fijado como ureidos, conformando una red más o menos continua de plasmodesmos

entre estas células, a través del tejido central, que facilita el transporte al tejido vascular (Selker 1988). En contraste, no existe evidencia de que en nódulos indeterminados, las células no infectadas presenten una especialización similar (Nap & Bisseling 1990).

#### **Nódulos después del día 14**

Debido a la presencia de un meristemo persistente, en los nódulos de tipo determinado una vez que se llega a esta etapa, la fijación del nitrógeno continúa indefinidamente. Por el contrario, en nódulos determinados donde no existe este meristemo, a partir del día 20-25, comienza la senescencia del nódulo (Sánchez *et al.*, 1991; Mylona *et al.*, 1995).

#### **ANTECEDENTES**

El proceso de diferenciación del nódulo simbiótico, ofrece una oportunidad de estudiar los procesos morfogénicos y de regulación génica que ocurren en este órgano. En este sentido, en el laboratorio del Dr. Federico Sánchez se han abordado diversos aspectos de este proceso (Sánchez *et al.*, 1987; Sánchez *et al.*, 1991; Campos *et al.*, 1995). Más recientemente, este grupo volviendo a una vieja querencia (Sánchez *et al.*, 1983), se ha interesado en el papel funcional de las proteínas y genes relacionados con el citoesqueleto durante el desarrollo y morfogénesis del nódulo (Pérez *et al.*, 1994; Vidali *et al.*, 1995; Sánchez *et al.*, 1996). De una manera análoga a cuando John Ellis purificaba a la Rubisco (*vide supra*); durante las etapas iniciales de purificación de la actina de raíces y nódulos de frijol, el Dr. Héctor Pérez detectó una proteína que copurificaba con la actina. Columnas de exclusión molecular (TSK 4000) sugerían que en estado nativo esta proteína podría tener más de 400 kDa. En geles de SDS-PAGE se detectaba una banda de alrededor de 60 kDa, que sugería que esta proteína pudiera ser un homo-oligómero. Extrañamente, su presencia era revelada por un anticuerpo monoclonal anti-actina. Aunque estos datos no fueron incluidos en el reporte final (Pérez *et al.*, 1994); las similitudes con los trabajos originales de la Rubisco Subunit Binding Protein, (excepto desde luego la unión del anticuerpo), eran obvias (Barraclough & Ellis 1980). Por otro lado, durante la caracterización electroforética de las isoformas de actina en plantas, en particular de soya, el grupo de Richard B. Meagher, usando diferentes anticuerpos contra actina, también había reportado la presencia de lo que llamaron "epítopes relacionados a actina" (ARE) en extractos de hipocotilo/raíz y hoja. Estos, fueron parcialmente caracterizados como polipéptidos de alto peso molecular (52-62 kDa) y se sugería que podrían estar relacionados con las chaperoninas del cloroplasto, si bien éstas no tienen ninguna relación de secuencia o estructural conocida con la actina (McLean *et al.*, 1990). Sin embargo, se sabía que a su vez las chaperoninas de mitocondria humana presentaban

reactividad inmunológica con anticuerpos anti-actina (Varma *et al.*, 1987; McLean *et al.*, 1990). El posible significado biológico de esta reactividad cruzada chaperonina-actina no ha sido aclarado. Vale la pena aclarar que la similitud estructural reportada entre el dominio amino terminal de la Hsp70 (ATPasa) y la actina (Flaherty *et al.*, 1991; Bork *et al.*, 1992), no parece ser de relevancia en este caso. De cualquier manera, en particular por la forma en que Meagher generó los anticuerpos anti-actina de soya (Sac1:257-272), usando el oligopéptido sintético **FGCPGVLYQPSMVGME**, los anticuerpos deberían reconocer secuencias similares. Una búsqueda rápida de secuencias similares en 18 secuencias de chaperonina a lo largo de la escala filogenética incluyendo TF55 y TCP-1, indica que solo en la cpn60- $\alpha$  de trigo y de *Brassica napus*, existe la secuencia **VGPDGVLSIESSSFE** (aa's 167-182), que muestra una baja identidad con la secuencia de actina de soya en particular el tripéptido **GVL** (las identidades se muestran en **negritas**). Otros dos anticuerpos contra las secuencias (isoformas) de actina de soya Sac4:257 y Sac3:257 de secuencias **FRCPEVLFQPSMIGME** / **FRCPKILFQPSMIFME**, respectivamente, tienen menor identidad con cpn60- $\alpha$ , no tienen el tripéptido **GVL**, y aparentemente **no** detectan a la chaperonina (ARE). Desde luego, muchas otras proteínas podrían contener dicho tripéptido, por lo cual si acaso la secuencia detectada fuera el epítipo "común" entre Sac1 y chaperonina; debe de haber otros factores involucrados. Estructural y/o funcionalmente, es imposible derivar ninguna similitud. En el caso de los otros anticuerpos anti-actina que detectan chaperoninas (Varma *et al.*, 1987; Pérez *et al.*, 1994), no es posible detectar la menor similitud de secuencia. Curiosamente, actina no es la única proteína del citoesqueleto que cruza inmunológicamente con cpn60. En este sentido, se ha reportado que una proteína de 65 kDa, de *Spirochaeta bajacaliforniensis*, inicialmente caracterizada como "tubulin-like", por detección con al menos cuatro diferentes anticuerpos contra tubulina, es en realidad una chaperonina (Munson *et al.*, 1993).

De cualquier manera, la detección en el laboratorio de una proteína, que copurificaba con actina y que podría corresponder a una chaperonina, valía la pena de ser investigada.

Por otro lado, existen reportes que demuestran que en varias circunstancias donde hay una parasitosis o simbiosis intracelulares; por un lado el citoesqueleto juega un papel primordial y por el otro en el microsimbionte las chaperoninas tienen un nivel de expresión elevado y desempeñan importantes funciones.

Por ejemplo, en el caso de la bacteria parasítica intracelular *Listeria monocytogenes*, después de ser fagocitada por macrófagos, ésta se encapsula por filamentos de actina. A medida que la bacteria crece y se divide, los filamentos de actina se rearreglan dando la apariencia de "colas de cometa", cuya velocidad de crecimiento concuerda con el desplazamiento de la bacteria

en el citoplasma de la célula infectada. Esta capacidad de la bacteria para utilizar al citoesqueleto de actina, le permite moverse por el citoplasma de una célula a otra y así evitar la respuesta inmune del huésped (**Tilney & Portnoy 1989**). *Shigella flexneri*, presenta una conducta similar dentro de la célula huésped. Inclusive el virus de vaccinia comparte esta habilidad de enrolar al citoesqueleto de actina. Los genes responsables de reclutar a la actina comienzan a ser caracterizados. Curiosamente en el caso de *Listeria* y *Shigella*, las proteínas involucradas, ActA y IcsA respectivamente, no tienen ninguna similitud de secuencia entre sí y vaccinia no contiene ningún gen parecido a éstos lo cual sugiere que existe más de un mecanismo para orientar al citoesqueleto a favor del patógeno intracelular (**Strauss & Falkow 1997**).

Por otro lado, *Salmonella typhimurium* cuando infecta macrófagos, permanece dentro de una vacuola durante toda su fase intracelular (**Strauss & Falkow 1997**). Durante esta etapa se incrementa selectivamente la síntesis de más de 30 proteínas de la bacteria. Dentro de éstas, las dos que mayor inducción alcanzan son GroEL y DnaK. Si bien deben ser otros genes los responsables de la patogenia, la sobreexpresión de estas proteínas de stress claramente le confieren una ventaja al patógeno para sobrevivir el ambiente hostil dentro de los macrófagos (**Buchmeyer & Heffron 1990**). *Amoeba proteus* cepa D, adquirió en el laboratorio, una bacteria endosimbionte no identificada (bacteria X). Aparentemente esta bacteria no es dañina a la ameba que acumula 42,000 bacterias intracelulares. Sin embargo, si bien se desconoce el papel que juega la bacteria, actualmente este protozooario se ha hecho dependiente de estos endosimbiontes y no puede sobrevivir sin ellos. Estudios de las proteínas bacterianas involucradas en esta relación, detectaron altos niveles de un homólogo de la familia GroEL, que sugieren que la condición de endosimbionte es una situación de stress y que por lo tanto la bacteria X, mantiene un alto nivel de expresión de GroEL para sobrevivir dentro de una célula eucarionte (**Choi et al., 1991**). El punto de vista de que el interior de una célula podría ser un ambiente hostil fue propuesto originalmente en 1974 y desarrollado posteriormente (**Moulder 1974; Moulder 1985**). Este enfoque está basado en la conceptualización de las barreras que se le van a presentar a los parásitos intracelulares para interiorizarse, sobrevivir y multiplicarse dentro del huésped. Estas condiciones intracelulares pueden ser comparables a las condiciones abióticas de stress como el calor, el frío, la salinidad o la desecación. Se podría pensar que el enfoque de que el citoplasma de la célula huésped podría ser un medio ambiente hostil, podría no aplicarse en el caso de simbiosis establecidas. El caso de la bacteria X y otros casos revisados más adelante señalan que aun en el caso de relaciones mutuamente benéficas esta percepción del citoplasma como un medio agresivo se sostiene.

Desde luego todos los parásitos intracelulares tienen genes de stress, pero no siempre se ha cuantificado si existe una diferencia en su nivel de expresión comparativamente entre el estadio

intracelular y el de vida libre. Por ejemplo, se ha reportado que *Mycobacterium tuberculosis* y *Mycobacterium leprae* expresan un homólogo de GroEL, que se sabe que éste es uno de los principales inmunógenos de estas bacterias ya sea en la infección o después de una inmunización. Inclusive, este antígeno podría estar involucrado en los fenómenos de autoinmunidad y en la artritis (**Shinnick et al., 1988**). En *Mycobacterium tuberculosis*, esta proteína está codificada por dos genes. A nivel de aminoácidos las proteínas muestran solo 61% de identidad, lo cual sugiere una divergencia funcional de los polipéptidos. Aunque no se ha demostrado, se ha sugerido que uno de los polipéptidos (cpn60-2), podría estar más relacionado con la patogenicidad que el otro (cpn60-1) (**Kong et al., 1993**).

Otro ejemplo es el de los Afidos (Insecta; Homoptera). Estos, poseen procariontes intracelulares simbiotróficos del género *Buchnera*, en células especializadas del abdomen denominadas bacteriocitos. Las bacterias endosimbiontes no pueden vivir fuera del citoplasma de su célula huésped y no se les conoce un estado de vida libre. Afidos aposimbióticos (sin bacterias endosimbiontes), muestran un crecimiento lento y son estériles. Se ha sugerido, que estos endosimbiontes aportan aminoácidos esenciales y vitaminas a su huésped, lo cual significa que existe una relación mutualista. Un aspecto importante en esta relación, es que el endosimbionte sintetiza predominantemente una sola proteína: la **simblonina**. Esta ha sido purificada del áfido del chícharo, y se ha demostrado que estructural y funcionalmente es un homólogo de GroEL (**Kakeda & Ishikawa 1991**). En este caso, basándose en las características morfológicas, metabólicas y genómicas del endosimbionte, se ha propuesto que se les podría considerar como formas intermediarias entre bacterias de vida libre y organelos (**Fukatsu & Ishikawa 1993**), pero *vide supra*. El genoma de este endosimbionte tiene una alta proporción de adenina-timina. Esta característica también la poseen los genomas mitocondriales y apoya la idea de que su genoma está sufriendo cambios concertados con su condición de endosimbionte. Pero lo más notable es que el tamaño de su genoma es mayor que el de *E. coli*, i. e., en potencia tiene capacidad codificante para miles de proteínas y sin embargo se concentra en la síntesis de una sola (**Ishikawa & Hashimoto 1986**). A nivel de secuencia, la simblonina está codificada por el operón symSL que codifica a los homólogos de GroES y GroEL y muestran respectivamente 79.6% y 85.5% de identidad con los homólogos de *E. coli*. Experimentos de complementación muestran que mutantes de *E. coli* pueden funcionar con los genes del endosimbionte (**Ohtaka et al., 1992**). Sin embargo, en el endosimbionte la expresión del homólogo de GroES es mucho más baja que la de GroEL. El significado evolutivo y funcional de esta asimetría no ha sido esclarecido (**Fukatsu & Ishikawa 1993**). Una posibilidad sería que la fisiología del endosimbionte ha cambiado de acuerdo a su situación de comensal obligado. En este sentido es interesante que se ha reportado que si bien el operón symSL en potencia podría responder al heat-shock (el promotor contiene las

secuencias regulatorias) (Ohtaka et al 1992) , la cantidad de mRNA de symSL, no se incrementa en respuesta a este stress. Posiblemente esta falta de respuesta es debida a defectos en la expresión y/o estructura del factor sigma 32, que regula este tipo de operones. Por el otro lado, la evolución de las vías regulatorias de estas bacterias deben de haber llevado a suprimir una regulación hasta cierto punto innecesaria y a facilitar la expresión de la simbionina aun en ausencia de stress calórico, pero en respuesta al medio ambiente intracelular del bacteriocito (Sato & Ishikawa 1997). Sin embargo no en todos los casos de parasitismo intracelular se ha perdido la capacidad de respuesta al stress calórico. Por ejemplo, en *Legionella pneumonia*, se ha reportado que efectivamente la sobreexpresión de su Hsp60 se presenta durante la infección de monocitos humanos y de ratón. Sin embargo, una cepa avirulenta de *Legionella*, es capaz de responder al heat shock, sugiriendo que en esta cepa, la vía regulatoria de stress funciona bien, pero la regulación durante la infección está defectuosa (Fernández et al., 1996). Finalmente, entre los insectos, el caso de los áfidos no parece ser único, ya que se han reportado bacterias endosimbiontes, que también tienen niveles elevados de SymL (GroEL), en moscas tse-tse y en tres especies de gorgojos (Aksoy 1995; Charles et al., 1995).

Por lo anterior, se puede apreciar que las chaperoninas están íntimamente asociadas al fenómeno del endoparasitismo y la endosimbiosis.

Por lo que respecta a la relación de chaperoninas y fijación del nitrógeno, el primer reporte que indicaba una relación entre éstas proviene no de una bacteria del genero *Rhizobia*, sino de la bacteria diazotrófica de vida libre *Klebsiella pneumoniae*. En esta bacteria se reportó que GroEL estaría involucrada en la biogénesis y plegamiento correcto de la nitrogenasa y de NifA (Govezensky et al., 1991). Dado que NifA es un activador transcripcional de la expresión de los genes *nif*, se ha considerado a la chaperonina como un regulador potencial de la expresión génica, aunque de segunda mano diría yo (Govezensky et al., 1994). Consecuentemente, en ambos aspectos (endosimbiosis y fijación del nitrógeno), la bacterias del genero *Rhizobia* no podían ser la excepción. En *Bradyrhizobium japonicum*, se ha reportado que GroEL, se incrementa 7 veces cuando la bacteria se encuentra como bacteroide en los nódulos de soya, en comparación con la bacteria de vida libre (Choi et al., 1991). Adicionalmente, en contraste con *E. coli* o *Mycobacterium tuberculosis*, que tienen 1 y 2 copias de GroEL respectivamente, se ha determinado que en *Bradyrhizobium japonicum*, el operon *groESL*, se localiza en una familia multigénica de cinco miembros. En particular el llamado *groESL<sub>3</sub>*, no se induce por stress calórico, sino que es co-regulado simultáneamente con los genes de la fijación del nitrógeno y su expresión es estrictamente dependiente de la proteína regulatoria NifA ya sea en el bacteroide o en condiciones de microaerofilia. Curiosamente, mutaciones en este gen *groESL<sub>3</sub>*, no afectan ni la

viabilidad, ni la nodulación, ni la fijación del nitrógeno. Esto parece indicar que los productos polipeptídicos pueden funcionalmente ser reemplazados uno por otro, pero que uno es preferencialmente inducido durante la fijación del nitrógeno, aunque en el mutante de éste la cantidad de GroEL sintetizada a partir de los otros genes sería suficiente para compensar la falta de la proteína GroEL3. Dos miembros de esta familia, parecen tener un patrón de expresión constitutivo y al menos uno de los operones es fuertemente inducido por el stress calórico. Los genes *groEL<sub>3</sub>* y *groEL<sub>4</sub>* están localizados en el cromosoma principal. (Fischer *et al.*, 1993). En *Rhizobium meliloti*, utilizando Southern blot, se han reportado entre cuatro y cinco copias de *groEL*. En este caso, se ha determinado que algunas de las copias están localizadas en sus dos de los megaplásmidos (Rusanganwa & Gupta 1993). Por su lado, *Rhizobium leguminosarum*, parece tener al menos tres copias de *groEL* ninguna de las cuales está en el plásmido simbiótico (Wallington & Lund 1994). Finalmente, se ha demostrado que en *Rhizobium meliloti*, GroEL está involucrado en el plegamiento y el ensamblaje del activador transcripcional NodD. Mutaciones en la copia cromosómica (*groEL<sub>c</sub>*) afectan notablemente la actividad de NodD1, además de la de NodD3 y SyrM. A diferencia del caso de *Bradyrhizobium*, estas mutantes son lentas en la formación de nódulos y presentan el fenotipo Fix<sup>-</sup>. Sin embargo la actividad de NifA no se ve disminuida, por lo cual se ha sugerido que el efecto de la mutación debe de estar a otro nivel de la fijación del nitrógeno, por ejemplo en el ensamblaje de la nitrogenasa. Aparentemente, el producto polipeptídico de otra de las copias de la chaperonina (*groEL<sub>a</sub>*), localizada en el megaplásmido pSyma, puede complementar las mutaciones en la copia cromosómica, lo cual sugiere una divergencia regulatoria de estos genes *i. e.*, desde el punto de vista estructural funcional ambos polipéptidos son similares; sin embargo presentan un patrón de expresión temporal diferente de manera que la copia en pSyma no puede compensar la mutación en la copia cromosómica. Aunque se han detectado otras copias adicionales de *groEL*, éstas no parecen codificar productos polipeptídicos relevantes ya que dobles mutantes en *groEL<sub>c</sub>* y *groEL<sub>a</sub>* no son viables ya sea por razones funcionales o regulatorias (Ogawa & Long 1995). En resumen, el genero Rhizobia presenta características particulares primero en el sentido de que a diferencia de otros procariontes, GroEL está codificado por pequeñas familias multigénicas y segundo porque al menos alguna de las isoformas está relacionada específicamente con la nodulación y/o la fijación del nitrógeno.

## **OBJETIVO GENERAL**

Por todo lo anterior, es claro que las chaperoninas ocupan una posición distintiva en los fenómenos de endosimbiosis en general y muy particularmente en la simbiosis Rhizobia-planta. Sin embargo, hasta donde sabemos, no existe ningún reporte en la literatura que haya estudiado a la contraparte del citoplasma de la célula eucarionte huésped ya sea animal o vegetal.

Así, por las razones antes expuestas, nos pareció interesante preguntarnos si desde el punto de vista de la planta, la simbiosis también constituye un evento estresante y si las chaperoninas vegetales también están implicadas en el proceso de morfogénesis del nódulo simbiótico.

Aparentemente, este sería el primer reporte en ese sentido.

## **OBJETIVOS PARTICULARES**

- Aislar, secuenciar y caracterizar clonas de chaperonina (cpn60) del frijol común *Phaseolus vulgaris*.
- Determinar la composición genómica del gen o genes aislados.
- Establecer si algún miembro de ésta familia multigénica, (en caso de que así se determinara), participa específicamente en alguna de las etapas de morfogénesis del nódulo simbiótico.

## **OBJETIVO GENERAL**

Por todo lo anterior, es claro que las chaperoninas ocupan una posición distintiva en los fenómenos de endosimbiosis en general y muy particularmente en la simbiosis Rhizobia-planta. Sin embargo, hasta donde sabemos, no existe ningún reporte en la literatura que haya estudiado a la contraparte del citoplasma de la célula eucarionte huésped ya sea animal o vegetal.

Así, por las razones antes expuestas, nos pareció interesante preguntarnos si desde el punto de vista de la planta, la simbiosis también constituye un evento estresante y si las chaperoninas vegetales también están implicadas en el proceso de morfogénesis del nódulo simbiótico.

Aparentemente, este sería el primer reporte en ese sentido.

## **OBJETIVOS PARTICULARES**

- Aislar, secuenciar y caracterizar clonas de chaperonina (cpn60) del frijol común *Phaseolus vulgaris*.
- Determinar la composición genómica del gen o genes aislados.
- Establecer si algún miembro de ésta familia multigénica, (en caso de que así se determinara), participa específicamente en alguna de las etapas de morfogénesis del nódulo simbiótico.



maneras. Primero; podría ser que la abundancia del transcrito fuera marginal, con la consecuencia de que la representación de las clonas correspondientes en una biblioteca de cDNA de nódulos también sería muy baja y sería muy difícil, utilizando la sonda heteróloga de *Arabidopsis*, aislar clonas positivas. La segunda posibilidad, sería que dado que en *Brassica napus* y en *Arabidopsis*, cpn60- $\beta$  ha sido reportada como codificada por una familia multigénica (Martel et al., 1990; Zabaleta et al., 1992), existía la posibilidad de ocurriera una expresión diferencial de miembros de esta supuesta familia y que la isoforma que se expresara en nódulo, tuviera una identidad baja con la sonda heteróloga que estábamos usando y que por lo tanto la detección fuera sumamente difícil, si no imposible. Por el otro lado el experimento del Southern genómico indicaba que podíamos detectar a las secuencias homólogas del frijol. Sin embargo, esas secuencias detectadas podrían corresponder a los genes que se expresan en hoja. Considerando las razones antes expuestas, el razonamiento entonces fue el siguiente: Si la chaperonina está codificada por una familia multigénica, sabiendo que la representación de clonas de cDNA en una biblioteca de hoja de frijol es alta y que la sonda de *Arabidopsis* representa a la supuesta chaperonina de hoja, entonces el mejor camino sería primero efectuar un tamizado de una biblioteca de cDNA de hoja. En este razonamiento, se veía más probable aislar clonas de hoja. La segunda parte del razonamiento era más ingenua. Si logramos aislar clonas que representen a alguno de los miembros de la familia multigénica, al menos será una secuencia de frijol, que, en teoría nos permitirá detectar secuencias homólogas en bibliotecas de nódulo. La ingenuidad consiste desde luego en que cuando los miembros de una familia multigénica de un organismo tienen un patrón de expresión diferencial, habitualmente también ha divergido su secuencia (McElroy et al., 1990), de tal manera que podría ser tan difícil la hibridación / detección de los cDNA's de nódulo usando la sonda de *Arabidopsis*, que con una clona de hoja de frijol. Pero tal vez por eso dicen que los biólogos moleculares somos los científicos más creyentes: Siempre estamos esperando el milagro. Nuestra esperanza era que la identidad reportada a nivel de nucleótidos entre los genes de cpn60 en la misma especie es más alta (cerca de 90% de identidad), que entre los homólogos de diferentes plantas (alrededor de 72%) (Valdés-López & Herrera-Estrella, 1992), de tal manera que metodológicamente el razonamiento por el que optamos era justificado.

#### **TAMIZADO DE LA BIBLIOTECA DE HOJA Y COMPARACION DE SECUENCIAS DE LAS CLONAS AISLADAS**

Así, utilizando las condiciones de hibridación antes determinadas, se tamizó una biblioteca cDNA de hoja de *Phaseolus*, construida en  $\lambda$ gt22A. Como resultado se obtuvieron dos clonas positivas (de un total de 10,000 placas). Ambas resultaron ser clonas truncadas. Sin embargo, la mayor

(≈1,200 pb), fue subclonada en el vector pBluescript SK+ para su mejor manejo y designada **chapphas-6.1**. Las comparaciones de secuencia de nucleótidos y la secuencia derivada de aminoácidos, con las secuencias correspondientes de la sonda de *Arabidopsis*, indicaron claramente que la clona aislada en efecto correspondía a un homólogo de cpn60-β.

#### **TAMIZADO DE LA BIBLIOTECA DE NODULO**

La clona de frijol aislada de hoja, fue utilizada como sonda para efectuar ahora el tamizado de una biblioteca de cDNA construida en el vector λZAP, a partir de RNA de nódulos de frijol de 12 días. Las condiciones de hibridación fueron más severas que en el primer tamizado, pero no se llegó a lavados de alta severidad. De un total de 80,000 placas se aislaron 19 clonas positivas *i. e.*, la representación de clonas en esta biblioteca, es similar a la representación de clonas en la biblioteca de hoja, lo cual sugiere que el nivel de expresión de los genes correspondientes es similar. Después de convertir al fasmido en fagémido, se realizó el patrón de restricción de las clonas aisladas que indicó que la mayoría, nuevamente eran clonas truncadas. La mayor (≈1900 bp), designada **chapnod-4.2**, se utilizó para determinar la secuencia de nucleótidos. Cabe indicar que, sin embargo, el patrón de restricción de las diferentes clonas era muy similar, si no idéntico, entre ellas mismas y con respecto a la clona aislada de hoja (Chapphas-6.1).

#### **SECUENCIACION Y ANALISIS DE LA CLONA DE NODULO**

Como era de esperarse, la secuencia de nucleótidos y la secuencia derivada de aminoácidos, confirmaron que la clona aislada de nódulo correspondía a la chaperonina cpn60-β. Específicamente, se realizó un alineamiento múltiple con la secuencia de aminoácidos correspondiente al péptido maduro (péptido señal excluido), contra secuencias a todo lo largo de la escala filogenética y particularmente con secuencias de algunas plantas, derivadas de los polipéptidos correspondientes a la mitocondria y a las llamadas subunidades alfa y beta del cloroplasto. Como se indicó, la secuencia aislada se agrupó con esta última, indicando que en los nódulos radiculares de *Phaseolus vulgaris* ésta tendría una localización plastídica. Por otro lado, la secuencia de nucleótidos señaló que a pesar del tamaño de la clona, no era una clona completa en el sentido de que dada su localización plastídica, el mRNA tendría que codificar no solo para la secuencia de aminoácidos de la proteína madura, sino, además, para la secuencia del péptido señal que dirige al polipéptido a su sitio definitivo, en este caso el plástido (Hemmingsen *et al.*, 1988; Martel *et al.*, 1990; Zabaleta *et al.*, 1992; Cole *et al.*, 1994). Dicho de otro modo, la clona aislada contenía toda la parte codificante de la proteína madura, pero solo

parte del péptido señal. Como era importante determinar la secuencia de este péptido líder para confirmar que tuviera las características esperadas para una proteína que se aloja en el plástido decidimos intentar aislar una clona completa. Se tamizó nuevamente la biblioteca de nódulo, pero utilizando ahora solo la región hacia el extremo 5' de nuestra clona de frijol (restricción NcoI/EcoRV-743 bp), para tratar de detectar en lo posible las clonas más completas hacia dicho extremo. Las condiciones de hibridación y lavado de los filtros fueron severas. Como resultado de este tamizado se aislaron 6 clonas. El patrón de restricción de la mayor, indicaba que contenía alrededor de 70 nucleótidos adicionales. De esta clona, se secuenciaron 700 bp de cada extremo (1,400 bp en total), observándose que la secuencia correspondiente era 100% idéntica a nivel de nucleótidos a la clona aislada anteriormente y, además, confirmando el tamaño determinado con el patrón de restricción, contenía **sesenta y seis** nucleótidos adicionales en el extremo 5', lo cual fue suficiente para llegar al codón de inicio y derivar la secuencia completa del péptido señal. El análisis de esta región, mostró que las características de este péptido efectivamente corresponden a un péptido señal de exportación al plástido. Dado que por lo demás, las clonas de nódulo eran idénticas en el resto de los nucleótidos, para la secuencia reportada completa se fusionó la información de ambas clonas, en una sola.

#### **DETERMINACION DEL PATRON DE EXPRESION (NORTHERN BLOT)**

Se inocularon plantas de frijol con *Rhizobium tropici* cepa CIAT 899. Se aisló RNA total de raíces y nódulos de 7-41 días. Después de la transferencia, la hibridación se efectuó utilizando como sonda a la región codificante completa de la clona de nódulo. Los resultados muestran que en raíces y nódulos de siete días, no se observan diferencias en la señal, lo cual indica que la chaperonina tiene un nivel basal de expresión similar en estas muestras. Después del día 10, la señal de chaperonina se incrementa hasta el día 14. Posteriormente la señal disminuye manteniéndose aproximadamente en los mismos niveles de expresión que en la raíz. Estos experimentos muestran que durante el momento que las células de la planta son penetradas por *Rhizobium*, este se diferencia en bacteroide, se induce la síntesis de nodulinas tardías y comienza la fijación del nitrógeno, se presenta una sobreexpresión de cpn60- $\beta$ .

#### **SOUTHERN BLOT GENOMICO. CARACTERIZACION DEL GEN DE CHAPERONINA**

Cuando se comparó la secuencia de nucleótidos de las clonas obtenidas de las bibliotecas de hoja y de nódulo de frijol, (**Chapphas-6.1 vs Chapnod-4.2**), se estableció que éstas eran 100% idénticas. Este era un resultado inesperado debido a lo expuesto anteriormente, respecto a los

reportes de que *cpn60-β* en *Arabidopsis* y en *Brassica* están codificados por una pequeña familia multigénica de tres o cuatro miembros. Si el razonamiento de la expresión diferencial de genes parálogos se aplicara en este caso, lo que deberíamos de haber encontrado era que una isoforma se expresara mayoritariamente en hoja y la otra mayoritariamente en raíz y nódulo, de tal manera que aislar de las bibliotecas correspondientes exactamente a la misma clona era muy poco probable. Adicionalmente, datos parciales de secuencia de otras clonas, así como los mapas de restricción, sugerían que no solo estas dos clonas eran idénticas, sino que todas las clonas en ambas bibliotecas, parecían serlo. Además, hibridaciones de tipo Southern de la región 3' no traducida de la clona completa de nódulo hibridaban con todas las clonas. Habitualmente, esta región trailer es considerada como gen específica y es capaz de distinguir entre transcritos de genes parálogos, si los hay (**Michelson & Orkin 1980; Valdés-López & Herrera-Estrella 1992; Kislauks et al., 1993; An et al., 1996**). Surgió entonces la hipótesis menos convencional de que posiblemente en *Phaseolus vulgaris*, *cpn60-β*, podría estar codificada por un solo gen y no por una familia multigénica. Para abordar esta pregunta, se aisló DNA genómico total de la planta (se manejaron muestras de plantas individuales, para evitar posibles problemas derivados de polimorfismos), se realizaron restricciones con varias enzimas y se realizaron hibridaciones de tipo Southern utilizando como sondas: la región codificante completa, las regiones correspondientes al extremo 5' codificante, al extremo 3' codificante y al extremo 3' no codificante. El resultado de estos experimentos efectuados tanto en alta como en baja severidad es compatible con el hecho de que en el frijol *cpn60-β*, está codificada por un solo gen.

## MATERIAL Y METODOS

### AISLAMIENTO DE DNA DE PLASMIDO

El aislamiento de DNA derivado de plásmidos de las diferentes clonas y subclonas se efectuó por una combinación de los siguientes métodos: **Birnboim, 1993; Zhou, et al., 1990; Stephen, et al., 1990**. De manera resumida: 1.5 ml de cultivo de *E.coli*, crecido 16 hrs en medio LB a 37 °C, es centrifugado por 20-30 segundos (la temperatura es intrascendente; puede hacerse a 4 °C o a temperatura ambiente). Es posible agregar nuevamente 1.5 ml de cultivo al mismo tubo, con lo que la recuperación de plásmido se duplica. Después de descartar el sobrenadante, el pellet es resuspendido en 1 ml de 100 mM NaCl en TE (10 mM Tris-HCl pH 7.4, 1 mM EDTA). Aquí es recomendable centrifugar nuevamente y repetir este paso de lavado. Se centrifuga nuevamente por 20-30 segundos y se decanta el sobrenadante dejando 50-100 µl de la solución. Después de resuspender vigorosamente, se adicionan 450 µl de TENS (TE, adicionado de 0.1 N NaOH y 0.5% de SDS). Se agita suavemente por inversión y se deja reposar 5 minutos en hielo. Se agregan 225 µl de acetato de sodio 3 M pH 4.8-5.2 y nuevamente se agita suavemente por inversión hasta que aparece un floculado blancuzco y se deja reposar 5 minutos en hielo. Pasado este tiempo se centrifuga por 7 minutos a 18° C (aproximadamente a temperatura ambiente, aunque se puede centrifugar a 4° C). Después de pasar el sobrenadante a un nuevo tubo, se agrega 1 volumen de etanol absoluto (aproximadamente 700 µl). Normalmente los métodos de precipitación de ácidos nucleicos emplean 2 volúmenes de etanol. Sin embargo, usando 1 volumen, la calidad aumenta notablemente, si bien lógicamente, la cantidad disminuye. Después de mezclar por inversión, la mezcla se deja en hielo por no más de 5 minutos y se centrifuga por 7 minutos. El pellet se lava con etanol 70%. Después de descartar el sobrenadante, se pulsa en la microfuga y se aspira el poco líquido remanente. El tubo se invierte y se deja evaporar el resto del alcohol por 10 minutos. Es importante no sobre-secar el pellet ya que como se sabe desde hace tiempo (**Avery et al., 1944**), el DNA sobre-deshidratado, pierde actividad biológica, es difícil de resuspender y puede ser difícil de manipular posteriormente, en particular si subsecuentemente se va a efectuar algún procedimiento con enzimas. Finalmente el pellet se resuspende en 10-100 µl de TE.

### TAMIZADO DE LAS BIBLIOTECAS DE cDNA

El tamizado inicial se hizo a partir de una biblioteca de cDNA de hoja de frijol amablemente proporcionada por el laboratorio de la Dra. Alejandra Covarrubias, construida en  $\lambda$ gt22A, usando los sitios *Sal* I (5') / *Not* I (3') del "polylinker" (GIBCO BRL). Se plaquearon dos cajas de Petri (90 mm)

con 5.000 unidades formadoras de placa (pfu), cada una. Como sonda se utilizó el cDNA de cpn60- $\beta$  de *Arabidopsis thaliana* (Zavaleta et al., 1992) y las condiciones de hibridación y lavado fueron laxas (vide infra). Se detectaron dos clonas positivas. Una de estas fue aislada, se aisló el DNA (Chisholm, 1989; Ausubel, et al., 1992) se restringió con Eco RI / Not I y el inserto fue subclonado en el plásmido (fagémido), Bluescript SK +.

Para aislar clonas de nódulos, se utilizó una biblioteca de cDNA construida en  $\lambda$  ZAP (Stratagene) a partir de nódulos de frijol de 12 días. Para el tamizado se plaquearon 4 cajas de Petri de 150 mm con  $\approx$ 20.000 unidades formadoras de placa (pfu), por caja. Como sonda se utilizó la clona aislada de la biblioteca de hoja y las condiciones de hibridación y lavado fueron de severidad media (vide infra). Después del tercer tamizado, las clonas positivas fueron aisladas y convertidas al plásmido Bluescript SK- por medio del protocolo de escisión *in vivo* con el fago ayudador ExAssist, de acuerdo a las instrucciones del fabricante (Stratagene). Se aislaron los plásmidos y se restringieron con EcoRI y XhoI para determinar el tamaño de los insertos. La clona mayor fue escogida para su posterior caracterización (Ausubel, et al., 1992).

## INDUCCION DE NODULOS

Semillas del frijol común *Phaseolus vulgaris* L. cv Negro Jamapa, fueron obtenidas de Pronase (México). Las semillas fueron germinadas por 2 días, transferidas a macetas e inoculadas con *Rhizobium tropici* cepa CIAT 899 (Martínez-Romero et al., 1991). Las plantulas se cultivaron en invernadero y los nódulos se colectaron después de diferentes días de la inoculación (entre 7 y 41). Raíces no infectadas fueron obtenidas de plantas cultivadas en condiciones similares. Ambos tejidos fueron congelados (-70 °C) inmediatamente después de colectados (Lara et al., 1984).

## AISLAMIENTO DE DNA DE PLANTA

El DNA de planta fue obtenido a partir de hojas de frijol por medio del método de Dellaporta (Dellaporta et al., 1983); del método de Doyle y Doyle (Doyle & Doyle, 1990); o por una combinación de ambos. En este último caso se siguió el método de Doyle y al DNA aislado se le reprecipitó con CTAB (cetyltrimethylammonium bromide) de acuerdo con mejoras propuestas por el método de Dellaporta. En este caso el DNA aislado fue de la mejor calidad.

## **SOUTHERN BLOTS GENOMICOS (Southern, 1975)**

A fin de evitar posibles problemas de polimorfismos, el aislamiento del DNA de planta se hizo a partir de hojas de individuos. Habitualmente 5-7 g de hoja resultan en una extracción de 300-500 µg de DNA. En el Southern genómico que se hizo utilizando la sonda homóloga de *cpn60-β* de *Arabidopsis thaliana*, el experimento se hizo con 16 µg de DNA. En los Southern que se hicieron utilizando como sonda secuencias homólogas del cDNA de la clona de frijol se utilizaron únicamente 2-5 µg de DNA. El DNA fue digerido con las enzimas de restricción adecuadas (Boehringer Mannheim, Mannheim, Germany), y se corrió una electroforesis en gel de agarosa al 1.2%. Después de tñido, el DNA fue transferido a membranas de Nylon (Hybond-N+ / Amersham) por el método de transferencia alcalina que posibilita la transferencia y el fijado simultáneos a la membrana sin necesidad de homeado o "crosslinking" (Sambrook, *et al.*, 1989). Como sonda se utilizaron la clona de *Arabidopsis thaliana* o la clona de frijol de *cpn60-β*, marcadas radioactivamente con <sup>32</sup>P por el método de primers al azar (Feinberg & Vogelstein, 1983). La actividad específica de las sondas en éste y en otros marcajes fue de alrededor de 5 X 10<sup>8</sup> cpm / µg. En la hibridación, la condición suave fue a 42° C, la condición media a 50-55° C y la condición severa fue a 65° C. Como bloqueador se utilizó heparina (Singh & Jones, 1984). En los lavados, la condición severa fue a 65° C en 0.1X SSC, 1% SDS; la condición de severidad media fue en 0.5 X SSC, 1% SDS a 55° C y la condición suave en 2-1 X SSC, 1% SDS a 42° C. Las membranas se expusieron con película X-Omat para rayos x (Kodak), con dos pantallas intensificadoras a -70° C por 1-7 días.

## **NORTHERN BLOT (Thomas, 1980)**

Se aisló RNA total de raíces y nódulos de frijol de 7-41 días por el método del fenol caliente (de Vries, *et al.*, 1991). 5 µg de este RNA total fueron corridos en gel de agarosa al 1.5% con formaldehído (Logemann, *et al.*, 1987;) y transferidos a membranas de Nylon (Sambrook *et al.*, 1989). Como sonda se utilizó la clona de frijol de *cpn60-β*. Las condiciones de hibridación y lavado fueron de alta severidad.

## **OBTENCION Y ANALISIS DE LAS SECUENCIAS DE DNA**

Se efectuó un análisis de restricción de las clonas aisladas. Debido al tamaño de las clonas (< 1,000 bp), para obtener la secuencia completa, se realizó una restricción *Xba* I / *Eco* RI, o *Kpn* I / *Xho* I a fin de generar una colección de delecciones anidadas por el método de exonucleasa III

(Putney, *et al.*, 1981; Jasin, *et al.*, 1983; Henikoff, 1984). Tanto las deleciones como las clonas originales, fueron secuenciadas utilizando cadena sencilla o cadena doble, por el método de los dideoxi-terminadores (Sanger, *et al.*, 1977). Como iniciadores, se usaron los "primers" T3 y/o T7 que se ubican en el propio vector de clonación (Bluescript SK (+/-)). En un caso, la clona se secuenció comercialmente (Molecular Genetics Core Facility, Dep. of Microbiology & Molecular Genetics, UT-Houston Medical School / Co. David S. Needleman, Ph. D.). Las clonas relevantes se secuenciaron varias veces en ambas cadenas.

Para el análisis y comparación de la secuencia se utilizaron los programas GeneWorks, versión 2.31; PCGENE, versión 6.01 (IntelliGenetics, Mountain View, CA) y bancos de datos asociados.

## **OTROS METODOS DE DNA RECOMBINANTE**

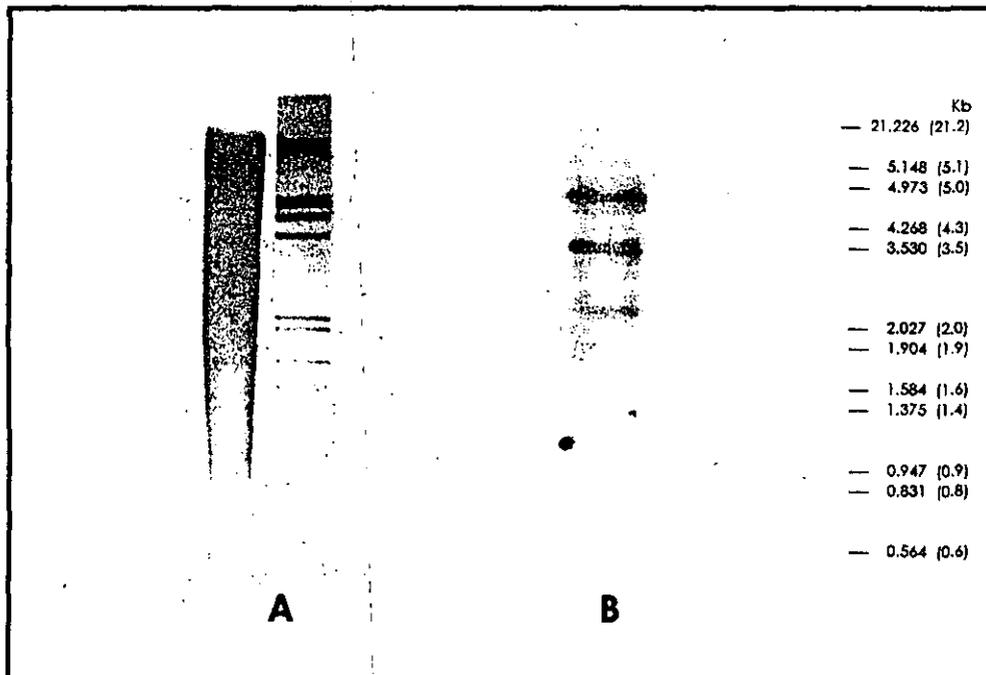
El aislamiento y purificación de fragmentos de restricción para ser subclonados o para marcaje radioactivo se hizo por el protocolo de polvo de vidrio comercializado como GENE CLEAN (BIO 101, Inc., La Jolla, CA), (Felberg & Vogelstein, 1983), o por el método de congelamiento y centrifugación (Koenen, 1989). La transformación de células de *E. coli*, se hizo por el método de cloruro de calcio (Mandel & Higa, 1970; Cohen *et al.*, 1972; Oishi & Cosloy, 1972; Hanahan, 1983); por el método de dimethyl sulfoxido (Chung, *et al.*, 1989) y ocasionalmente por electroporación (Dower *et al.*, 1988; Solfoz & Bienz, 1990). Las cepas de *E. coli* que se usaron en la manipulación de plásmidos, fásmidos, fagémidos y fagos fueron: XLI-Blue, XLI-Blue MRF, SOLR (Stratagene); Y1090*hsdR* y DH5 $\alpha$  (GIBCO-BRL). Todas las enzimas (endonucleasas de restricción, polimerasas, nucleasas, fosforilasas, fosfatasa y ligasas), así como los reactivos, fueron de la más alta calidad. El resto de los procedimientos (ligaciones, defosforilaciones, etc.), fueron realizados de acuerdo a protocolos establecidos de DNA recombinante (Ausubel, *et al.*, 1992; Sambrook, *et al.*, 1989).

## RESULTADOS Y DISCUSION

### Determinación de las condiciones de hibridación

#### Southern genómico de *Phaseolus vulgaris* con la sonda de *cpn60-β* de *Arabidopsis thaliana*.

El DNA total aislado de *Phaseolus vulgaris*, se digirió con *Hind*III, se corrió en un gel de agarosa al 1.2% y se tiñó con bromuro de etidio. En la **figura 1A** se muestra este gel. Este DNA se transfirió a una membrana de nylon y se hibridó con la clona completa de *cpn60-β* de *Arabidopsis thaliana* (**figura 1B**). Las condiciones de hibridación y lavado permitieron observar bandas de hibridación, alguna de las cuales podrían corresponder al gen o genes homólogos del frijol.

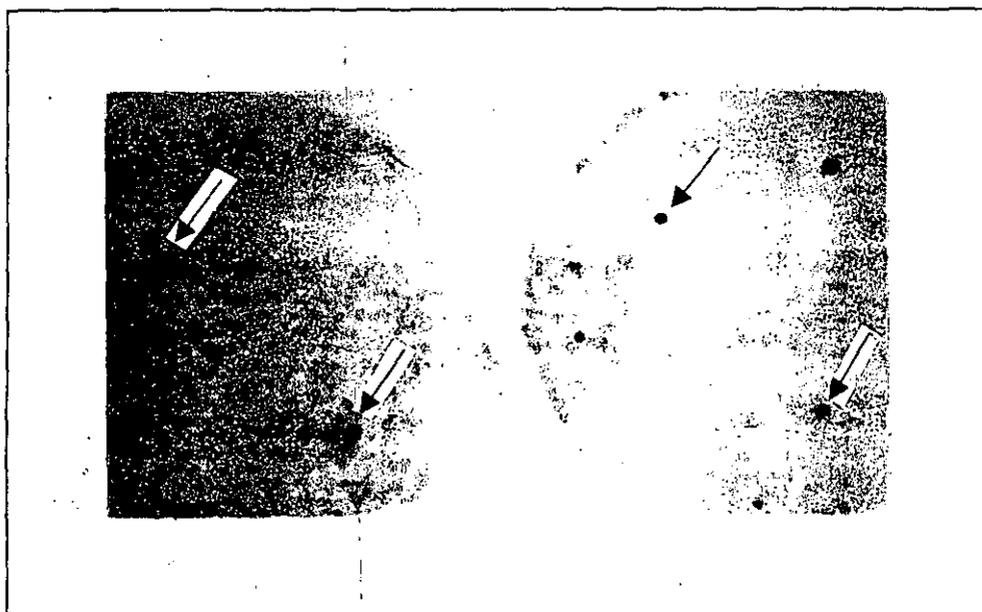


**Figura 1. A)** Gel de agarosa al 1.2% del DNA total de *Phaseolus vulgaris* (16 µg) digerido con *Hind*III. El marcador de peso molecular es el genoma del fago λ digerido con *Hind*III / *Eco*RI. **B)** Southern genómico del DNA total de *Phaseolus vulgaris*, hibridado con la clona completa de *cpn60-β* de *Arabidopsis thaliana* (Zabaleta *et al.*, 1992). La hibridación se efectuó a 42° C con 50% formamida y los lavados a 60° C con 2 X SSC, 0.1% SDS.

#### Tamizado de la biblioteca de hoja de frijol con la sonda *cpn60-β* de *Arabidopsis thaliana*.

La biblioteca de cDNA de hoja de frijol construida en λgt22A, fue plaqueada en dos cajas de 90 mm usando *E. coli* Y1090, a una densidad de aproximadamente 5.000 pfu por caja. Para la

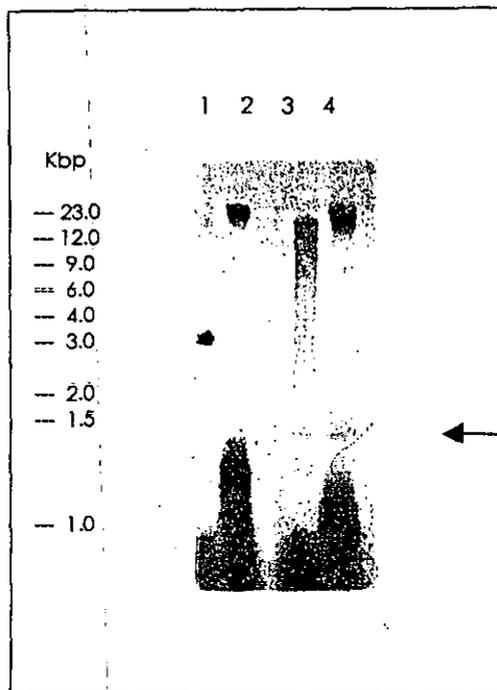
transferencia se utilizaron filtros de Nylon Hybond N+ (original y duplicado). Estos, fueron hibridados con la misma sonda y en las mismas condiciones establecidas anteriormente. Como se ve en la **figura 2**, se detectaron dos señales muy claras en uno de los filtros. La otra caja no mostró ninguna señal, por lo que tenemos una relación señales/pfu de 1:5,000. Este valor, de alguna manera indica que la representación de este transcrito en las células de la hoja equivale aproximadamente al 0.02% del mRNA total ( $\approx$  40-60 moléculas / célula, asumiendo que cada célula contiene 200,000-300,000 moléculas de mRNAs), lo cual lo cataloga en un nivel de abundancia baja moderada (Hightower & Meagher, 1985; Lewin, 1994; Sambrook, et al., 1989; Adams et al., 1995; Velculescu, et al., 1995; Zhang, et al., 1997)



**Figura 2.** Clonas positivas de hoja de frijol detectadas en el primer tamizado de la biblioteca de cDNA. Como sonda se utilizó la clona completa de cpn60- $\beta$  de *Arabidopsis thaliana*. Se muestra el filtro que dió 2 señales positivas que se identifican claramente con las flechas en el filtro y su duplicado.

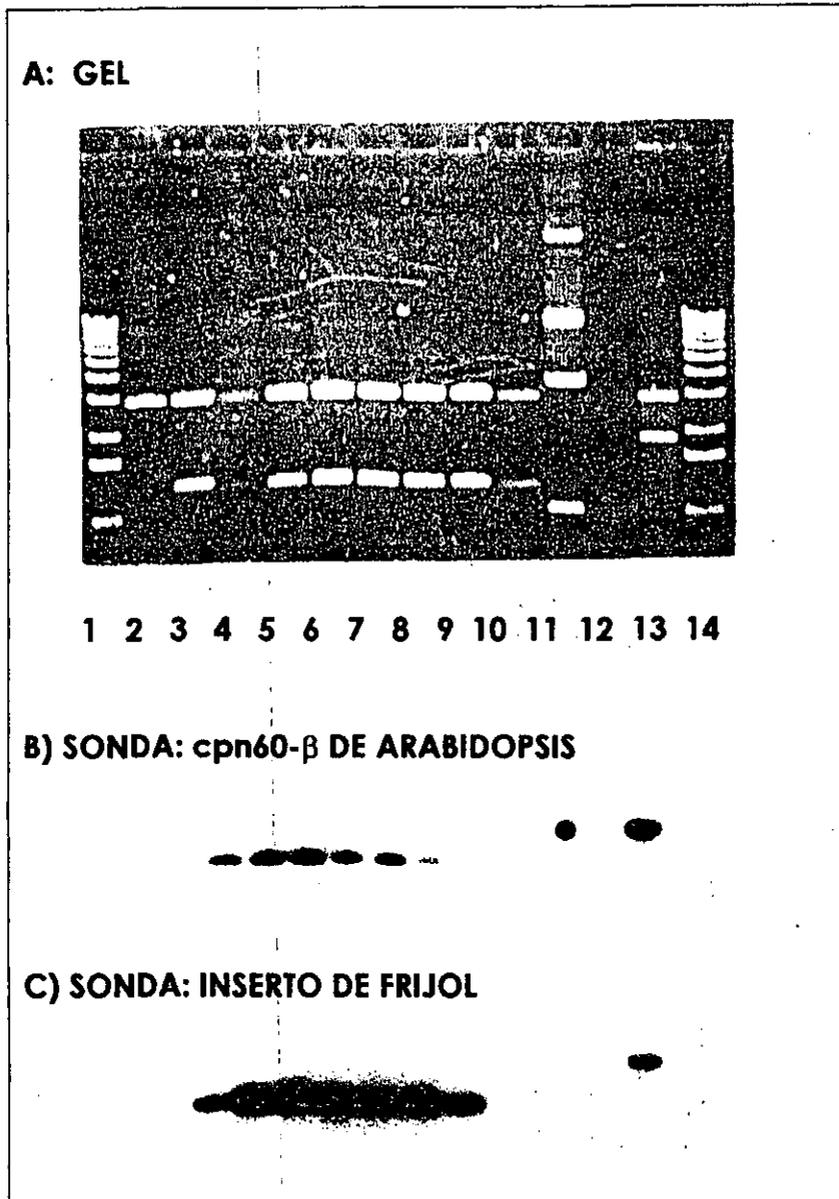
#### **Subclonación del inserto de la clona de hoja en pBluescript SK.**

Después del tercer tamizado, una de las placas, fue separada, cultivada y su DNA asilado y digerido para liberar el inserto y subclonarlo. En la **figura 3** se muestra que el inserto contenido en la clona es de alrededor de 1,200 bp, lo cual indica que es una clona truncada, circunstancia que no impide que pueda ser utilizada como sonda homóloga.



**Figura 3.** Restricción Eco RI / Not I de la clona ( $\lambda$ gt22A) aislada de la biblioteca de hoja de frijol. La flecha indica la posición del inserto liberado. El carril 1 contiene un plásmido de 3 Kbp; el carril 2 a la clona sin digerir y los carriles 3 y 4 a la clona digerida con las enzimas indicadas.

El inserto fue aislado y subclonado en el vector pBluescript SK+. Después de aislar varias transformantes individuales, se aislaron los plásmidos, se digirieron con las mismas enzimas que sirvieron para subclonarlo (Eco RI / Not I) y se efectuó una transferencia de tipo Southern. Para comprobar su identidad las subclonas se hibridaron contra la sonda original cpn60- $\beta$  de *Arabidopsis thaliana*. Como control adicional el inserto de frijol fue utilizado a su vez como sonda para efectuar la hibridación recíproca. Como se puede apreciar en la **figura 4**, las hibridaciones indican que la clona de frijol es homóloga al gen cpn60- $\beta$  de *Arabidopsis*.



**Figura 4. A)** Gel de agarosa de las clonas aisladas de la biblioteca de hoja de frijol, digeridas con *Eco* RI / *Not* I. Los carriles 1 y 14 contienen el marcador de peso molecular (escalera de 1 kb). El carril 2 contiene al vector sin inserto. Como control negativo, el carril 11 contiene una clona de actina de *Acanthamoeba* digerida con *Bam* HI. Como control positivo, el carril 13 contiene a la clona de cpn60-β de *Arabidopsis thaliana* digerida con *Eco* RI. Los carriles 3-10, contienen a las clonas aisladas de transformantes individuales de cpn60-β de frijol. **B)** Este gel fue transferido a una membrana de Nylon (Hybond N+) e hibridado con la sonda cpn60-β de *Arabidopsis thaliana*. **C)** Después de remover la marca radioactiva, la misma membrana fue hibridada utilizando ahora el propio inserto liberado de las clonas de frijol. En ambos casos la hibridación se efectuó a 42° C en 50% formamida y los lavados hasta 0.2 x SSC, 0.2% SDS, 65° C.

## Secuencia y comparación de la clona chapphas-6.1 de hoja de frijol.

Como todas las subclonas provienen de la misma clona original, una de éstas, designada a partir de aquí **chapphas-6.1**, fue secuenciada para confirmar su identidad. En la **figura 5**, se muestra la secuencia de nucleótidos y la secuencia derivada de aminoácidos. Como se puede apreciar, la clona consta de 1,189 nucleótidos incluyendo 22 adenilatos correspondientes al Poly-A. Así mismo, la secuencia muestra un marco de lectura con capacidad de codificar 322 residuos de aminoácidos. Obviamente, por la longitud de la clona, se trata de una clona incompleta (truncada). En la **figura 6**, se muestra el alineamiento de nucleótidos de esta clona de hoja, con la secuencia completa de la clona B3 de cpn60-β de *Arabidopsis thaliana*. A partir del porcentaje global de identidad (74.0 %), claramente se puede inferir que ambas secuencias son homólogas. En particular desde el inicio del alineamiento hasta los codones de término respectivos, el porcentaje es de 79.77 %, sin un solo "gap" (773 identidades / 969 bp). Por el otro lado, como era de esperarse, en la región 3' no traducida la identidad es de 28.9 %, además de que en esta región se localizan todos los "gaps" del alineamiento. La comparación de las secuencias derivadas de aminoácidos (88.8% de identidad), también confirma que la clona 6.1 de hoja de frijol corresponde a una chaperonina cpn60-β.

```

*****
* TRANSLATION OF A NUCLEIC ACID SEQUENCE *
*****
Done on DNA sequence CHAPHOJA.
Total number of bases is: 1189.
Analysis done on the complete sequence.
Done on (absolute) phase(s): 1.
Using the Universal genetic code.

AAAAAGATAACCAATGCTAGGGATCTTTTCAACATACTGGAGGATGCTATTAGAAGTGGATATCCTATATTGGTTATTGCGAGGATATTGAACAGGAAGCTCTAGCAACTCTAGTGGTG 120
K K I T N A R D L F N I L E D A I R S G Y P I L V I A E D I E Q E A L A T L V V 40

AACAGACTTAGAGGATCACTGAAGATTGCAGCGCTTAAGGCCCTGGGTTGGTGAACGCAAGAGCCAGTACCTTGATGACATTGCAATTTGACTGGAGTACTGTAATCAGAGAAGAG 240
N R L R G S L K I A A L K A P G F G E R K S Q Y L D D I A I L T G G T V I R E E 80

GTGGCCCTACTTTGGCAAAGCTGGGAAAGAGGTTCTCGGATATGCGTCCAAGGTGGTACTCACCAGGATACAACAACAATTGTTGGTGATGGAAGTACCAGGAGGCAGTGAACAAG 360
V G L T L D K A G K E V L G Y A S K V V L T K D T T T I V G D G S T Q E A V N K 120

AGAGTTGCACAAATTAAGATCCAAATTGAGGCTGCAGAGCAAGATTACGAGAGGGAGAAGCTGAATGAGAGAATTGCTAAATTTGCTGGTGGTGGCCGATACAGGTTGGTGCACAA 480
R V A Q I K I Q I E A A E Q D Y E R E K L N E R I A R L S G S V A V I Q V G A Q 160

ACTGAAACAGAGCTCAAGGAAAGAAAATTGAGAGTTGAAGATGCTCTTAATGCTACAAGCCAGCTSTAGAGGAAGTATTGTAGTTGGAGGTGGCTGCACCTCTGAGACTTGCATCA 600
T E T E L K E K K L R V E D A L N A T K A A V E E G I V V G G G C T L L R L A S 200

AAGTGGATGCCATCAGAGATAGTCTTGATAATGATGAAGAAAAGTTGGAGCTGATATTGTAAGAGAGCTCTTAGTTACCCCTGAAATTAATGCCAAGATGCTGGTGTCAATGGT 720
K V D A I R D S L D N D E E K V G A D I V K R A L S Y P L K L I A K N A G V N G 240

AGTGTGCTCAGTGAGAAGTATTGTCAGCGCAATGCAAGATATGGATATAATGCTGCCACCGGAACCTATGAAGATTTGATGCTGCCGGGATCATTGACCCAACAAGGTGTCAGA 840
S V V S E K V L S S D N A R Y G Y N A A T G N Y E D L H S A G I I D P T K V V R 280

TGTGCTGCAAGATGCAAGCTTCAAGTCCCAAGACCTTCTTAATGTCAGATTGCGTGGTGTGAGATAAAGGACCTGAGCCGCTACCTGCTGGAACCCCATGGAGCTTCAAGGATAT 960
C C L E H A A S V A K T F L H S D C V V V E I K E P E A V P A G N P H D A S G Y 320

GGTATGTAGAGAGCAACCGGAGAAGTTGAGTGAGAACAAGATTTTGTCTCTCAGAGAAATGAGAAATTTGGAAGGCATTATTATAGCAACAGGTCCAAATTTACGATGGAAACCTGT 1080
G H * 322

AGTADGTGTGATGAACAGATAGTGATGCACCCCCATCCCATCCCAATTTTTTTTAAAGTATTAAATAATTTTTTGGCTATTTCTAAAAAAAAAAAAAAAAAAAAAAAAAA 1189

```

**Figura 5.** Secuencia de nucleótidos y secuencia derivada de aminoácidos de la clona de hoja de frijol **chapphas- 6.1**. El codón de término y el sitio de poliadenilación están subrayados.

```

*****
* ALIGNMENT OF TWO NUCLEOTIDE SEQUENCES. *
*****
The two sequences to be aligned are:
61CHAHOJ. Cpn60B. Hoja Phaseolus vulgaris.
Total number of bases: 1189.
SEQATH60B. Cpn60B. Arabidopsis thaliana.
Total number of bases: 2054.
Open gap cost      : 10
Unit gap cost      : 10
The character to show that two aligned residues are identical is ':'

SEQATH60B - CTTCTTCTCTGCTGCTCAGCTCCCCTTTTCTTATCCGSAACCAAAAACCTTCGGTGATAAAAACCTCTCTGTCATCTCTACTTCTGGATAAAATGGCGTCAACTTT -110
SEQATH60B - CACTGCCACATCTTCTAATGGTCAATGGTTGCTCCGAATGGTCACAAATCTGATAAGAAACTATCAGCAAGTGTCTCTCAAGCTCATTTGGAAGGAGACAAAGTGTGT -220
SEQATH60B - GTCCTAGACCCAGAAGATCCAGTTCTGCAATTTGTGCGCAGCAAGGAATTACATTTCAACAAGATGGGACTACCATTAGGAGACTTCAAGCTGGTGTCAACAAGCTC -330
SEQATH60B - GCAGATCTTGTGTTTACACTTGGACCCAAAGGACGAAATGTTGTTCTTGAGAGCAAGTATGGATCACCAGAAATGTTAATGATGGTGTGACTGTTGCAAGAGAGGT -440
SEQATH60B - TGAGTTGGAAGACCCGGTTGAGAACATTTGGCGCTAAGCTTGTGAGGCAAGCGGCTGCCAAAACCAATGACTTGGCTGGTACCGTACACAACATCTGTTGTTCTTGAC -550
SEQATH60B - AAGGTTTTATCGCTGAGGGTGTCAAGTGGTAGCTGCTGGTGGAAACCTGTTTGTATCACCAGAGTATGAGAAGACAGCAAGGCTTTGGTTACCCGAGCTCAAGAAA -660
SEQATH60B - ATGCTAAGAGSSTGAAGACAGTGAACCTTGCAGATGTTGTCAGCTGTTAGTGGGGTAAACAAGCAAGTGGAAATGATTTGCTGAAGCAATGAGCAAGTGGGCGAG -770
SEQATH60B - GAAAGGTGTGGTACCCTTGAGGAGGTAAGAGTGGCGAGAACAACCTCTACGCTGTAGAAGGAATGCAATTTGATCGTGGTTATATCTCTCTTACTTTGTCAGACAGA -880

61CHAHOJ - AAAAAAGATAACCAATGCTAGGGATCTTTTCAACATACTGGAGGATGCTATTAGAAGTGGGA - 60
SEQATH60B - GCGAGAAAATGCCGTGAGTTCGACAATTGCAAGTTACTTCTCGTTGACAGAAAATTACCAATGCAAGGATCTTGTGGTGTTTTGGAGGATGCAATTAGAGGGCGGA -990

61CHAHOJ - TATCCTATATTTGTTTATGTCAGAGGATATTGAACAGGAAGCTCTAGCAACTCTAGTGGTGAACAGACTTAGAGGATCAGTGAAGATGTCAGCGCTTAAGGCCCTGGGTT -170
SEQATH60B - TACCCAATATAAATTATGCTGAAGACATTGAACAAGAGCTTTAGCGACTTGTGTTGAAACAGCTTAGAGGTACACTGAAGATAGGAGCTCTCAGAGCTCCAGGGTT -1100

61CHAHOJ - TGGTGAACGCAAGAGCCAGTACCTTGATGACATTGCAATTTGACTGGAGGACTGTAATCAGAGAAGAGGTTGGCCTTACTTTGGACAAGCTGGGAAAGAGGTTCTCG -280
SEQATH60B - TGGAGAGCGCAAGAGCCAATACCTTGATGATATTGCCATTCTAACTGGAGCAACTGTGATCGGTAGGAAAGTGGTCTTTCACTTGACAAGCTGGAAAAGAGGTTCTTG -1210

61CHAHOJ - GATATCGTCCAAGTGGTACTCACAAGGATACAACAACAAATGTTGGTGTGATGGAAGTACCAGGAGGAGTGAACAAGAGGTTGCAAAAATAAGATCCAAATTTGAG -390
SEQATH60B - GTAATGCTTCAAAGGTTGTCCTCAGCAAGGAGACCTCGACCATAGTGGGTGATGGAAGCACACAGGATGCAGTGAAGAAGCGTGTACACAGATTAATAATCTTATGAG -1320

61CHAHOJ - GCTGCAGAGCAAGATTACGAGAGGGAGAAGCTGAATGAGAGAATTGCTAAATGCTGGTGGTGTGGCCGTGATACAGGTTGGTGCACAACTGAAACAGAGCTCAAGGA -500
SEQATH60B - CAAGCGGACGAAGATTACGAGAAGGAGAAAACCTCAATGAGAGAATTGCAAGCTCTCTGGTGGAGTTGCTGTGATTGAGTTGGAGCACAACTGAGACAGAACTCAAAGA -1430

61CHAHOJ - AAAGAAATGAGAGTTGAAGATGCTCTAATGCTACAAGGCACTGTAGAGAAAGSTATTGATGAGGAGTGGTGCACCTCTGCTGAGACTTGATCAAGGTTGGATG -610
SEQATH60B - GAAGAGTTGAGAGTTGAAGATGCTCTAATGCAACAAGGCGCCTGTGGAGGAAAGTATTGTTGTTGGTGGTGGTGCACCTCTGCTTGGTCTTCCAGGTTGATG -1540

61CHAHOJ - CCATCAGAGATAGTCTTGATAATGATGAAGAAAAGTTGGAGCTGATATTGTAAGAAAGAGCTCTTAGTACCCTGAAAATAATGCAAGAATGCTGGTGTCAATGGT -720
SEQATH60B - CCATTAAAGCCACGCTTGATAATGATGAAGAAAAGTTGGAGCGGATATAGTTAAAGAGCACTAAGTACCCTGAAAATAATGCAAGAATGCTGGTGTAAATGGA -1650

61CHAHOJ - AGTGTGTGTCAGTGAAGGATTTGTCAGGACAAATGCAAGATATGGATATAATGCTGCCACCGGAACTATGAGATTTGATGCTGCCGGGATCATTGACCCAAACAAA -830
SEQATH60B - AGTGTAGTTAGCGAGAAGGCTTTTCCACGATAACGTGAAATTCGGTTACAATGCTGCAACCGCAAGTACGAGGATTTGATGGCTGCAGGAATCATCGATCCAACAAA -1760

61CHAHOJ - GGTGGTCAAGTGTGCTTGGACATGCGACTTCAAGTTCAGTTGCGCAAGACCTTCTTAATGTCAGATTGCGTGGTGGTTGAGATAAAGAACTGAGGCGGTACCTGCTGGAAACC -940
SEQATH60B - GGTGTGAGATGTTGCTTGGAAACCGCAGCTTCGGTTGCAAGACTTCTTGTGTCGACTGCGTGGTGGTGGTGGATCAAGGAACCTGAACCAAGTCCCCTCGGCAACC -1870

61CHAHOJ - CCATGGACGCTTCAAGGATATGTTATGAG---AGAGCAACGGAGAAGTTGAGTGAGAACAGAAATTTTTGTTTTTCTCAGAGAAATGCAAGATTTGGAAGGCATTTTATG -1046
SEQATH60B - CAATGGACAACCTCAGGATTAGGATCTAAGTATGGA-CATAGGCTAAAT--ATTGA-AACGGA-TTTATTCTTAGCTTTCGAG-TGTA-AATGT-----ATTG---G -1963

61CHAHOJ - TAGCAACAGGTCCTAAATTTACGATGGAAACCTGTAGTAGTGTGATGAACAGATAGTATGACCCCCATCCCATCCCAATTTTTTTTTAAGTTATTAATAATTTTTT -1156
SEQATH60B - AA-CAAGG-----AAT-----AGTGT-----AATCTGAAAATGA---A-----ATTTGAATTAATCATCA--TATGAATAAATCATGT -2027

61CHAHOJ - TGCTATTTTCTAAAAA----- -1189 (1167 WO/POLY A)
SEQATH60B - TAC-----CTAAAAA----- -2054 (2032 WO/POLY A)

Identity (WO/POLY A) : 864 ( 74.08)
Number of gaps inserted in 61CHAHOJ: 1
Number of gaps inserted in SEQATH60B (all in 3' UTR): 16

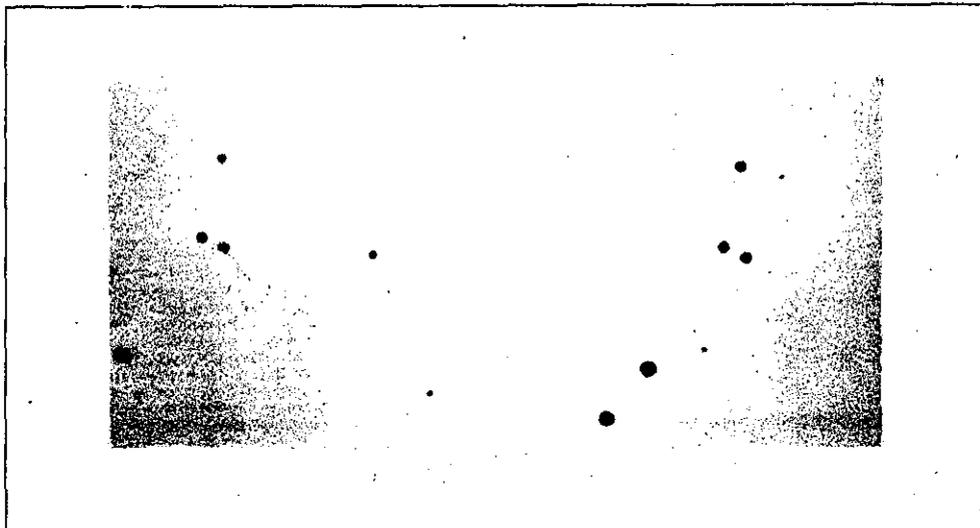
```

**Figura 6.** Alineamiento pareado de la secuencia de nucleótidos de la clona de frijol **chapphas 6.1**, y la secuencia de la clona **b3** de *Arabidopsis thaliana*. Los codones de término y los sitios de poliadenilación están subrayados. El porcentaje global de identidad no incluye las regiones de poly-A. La región de identidad de las clonas se inicia en el nucleótido 931 de la secuencia de *Arabidopsis thaliana*.

### Tamizado de la biblioteca de nódulo utilizando como sonda la clona chapphas-6.1.

Dado que la clona aislada de frijol, corresponde a la secuencia buscada, el inserto fue aislado y usado para tamizar una biblioteca de nódulos de frijol.

La biblioteca de cDNA construida en  $\lambda$  ZAP (Stratagene) a partir de nódulos de frijol de 12 días se plaqueó en 4 cajas de Petri de 150 mm con 20,000 unidades formadoras de placa (pfu), por caja. Como sonda se utilizó la clona aislada de la biblioteca de hoja y las condiciones de hibridación y lavado fueron de severidad media. En la **figura 7**, se muestra el resultado de una de las cajas después del primer tamizado.

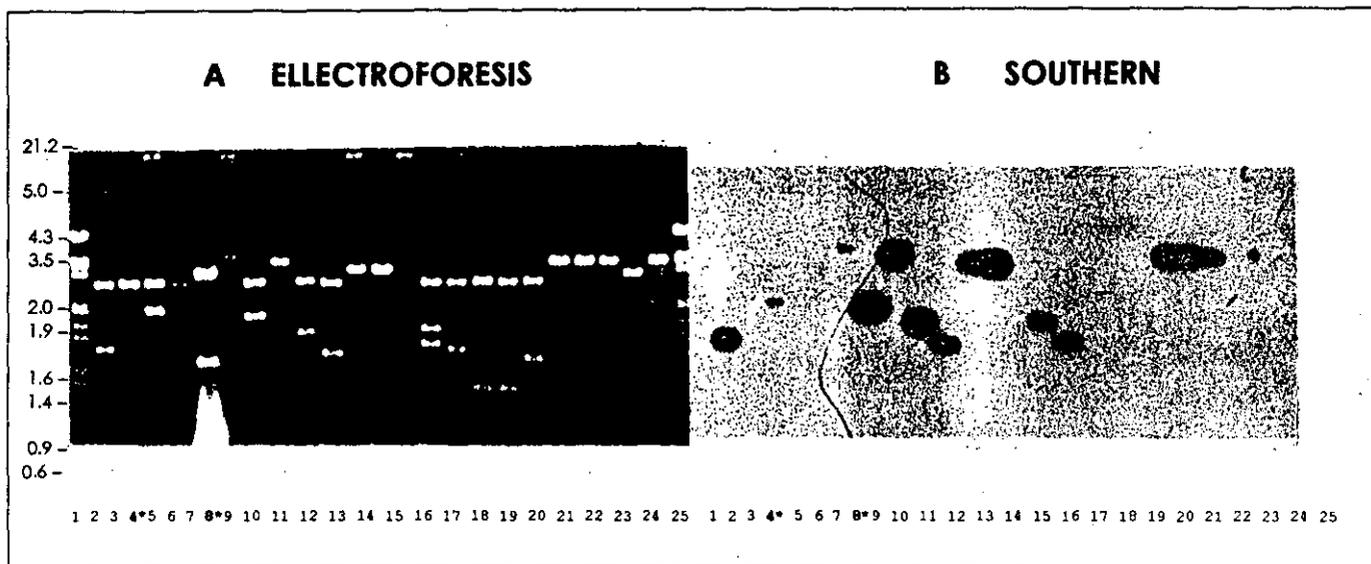


**Figura 7.** Clonas positivas de nódulo de frijol detectadas en el primer tamizado de la biblioteca de cDNA. Como sonda se utilizó la clona completa de cpn60- $\beta$  (chapphas-6.1) de *Phaseolus vulgaris*. Se muestra el filtro que dió 6 señales positivas que se identifican claramente en el filtro y su duplicado.

En total de las cuatro cajas (80,000 pfu totales), se identificaron 19 señales positivas, que dá una relación 1: 4,210. Este resultado sugiere que el nivel de expresión de este gen en nódulos de 12 días es muy similar al de hoja; indicando que al menos en este estado de diferenciación, cpn60- $\beta$  es tan abundante en nódulo como en hoja.

Después del tercer tamizado, las clonas positivas fueron aisladas y convertidas al plásmido Bluescrip SK – por medio del protocolo de escisión *in vivo* con el fago ayudador ExAssist. Una de las clonas no pudo ser recuperada por este protocolo, por lo cual se trabajó con las 18 clonas restantes. Después de aislar a los plásmidos de 18 de las clonas, éstos se restringieron con Eco RI y

Xho I para determinar el tamaño de los insertos. En la **figura 8** se muestra la restricción de las clonas así como la hibridación de tipo Southern. Como se puede apreciar, de las clonas aisladas la mayoría contienen insertos que, o son muy pequeños, o no hibridan con la sonda, o han perdido alguno de los sitios de restricción del "polylinker" e inclusive alguna contiene insertos dobles uno de los cuales hibrida y el otro no (carril 15). Sin embargo, una de las clonas (carril 8\*) designada a partir de aquí **chapnod-4.2**, contiene un inserto de más de 1.8 Kbp que hibrida claramente. Inclusive, comparando el tamaño de este inserto con la clona completa de *Arabidopsis thaliana*, parecería que esta clona pudiera contener la secuencia completa de la chaperonina.

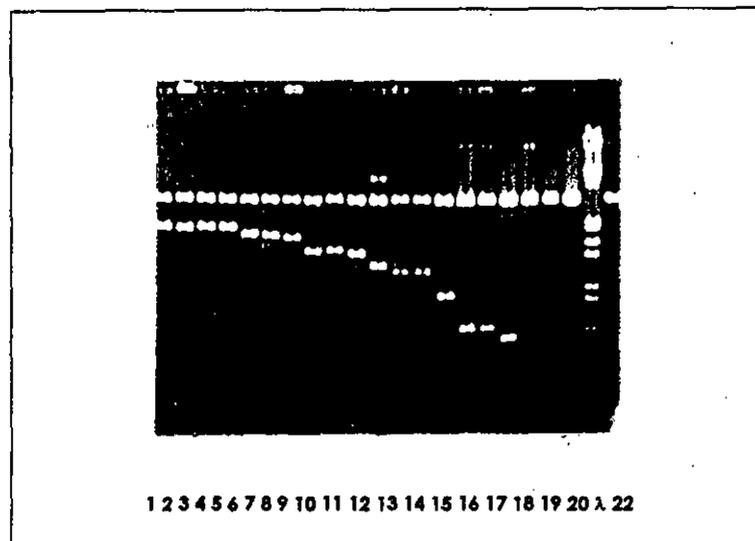


**Figura 8. A)** Gel de agarosa con el DNA de 18 de las clonas aisladas de la biblioteca de nódulo de frijol, digeridas con Eco RI / Xho I. Los carriles 1 y 25, contienen el marcador de peso molecular ( $\lambda$  Hind III / Eco RI). El carril 2 contiene como control positivo, a la restricción de la clona chapphas 6.1. El carril 3 contiene al vector de clonación sin inserto. El carril 4\* contiene como control "heterólogo", a la restricción de la clona completa de *Arabidopsis thaliana*. Los carriles 5 y 6 contienen como control negativo, a la restricción de dos clonas de actina de frijol. Los carriles 7-24, contienen a las restricciones de 18 de las clonas aisladas de la biblioteca de nódulo. En el carril 8\* se encuentra la clona designada **chapnod-4.2**. **B)** Hibridación de tipo Southern. Como sonda se uso el inserto completo de la clona de frijol chapphas 6.1: La hibridación se efectuó a 65° C y los lavados hasta 0.1 x SSC, 0.2% SDS a 65° C.

#### Secuencia y comparación de la clona chapnod-4.2 de nódulo de frijol.

Por el tamaño de la clona de nódulo, para obtener la secuencia completa se optó por generar deleciones con el protocolo de la exonucleasa III (Putney, et al, 1981; Jasin, et al., 1983; Henikoff, 1984). Como sitios de restricción se utilizaron Xba I (rellenado con  $\alpha$ -fionucleótidos) y Eco RI. En la

**figura 9.** se muestran varias subclonas después de efectuar el protocolo de deleciones restringidas con *Apa* I / *Not* I, para mostrar el tamaño de los insertos. De izquierda a derecha se puede apreciar que el tamaño de los insertos disminuye desde la clona completa hasta la deleción con el inserto mas pequeño de alrededor de 125 pares de bases. De éstas, se escogieron para secuenciar 10 subclonas que en promedio difieren en su longitud en 185 pares de bases (carriles 1, 5, 9, 10, 11, 12, 14, 15, 17 y 19).



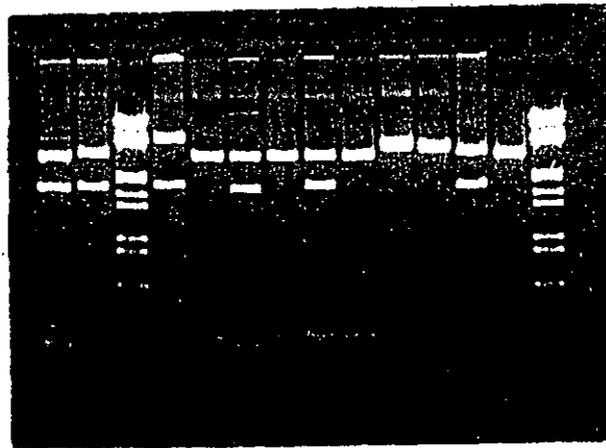
**Figura 9.** Subclonas seleccionadas de **chapnod-4.2**, después de efectuar deleciones progresivas por el protocolo de exonucleasa III. Los carriles 1-20, muestran los tamaños de las clonas después de una doble restricción *Apa* I / *Not* I. El carril 21 contiene a  $\lambda$  *Eco* RI / *Hind* III. En el carril 22 está el vector (pBluescript SK-) sin inserto y restringido con las mismas enzimas.

La secuencia completa de las clonas seleccionadas, incluyendo la original sin deleciones, mostró que esta clona de nódulo, contiene 1,965 pares de bases (no mostrados) y que tanto a nivel de nucleótidos, como de aminoácidos corresponde a una secuencia de *cpn60- $\beta$*  de frijol. Sin embargo el ORF de esta clona, si bien incluye la región codificante completa del péptido maduro, más 37 aminoácidos correspondientes al péptido líder, no llega al codón de inicio y por lo tanto no contiene la secuencia codificante completa que incluiría al péptido líder completo, ni la región 5' no traducida. Dicho de otro modo, la clona está truncada, si bien le faltan pocos nucleótidos. No obstante lo anterior se decidió tamizar nuevamente la biblioteca de nódulo para aislar a la clona completa.

### **Tamizado secundario de la biblioteca de nódulo para aislar una clona completa.**

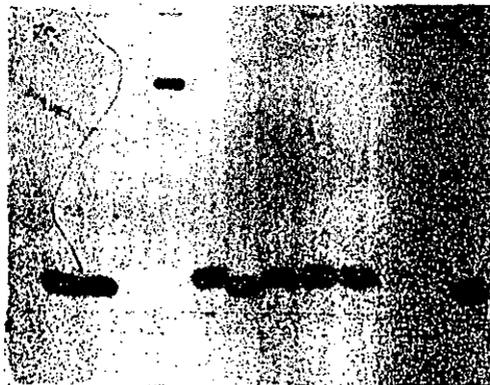
Para aumentar las posibilidades de aislar clonas completas hacia el extremo 5', se utilizó como sonda un fragmento *Eco* RI / *Xho* I de la clona chapnod-4.2, que contiene los primeros 223 pares de bases de esa clona. Se plaquearon 6 cajas de 150mm, con 15,000 pfu cada una de la misma biblioteca de nódulo, se prepararon los filtros y se hibridaron y lavaron en condiciones de alta severidad. Se identificaron 6 señales positivas (datos no mostrados). Después del tercer tamizado, se obtuvieron los plásmidos y se restringieron con *Eco* RI / *Xho* I y se hibridaron con el fragmento 5' de la clona chapnod-4.2. En la **Figura 10**, se puede apreciar que de las clonas aisladas todas hibridan con dicho fragmento. Sin embargo, la clona 1 no libera el fragmento completo, las clonas 2, 4 y 6 si liberan un fragmento 5' pero son clonas truncadas hacia el extremo 3'. La clona 3 libera un fragmento del mismo tamaño que la clona chapnod-4.2, de tal manera que solo la clona 5 corresponde a una clona completa que parece contener nucleótidos adicionales en el extremo 5' que la clona chapnod-4.2.

**A) GEL**



1 2 3 4 5 6 7 8\* 9 10 11 12 13 14

**B) SOUTHERN**

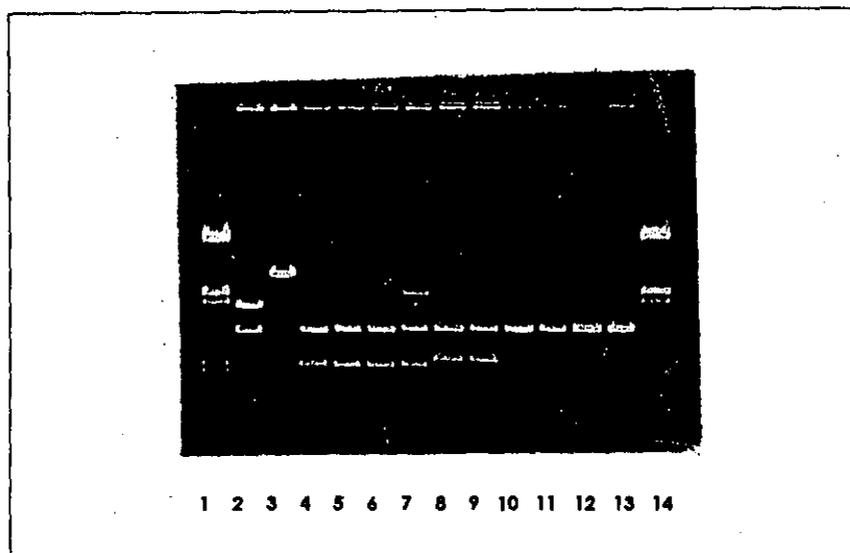


1 2 3 4 5 6 7 8\* 9 10 11 12 13 14

**Figura 10. A)** Electroforesis de agarosa de las clonas aisladas utilizando como sonda el fragmento *Eco* RI / *Xho* I de la clona chapnod-4.2, restringidas con las mismas enzimas. Los carriles 1, 2 y 12, contienen a la clona chapnod-4.2 como control positivo. Los carriles 3 y 14 contienen a  $\lambda$  (*Hind* III / *Eco* RI), como marcador de peso molecular. Los carriles 10, 11 y 13 contienen otras clonas como controles negativos. En los carriles 4-9 se encuentran las clonas aisladas. La clona 5 está en el carril 8\*. **B)** Southern blot usando como sonda el fragmento *Eco* RI / *Xho* I de la clona chapnod-4.2.

## Restricción de las clonas para visualizar el tamaño de los insertos

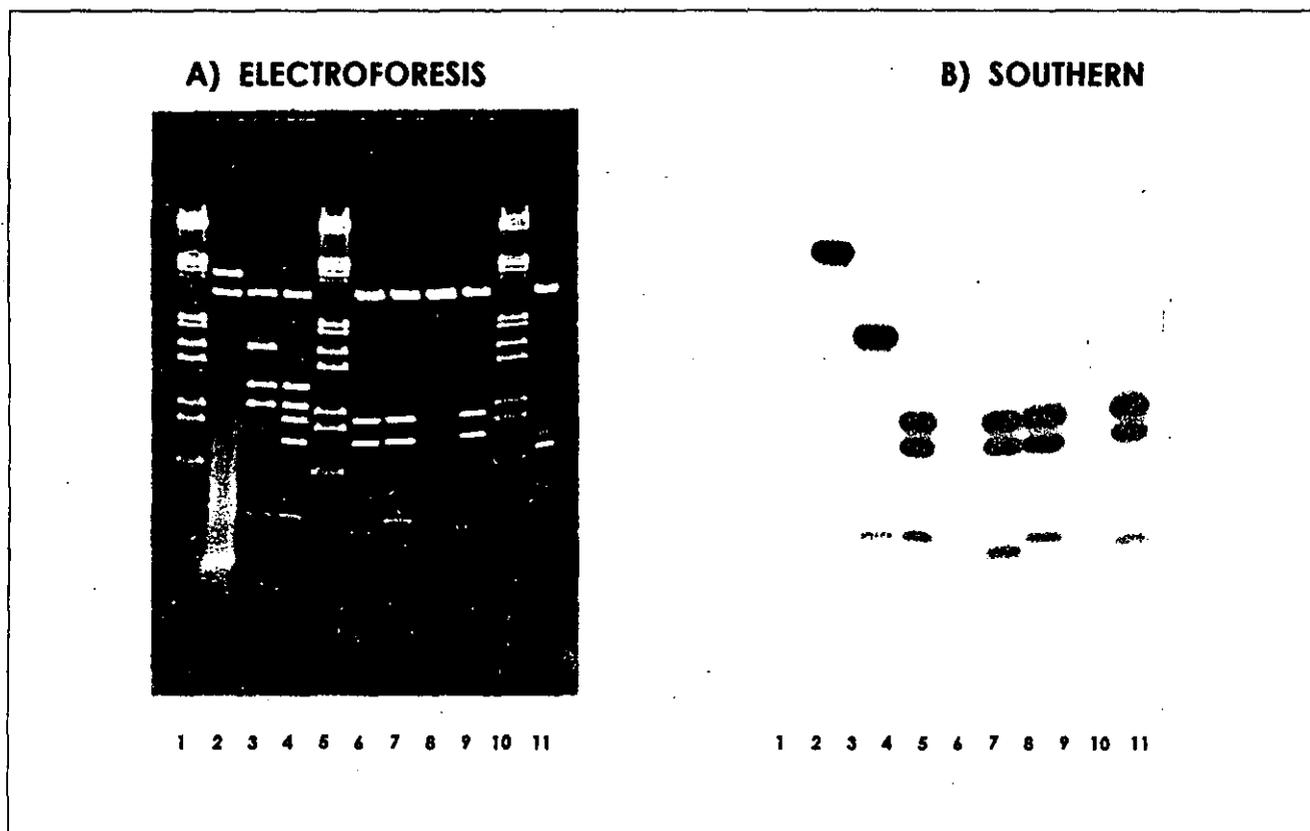
Para confirmar el tamaño de los insertos las clonas 1, 3 y 5, éstas se digirieron con *Sma* I / *Apa* I y *Eco* RI / *Apa* I. Estas enzimas no tienen sitios de corte interno en las clonas sino solo en el polylinker. Como se puede apreciar en la **figura 11**, solo la clona 5 tiene un inserto de mayor tamaño que las clonas 3 y 4.2. Por otro lado la restricción *Sma* I / *Apa* I de la clona 1, libera un inserto de alrededor de 3.9 kbp, mientras que la restricción *Eco* RI / *Apa* I, solo lineariza a la clona. Este resultado sugiere que la clona 1 es un artefacto de clonación.



**Figura 11.** Gel de agarosa de las clonas de nódulo aisladas con el extremo 5' de la clona chapnod-4.2. En todos los casos primero se presenta la restricción *Sma* I / *Apa* I y después la restricción *Eco* RI / *Apa* I. Carriles 1 y 14,  $\lambda$  *Hind* III / *Eco* RI. Carriles 2 y 3, clona 1. Carriles 4 y 5, clona 3. Carriles 6 y 7, clona chapnod-4.2. Carriles 8 y 9, clona 5. Carriles 10 – 13, pBluescript SK+.

Continuando la caracterización de las clonas 1, 3 y 5 se efectuó una restricción múltiple *Sma* I / *Nco* I / *Eco* RV / *Apa* I de estas clonas. (Para visualizar los productos esperados, ver el mapa de restricción de la clona completa más adelante). En el caso de la clona 1 las restricciones fueron secuenciales (ver pie de figura). En la **figura 12**, se muestra el resultado de estas restricciones y el resultado de hibridarlas con el inserto completo de la clona chapnod-4.2. Para la clona chapnod-4.2, se espera, en orden decreciente, los siguientes fragmentos: el vector de clonación en 3 kb (*Sma* I / *Apa* I); la región codificante correspondiente al carboxilo terminal de 880 pb (*Eco* RV / *Nco* I); la región codificante correspondiente al amino terminal de 774 bp (*Nco* I / *Eco* RV); la región 3' UTR de 241 bp (*Nco* I / *Apa* I) y la región hacia el extremo 5' de la clona que codifica parte del péptido líder de 110 bp (*Sma* I / *Nco* I). Como se ve en la **figura 12**, las clonas 3 y 5 muestran un patrón de restricción muy parecido, si no idéntico con la clona chapnod-4.2 y

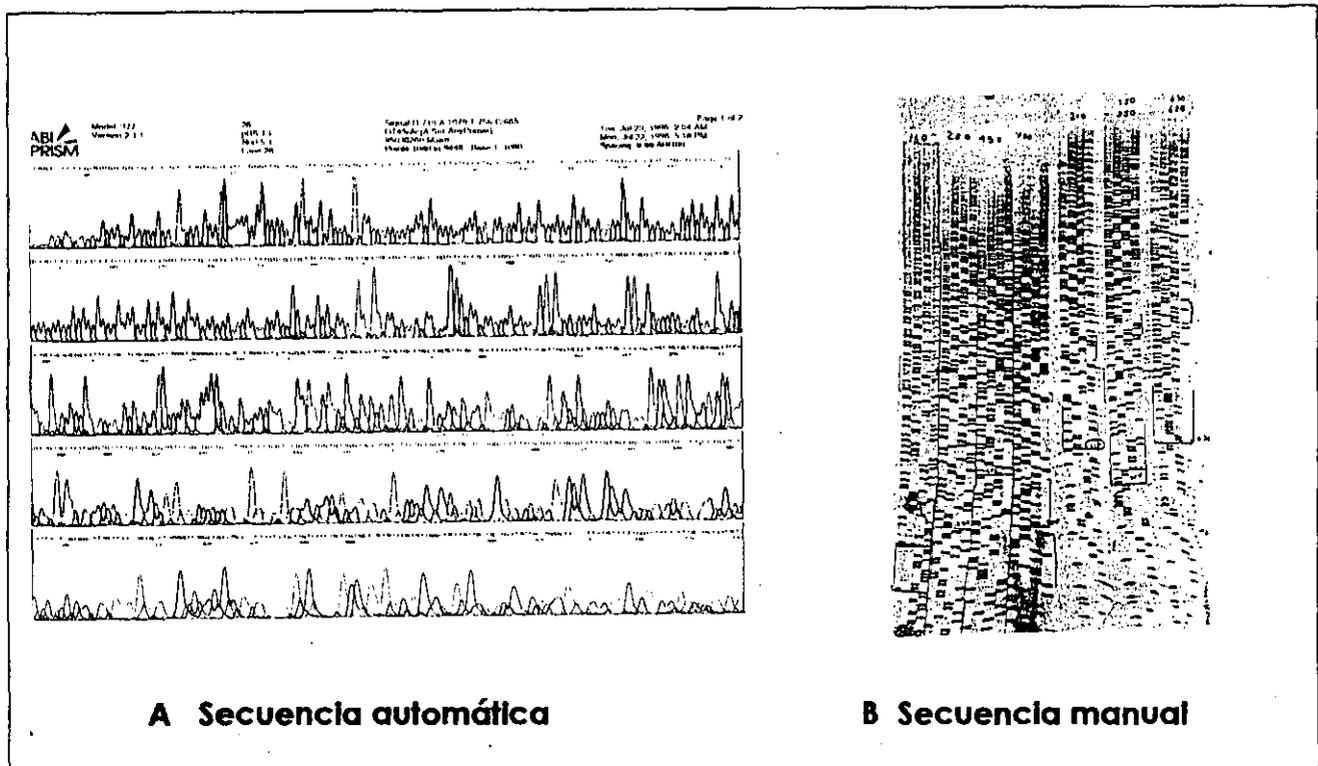
además la hibridación corrobora este parecido. En particular los fragmentos de la región codificante son idénticos. Adicionalmente, si bien se aprecia una pequeña diferencia en el tamaño de la región 3' UTR que es un poco menor en la clona 3 (posiblemente debido a la longitud del Poly-A), **la intensidad de la hibridación sugiere que esta región tiene la misma secuencia en todas las clonas aisladas, (Inclusive la clona 1).** Este resultado indicaría que todas las clonas provienen del mismo gen, ya que justamente ésta es la región más divergente en genes parálogos. Respecto a la clona 1, estos resultados confirman que es un artefacto de clonación, que si bien contiene regiones homólogas a la región codificante y a la región 3' UTR de cpn60- $\beta$ , se aprecian otras dos bandas de 1100 y 1250 bp aproximadamente que no hibridan con la sonda. En la figura presentada, por el tiempo de exposición del filtro (1.5 hrs), en el Southern blot, no se alcanza a apreciar la hibridación del extremo 5' de las clonas. Sin embargo una sobreexposición del filtro (18 hrs), confirma la posición e identidad de este fragmento de 110 bp (datos no mostrados).



**Figura 12. A)** Electroforesis en gel de agarosa al 1.4 % de las clonas 1, 3, y 5 aisladas de una biblioteca de nódulo. Carriles 1, 5 y 10,  $\lambda$  *Hind* III / *Eco* RI como marcador de peso molecular. carriles 2, 3 y 4 clona 1 digerida secuencialmente *Apa* I / *Sma* I; *Apa* / *Nco* I; *Apa* / *Eco* RV, respectivamente. Carriles 6, 7 y 9, clonas 3, 5 y chapnod-4.2 respectivamente, digeridas *Sma* I / *Nco* I / *Eco* RV / *Apa* I. Carriles 8 y 11 controles negativos. **B)** Southern blot utilizando como sonda el inserto completo (*Sma* I / *Apa* I), de la clona chapnod-4.2.

## Secuencia de la clona completa aislada de la biblioteca de nódulo (chapnod-5.1).

Por sus características, la clona 5 (designada a partir de aquí **chapnod-5.1**), fue escogida como el mejor candidato para contener la secuencia completa de la chaperonina de frijol. En la **figura 13**, se muestra un ejemplo de la comparación de la secuenciación manual y la secuenciación automática de los nucleótidos del extremo 5' de la clona chapnod 5.1.



**Figura 13.** Comparación de la secuenciación de la región 5' de la clona completa de nódulo chapnod-5.1. **A)** Impresión de la secuencia automática. **B)** Autorradiografía del gel de secuencia manual. Se muestran diferentes corridas de las deleciones de la clona de nódulo. El orden de cargado del gel es GATC. Las marcas y los números indican las regiones que se corresponden con la numeración de la secuencia automática, si bien esta numeración no corresponde a la numeración final de la clona, que comienza en el nucleótido 60 de la secuencia automática. ∴ Los primeros 59 nucleótidos corresponden al polylinker del vector de clonación.

### Secuencia de la clona 5.1

La secuencia de nucleótidos obtenida por los dos métodos es idéntica a la de la clona chapnod-4.2, excepto que la nueva clona contiene 66 nucleótidos adicionales en el extremo 5'. En esta región, se localiza el ATG de inicio, lo que significa que esta clona sí es una clona completa, que

incluye toda la región codificante (vide *infra*). En la **figura 14**, se muestra la secuencia de nucleótidos y la secuencia derivada de aminoácidos.

```

-----20-APR-1998-----PC/GENE-----
* TRANSLATION OF A NUCLEIC ACID SEQUENCE *
Done on DNA sequence 51CHANOD.
Total number of bases is: 2031.
Analysis done on the complete sequence.
Done on (absolute) phase(s): 1.
Using the Universal genetic code.

GCACGAGCCACAACAAACCCCTCCTAAACCCCTAGCCATGGCTTCCACCTTCTCGCAATGTCTCTCTAGCAAGCTCTCTCTCCGCGCCATCTCTCTCTCCGCTCTCCGCTCCCGC 120
      M A S T F S A M S S Y K L S S S A A I S S F P F S G S R -20

CGTCAATCCAATGGGTCGTCTCTCGCCAGAAGAAGCCGCAAGCCAGGGTCTCGCCATGGCCAAGGAGTTGCACTTCAACAAGGACGGTACTGCAATCAAGAAGCTCCAGAGCGGTGTT 240
R Q S N G V V L A R R S R N A R V S A M A K E L H F N K D G T A I K K L Q S G V 21

AACAAGCTCGCGGACTCGGTGCGGTCAACGCTTGGTCCCAAGGGAAGGAATGTTGTTCTCGAGAGCAAGTATGGTTCGCGCAAAATGGTTAACGACGGTGTCACTGTGCCAAAGAGGTT 360
N K L A D L V A V T L G P K G R N V V L E S K Y G S P K I V N D G V T V A K E V 61

GAGTTGAGGATCCCGTTGAGAATATGGTGTCTAGTGGTGGAGACAGCTGCTGCAAGACCAAGCACTGGCTGGTGTGAACACGACATCTGTTGTTCTAGCTCAAGGCTTATC 480
E L E D P V E N I G A K L V R Q A A A K T N D L A G D G T T T S V V L A Q G L I 101

GCGAAGGTTAAAGTTGTGCTGCCGGTCCCAACCCCTGCTCATCACTCGTGGCATTGAAAGACAGCCAAAGGCTCTGGTGTGAACTTAAAAAGATGTCAAAGAGGTTGAAGAT 600
A E G V K V V A A G A N P V L I T R G I E K T A K A L V S E L K K M S K E V E D 141

AGTGAGTTGGCAGATGTGGCAGCTGTAGTGTCTGGGAACAATTATGAAGTAGGAAATATGATTGCTGAAGCATTGAGTAAAGTTGGTAGAAGGGTGTGTGACCCCTGAAGAGGGAAG 720
S E L A D V A A V S A G N N Y E V G N M I A E A L S K V G R K G V V T L E E G K 181

AGTCCGCAACAGCTTTATATGTTGTGAGGGAATGCAGTTTGCAGCGTACTATTTCCATACTTGTGACTGACAGTGAAGAAATGGCTGTGAGTATGAGAACTGTAAGTTGCTG 840
S A D N S L Y V V E G M Q F D R G Y I S P Y F V T D S E K M A V B Y E N C K L L 221

TTGGTTGACAAAAGATAACCAATGCTAGGGATCTTTTCAACATACTGGAGGATGCTATTAGAAGTGATATCCATATTTGGTTATTGACAGGATATTGAACAGGAAGCTTAGCAACT 960
L V D K K I T N A R D L F N I L E D A I R S G Y P I L V I A E D I E Q E A L A T 261

CTAGTGGTGAACAGACTTAGAGGATCACTGAAGATTGCAGCGCTTAAGGCCCTCGGGTTTGGTGAACGCAAGAGCCAGTACCTTGATGACATTGCAATTTTACTGGAGGACTGTAATC 1080
L V V N R L R G S L K I A A L K A P G F G E R K S Q Y L D D I A I L T G G T V I 301

AGAGAAGAGGTTGGCCTTACTTTGACAAAGCTGGGAAGAGGTTCTCGGATATGGTCCAAGGTGGTACTCACCAAGGATACAACAACATGTTGGTGAAGTACCCAGGAGGCA 1200
R E E V G L T L D K A G K E V L G Y A S K V V L T K D T T T I V G D G S T Q E A 341

GTGAACAGAGAGTTGCACAAATTAAGATCCAAATGAGGCTGCAGAGCAAGATTACGAGAGGAGAGCTGAATGAGAGAAATGCTAAATGTCTGGTGGTGGCGGTGATACAGGTT 1320
V N K R V A Q I K I Q I E A A E Q D Y E R E K L N E R I A K L S G G V A V I Q V 381

GGTGCAAAAAGTGAACAGAGCTCAAGGAAAAGAAATGAGAGTTGAAGATGCTTAAATGCTACAAGGAGCTGTAGAGGAAGGATTGTAGTTGGAGGTGGCTGCACTCTGCTGAGA 1440
A Q T E T E L K E K K L R V E D A L N A T K A A V S E G I V V G G C T L L R 421

CTTGCAATCAAGTGGATGCATCAGAGATAGTCTTGATAATGATGAGAAAAAGTTGGAGCTGATATTGAAAAAGAGCTCTTAGTACCCCTGAAATTAATGCCAAGAATGCTGGT 1560
L A S K V D A I R D S L D N D E E K V G A D I V K R A L S Y P L K L I A K N A G 461

GTCATGGTGTGTTGTCAGTGAGAAGGTTATGTCAGCGACAATGCAAGATATGATATAATGCTGCCACCGGAACTATGAAGATTGATGCTCGCGGGATCATTGACCCAAACAAG 1680
V N G S V V S E K V L S S D N A R Y G Y N A A T G N Y E D L M S A G I I D P T K 501

GTGGTCAGATGTTGCTTGGAAATGACAGCTTCAAGTGGCAAGACCTTCTTAATGTCAGATGGGTGGTGTGTTGAGATAAAGAACTGAGGCGGTACCTGCTGGAACCCCATGGAGCT 1800
V V R C C L E H A A S V A K T F L M S D C V V V E I K E P E A V P A G N P M D A 541

TCAGGATATGGTATGATAGAGAGCAACGGAGAAGTTGAGTGAGAACAGAAATTTTTTTTCTCAGAGAAATGCAGAATTTGGAAGGCATTTTATGTAGCAACAGGTCCAAATTTACGATG 1920
S G Y G M * 546

GAAACCTGTAGTGTGATGAACCCAGATAGTGTGCACCCCCATCCCATCCCAATTTTTTTTAAAGTTATTAAATAATTTTTTGTAAAAAAAAAAAAAAAAAAAAAAAAA 2031

```

**Figura 14.** Secuencia de nucleótidos y secuencia derivada de aminoácidos de la clona de nódulo de frijol **chapnod-5.1**. El codón de inicio, el codón de término y el sitio de poliadenilación están subrayados. Arriba del codón de inicio, se marca en negrita un triplete TAA, en la misma fase de lectura del ORF (nucleótidos 25-27). La secuencia del péptido líder, está subrayada. La numeración de la secuencia derivada de aminoácidos inicia con números negativos en la región correspondiente al péptido líder (47 residuos de aminoácidos) y continúa con numeración positiva a partir de la secuencia correspondiente al péptido maduro (546 residuos de aminoácido).

La secuencia sin incluir el Poly-A, consta de 2,009 pares de bases y contiene un ORF del nucleótido 37 al 1,818 que va del codón de inicio al codón de término. La capacidad codificante de este ORF es de 593 residuos de aminoácido. De éstos, los primeros 47 corresponden al péptido señal y los 546 restantes al péptido maduro. Arriba del codón de inicio (nucleótidos 25-27), existe un triplete TAA en la misma fase de lectura del ORF, lo cual sugiere que la fase de lectura correcta se inicia en el primer ATG localizado en la posición 37 (Joshi, 1987). Adicionalmente, como se muestra en la **figura 15**, la secuencia alrededor del codón de inicio se adapta al contexto reportado para otros genes de plantas dicotiledóneas: AAA(A/C)AATGGCT (Joshi et al., 1997).

plants	c	a	A	(A/C)a	<b>AUG</b>	GCg
monocot	c	(a/c)	(A/G)	(A/C)c	<b>AUG</b>	GCG
dicot	a	a	A	(A/C)a	<b>AUG</b>	GCu
CHAPNOD-5.1	T	A	G	C C	<b>ATG</b>	GCT

**Figura 15.** Comparación de la secuencia alrededor del codón de inicio reportado para plantas en general, monocotiledóneas y dicotiledóneas, con la región equivalente de la clona chapnod-5.1

Por lo demás, la región 5' UTR, no muestra ninguna característica peculiar. Abajo del codón de término se extiende una región 3' UTR de 202 pares de bases donde se ubica la secuencia de poliadenilación canónica AATAAA. Sin embargo, en esta región no se localiza una estructura tallo-asa, similar a la que se ha reportado para otros genes de plantas (Heldecker & Messing, 1986).

## CARACTERÍSTICAS GENERALES DE LA SECUENCIA

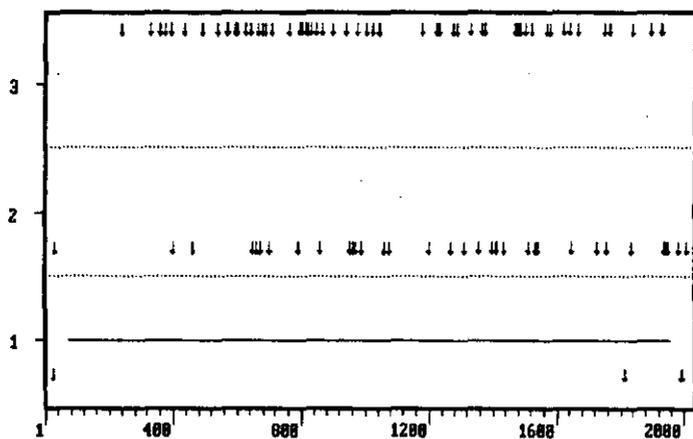
### RNY

A nivel de nucleótidos, la clona chapnod-5.1, muestra una distribución asimétrica de purinas (R) y pirimidinas (Y), particularmente en la región codificante. La gráfica de la **figura 16**, presenta el análisis denominado RNY (Shepherd, 1981; Shepherd, 1990). El análisis se basa en que en una secuencia de nucleótidos al azar, la probabilidad de que un nucleótido particular sea purina (R) o pirimidina (Y) es de 50%. Sin embargo, al analizar las regiones codificantes, Shepherd encontró que esta relación no es simétrica. Al considerar codones de genes específicos, la probabilidad de

que la primera base del codón sea R, la segunda N (R/Y 50%) y la tercera Y, excede los límites del azar. La inferencia de que la distribución RNY, pudiera reflejar un código ancestral, ha sido cuestionada (Jukes, 1996). No obstante, el análisis de Shepherd enfatiza una característica intrínseca de las regiones codificantes, que las distingue de las regiones no-codificantes de los genomas. De hecho, otros métodos para detectar regiones potencialmente codificantes (Staden, 1990), implícitamente incluyen la distribución RNY. Para la clona chapnod-5.1, la figura 16 muestra que en la fase de lectura 1, se localiza el ORF y que en esta fase de lectura se cumple la predicción RNY a lo largo de toda la región codificante. Como era de esperarse, las otras dos fases de lectura contiene un gran número de tripletes TAA, TAG y TGA. En particular, en la región codificante del péptido maduro de la clona la proporción R : Y para la primera, segunda y tercera posición de los codones es: 77.8 : 22.2 / 49.4 : 50.6 / 46.5 : 53.5, respectivamente.

```

Program COD RNY
Search for protein coding region by the Shepherd's method
On DNA sequence 51CHANOD
Total number of basen is: 2031
  
```



Plot of the RNY frame analysis for sequence 51CHANOD.  
On bases 1 to 2031 computed by length of 147 bp, with a step of 3 bp.

**Figura 16.** Análisis de distribución de purinas (R) y pirimidinas (Y) en la secuencia de la clona chapnod-5.1. Los números señalan las tres fases de lectura. Las flechas indican posibles codones de término. La línea horizontal continua marca la fase de lectura en la que se cumple la distribución asimétrica RNY para 1a, 2a y 3a posición respectivamente.

### Utilización de codones

Se ha demostrado, que en diferentes organismos la utilización preferencial de codones sinónimos es una característica específica, relacionada con el nivel de expresión génica (Ikemura, 1985; Li & Luo, 1996). En organismos unicelulares, claramente se puede definir un dialecto del código genético. En organismos pluricelulares, el dialecto es más difícil de definir ya que aparentemente

existen diferencias tejido específico. Sin embargo, se han publicado patrones de utilización preferencial de codones para diversos grupos de organismos incluyendo plantas (**Ausubel, et al., 1992**). La **figura 17**, muestra el análisis de utilización de codones de la porción codificante del péptido maduro de la clona chapnod-5.1. La región correspondiente al péptido líder se omitió del análisis ya que las presiones de selección en esta región podrían ser diferentes. Sin embargo, el análisis de todo el ORF, no muestra diferencias significativas con los datos presentados.

En la primera parte de la **figura 17**, se muestra la tabla de los 64 codones y su frecuencia de utilización en la clona chapnod-5.1. Subrayados se muestran los codones para los que claramente se aprecia una utilización preferencial (11 codones en total). Estos datos sugieren que en esta clona existe un "dialecto" del código genético tal cual ha sido descrito para otros organismos (**Ikemura, 1985**). Este resultado se puede apreciar más claramente en la parte final de la figura, donde se muestra la frecuencia de utilización de los diferentes codones para cada aminoácido. El primer valor indica el número de veces que cada codón es utilizado en la clona. El segundo valor (porcentual), presenta la proporción de utilización de los diferentes codones sinónimos como ha sido descrito después de analizar un total de 71, 408 codones de plantas dicotiledóneas (**Ausubel, et al., 1992**). Vale la pena enfatizar que los valores son particulares para dicotiledóneas y que los valores para monocotiledóneas difieren de éstos. En el caso de la clona chapnod-5.1, se puede apreciar que en la gran mayoría de los casos, tanto los codones preferenciales, como los codones menos utilizados coinciden en esta tabla. Las excepciones se encuentran en los aminoácidos glutamina, prolina y serina y en el codón de término. No obstante, en lo general se puede inferir que esta clona de frijol presenta una utilización preferencial de codones que se enmarca dentro del dialecto general de plantas dicotiledóneas. En total de los 546 codones, 267 corresponden a codones preferenciales de dicotiledóneas, lo cual significa que 21 codones corresponden al 48.9% del total de codones y los 40 codones restantes corresponden al 51.1%.

Se ha encontrado, que la utilización de codones preferenciales y no preferenciales puede mostrar una distribución no aleatoria en las regiones codificantes de algunos genes y que en particular los codones no-preferenciales podrían servir como pausas traduccionales que facilitarían el plegamiento de dominios polipeptídicos particulares (**Cromble, et al., 1992; Netzer & Hartl, 1997**). En el caso de la clona chapnod-5.1, la distribución de estos dos tipos de codones a lo largo de la clona, no muestra un patrón, sino que se observa una alternancia al azar (datos no mostrados). El significado de estos resultados, es que finalmente las pausas traduccionales parecen ser características de proteínas citoplasmáticas de sistemas eucariontes. En el caso de esta chaperonina de frijol, su localización es plastídica (*vide infra*), queriendo decir con esto que su plegamiento dentro del plástido ocurriría de acuerdo a las características de plegamiento de

un sistema de tipo procarionte y por lo tanto dependiente de eventos pos-traduccionales, posiblemente supeditado a la participación de chaperonas moleculares (Gething, 1997).

\*\*\*\*\*  
 \* CODON USAGE STATISTICS ANALYSIS \*  
 \*\*\*\*\*

Done on DNA sequence 51ORF.  
 DE ORF. CPN60 BETA PHASEOLUS. FROM START TO STOP  
 Total number of bases is: 1641.  
 Analysis done on the complete sequence.  
 Using the Universal genetic code.  
 Total number of codons: 547  
 No unidentified codon.  
 Base composition: A= 477 ; T= 412 ; G= 475 ; C= 277 .  
 29% ; 25.1% ; 28.9% ; 16.8%.

-----  
 Normal codon usage table.  
 =====

TTT Phe	2	.3%	TCT Ser	5	.9%	<b>TAT Tyr</b>	<b>10</b>	<b>1.8%</b>	TGT Cys	2	.3%
TTC Phe	4	.7%	TCC Ser	2	.3%	TAC Tyr	5	.9%	TGC Cys	3	.5%
TTA Leu	3	.5%	TCA Ser	6	1%	TAA ---	0	0%	TGA ---	0	0%
<b>TTG Leu</b>	<b>16</b>	<b>2.9%</b>	TCG Ser	1	.1%	TAG ---	1	.1%	TGG Trp	0	0%
<b>CTT Leu</b>	<b>13</b>	<b>2.3%</b>	CCT Pro,	5	.9%	CAT His	1	.1%	CGT Arg	2	.3%
CTC Leu	7	1.2%	CCC Pro	4	.7%	CAC His	1	.1%	CGC Arg	1	.1%
CTA Leu	3	.5%	CCA Pro	2	.3%	CAA Gln	6	1%	CGA Arg	0	0%
CTG Leu	9	1.6%	CCG Pro	1	.1%	CAG Gln	6	1%	CGG Arg	0	0%
<b>ATT Ile</b>	<b>18</b>	<b>3.2%</b>	ACT Thr	10	1.8%	AAT Asn	12	2.1%	<b>AGT Ser</b>	<b>12</b>	<b>2.1%</b>
ATC Ile	8	1.4%	ACC Thr	8	1.4%	AAC Asn	13	2.3%	AGC Ser	4	.7%
ATA Ile	5	.9%	ACA Thr	8	1.4%	AAA Lys	14	2.5%	<b>AGA Arg</b>	<b>14</b>	<b>2.5%</b>
ATG MET	9	1.6%	ACG Thr	2	.3%	<b>AAG Lys</b>	<b>34</b>	<b>6.2%</b>	AGG Arg	3	.5%
<b>GTT Val</b>	<b>32</b>	<b>5.8%</b>	<b>GCT Ala</b>	<b>27</b>	<b>4.9%</b>	<b>GAT Asp</b>	<b>20</b>	<b>3.6%</b>	<b>GGT Gly</b>	<b>22</b>	<b>4%</b>
GTC Val	7	1.2%	GCC Ala	15	2.7%	GAC Asp	12	2.1%	GGC Gly	5	.9%
GTA Val	8	1.4%	GCA Ala	16	2.9%	GAA Glu	20	3.6%	GGA Gly	15	2.7%
GTG Val	15	2.7%	GCG Ala	5	.9%	GAG Glu	28	5.1%	GGG Gly	5	.9%

-----  
 Sorted list by codon frequency.  
 =====

(01) AAG Lys	34	(23) ATG MET	9	(45) TGC Cys	3
(02) GTT Val	32	(24) CTG Leu	9	(46) TTA Leu	3
(03) GAG Glu	28	(25) ACC Thr	8	(47) CTA Leu	3
(04) GCT Ala	27	(26) ACA Thr	8	(48) TGT Cys	2
(05) GGT Gly	22	(27) GTA Val	8	(49) CGT Arg	2
(06) GAT Asp	20	(28) ATC Ile	8	(50) ACG Thr	2
(07) GAA Glu	20	(29) CTC Leu	7	(51) TCC Ser	2
(08) ATT Ile	18	(30) GTC Val	7	(52) CCA Pro	2
(09) GCA Ala	16	(31) TCA Ser	6	(53) TTT Phe	2

(10) TTG	Leu	16	(32) CAG	Gln	6	(54) CGC	Arg	1
(11) GGA	Gly	15	(33) CAA	Gln	6	(55) TCG	Ser	1
(12) GCC	Ala	15	(34) GGC	Gly	5	(56) CCG	Pro	1
(13) GTG	Val	15	(35) GGG	Gly	5	(57) CAT	His	1
(14) AGA	Arg	14	(36) GCG	Ala	5	(58) CAC	His	1
(15) AAA	Lys	14	(37) TCT	Ser	5	(59) TAG	---	1
(16) CTT	Leu	13	(38) CCT	Pro	5	All others are not found.		
(17) AAC	Asn	13	(39) ATA	Ile	5			
(18) AGT	Ser	12	(40) TAC	Tyr	5			
(19) AAT	Asn	12	(41) AGC	Ser	4			
(20) GAC	Asp	12	(42) CCC	Pro	4			
(21) ACT	Thr	10	(43) TTC	Phe	4			
(22) TAT	Tyr	10	(44) AGG	Arg	3			

Codon usage table by amino acid.

<u>cpn60</u>		<u>DICOT</u>		<u>cpn60</u>	<u>DICOT</u>			
Ala	GCT	27	<b>(44.6%)</b>	CTC	7	(16.9%)		
	GCC	15	(21.8%)	CTA	3	(9.8%) *		
	GCA	16	(27.0%)	CTG	9	(11.0%) *	Σ => 51	
	GCG	5	(6.6%) *	Lys	AAA	14	(42.3%)	
Arg	CGT	2	(18.8%)		AAG	34	<b>(57.7%)</b>	Σ => 48
	CGC	1	(9.0%)	MET	ATG	9	(100%)	Σ => 9
	CGA	0	(9.5%)	Phe	TTT	2	(49.5%)	
	CGG	0	(5.0%) *		TTC	4	(50.5%)	Σ => 6
	AGA	14	<b>(32.6%)</b>	Pro	CCT	5	(34.1%)	
	AGG	3	(25.1%)		CCC	4	(15.9%)	
Asn	AAT	12	(45.9%)		CCA	2	<b>(41.9%)</b>	
	AAC	13	<b>(54.1%)</b>		CCG	1	(8.2%) *	Σ => 12
Asp	GAT	20	<b>(60.1%)</b>	Ser	TCT	5	<b>(23.6%)</b>	
	GAC	12	(39.9%)		TCC	2	(17.2%)	
Cys	TGT	2	(49.3%)		TCA	6	(19.7%)	
	TGC	3	(50.7%)		TCG	1	(5.8%) *	
Gln	CAA	6	<b>(60.3%)</b>		AGT	12	(16.4%)	
	CAG	6	(39.7%)		AGC	4	(17.3%)	Σ => 30
Glu	GAA	20	<b>(51.1%)</b>	Thr	ACT	10	<b>(35.0%)</b>	
	GAG	28	(48.9%)		ACC	8	(27.2%)	
Gly	GGT	22	(35.6%)		ACA	8	(29.3%)	
	GGC	5	(14.9%)		ACG	2	(8.5%) *	Σ => 28
	GGA	15	<b>(37.9%)</b>	Trp	TGG	0	(100%)	Σ => 0
	GGG	5	(11.6%) *	Tyr	TAT	10	(46.9%)	
His	CAT	1	<b>(53.0%)</b>		TAC	5	<b>(53.1)</b>	Σ => 15
	CAC	1	(47.0%)	Val	GTT	32	<b>(40.0%)</b>	
Ile	ATT	18	<b>(46.2%)</b>		GTC	7	(17.4%)	
	ATC	8	(34.9%)		GTA	8	(14.0%) *	
	ATA	5	(18.9%) *		GTG	15	(28.6%)	Σ => 62
Leu	TTA	3	(11.1%) *	Stop	TAA	0	(30.6%)	
	TTG	16	<b>(26.1%)</b>		TGA	0	(51.4%)	
	CTT	13	<b>(25.2%)</b>		TAG	1	(18.0%) *	Σ => 1

**Figura 17.** Utilización de codones en la región codificante de la clona chapnod-5.1. La tabla normal de utilización de codones, muestra la frecuencia con la que se usa cada uno de los 64 codones. En **negrita y subrayado**, se indican los codones que claramente muestran una utilización preferencial en la clona. La lista de frecuencia de codones muestra en orden descendente la frecuencia con la que se utiliza cada codón. La tabla de utilización de codones, muestra la frecuencia de utilización de codones sinónimos para cada aminoácido. Los valores porcentuales en paréntesis señalan el valor de utilización de los codones en plantas dicotiledóneas (**Ausubel, et al., 1992**). De estos valores, en **negrita y subrayados** se señalan los codones preferenciales y con asterisco (\*), los codones menos utilizados en dicotiledóneas. La suma se refiere al número de veces que cada aminoácido es codificado por los diferentes codones sinónimos.

### **Características del péptido líder**

Respecto al polipéptido codificado por la clona chapnod-5.1, los primeros 47 residuos de aminoácido codifican un péptido señal. En la **figura 18**, se muestra la comparación de esta secuencia con la de otros péptidos de exporte a cloroplasto y a mitocondria (principalmente de plantas). Si bien no es posible derivar homologías, ya que no existe conservación de la secuencia de aminoácidos, las diferencias entre estos dos tipos de péptidos han sido definidas. En primer lugar, los péptidos de exporte a cloroplasto tienden a ser de mayor tamaño que los de exporte a mitocondria. En segundo lugar, ambos tipos de péptidos son básicos con pocos o ningún residuo ácido. Sin embargo, en el caso de los péptidos de cloroplasto se define un primer dominio ( $\approx 10$  aa's) hacia el extremo amino que no tiene aminoácidos cargados, mientras que lo contrario ocurre en los mitocondriales. Adicionalmente, las secuencias de cloroplasto inician con MA. Sin embargo, la característica más distintiva de los péptidos de tránsito plastídicos es su inusualmente alto contenido de serina (**von Heijne, et al., 1989**). Como se puede apreciar en la **figura 18**, en el caso de la clona chapnod-5.1, todas las características del péptido topogénico corresponden con las características de una secuencia de exporte al plástido si bien con esta información no es posible definir si es de tipo  $\alpha$  o  $\beta$  (pero *vide infra*).

# PEPTIDOS LIDER DE CHAPERONINAS PLASTIDICAS Y MITOCONDRIALES.

## PLASTIDICAS (CLOROPLASTO)

>gi 1223910  <u>FLASTID</u> chaperonin-60 alpha subunit <i>Arabidopsis thaliana</i>	MASANALSSASVLCSSRQSKLGGGNQQQGQRVSYNKRTIRRFVRA	46 aa
>gi 1710807 sp P08926 RUBA_PEA Pisum sativum <u>FLASTID</u> (PEA)RUBISCO SUBUNIT BINDING-PROTEIN ALPHA SUBUNIT PRECURSOR (60 KD CHAPERONIN ALPHA SUBUNIT) (CPN-60 ALPHA)	MASTNALSSSTILRSPTNQAQTSLSKKVKQHGRVNFROKPNRFVKA	47 aa
>gi 81601 pir JF0901 <u>FLASTID</u> chaperonin 60 beta - <i>Arabidopsis thaliana</i>	MASTFTATSSIGSMVAPNGHKSDDKLLSKLSSSSFGRRQSVCPRRRSSSAIVC	54 aa
BRA60BDNA chaperonin 60 beta <i>Brassica napus</i> <u>FLASTID</u> ACC M35600	MASTFTATSSIGSMVAPNAHKTDDKLMNKLSSSSFGRRQNVFPLRRSSPAVVC	53 aa
<i>Phaseolus vulgaris</i> , chapnod 5.1 chaperonin 60 beta (this work) <u>FLASTID</u>	MASTFSAMSSYKLLSSAAISSFPFSGSRRQNGVVLARRSRNARVSA	47 aa
heat shock protein 26.6B. <i>Triticum aestivum</i> ACC 21811 TriMs26 <u>FLASTID</u>	MAAANAPFALSRLS-PAARLPFRAWRAARPAPVWTGRTRPLSVASAA	46 aa
Ara $\alpha$ -MASANALSSASVLCSSRQSKLGGGNQQQGQRVSYNKRTIRRFVRA		46 aa
Pea $\alpha$ -MASTNALSSSTILRSPTNQAQTSLSKKVKQHGRVNFROKPNRFVKA		47 aa
Ara $\beta$ -MASTFTATSSIGSMVAPNGHKSDDKLLSKLSSSSFGRRQSVCPRRRSSSAIVC		54 aa
Bra $\beta$ -MASTFTATSSIGSMVAPNAHKTDDKLMNKLSSSSFGRRQNVFPLRR-SSPAVVC		53 aa
Pha $\beta$ -MASTFSAMSSYKLLSSAAISSFPFSGSRRQNGVVLARRSRNARVSA		47 aa
Tri H -MAAANAPFALSRLS-PAARLPFRAWRAARPAPVWTGRTRPLSVASAA		46 aa

## MITOCONDRIALES

>gi 116229 sp P29197 CH60_ARATH <i>Arabidopsis thaliana</i> <u>MITOCHONDRIAL</u> CHAPERONIN HSP60 PRECURSOR	MYRFASNLASKARIAQNARQVSSRMSWSRNY	31 aa
>gi 2506275 sp P29185 CH61_MAIZE <i>Zea mays</i> <u>MITOCHONDRIAL</u> CHAPERONIN HSP60-1 PRECURSOR	MYRAAASLASKARQAGNSLATRQVGSRLAWSRNY	34 aa
>gi 543992 sp P35480 CH60_BRANA <u>MITOCHONDRIAL</u> <i>Brassica napus</i> CHAPERONIN CPN60 PRECURSOR	MYRLISSIASKARVARNCTSQIGSRLSSTRNY	32 aa
HUMCAPDNA <u>MITOCHONDRIAL</u> ACC. M22383	MLRLPTVFRQMRPVSRLAPHLTRA	25 aa
Ara Mit -MYRFASNLASKARIAQN---ARQVSSRMSWSRNY		31 aa
Zea Mit -MYRAAASLASKARQAGNSLATRQVGSRLAWSRNY		34 aa
Bra Mit -MYRLISSIASKARVARN--CTSQIGSRLSSTRNY		32 aa
Hum Mit -MLRLPTVFRQMRPVSRLAPHLTRA		25 aa

Figura 18. Comparación del péptido líder de la clona chapnod-5.1, con secuencias equivalentes de exporte a cloroplastos y mitocondrias. Se muestra la fuente de cada secuencia y la comparación de las secuencias entre sí.

## Composición de aminoácidos

Por lo que respecta a las características del péptido maduro codificado por la clona de frijol, en la **figura 19**, se presenta la composición de residuos de aminoácido. En general, tiene una composición promedio para una proteína globular *i. e.* los aminoácidos abundantes y los poco abundantes corresponden con el promedio de muchas proteínas. Contiene 5 cisteínas y no tiene triptofano. No se ha reportado que en ninguna chaperonina las cisteínas estén formando enlaces disulfuro y en general la posición de estos residuos no está conservada filogenéticamente (*vide infra*). De hecho, el modelaje molecular de este péptido con la estructura tridimensional publicada para *Escherichia coli*, muestra que la distancia de los residuos de cisteína entre si, no es favorable para la formación de enlaces disulfuro debido a la distancia (datos no mostrados). El peso molecular de esta proteína, sería de 58, 059 Daltons y su punto isoeléctrico de 4.91. Comparativamente con otras chaperoninas no muestra desviaciones, ni particularidades distintivas en todas estas características.

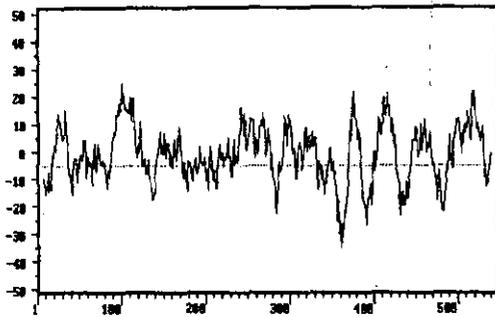


## Hidropatía

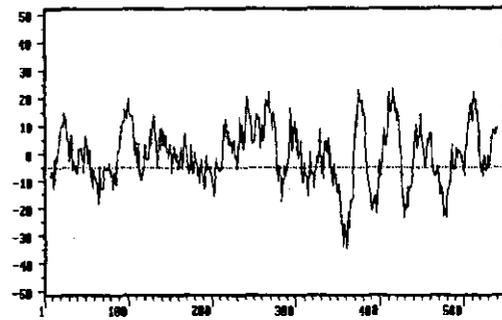
En la **figura 20**, de manera ilustrativa, se muestra comparativamente el perfil de hidrofobicidad del péptido maduro de frijol, de GroEL de una bacteria (*E. coli*), GroEL de una cianobacteria (*Syneccocystis*) y de una HSP60 mitocondrial (*Homo sapiens*), sin considerar el péptido señal. A pesar de los orígenes filogenéticos tan diversos de estas proteínas, se puede apreciar que el grado de parecido en estos patrones es notable. En particular en la región comprendida entre los aminoácidos 360-420, se aprecian dos transiciones muy marcadas de hidrofílico a hidrofóbico, que es prácticamente idéntica en todas las secuencias de chaperoninas (resultados no mostrados). En esta región, se localizan las secuencias altamente conservadas GGVAV (374-378) y GGG (414-416), consideradas las regiones de firma de las chaperoninas (*vide infra*). Sin embargo, es notable que la conservación a nivel del perfil de hidrofobicidad en esta región, sea mayor que la conservación evolutiva a nivel de residuos de aminoácido. Sorprendentemente, estudios de mutagénesis sitio dirigido en aminoácidos puntuales en esta región no parecen afectar drásticamente la función de la chaperonina de *E. coli*, cuando se considera el efecto de mutaciones en otras regiones. Aparentemente el mayor efecto de mutaciones en esta región, consiste en que los polipéptidos resultantes no pueden producir oligómeros adecuados (**Fenton, et al., 1994**). A nivel de estructura tridimensional esta región abarca principalmente la segunda mitad del dominio intermedio (residuos 377-408) y las regiones que lo preceden o siguen en los dominios apical y ecuatorial respectivamente. Dicho de otro modo, desde el punto de vista estructural, esta zona parece solo estar participando en la conexión entre los dos dominios extremos de las chaperoninas (**Braig et al., 1994**), de manera que su relevancia estructural y/o funcional está por establecerse.

**A) chapnod-5.1 (cpn60- $\beta$ /Phaseolus vulgaris)**

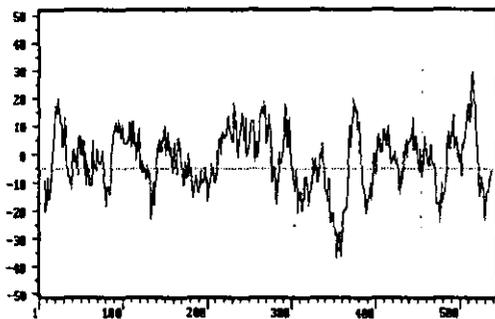
**B) Escherichia coli**



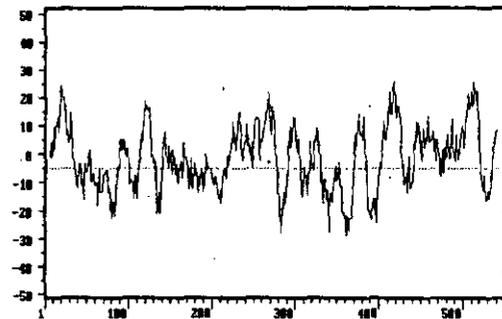
Hydropathic index of CPN60BETA from amino acid 1 to amino acid 546. Computed using an interval of 15 amino acids. (GRANV  $\bar{x}$  = -75).



Hydropathic index of ECOCHELE from amino acid 1 to amino acid 548. Computed using an interval of 15 amino acids. (GRANV  $\bar{x}$  = -93).



Hydropathic index of SYNECHELE from amino acid 1 to amino acid 541. Computed using an interval of 15 amino acids. (GRANV  $\bar{x}$  = -133).



Hydropathic index of PGMHORO from amino acid 1 to amino acid 548. Computed using an interval of 15 amino acids. (GRANV  $\bar{x}$  = -84).

**C) Synechococcus sp. (cianobacteria)**

**D) Homo sapiens (mitochondrial)**

**Figura 20.** Perfiles de hidrofobicidad de chaperoninas seleccionadas (Kyte & Doolittle, 1982). **A)** cpn60- $\beta$  de frijol. Péptido maduro codificado por la clona chapnod-5.1. **B)** GroEL de *Escherichia coli*. **C)** GroEL de la cianobacteria *Synechococcus* sp. strain PCC 7942. **D)** HSP60 (mitochondrial), péptido maduro de *Homo sapiens*. Los perfiles se hicieron tomando una ventana de 15 aminoácidos. Las regiones de mayor hidrofobicidad promedio se ubican arriba del promedio de la línea horizontal punteada.

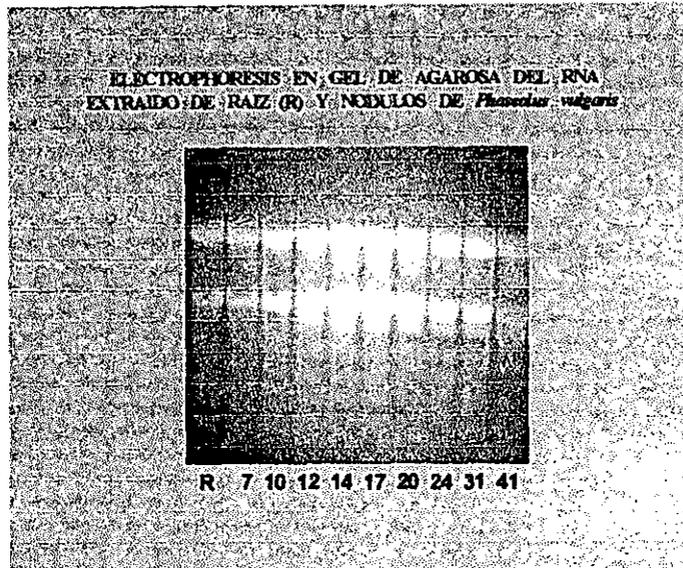
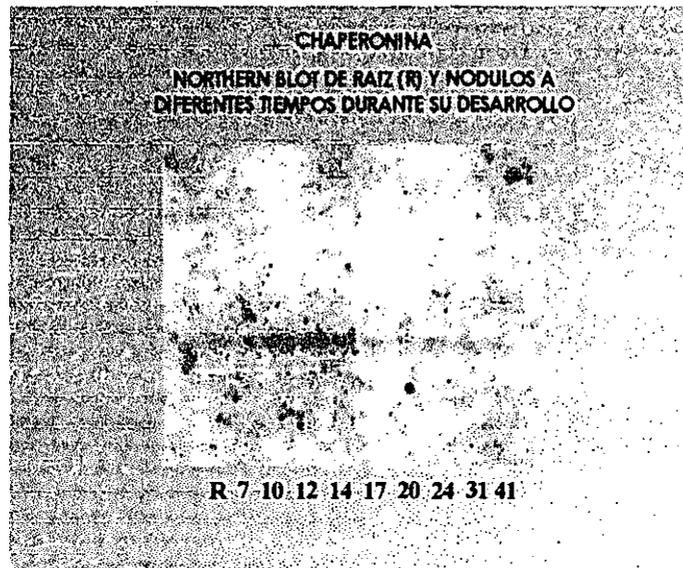
### Comparación de la secuencia derivada de aminoácidos del péptido maduro

La **figura 21**, muestra el alineamiento y el dendograma derivado de la comparación de la chaperonina de frijol y otras chaperoninas seleccionadas que incluyen chaperoninas plastídicas de tipo  $\alpha$  y  $\beta$ , chaperoninas mitocondriales, de bacteria y de cianobacteria. El alineamiento muestra 146 residuos invariantes a lo largo de la escala filogenética. Este valor es muy alto considerando lo diverso de las secuencias. Así mismo se puede apreciar varias regiones a lo largo del alineamiento que definen regiones de alto grado de similitud, en particular la región 360-420. En el dendograma claramente muestra que la secuencia de frijol tiene una relación más cercana con su homólogo plastídico cpn60- $\beta$  de *Arabidopsis thaliana* (89.9% de identidad) y que está más distantemente relacionada con los homólogos de planta cpn60- $\alpha$  y mitocondrial ( $\approx$  54% de identidad). La inferencia obvia, es que la clona aislada de nódulos de frijol codifica para una chaperonina plastídica de tipo  $\beta$ .



### **Determinación del patrón de expresión de la clona chapnod-5.1 (northern blot)**

Se cultivaron plantas de frijol, se inocularon con *Rhizobium tropici* y se aisló el RNA de raíces no infectadas y de nódulos obtenidos de 7 a 41 días después de la infección. En la **Figura 22**, se muestra la electroforesis del RNA aislado así como el northern blot correspondiente. Como se puede apreciar en el gel el RNA aislado no muestra señales de degradación, e inclusive, a partir del día 12 se observa la aparición de los RNAs ribosomales correspondientes a las bacterias, que migran un poco por debajo de los RNA ribosomales de la planta. El northern blot muestra que la señal de raíz y nódulos de 7 días es equiparable, mientras que a partir del día 10 la señal se incrementa hasta un 100% en el día 12, para decrecer después del día 14 disminuyendo progresivamente. Este valor fue obtenido con un analizador de imágenes. Cabe indicar que por las condiciones de alta severidad de la hibridación y de los lavados del filtro, la señal que podría corresponder a la GroE del bacteroide no se detecta en estos experimentos. Estos resultados indican que como respuesta de la planta a la penetración intracelular de los bacteroides, se requiere de la sobreexpresión de estas proteínas de stress. Esta respuesta puede ser explicada debido a que justamente cuando la bacteria se establece como endosimbionte, la célula vegetal procede a sintetizar varias nodulinas tardías muy abundantes (como la leghemoglobina), necesarias para el establecimiento de la función propiamente dicha del nódulo que es la fijación del nitrógeno (**Bennet et al., 1989; Sánchez et al., 1991; Temple et al., 1995; Losick & Kaiser 1997**).

**A****B**

**Figura 22. A)** Gel de agarosa al 1.5 % con formaldehído del RNA aislado de raíces y nódulos simbiotes de *Phaseolus vulgaris* inoculados con *Rhizobium tropici* cepa CIAT 899. Cada carril contiene 5  $\mu$ g de RNA total. R: Raíz. Los números indican el día después de la infección en que fueron colectados los nódulos. Los datos obtenidos con un analizador de imágenes indican que las variaciones en la cantidad de ácido nucleico en los carriles es mínima **B)** Northern blot de la membrana utilizando como sonda a la clona completa aislada de frijol que corresponde a la chaperonina cpn60- $\beta$ . Los resultados obtenidos con un analizador de imágenes indican que la expresión máxima en el día 12 corresponde a un aumento de 100 % respecto a la de raíz o nódulos en el día 7. La hibridación se efectuó a 42° C en 50 % formamida y los lavados a 50° C con 1 x SSC, 0.1 % SSC.

## Comparación de la secuencia de las clonas de hoja y nódulo de frijol

Como quedó indicado anteriormente se aislaron dos clonas de cDNA de cpn60-β de frijol; una de hoja y la otra de nódulo. En la **figura 23** se presenta el alineamiento de las secuencias respectivas a nivel de nucleótidos. Como se puede apreciar, éstas son prácticamente idénticas, excepto por siete nucleótidos antes del poly-A que no se encuentran en la clona de hoja. Esta pequeña diferencia puede ser explicada simplemente por ligeras diferencias de procesamiento del transcrito en cada caso. De cualquier manera, se puede concluir que ambas clonas están codificadas por el mismo gen.

```

*****
* ALIGNMENT OF TWO NUCLEOTIDE SEQUENCES. *
*****
The two sequences to be aligned are:
51CHANOD.
DE CPN60 BETA PHASEOLUS FULL LENGHT (CHAPERONIN VICTOR BEAN)
DE 5'UTR, LEADER PEPTIDE, ORF (MATURE PEOTIDE) 3' UTR, POLY-A
Total number of bases: 2031.

CHAPHOJA. ORIGINAL NAME: 61CHAHOJ. Cpn60β. LEAF Phaseolus vulgaris.
Total number of bases: 1189.
Open gap cost      : 20
Unit gap cost      : 20
The character to show that two aligned residues are identical is ':'

51CHANOD - GCACGAGCCACAACAAACCCTCCTAAACCCCTAGCCATGGCTTCCACCTTCTCGCCAAATGTCCTCTACAAGCTCTCCTCCTCGCGCCCATCTCCTCCTTCCCTTCTC -110
51CHANOD - CCGCTCCCGCCGTCAATCCAATGGGTCGTTCTCGCCAGAAGAAGCCGCAACGCCAGGGTCTCCGCCATGGCCAGGAGTTGCACTTCAACAAGGACGGTACTGCAATCA -220
51CHANOD - AGAAGCTCCAGAGCGGTGTAACAAGCTCGCGGATCTGGTTGCCGTCAAGCTTGGTCCCAAGGGAAGGAATGTTGTTCTCGAGAGCAAGTATGGTTCCGCCGAAAATCGTT -330
51CHANOD - AAGCAGCGTGTCACTGTGGCCAAAGAGTTGAGTTGGAGGATCCCGTTGAGAATATGGTGTCTANGTTGGTGAGACAAGCTGTGCCAAGACCAAGCACTTGGCTGGTGA -440
51CHANOD - TGAACCAACGACATCTGTTGTTCTAGCTCAAGCCCTATCGCGGAAGGTGTAGGTTGTTGCTGCCCGGTGCCAACCCCTGTGCTCATCTACTGTGGCATTGAAAAGCAG -550
51CHANOD - CCAAGGCTCTGGTGTCTGAAGTAAAGATGTCAAAGAGGTTGAAGATAGTGAAGTGGCAGATGTGCCAGCTGTTAGTGTGGGAAACAATTGAAGTAGGAAATATG -560
51CHANOD - ATTGCTGAAGCATTGAGTAAAGTTGGTAGAAAGGTTGTGTGACCCCTGAAGAGGGAAGAGTGGCAGACAACAGTTTATATGTTGTCGAGGGAATGCAGTTTGACCGTGG -770
51CHANOD - CTACATTTCTCCATACTTCGTGACTGACAGTGAGAAAATGGCTGTTGAGTATGAGAAGCTGAAGTTGCTGTTGGTTGACAAAAAGATAACCAATGCTAGGGATCTTTTCA -880
CHAPHOJA - AAAAAATAACCAATGCTAGGGATCTTTTCA -31

51CHANOD - ACATACTGGAGGATGCTATTAGAAGTGGATATCCTATATTGGTTATTGACAGGATATTGAACAGGAAGCTCTAGCAACTCTAGTGGTGAACAGACTTAGAGGATCACTG -990
CHAPHOJA - ACATACTGGAGGATGCTATTAGAAGTGGATATCCTATATTGGTTATTGACAGGATATTGAACAGGAAGCTCTAGCAACTCTAGTGGTGAACAGACTTAGAGGATCACTG -141

51CHANOD - AAGATTGCAGCGCTTAAGGCCCTGGGTTTGGTGAACGCAAGAGCCAGTACCTTGATGACATTGCAATTTTGACTGGAGGTAAGTAAATCAGAGAAGAGGTTGGCCCTAC -1100
CHAPHOJA - AAGATTGCAGCGCTTAAGGCCCTGGGTTTGGTGAACGCAAGAGCCAGTACCTTGATGACATTGCAATTTTGACTGGAGGTAAGTAAATCAGAGAAGAGGTTGGCCCTAC -251

51CHANOD - TTTGGACAAAGCTGGGAAAGAGTTCTCGGATATCGTCCAAGGTGGTACTCACCAAGGATACAACAACAATTGTTGGTATGGAAGTACCCAGGAGGCGAGTGAACAAGA -1210
CHAPHOJA - TTTGGACAAAGCTGGGAAAGAGGTTCTCGGATATCGTCCAAGGTGGTACTCACCAAGGATACAACAACAATTGTTGGTATGGAAGTACCCAGGAGGCGAGTGAACAAGA -361

51CHANOD - GAGTTGCACAAATTAAGATCCAAATTGAGGCTGCAGAGCAAGATTAAGAGAGGGAAGCTGAATGAGAGAATTGCTAAATTTGCTGGTGGTGTGGCCGTTGATACAGGTT -1320
CHAPHOJA - GAGTTGCACAAATTAAGATCCAAATTGAGGCTGCAGAGCAAGATTAAGAGAGGGAAGCTGAATGAGAGAATTGCTAAATTTGCTGGTGGTGTGGCCGTTGATACAGGTT -471

51CHANOD - GGTGCACAACTGAAACAGAGCTCAAGGAAAAGAAATTGAGAGTTGAAGATGCTCTTAATGCTACAAAGCCAGCTGTAGAGGAAGGTAATTGATGTTGGAGGTGGCTGCAC -1430
CHAPHOJA - GGTGCACAACTGAAACAGAGCTCAAGGAAAAGAAATTGAGAGTTGAAGATGCTCTTAATGCTACAAAGCCAGCTGTAGAGGAAGGTAATTGATGTTGGAGGTGGCTGCAC -581

51CHANOD - TCTGCTGAGACTTGCATCAAAGTGGATGCCATCAGAGATAGTCTTGATAATGATGAAGAAAAAGTTGGAGCTGATATTGTAAGAAAGAGCTCTTAGTTACCCCTGAAAT -1540
CHAPHOJA - TCTGCTGAGACTTGCATCAAAGTGGATGCCATCAGAGATAGTCTTGATAATGATGAAGAAAAAGTTGGAGCTGATATTGTAAGAAAGAGCTCTTAGTTACCCCTGAAAT -691

51CHANOD - TAAATGCCAAGAATGCTGGTGTCAATGGTAGTGTGTGTCAGTGAAGAGTATTGTCAGGCAAGATGGAATATAATGCTGCCACCGGAACTATGAAGATTG -1650
CHAPHOJA - TAAATGCCAAGAATGCTGGTGTCAATGGTAGTGTGTGTCAGTGAAGAGTATTGTCAGGCAAGATGGAATATAATGCTGCCACCGGAACTATGAAGATTG -801

51CHANOD - ATGCTCGCGGATCATGACCCAAAGAGGTTGTCAGATGTTGCTTGGAAATGCAAGCTTCAAGTGGCAAGACCTTCTAATGTCAGATTGGCTGGTGTGAGATAAA -1760
CHAPHOJA - ATGCTCGCGGATCATGACCCAAAGAGGTTGTCAGATGTTGCTTGGAAATGCAAGCTTCAAGTGGCAAGACCTTCTAATGTCAGATTGGCTGGTGTGAGATAAA -911

```

```

51CHANOD - GGAACCTGAGGCCGTACCTGCTGGAAACCCCATGGAOCCTTCAGGATATGGTATGTAGAGAGCAACGGAGAAGTTGAGTGAGAACAGAATTTTTGTTTTCTCAGAGAAA -1870
          .....
CHAPHOJA - GGAACCTGAGGCCGTACCTGCTGGAAACCCCATGGAOCCTTCAGGATATGGTATGTAGAGAGCAACGGAGAAGTTGAGTGAGAACAGAATTTTTGTTTTCTCAGAGAAA -1021

51CHANOD - TGCAGAATTTGGAAGGCATTTTATGTAGCAACAGGTCCAAATTTAOCGATGGAAACCTGTAGTAGTGTGATGAACCAGATAGTGATGCACCCCCATCCCATCCCAAATTT -1980
          .....
CHAPHOJA - TGCAGAATTTGGAAGGCATTTTATGTAGCAACAGGTCCAAATTTAOCGATGGAAACCTGTAGTAGTGTGATGAACCAGATAGTGATGCACCCCCATCCCATCCCAAATTT -1131

51CHANOD - TTTTAAAGTTATTAAATAATTTTTTTGCT-----AAAAAAAAAAAAAAAAAAAAAAAA -2031
          .....
CHAPHOJA - TTTTAAAGTTATTAAATAATTTTTTTGCTATTTCTAAAAAAAAAAAAAAAAAAAAAAAA -1189

```

```

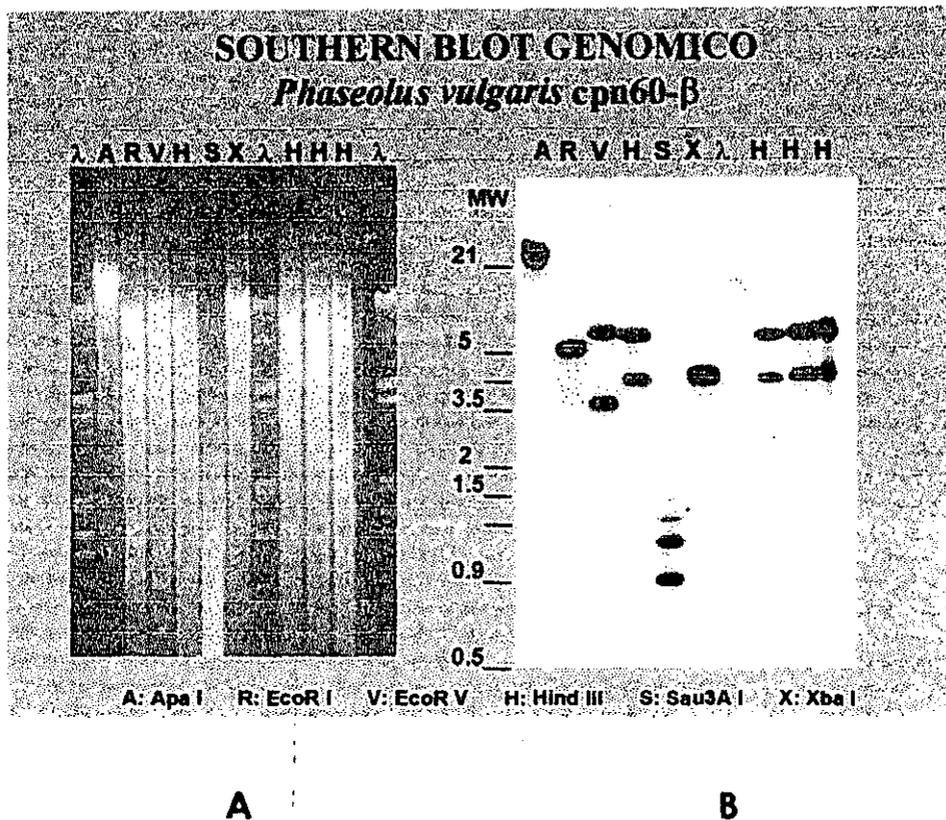
Identity : 1182 ( 99.4%)
Number of gaps inserted in 51CHANOD: 1
Number of gaps inserted in CHAPHOJA: 0

```

**Figura 23.** Alineamiento pareado de la secuencia de nucleótidos de las clonas de frijol **chapnod-5.1** y **chapphas 6.1**, aisladas de nódulos y hojas respectivamente. Los codones de término y los sitios de poliadenilación respectivos están subrayados. Asimismo, en la secuencia de chapnod-5.1, se indican el codón de inicio y el codón GCC, donde termina la región codificante correspondiente al péptido líder. La identidad es menor al 100%, debido a que en la clona de hoja, la poliadenilación se inicia 7 nucleótidos mas allá que en la clona de nódulo. La identidad de las clonas se inicia en el nucleótido 850 de la clona de nódulo.

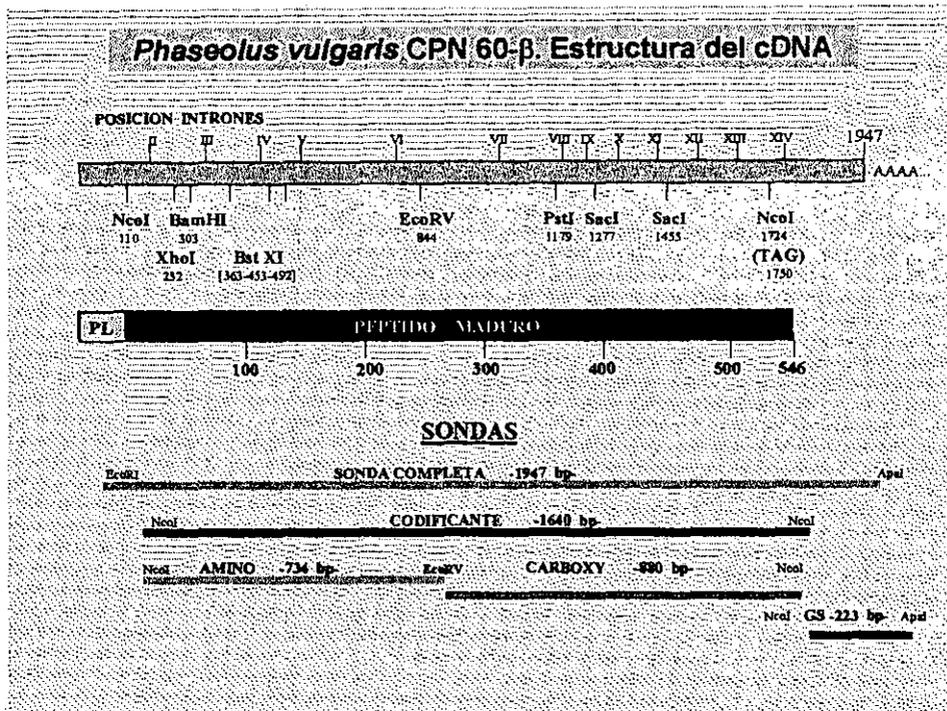
**Caracterización genómica de cpn60-β de Phaseolus vulgaris.**

El resultado anterior no es congruente con nuestra hipótesis original de que a nivel genómico, esta chaperonina podría estar codificada por una familia multigénica, en la que habría miembros individuales que presentarían una expresión mayoritaria de tipo tejido específico como ocurre en otras plantas (Martel et al., 1990; Zabaleta et al., 1992). A fin de obtener datos sobre la estructura genómica del gen o genes correspondientes, se aisló DNA total a partir de hojas y se efectuaron diversas hibridaciones de tipo Southern. En la **figura 24**, se muestra el experimento básico. En particular, se puede apreciar que para ninguna de las enzimas de restricción utilizadas se aprecian muchas bandas. Inclusive, para algunas enzimas se observa una sola banda. Este resultado normalmente se puede interpretar para sugerir que a nivel genómico, cpn60-β no está codificado por una gran familia multigénica (Huang et al., 1996; Guillen et al., 1999). Adicionalmente, en la **figura 24**, los carriles finales presentan la restricción con HindIII de el DNA obtenido de tres plantas adicionales. El propósito de este experimento, fue visualizar mínimamente el posible nivel de polimorfismos que pudieran ocurrir en diferentes alelos presentes en diferentes individuos. Como se puede ver el patrón de restricción con esta enzima para el DNA de cuatro individuos diferentes es idéntico, indicando que a nivel genómico el polimorfismo es bajo.



**Figura 24. A)** Gel de agarosa al 1.2 % de DNA genómico total de *Phaseolus vulgaris*. Cada carril contiene 3  $\mu$ g de DNA digerido con las enzimas que se indican en la figura. Los últimos tres carriles de DNA genómico corresponden a DNA aislado de diferentes individuos. Southern blot genómico de *Phaseolus vulgaris* usando como sonda la clona completa de cpn60- $\beta$ . **B)** Southern blot hibridado con la sonda completa de cpn60- $\beta$  de *Phaseolus vulgaris*. La hibridación se efectuó a 65° C y los lavados a 60° C en 0.1 x SSC, 0.2 % SDS.

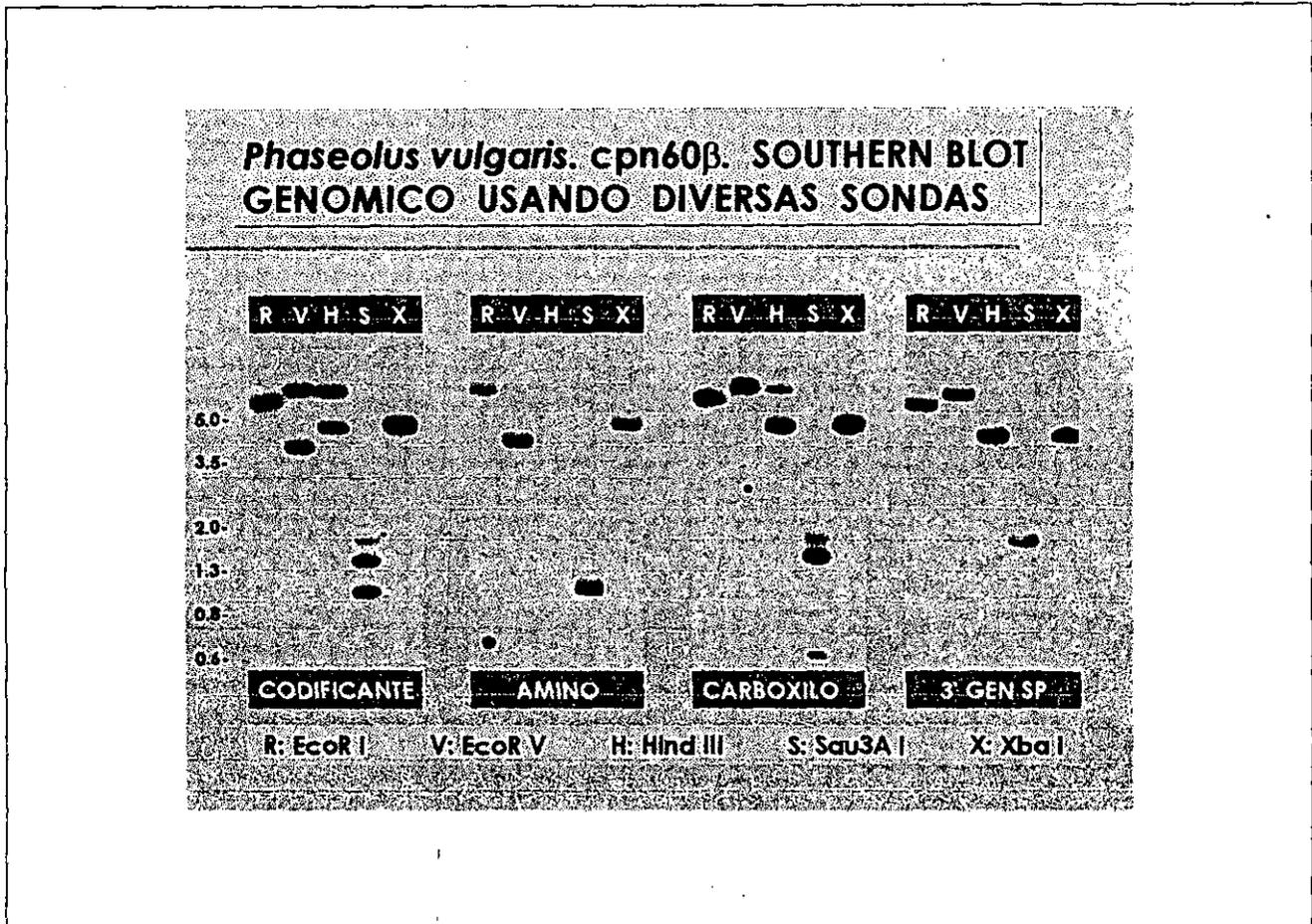
A fin de tener mas elementos sobre la composición genómica de cpn60- $\beta$ , se efectuaron diversas restricciones utilizando enzimas que cortan (*EcoRV* / *Sau3AI*) o que no cortan (*EcoRI* / *HindIII* / *XbaI*) la secuencia nucleotídica del cDNA correspondiente. En la **Figura 25**, se muestra el mapa de restricción así como la estrategia para generar las diversas sondas. Se utilizaron cuatro sondas. La primera corresponde a la región codificante completa; la segunda corresponde a la primera mitad del cDNA en la región que codifica la porción amino terminal del polipéptido; la tercera corresponde a la segunda mitad del cDNA en la región que codifica la porción carboxilo terminal del polipéptido y la cuarta corresponde a la región 3' no traducida del transcrito y que correspondería a lo que se considera la región gen específica.



**Figura 25.** Estructura del cDNA de cpn60-β de *Phaseolus vulgaris*. La primera línea muestra el mapa de restricción de la clona. Los nombres de las enzimas se indican en la parte de debajo de la línea. En la parte superior se marca la posible posición que tendrían los intrones, tomando como base la estructura genómica de uno de los genes de *Arabidopsis thaliana* (Zabaleta et al., 1992). La segunda línea, muestra la correspondencia con el polipéptido codificado, indicando la porción del péptido líder (PL) y la del péptido maduro. En las líneas correspondientes a las sondas se indican las enzimas de restricción utilizadas para liberar las diferentes regiones. EcoRI y ApaI, se encuentran en el polylinker del vector de clonación.

La **figura 26**, muestra el resultado de las hibridaciones respectivas. La hibridación con EcoRI y XbaI, son idénticas en todos los casos, sugiriendo que la restricción del gen siempre genera una única banda que contiene las regiones correspondientes a todas las sondas. La posibilidad de que existieran dos genes no alélicos que presentaran este mismo patrón es baja. La restricción con Sau3AI, no es particularmente útil dado que tanto los exones como los posibles intrones tienen varios sitios para esta enzima. Sin embargo esta restricción también presenta características que excluyen la posibilidad de dos copias parálogas del gen. La restricción con HindIII, es problemática ya que no se detecta señal en el Southern hibridado con la sonda correspondiente al amino terminal. Una explicación podría ser que en la primera mitad del gen, los intrones tuvieran muchos

sitios para esta enzima, lo cual generaría fragmentos demasiado pequeños para ser detectados. La restricción con *EcoRV*, es la más informativa. Esta enzima corta el cDNA justo a la mitad. La hibridación con la región codificante muestra solo dos bandas. Cuando se hibrida con la sonda del amino, una de las bandas desaparece. Por el contrario, en la restricción con las sondas correspondientes al carboxilo terminal y a la región 3' no traducida es la otra banda la que se detecta y no se observa la que hibridaba con el amino. Este resultado sería muy improbable si existieran dos o mas copias no alélicas del gen.



**Figura 26.** Southern blot genómico de *Phaseolus vulgaris* usando como sondas diversas regiones del cDNA de cpn60- $\beta$ . Los nombres de las enzimas se indican en la figura. La hibridación se efectuó a 65° C y los lavados a 60° C en 0.1 x SSC en 0.2 % SDS.

Así, en su conjunto, estos resultados indican que en *Phaseolus vulgaris*, cpn60- $\beta$  es un gen de copia única. La región regulatoria de este único gen le debe permitir variar su expresión en diferentes condiciones. Dicho de otro modo, el promotor debe tener señales que permiten la expresión constitutiva de este gen y por otro lado debe contener secuencias que como en el caso del nódulo le permiten incrementar su expresión de acuerdo a las demandas de la célula.

## CONCLUSIONES FINALES

- Se aislaron clonas de cDNA idénticas tanto de hoja como de nódulos de frijol. La abundancia de las clonas en las bibliotecas correspondientes es similar (0.02 % del mRNA total). Considerando que la biblioteca de nódulo se hizo con RNA de nódulos de 12 días y sabiendo que la expresión de este gen en raíz es baja comparativamente con hoja, se puede concluir que durante esta etapa de diferenciación del nódulo este gen está sobre expresado.
- Comparaciones de secuencia de las clonas indican que éstas corresponden a una chaperonina (cpn60- $\beta$ ), que se localiza en los plástidos. El requerimiento de estos asistentes del plegamiento en estos organelos debe de incrementarse de acuerdo a las demandas impuestas por la endosimbiosis de los *Rhizobia* y a la síntesis de proteínas involucradas en la fijación del nitrógeno.
- Análisis de tipo northern, corroboran que durante la morfogénesis del nódulo, el mRNA de esta chaperonina se incrementa alcanzando un máximo de expresión de hasta 100 % entre los días 10 - 14 después de la infección.
- Análisis de tipo Southern genómico en distintas condiciones indican que muy probablemente en el genoma de *Phaseolus vulgaris*, cpn60- $\beta$  se encuentra representado por un solo gen.

## PERSPECTIVAS

Los resultados obtenidos plantean varias preguntas que serían interesantes de ser abordadas experimentalmente. En primer lugar, se podría determinar a nivel de la proteína si el aumento de expresión es de igual al detectado a nivel del mRNA. Otro aspecto sería establecer si el incremento en la expresión ocurre sólo en la células infectadas, en las no infectadas o en ambas. Aclarado este punto, se debería abordar el estudio de cuales proteínas relevantes en la fijación del nitrógeno de localización plastídica requieren asistencia en el plegamiento. Este es un punto relevante para establecer algún factor limitante en este proceso. Por otro lado, esta el problema de caracterizar las señales regulatorias de este gen para identificar las secuencias responsables de la respuesta constitutiva, la respuesta a la síntesis de la Rubisco y de relevancia a los datos presentados, las señales de regulación durante la nodulación. También, valdría la pena caracterizar en otras leguminosas si también tienen un gen cpn60- $\beta$  de copia única o si es una característica particular de *Phaseolus vulgaris*. En este sentido, se ha demostrado que en el caso de la profilina el frijol solo tiene dos genes mientras que en otras plantas esta proteína esta codificada por una familia multigénica (Guillen et al., 1999), lo cual sugiere que el frijol utiliza mecanismos de regulación postranscripcionales para producir isoformas de proteínas particulares.

## BIBLIOGRAFIA

1. **Adams, M. D., Kerlavage, A. R., Fleischmann, R. D., Fuldner, R. A., Bult, C. J., et al.:** Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* 377 Suppl.: 3-174, 1995.
2. **Aksoy, S.:** Molecular analysis of the endosymbionts of tsetse flies: 16S rDNA locus and over-expression of a chaperonin. *Insect. Mol. Biol.* 4: 23-29, 1995.
3. **An, Y. -Q., McDowell, J. M., Huang, S., McKinney, E. C., Chambliss, S., and Meagher, R. B.:** Strong, constitutive expression of the *Arabidopsis* ACT2/ACT8 actin subclass in vegetative tissues. *Plant J.* 10: 107-121, 1996.
4. **Ananthan, J., Goldberg, A. L., and Voellmy, R.:** Abnormal proteins serve as eukaryotic stress signals and trigger the activation of heat shock genes. *Science* 232: 522-524, 1986.
5. **Anderson, C. R., Jensen, E. O., Llewellyn, D. J., Dennis, E. S., and Peacock, W. J.:** A new hemoglobin gene from soybean: a role for hemoglobin in all plants. *Proc. Natl. Acad. Sci. USA* 93: 5682-5687, 1996.
6. **Anfinsen, C. D.:** Principles that govern the folding of protein chains. *Science* 181: 223-230, 1973.
7. **Ang, D., Ilberek, K., Skowrya, D., Zylicz, M., and Georgopoulos, G.:** Biological role and regulation of the universally conserved heat shock proteins. *J. Biol. Chem.* 266: 24233-24236, 1991.
8. **Antoine, M., and Niessing, J.:** Intron-less globin genes in the insect *Chironomus thummi thummi*. *Nature* 310: 795-798, 1984.
9. **Appleby, C. A.:** Leghemoglobin and *Rhizobium* respiration. *Annu. Rev. Plant Physiol.* 35: 443-478, 1984.
10. **Arredondo-Peter, R., and Escamilla, E.:** A consensus sequence of plant hemoglobins. *Plant Mol. Biol. Rep.* 9: 195-207, 1991.
11. **Ashburner, M., and Bonner, J. J.:** The induction of gene activity in *Drosophila* by heat shock. *Cell* 17: 241-254, 1979.
12. **Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. D., Smith, J. A., and Struhl, K. (Eds):** Short Protocols in Molecular Biology. A Compendium of Methods from Current Protocols in Molecular Biology. Greene Publishing Associates and John Wiley and Sons. New York, 1992.
13. **Avery, O. T., Macleod, C. M., and McCarty, M.:** Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J. Exp. Med.* 79: 137-158, 1944.

14. **Azem, A., Kessel, M., and Goloubinoff, P.:** Characterization of a functional GroEL<sub>14</sub>(GroES<sub>7</sub>)<sub>2</sub> chaperonin hetero-oligomer. *Science* 265: 653-656, 1994.
15. **Baldwin, R. L.:** Matching speed and stability. *Nature* 369: 183-184, 1994.
16. **Baneyx, F., Bertsch, U., Kalbach, C. E., van der Vies, S. M., Soll, J., and Gatenby, A. A.:** Spinach chloroplast cpn21 co-chaperonin possesses two functional domains fused together in a toroidal structure and exhibits nucleotide-dependent binding to plastid chaperonin 60. *J. Biol. Chem.* 270: 10695-10702, 1995.
17. **Bardwell, J. C. A., and Craig, E. A.:** Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 84: 5177-5181, 1987.
18. **Bardwell, J. C. A., and Craig, E. A.:** Ancient heat shock protein is dispensable. *J. Bacteriol.* 170: 2977-2983 1988.
19. **Barracough, R., and Ellis, R. J.:** Protein synthesis in chloroplasts. IX. Assembly of newly-synthesized large subunits into ribulose biphosphate carboxylase in isolated intact pea chloroplasts. *Biochim. Biophys. Acta* 608: 19-31, 1980.
20. **Basler, K., Oesch, B., Scott, M., Westaway, D., Wälchli, M., et al.:** Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene. *Cell* 46: 417-428, 1986.
21. **Becker, J., and Craig, E. A.:** Heat-shock proteins as molecular chaperones. *Eur. J. Biochem.* 219: 11-23, 1994.
22. **Bennet, M. J., Lightfoot, D. A., and Cullimore, J. V.:** cDNA sequences and differential expression of the gene encoding the glutamine synthetase  $\gamma$  polypeptide of *Phaseolus vulgaris* L. *Plant Mol. Biol.* 12: 553-565, 1989.
23. **Bertsch, U., Soll, J., Seetharam, R., and Viitanen, P. V.:** Identification, characterization, and DNA sequence of a functional 'double' groES-like chaperonin from chloroplasts of higher plants. *Proc. Natl. Acad. Sci. USA* 89: 8696-8700, 1992.
24. **Bertsch, U., and Soll, J.:** Functional analysis of isolated cpn10 domains and conserved amino acid residues in spinach chloroplast co-chaperonin by site-directed mutagenesis. *Plant Mol. Biol.* 29: 1039-1055, 1995.
25. **Bessen, R. A.:** Neurodegenerative prion diseases. *Sci. Med.* 3(5): 12-21, 1996.
26. **Beyreuther, K., and Masters, C. L.:** Catching the culprit prion. *Nature* 370: 419-420, 1994.
27. **Bhuvaneward, T. V., Bhagwat, A. A., and Bauer, W. D.:** Transient susceptibility of root cells in four common legumes to nodulation by rhizobia. *Plant Physiol.* 68: 1144-1149, 1981.
28. **Bird, A. P.:** Gene number, noise reduction and biological complexity. *Trends Genet.* 11: 94-100, 1995.
29. **Birnboim, H. C.:** A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods Enzymol.* 100: 243-255, 1983.

30. **Bisseling, T., Franssen, H., van de Sande, K., Pawlowski, K., and Walden, R.:** Recruitment of common development in nodule formation. (Abstract No. 30004). *Plant Physiol.* 111 (Supp.): 10, 1996.
31. **Blake, C.:** Exons-present from the beginning? *Nature* 306: 535-537, 1983
32. **Bloom, M. V., Milos, P., and Roy, H.:** Light-dependent assembly of ribulose-1,5-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. USA* 80: 1013-1017, 1983.
33. **Bochkareva, E. S., Lissin, N. M., and Girshovich, A. S.:** Transient association of newly synthesized unfolded proteins with the heat-shock GroEL protein. *Nature* 336: 254-257, 1988.
34. **Bogusz, D., Appleby, C. A., Landsmann, J., Dennis, E. S., Trinick, M. J., and Peacock, W. J.:** Functioning haemoglobin genes in non-nodulating plants. *Nature* 331: 178-180, 1988.
35. **Bond, U., and Schlesinger, M. J.:** Heat-shock proteins and development. *Adv. Genet.* 24: 1-28, 1987.
36. **Boorstein, W. R., Ziegelhoffer, T., and Craig, E. A.:** Molecular evolution of the HSP70 multigene family. *J. Mol. Evol.* 38: 1-17, 1994.
37. **Bork, P., Sander, C., and Valencia, A.:** An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin and hsp70 heat shock proteins. *Proc. Natl. Acad. Sci. USA* 89: 7290-7294, 1992.
38. **Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B.:** The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371: 578-586, 1994.
39. **Brenner, S.:** That lonesome grail. *Nature* 358: 27-28, 1992.
40. **Brewin, N. J.:** Development of the legume root nodule. *Annu. Rev. Cell Biol.* 7: 191-226, 1991.
41. **Buchmeier, N. A., and Heffron, F.:** Induction of *Salmonella* stress proteins upon infection of macrophages. *Science* 248: 730-732, 1990.
42. **Buchner, J.:** Supervising the fold: Functional principles of molecular chaperones. *FASEB J.* 10: 10-19, 1996.
43. **Büeler, H., Fischer, M., Lang, Y., Bluethmann, H., Lipp, H-P., DeArmond, S. J., Prusiner, S. B., Aguet, M., Weissmann, C.:** Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature* 356: 577-582, 1992.
44. **Bui, E. T., Bradley, P. J., and Johnson, P. J.:** A common evolutionary origin for mitochondria and hydrogenosomes. *Proc. Natl. Acad. Sci. USA* 93: 9651-9656, 1996.
45. **Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., et al.:** Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273: 1058-1073, 1996.
46. **Burston, S. G., Weissman, J. S., Fenton, W. A., and Horwich, A. L.:** Release of both native and non-native proteins from a *cis*-only GroEL ternary complex. *Nature* 383: 96-99, 1996.

47. **Campos, F., Carsollo, C., Kuhn, H., Bisseling, T., Rocha-Sosa, M., and Sánchez, F.:** Characterization and gene expression of nodulin Npv30 from common bean. *Plant Physiol.* 109: 363-370, 1995.
48. **Cloney, L. P., Wu, H. B., and Hemmingsen, S. M.:** Expression of plant chaperonin-60 genes in *Escherichia coli*. *J. Biol. Chem.* 267: 23327-23332, 1992 a.
49. **Cloney, L. P., Bekkaoui, D. R., Wood, M. G., and Hemmingsen, S. M.:** Assessment of plant chaperonin-60 gene function in *Escherichia coli*. *J. Biol. Chem.* 267: 23333-23336, 1992 b.
50. **Cloney, L. P., Bekkaoui, D. R., and Hemmingsen, S. M.:** Co-expression of plastid chaperonin genes and a synthetic plant Rubisco operon in *Escherichia coli*. *Plant Mol. Biol.* 23: 1285-1290, 1993.
51. **Cloney, L. P., Bekkaoui, D. R., Feist, G. L., Lane, W. S., and Hemmingsen, S. M.:** *Brassica napus* Plastid and mitochondrial chaperonin-60 proteins contain multiple distinct polypeptides. *Plant Physiol.* 105: 233-241, 1994.
52. **Cohen, S. N., Chang, A. C. Y., and Hsu, L.:** Nonchromosomal antibiotic resistance in bacteria: Genetic transformation of *Escherichia coli* by R-factor DNA. *Proc. Natl. Acad. Sci. USA* 69: 2110-2114, 1972.
53. **Cole, K. P., Blakeley, S. D., and Dennis, D. T.:** Isolation of a full-length cDNA encoding *Brassica napus* mitochondrial chaperonin-60. *Plant Physiol.* 105: 451, 1994.
54. **Collinge, J., Whittington, M. A., Sidle, K. C. L., Smith, C. J., Palmer, M. S., Clarke, A. R., and Jefferys, J. G. R.:** Prion protein is necessary for normal synaptic function. *Nature* 370: 295-297, 1994.
55. **Cook, D., Dreyer, D., Bonnet, D., Howell, M., Nony, E., and VanderBosch, K.:** Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. *Plant Cell* 7: 43-55, 1995.
56. **Coppo, A., Manzi, A., Pulitzer, J. F., and Takahashi, H.:** Abortive bacteriophage T4 head assembly in mutants of *Escherichia coli*. *J. Mol. Biol.* 76: 61-67, 1973.
57. **Cox, B.:** Prion-like factors in yeast. *Curr. Biol.* 4: 744-748, 1994.
58. **Craig, E. A., Weissman, J. S., and Horwich, A. L.:** Heat shock proteins and molecular chaperones: mediators of protein conformation and turnover in the cell. *Cell* 78: 365-372, 1994.
59. **Crick, F. H. C.:** On protein synthesis. In *Symp. Soc. Exp. Biol., The Biological Replication of Macromolecules*. XII, 138-163, 1958.
60. **Crick, F.:** Central dogma of molecular biology. *Nature* 227: 561-563, 1970.
61. **Crichton, M.:** *The Lost World*. pp. xi-xiv. Ballantine Books, New York, 1995.

62. **Cromble, T., Swaffield, J. C., and Brown, A. J. P.:** Protein folding within the cell is influenced by controlled rates of polypeptide elongation. *J. Mol. Biol.* 228: 7-12, 1992.
63. **Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R., and Georgopoulos, C.:** Purification and properties of the groES morphogenetic protein of *Escherichia coli*. *J. Biol. Chem.* 261: 12414-12419, 1986.
64. **Charles, H., Ishikawa, H., Nardon, P.:** Presence of a protein specific of endocytobiosis (symbionin) in the weevil *Sitophilus*. *C. R. Acad. Sci. III* 318: 35-41, 1995.
65. **Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R., and Saibil, H. R.:** Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature* 371: 261-264, 1994.
66. **Cheng, M. Y., Hartl, F. -U., Martin, J., Pollock, R. A., Kalousek, F., Neupert, W., Hallberg, E. M., Hallberg, R. L., and Horwich, A. L.:** Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. *Nature* 337: 620-625, 1989.
67. **Cheng, M. Y., Hartl, F. -U., and Horwich, A. L.:** The mitochondrial chaperonin hsp60 is required for its own assembly. *Nature* 348: 455-458, 1990.
68. **Chernoff, Y. O., Lindquist, S. L., Ono, B-i, Inge-Vechtomov, S. G., and Lieberman, S. W.:** Role of chaperone protein Hsp104 in propagation of the yeast prion-like factor [*psi*<sup>+</sup>]. *Science* 268: 880-884, 1995.
69. **Chisholm, D.:** A convenient moderate-scale procedure for obtaining DNA from bacteriophage Lambda. *BioTechniques* 7: 21-23, 1989.
70. **Choi, E. Y., Ahn, G. S., and Jeon, K. W.:** Elevated levels of stress proteins associated with bacterial symbiosis in *Amoeba proteus* and soybean root nodule cells. *BioSystems* 25: 205-212, 1991.
71. **Chung, C. T., Niemela, S. L., and Miller, R. H.:** One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl Acad. Sci. USA* 86: 2172-2175, 1989.
72. **Darwin, C.:** *The Origin of Species: By means of natural selection or the preservation of favoured races in the struggle for life.* Mentor Book from New American Library. New York, 1872
73. **de Vries, S., Hoge, H., and Bisseling, T.:** Isolation of total and polysomal RNA from plant tissues. In *Plant Molecular Biology Manual*. S. B. Gelvin, R. A. Schilperoot, and D. P. S. Verma, Eds. Kluwer Academic Publishers, Dordrecht, The Netherlands. Section B, chapter 6: 1-13, 1991.
74. **Dean, C., Pichersky, E., and Dunsmuir, P.:** Structure, evolution and regulation of RbcS genes in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 40: 415-439, 1989.

75. **Dellaporta, S. J., Wood, J., and Hicks, J. B.:** A plant miniprep: Version II. *Plant Mol. Biol. Rep.* 1: 19-21, 1983.
76. **Dorff, R. L., Schoenbach, L., and Gilbert, W.:** How big is the universe of exons? *Science* 250: 1377-1382, 1990.
77. **Dower, W. L., Miller, J. F., and Ragsdale, C. W.:** High efficiency transformation of *E. coli* by high voltage electroporation. *Nucl. Acids Res.* 16: 6127-6145, 1988.
78. **Downie, J. A.:** Signaling strategies for nodulation of legumes by rhizobia. *Trends Microbiol.* 2: 318-324, 1994.
79. **Doyle, J. J., and Doyle, J. L.:** Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15, 1990.
80. **Ecker, J. R.:** The ethylene signal transduction pathway in plants. *Science* 268: 667-675, 1995.
81. **Edenhofer, F., Rieger, R., Famulok, M., Wendler, W., Weiss, S., and Winnacker, E.L.:** Prion protein PrPc interacts with molecular chaperones of the Hsp60 family. *J. Virol.* 70: 4724-4728, 1996.
82. **Ehrhardt, D., Waiss, R., and Long, S. R.:** Calcium spiking in alfalfa root hairs responding to *Rhizobium meliloti* nodulation signals. (Abstract No. 61002). *Plant Physiol.* 111. (Supp.): 25, 1996.
83. **Ellis, R. J.:** The most abundant protein in the world. *Trends. Biochem Sci.* 4: 241-244, 1979.
84. **Ellis, R. J.:** Proteins as molecular chaperones. *Nature* 328-329, 1987.
85. **Ellis, R. J.:** Molecular chaperones: The plant connection. *Science* 250: 954-959, 1990 a.
86. **Ellis, R. J.:** The molecular chaperone concept. *Sem. Cell Biol.* 1: 1-9, 1990 b.
87. **Ellis, R. J., and van der Vies, S. M.:** Molecular chaperones. *Annu. Rev. Biochem.* 60: 321-347, 1991.
88. **Ellis, J.:** Cytosolic chaperonin confirmed. *Nature* 358: 191-192, 1992.
89. **Ellis, R. J.:** Chaperonin duet. *Nature* 366: 213-214, 1993.
90. **Ellis, R. J.:** Chaperoning nascent proteins. *Nature* 370: 96-97, 1994 a.
91. **Ellis R. J.:** Opening and closing the Anfinsen cage. *Curr. Biol.* 4: 633-635, 1994 b.
92. **Ellis, J.:** The "bio" in biochemistry: protein folding inside and outside the cell. *Science* 272: 1448-1449, 1996.
93. **Ellis, R. J., and Hartl, F. -U.:** Protein folding in the cell: competing models of chaperonin function. *FASEB J.* 10: 20-26, 1996.
94. **Engel, A., Hayer-Hartl, M. K., Goldie, K. N., Pfeifer, G., Hegerl, R., Müller, S., da Silva, A. C. R., Baumelster, W., and Hartl, F. U.:** Functional significance of symmetrical versus asymmetrical GroEL-GroES chaperonin complexes. *Science* 269: 832-836, 1995.
95. **Fayet, O., Louarn, J. -M., and Georgopoulos, C.:** Suppression of the *Escherichia coli* dnaA46 mutation by amplification of the *groES* and *groEL* genes. *Mol. Gen. Genet.* 202: 435-445, 1986.

96. **Fayet, O., Zeleghoffer, T., and Georgopoulos, C.:** The *groES* and *groEL* heat shock genes of *Escherichia coli* are essential for bacterial growth at all temperatures. *J. Bacteriol.* 171: 1379-1385, 1989.
97. **Feinberg, A. P., and Vogelstein, B.:** A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132: 6-13, 1983.
98. **Fenton, W. A., Kashi, Y., Furtak, K., and Horwich, A. L.:** Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* 371: 614-619, 1994.
99. **Fernández, R. C., Logan, S. M., Lee, S. H., and Hoffman, P. S.:** Elevated levels of *Legionella pneumophila* stress protein Hsp60 early in infection of human monocytes and L929 cells correlate with virulence. *Infect. Immun.* 64: 1968-1976, 1996.
100. **Fischer, H. M., Babst, M., Kaspar, T., Acuña, G., Arigoni, F., and Hennecke, H.:** One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. *EMBO J.* 12: 2901-2912, 1993.
101. **Fisher, R. F., and Long, S. R.:** *Rhizobium*-plant signal exchange. *Nature* 357: 655-660, 1992.
102. **Flaherty, K. M., DeLuca-Flaherty, C., and McKay, D. B.:** Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346: 623-628, 1990.
103. **Flaherty, K. M., McKay, D. B., Kabsch, W., and Holmes, K. C.:** Similarity of the three-dimensional structures of actin and the ATPase fragment of a 79-kDa heat shock cognate protein. *Proc. Natl. Acad. Sci. USA* 88: 5041-5045, 1991.
104. **Flynn, G. C., Chappell, T. G., and Rothman, J. E.:** Peptide binding and release by proteins implicated as catalysts of protein assembly. *Science* 245: 385-390, 1989.
105. **Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E.:** Peptide-binding specificity of the molecular chaperone BiP. *Nature* 353: 726-730, 1991.
106. **Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., et al.:** The minimal gene complement of *Mycoplasma genitalium*. *Science* 270: 397-403, 1995.
107. **Friedrich, G. A.:** Moving beyond the genome projects. *Nature Biotech.* 14: 1234-1237, 1996.
108. **Frydman, J., Nimmegern, E., Erdjument-Bromage, H., Wall, J. S., Tempst, P., and Hartl, F. U.:** Function in protein folding of TRiC, a cytosolic ring complex containing TCP-1 and structurally related subunits. *EMBO J.* 11: 4767-4778, 1992.
109. **Frydman, J., Nimmegern, E., Ohtsuka, K., and Hartl, F. U.:** Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* 370: 111-117, 1994.
110. **Frydman, J., and Hartl, F. U.:** Principles of chaperone-assisted protein folding: differences between in vivo and in vitro mechanisms. *Science* 272: 1497-1502, 1996.

111. **Fukatsu, T., and Ishikawa, H.:** Occurrence of chaperonin 60 and chaperonin 10 in primary and secondary bacterial symbionts of aphids: Implications for the evolution of an endosymbiotic system in aphids. *J. Mol. Evol.* 36: 568-577, 1993.
112. **Gage, D. J., Bobo, T., and Long, S. R.:** Use of green fluorescent protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*). *J. Bacteriol.* 178: 7159-7166, 1996.
113. **Gantt, J. S., Larson, R. J., Farnham, M. W., Pathirana, S. M., Miller, S. S., and Vance, C. P.:** Aspartate aminotransferase in effective and ineffective alfalfa nodules. Cloning of a cDNA and determination of enzyme activity, protein, and mRNA levels. *Plant Physiol.* 98: 868-878, 1992.
114. **Gao, Y., Thomas, J. O., Chow, R. L., Lee, G. -H., and Cowan, N. J.:** A cytoplasmic chaperonin that catalyzes  $\beta$ -actin folding. *Cell* 69: 1043-1050, 1992.
115. **Gao, Y., Valnberg, I. E., Chow, R. L., and Cowan, N. J.:** Two cofactors and cytoplasmic chaperonin are required for the folding of  $\alpha$ - and  $\beta$ -tubulin. *Mol. Cell Biol.* 13: 2478-2484, 1993.
116. **Gao, Y., Melki, R., Walden, P. D., Lewis, S. A., Ampe, C., Rommelaere, H., Vandekerckhove, J., and Cowan, N. J.:** A novel cochaperonin that modulates the ATPase activity of cytoplasmic chaperonin. *J. Cell Biol.* 125: 989-996, 1994.
117. **Gatenby, A. A., Lubben, T. H., Ahlquist, P., and Keegstra, K.:** Imported large subunits of ribulose biphosphate carboxylase/oxygenase, but not imported  $\beta$ -ATP synthase subunits, are assembled into holoenzyme in isolated chloroplasts. *EMBO J.* 7: 1307-1314, 1988.
118. **Gatenby, A. A., and Ellis, R. J.:** Chaperone function: The assembly of ribulose biphosphate carboxylase-oxygenase. *Annu. Rev. Cell Biol.* 6: 125-149, 1990.
119. **Georgellis, D., Sohlberg, B., Hartl, F. U., and von Gabain, A.:** Identification of GroEL as a constituent of an mRNA-protection complex in *Escherichia coli*. *Mol. Microbiol.* 16: 1259-1268, 1995.
120. **Georgopoulos, C. P., Hendrix, R. W., Kaiser, A. D., and Wood, W. B.:** Role of the host cell in bacteriophage morphogenesis: Effects of a bacterial mutation on T4 head assembly. *Nature New Biol.* 239: 38-42, 1972.
121. **Georgopoulos, C. P., Hendrix, R. W., Casjens, S. R., and Kaiser, A. D.:** Host participation in bacteriophage lambda head assembly. *J. Mol. Biol.* 76: 45-60, 1973.
122. **Georgopoulos, C. P.:** A new bacterial gene (*groPC*) which affects  $\lambda$  DNA replication. *Mol. Gen. Genet.* 151: 35-39, 1977.
123. **Georgopoulos, C., Ang, D., Ilberek, K., and Zyllicz, M.:** Properties of the *Escherichia coli* heat shock proteins and their role in bacteriophage  $\lambda$  growth. In *Stress Proteins in Biology and*

Medicine. R. I. Morimoto, A. Tissières, & C. Georgopoulos Eds. Cold Spring Harbor Laboratory Press. New York. pp. 191-221, 1990.

124. **Gething, M.-J., and Sambrook, J.:** Protein folding in the cell. *Nature* 355: 33-45, 1992.
125. **Gething, M.-J.:** The difference with prokaryotes. *Nature* 388: 329-331, 1997.
126. **Gibbons, D. L., and Horowitz, P. M.:** Exposure of hydrophobic surfaces on the chaperonin GroEL oligomer by protonation or modification of His-401. *J. Biol. Chem.* 270: 7335-7340, 1995.
127. **Girard, M. L., Flores, M., Brom, S., Romero, D., Palacios, R., and Dávila, G.:** Structural complexity of the symbiotic plasmid of *Rhizobium leguminosarum* biovar *phaseoli*. *J. Bacteriol.* 173: 2411-2419, 1991.
128. **Go, M.:** Correlation of DNA exonic regions with protein structural units in haemoglobin. *Nature* 291: 90-92, 1981.
129. **Goff, S. A., and Goldberg, A. L.:** Production of abnormal protein in *E. coli* stimulates transcription of *lon* and other heat shock genes. *Cell* 41: 587-595, 1985.
130. **Goloubinoff, P., Christeller, J. T., Gatenby, A. A., and Lorimer, G. H.:** Reconstitution of active dimeric ribulose biphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. *Nature* 342: 884-889, 1989.
131. **Goodsell, D. S.:** Inside a living cell. *Trends Biochem. Sci.* 16: 203-206, 1991.
132. **Govezensky, D., Greener, T., Segal, G., and Zamir, A.:** Involvement of GroEL in *nif* gene regulation and nitrogenase assembly. *J. Bacteriol.* 173: 6339-6346, 1991.
133. **Govezensky, D., Bochkareva, E. S., Zamir, A., and Girshovich, A. S.:** Chaperonins as potential gene regulatory factors. *In vitro* interaction and solubilization of NifA, the *nif* transcriptional activator, with GroEL. *J. Biol. Chem.* 269: 14003-14006, 1994.
134. **Grimm, R., Speth, V., Gatenby, A. A. and Schäfer, E.:** GroEL-related molecular chaperones are present in the cytosol of oat cells. *FEBS Lett.* 286: 155-158, 1991.
135. **Grimm, R., Donaldson, G. K., van der Vies, S. M. Scäfer, E., and Gatenby, A. A.:** Chaperonin-mediated reconstitution of the phytochrome photoreceptor. *J. Biol. Chem.* 268: 5220-5226, 1993.
136. **Guagliardi, A., Cerchia, L., Barolucci, S., and Rossi, M.:** The chaperonin from the archaeon *Sulfolobus solfataricus* promotes correct refolding and prevents thermal denaturation *in vitro*. *Protein Sci.* 3: 1436-1443, 1994.
137. **Guest, J. R., and Nice, H. M.:** Chromosomal location of the *mop* (*groE*) gene necessary for bacteriophage morphogenesis in *Escherichia coli*. *J. Gen. Microbiol.* 109: 329-333, 1978.
138. **Guillen, G., Valdés-López, V., Noguez, R., Ollvares, J., Rodríguez-Zapata, L. C., Pérez, H., Vidali, L., Villanueva, M. A., and Sánchez, F.:** Profilin in *Phaseolus vulgaris* is encoded by two genes

- (only one expressed in root nodules) but multiple isoforms are generated *in vivo* by phosphorylation on tyrosine residues. *Plant J.* 19: 1-12, 1999.
139. **Gupta, R. S.:** Sequence and structural homology between a mouse T-complex protein TCP-1 and the 'chaperonin' family of bacterial (GroEL, 60-65 kDa heat shock antigen) and eukaryotic proteins. *Biochem. Inter.* 20: 833-841, 1990 a.
  140. **Gupta, R.S.:** Mitochondria, molecular chaperone proteins and the *in vivo* assembly of microtubules. *Trends Biochem. Sci.* 15: 415-418, 1990 b.
  141. **Gupta, R. S.:** Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of eukaryotic cells. *Mol. Microbiol.* 15: 1-11, 1995.
  142. **Gus-Mayer, S., Brunner, H., Schneider-Poetsch, H. A. W., Lottspeich, F., Eckerskorn C., Grimm, R., and Rudiger, W.:** The amino acid sequence previously attributed to a protein kinase or a TCP-1 related molecular chaperone and co-purified with phytochrome is a  $\beta$ -glucosidase. *FEBS Lett.* 347: 51-54, 1994.
  143. **Hanahan, D.:** Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166: 557-580, 1983.
  144. **Hanks, J. F., Tolbert, N. E., and Schubert, K. R.:** Localization of enzymes of ureide biosynthesis in peroxisomes and microsomes of nodules. *Plant Physiol.* 68: 65-69, 1981.
  145. **Hartl, F. -U., and Neupert, W.:** Protein sorting to mitochondria: Evolutionary conservation of folding and assembly. *Science* 247: 930-938, 1990.
  146. **Hartl, F. U., and Martin, J.:** Protein folding in the cell: The role of molecular chaperones hsp70 and hsp60. *Annu. Rev. Biophys. Biomol. Struct.* 21: 293-322, 1992.
  147. **Hartl, F. U.:** Secrets of a double-doughnut. *Nature* 371: 557-559, 1994.
  148. **Hartl, F. U.:** Molecular chaperones in cellular protein folding. *Nature* 381: 571-580, 1996.
  149. **Harutyunyan, E. H., Safonova, T. N., Kuranova, I. P., Popov, A. N., Teplyakov, A. V., Obmolonga, G. V., Rusakov, A. A., Dodson, G. G., Wilson, J. C., et al.:** The structure of deoxy- and oxy-leghaemoglobin from lupin. *J. Mol. Biol.* 251: 104-115, 1995.
  150. **Hayer-Hartl, M. K., Martin, J., and Hartl, F. U.:** Asymmetrical interactions of GroEL and GroES in the ATPase cycle of assisted protein folding. *Science* 269: 836-849, 1995.
  151. **Heldecker, G., and Messing, J.:** Structural analysis of plant genes. *Annu. Rev. Plant Physiol.* 37: 439-466, 1986.
  152. **Hemmingsen, S. M., and Ellis, R. J.:** Purification and properties of ribulosebiphosphate carboxylase large subunit binding protein. *Plant Physiol.* 80: 269-276, 1986.
  153. **Hemmingsen, S. M., Woolford, C., van der Vlies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W., and Ellis, R. J.:** Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* 333: 330-334, 1988.

154. **Hemmingsen, S. M.:** What is a chaperonin? *Nature* 357: 650, 1992.
155. **Hendrick, J.P., and Hartl, F-U.:** Molecular chaperone functions of heat-shock proteins. *Annu. Rev. Biochem.* 62: 349-384, 1993.
156. **Henikoff, S.:** Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28: 351-359, 1984.
157. **Herrmann, B. G., Barlow, D. P., and Lehrach, H.:** A large inverted duplication allows homologous recombination between chromosomes heterozygous for the proximal t complex inversion. *Cell* 48: 813-825, 1987.
158. **Herskowitz, Y.:** Functional inactivation of genes by dominant negative mutations. *Nature* 329: 219-222, 1987.
159. **Hightower, L. E.:** Cultured animal cells exposed in amino acid analoges or puromycin rapidly synthesize several polypeptides. *J. Cell Physiol.* 102: 407-427, 1980.
160. **Hightower, L. E.:** Heat shock, stress proteins, chaperones and proteotoxicity. *Cell* 66: 191-197, 1991.
161. **Hightower, R. C., and Meagher, R. B.:** Divergence and differential expression of soybean actin genes. *EMBO J.* 4: 1-8, 1985.
162. **Hirsch, A. M., and Smith, C. A.:** Effects of *Rhizobium meliloti nif* and *fix* mutants on alfalfa root nodule development. *J. Bacteriol.* 169: 1137-1146, 1987.
163. **Hirsch, A. M., Bhuvaneshwar, T. V., Torrey, J. G., and Bisseling, T.:** Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. *Proc. Natl. Acad. Sci. USA* 86: 1244-1248, 1989.
164. **Hodgkin, J., Plasterk, R. H. A., and Waterston, R. H.:** The nematode *Caenorhabditis elegans* and its genome. *Science* 270: 410-414, 1995.
165. **Holland, S. K., and Blake, C. C. F.:** Proteins, exons and molecular evolution. *BioSystems* 20: 181-206, 1987.
166. **Horowitz, P. M., Hua, S., and Gibbons, D. L.:** Hydrophobic surfaces that are hidden in chaperonin cpn60 can be exposed by formation of assembly-competent monomers or by ionic perturbation of the oligomer. *J. Biol. Chem.* 270: 1535-1542, 1995.
167. **Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, Y. N., and Furtak, K.:** Folding in vivo of bacterial cytoplasmic proteins: Role of GroEL. *Cell* 74: 909-917, 1993.
168. **Horwitz, J.:** Alpha-crystallin can function as a molecular chaperone. *Proc. Natl. Acad. Sci. USA.* 89: 10449-10453, 1992.
169. **Hsiao, K., Baker, H.F., Crow, T. J., Poulter, M., Owen, F., Terwilliger, J. D., Westaway, D., Ott, J., and Prusiner, S. B.:** Linkage of prion protein missense variant to Gerstmann-Sträussler syndrome. *Nature* 338: 342-345, 1989.

170. **Hsiao, K. K. Scott, M., Foster, D., Groth, D. F., DeArmond, S. J., and Prusiner, S. B.:** Spontaneous neurodegeneration in transgenic mice with mutant prion protein. *Science* 250: 1587-1590, 1990.
171. [http://www.yahoo.com/Health/Diseases\\_and\\_Conditions/Creutzfeldt\\_Jakob\\_Disease\\_CJD/](http://www.yahoo.com/Health/Diseases_and_Conditions/Creutzfeldt_Jakob_Disease_CJD/)
172. **Huang, S., McDowell, J. M., Weise, M. J., and Meagher, R. B.:** The *Arabidopsis* profilin gene family. Evidence for an ancient split between constitutive and pollen-specific profilin genes. *Plant Physiol.* 111: 115-126, 1996.
173. **Huey, R. B., and Bennett, A. F.:** Physiological adjustments to fluctuating thermal environments: An ecological and evolutionary perspective. In *Stress Proteins in Biology and Medicine*. R. I. Morimoto, A. Tissières, & C. Georgopoulos Eds. Cold Spring Harbor Laboratory Press, New York. pp. 37-59, 1990.
174. **Hughes, A. L.:** Non linear relationships among evolutionary rates identify regions of functional divergence in heat-shock protein 70 genes. *Mol. Biol. Evol.* 10: 243-255, 1993.
175. **Hunt, J. H., Weaver, A. J., Landry, S. J., Gierasch, L., and Delsenhofer, J.:** The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. *Nature* 379: 37-45, 1996.
176. **Ikemura, T.:** Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* 2: 13-24, 1985.
177. **Ishikawa, H., and Hashimoto, H.:** The molecular biology of symbiotic bacteria of aphididae. *Microbiol. Sci.* 3: 117-120, 1986.
178. **Itoya, M.:** An estimation of the minimal genome size required for life. *FEBS Lett.* 362: 257-260, 1995.
179. **Jacob, F.:** *La Logique du Vivant. Une histoire de l'hérédité.* Éditions Gallimard. Paris, 1970.
180. **Jacobsen-Lyon, K., Jensen, E. O., Jörgensen, J. E., Marcker, K. A., Peacock, W. J., and Dennis, E. S.:** Symbiotic and nonsymbiotic hemoglobin genes of *Casuarina glauca*. *Plant Cell* 7: 213-223, 1995.
181. **Jakob, U., Gaestel, M., Engel, K., and Buchner, J.:** Small heat shock proteins are molecular chaperones. *J. Biol. Chem.* 268: 1517-1520, 1993.
182. **Jakob, U., and Buchner, J.:** Assisting spontaneity-the role of Hsp90 and small Hsps as molecular chaperones. *Trends Biochem. Sci.* 19: 205-211, 1994.
183. **Jakob, U., Lille, H., Meyer, Y., and Buchner, J.:** Transient interactions of Hsp90 with early unfolding intermediates of citrate synthase. Implications for heat shock *in vivo*. *J. Biol. Chem.* 270: 7288-7294, 1995.
184. **Jasln, M., Regan, L., and Scimmel, P.:** Modular arrangement of functional domains along the sequence of an aminoacyl tRNA synthetase. *Nature* 306: 441-447, 1983.

185. **Jindal, S., Dudani, A. K., Singh, B., Harley, C. B., and Gupta, R. S.:** Primary structure of a human mitochondrial protein homologous to the bacterial and plant chaperonins and to the 65-kilodalton mycobacterial antigen. *Mol. Cell. Biol.* 9: 2279-2283, 1989.
186. **Joshi, C. P.:** An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucl. Acid. Res.* 15: 6643-6653, 1987.
187. **Joshi, C. P., Zhou, H., Huang, X., and Chiang, V. L.:** Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* 35: 993-1001, 1997.
188. **Jukes, T. H.:** On the prevalence of certain codons ("RNY") in genes for proteins. *J. Mol. Evol.* 42: 377-381, 1996.
189. **Kahn, P.:** From genome to proteome: Looking at the cell's proteins. *Science* 270: 369-370, 1995.
190. **Kakeda, K., and Ishikawa, H.:** Molecular chaperon produced by an intracellular symbiont. *J. Biochem.* 110: 583-587, 1991.
191. **Kardar, M.:** Which came first, protein sequence or structure? *Science* 273: 610, 1996.
192. **Kellin, F. R. S., and Smith, J. S.:** Hemoglobin and nitrogen fixation in the root nodules of leguminous plants. *Nature* 159: 692-694, 1947.
193. **Kelley, W. L., and Georgopoulos, C.:** Chaperones and protein folding. *Curr. Op. Cell Biol.* 4: 984-991, 1992.
194. **Kim, J., and Rees, D. C.:** Structural models for metal centers in the nitrogenase molybdenum-iron protein. *Science* 257: 1677-1682, 1992.
195. **Kim, S., Willison, K. R., and Horwich, A. L.:** Cytosolic chaperonin subunits have a conserved ATPase domain but diverged polypeptide-binding domains. *Trends Biochem. Sci.* 19: 543-548, 1994.
196. **Kislauskis, E. H., Li, Z., Singer, R. H., and Taneja, K. L.:** Isoform-specific 3'-untranslated sequences sort  $\alpha$ -cardiac and  $\beta$ -cytoplasmic actin messenger RNAs to different cytoplasmic compartments. *J. Cell Biol.* 123: 165-172, 1993.
197. **Kluger, M. J.:** The febrile response. In *Stress Proteins in Biology and Medicine*. R. I. Morimoto, A. Tissières, & C. Georgopoulos Eds. Cold Spring Harbor Laboratory Press. New York. pp. 61-78, 1990.
198. **Knapp, S., Schmidt-Krey, Y., Herbert, H., Bergman, T., Jornvall, H., and Ladenstein, R.:** The molecular chaperonin TF55 from archaeon *Sulfolobus solfataricus*. A biochemical and structural characterization. *J. Mol. Biol.* 242: 397-407, 1994.
199. **Kocisko, D. A., Come, J. H., Priola, S. A., Chesebro, B., Raymond, G. J., Lansbury, P. T., and Caughey, B.:** Cell-free formation of protease-resistant prion protein. *Nature* 370: 471-474, 1994.
200. **Koenen, M.:** Recovery of DNA from agarose gels using liquid nitrogen. *Trends Genet.* 5: 137, 1989.

201. **Köhler, S., Delwiche, C. F., Denny, P. W., Tilney, L. G., Webster, P., Wilson, R. J. M., Palmer, J. D., and Roos, D. S.:** A plastid of probable green algal origin in apicomplexan parasites. *Science* 275: 1485-1489, 1997.
202. **Kong, T. H., Coates, A. R. M., Butcher, P. D., Hickman, C. J., and Shinnick, T. M.:** *Mycobacterium tuberculosis* expresses two chaperonin-60 homologs. *Proc. Natl. Acad. Sci. USA* 90: 2608-2612, 1993.
203. **Koonin, E. V., Mushegian, A. R., and Rudd, K. E.:** Sequencing and analysis of bacterial genomes. *Curr. Biol.* 6: 404-416, 1996.
204. **Kouchi, H., and Hata, S.:** Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol. Gen. Genet.* 238: 106-119, 1993.
205. **Kubota, H., Willison, K., Ashworth, A., Nozaki, M., Miyamoto, H., Yamamoto, H., Matsushiro, A., and Morita, T.:** Structure and expression of the gene encoding mouse *t*-complex polypeptide (*Tcp-1*). *Gene* 120: 207-215, 1992.
206. **Kubota, H., Hynes, G., Carne, A., Ashworth, A., and Willison, K.:** Identification of six *Tcp-1*-related genes encoding divergent subunits of the TCP-1-containing chaperonin. *Curr. Biol.* 4: 89-99, 1994.
207. **Kubota, H., Hynes, G., and Willison, K.:** The chaperonin containing *t*-complex polypeptide 1 (TCP-1). Multisubunit machinery assisting in protein folding and assembly in the eukaryotic cytosol. *Eur. J. Biochem.* 230: 3-16, 1995.
208. **Kuhlemeier, C., Green, P. J., and Chua, N.-H.:** Regulation of gene expression in higher plants. *Ann. Rev. Plant Physiol.* 38: 221-257, 1987.
209. **Kuwajima, K.:** The molten globule state of  $\alpha$ -lactalbumin. *FASEB J.* 10: 102-109, 1996.
210. **Kyte, J., and Doolittle, R. F.:** A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157: 105-132, 1982.
211. **Lakhotia, S. C., and Singh, B. N.:** Synthesis of a ubiquitously present new HSP60 family protein is enhanced by heat shock only in the Malpighian tubules of *Drosophila*. *Experientia* 52: 751-756, 1996.
212. **Lander, E. S.:** The new genomics: Global views of biology. *Science* 274: 536-539, 1996.
213. **Landry, S. J., and Gierasch, L. M.:** The chaperonin GroEL binds a polypeptide in an  $\alpha$ -helical conformation. *Biochemistry* 30: 7359-7362, 1991.
214. **Landry, S. J., Jordan, R., McMacken, R., and Gierasch, L. M.:** Different conformations for the same polypeptide bound to chaperones DnaK and GroEL. *Nature* 355: 455-457, 1992.
215. **Landry, S. J., and Gierasch, L. M.:** Polypeptide interactions with molecular chaperones and their relationship to in vivo protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 23: 645-669, 1994.

216. **Lanfranchi, G., Pallavicini, A., Lavender, P., and Valle, G.:** Ancestral hemoglobin switching in lampreys. *Dev. Biol.* 164: 402-408, 1994.
217. **Langer, T., Pfeifer, G., Martin, J., Baumeister, W., and Hartl, F.- U.:** Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. *EMBO J.* 11: 4757-4765, 1992 a.
218. **Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K., and Hartl, F. U.:** Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* 356: 693-689, 1992 b.
219. **Lara, M., Porta, H., Padilla, J. E., Folch, J., and Sánchez, F.:** Heterogeneity of glutamine synthetase polypeptides in *Phaseolus vulgaris* L. *Plant Physiol.* 76: 1019-1023, 1984.
220. **Laskey, R. A., Honda, B. M., Mills, A. D., and Finch, J. T.:** Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature* 275: 416-420, 1978.
221. **Lee, G. J., Basha, E., Suzuki, T., Waters, E. R., Krawitz, D., Azpiroz, R., Wehmeyer, N., and Vierling, E.:** Structure and molecular chaperone activity of small heat shock proteins. (Abstract 21001). *Plant Physiol.* 111 (Supp.): 18, 1996.
222. **Lehninger, A. L.:** Principles of Biochemistry. Worth Publishers, Inc. New York, 1982.
223. **Lesk, A. M., and Chothia, C.:** How different aminoacid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins. *J. Mol. Biol.* 136: 225-270, 1980.
224. **Lewin, B.:** Genes V. Chapter 22, pp 657-676, Oxford University Press, Oxford, 1994.
225. **Lewin, R.:** Surprise finding with insect globin genes. *Science* 226: 328, 1984.
226. **Lewis, V., Hynes, G. M., Zheng, D., Salbil, H., and Willison, K.:** T-complex polypeptide-1 is a subunit of a heteromeric particle in the eukaryotic cytosol. *Nature* 358: 249-252, 1992.
227. **Li, H., and Luo, L.:** The relation between codon usage, base correlation and gene expression level in *Escherichia coli* and yeast. *J. theor. Biol.* 181: 111-124, 1996.
228. **Li, H., Helling, R., Tang, C., and Wingreen, N.:** Emergence of preferred structures in a simple model of protein folding. *Science* 273: 666-669, 1996.
229. **Liang, P., and MacRae, T. H.:** Molecular chaperones and the cytoskeleton. *J. Cell Sci.* 110: 1431-1440, 1997.
230. **Lin, Z., Schwarz, F. P., and Eisenstein, E.:** The hydrophobic nature of GroEL-substrate binding. *J. Biol. Chem.* 270: 1011-1014, 1995.
231. **Lindquist, S.:** Regulation of protein synthesis during heat shock. *Nature* 293: 311-314, 1981.
232. **Lindquist, S.:** The heat shock response. *Annu Rev. Biochem.* 45: 39-72, 1986.
233. **Lindquist, S., and Craig, E. A.:** The heat shock proteins. *Annu. Rev. Genet.* 22: 631-677, 1988.

234. **Lingappa, J. R., Martin, R. L., Wong, M. L., Gamen, D., Welch, W. J., and Lingappa, V. R.:** A eukaryotic cytosolic chaperonin is associated with a high molecular weight intermediate in the assembly of hepatitis B virus capsid, a multimeric particle. *J. Cell Biol.* 125: 99-111, 1994.
235. **Lissin, N. M., Venyaminov, S. Yu., and Girshovich, A. S.:** (Mg-ATP)-dependent self-assembly of molecular chaperone GroEL. *Nature* 348: 339-342, 1990.
236. **Lissin, N. M.:** *In vitro* dissociation and self-assembly of three chaperonin 60s: The role of ATP. *FEBS Lett.* 361: 55-60, 1995.
237. **Logemann, J., Schell, J., and Willmitzer, L.:** Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* 163: 16-20, 1987.
238. **Lorimer, G.:** Optimizing protein folding with chaperonins. (Abstract No. 21003). *Plant Physiol.* 111 (Supp.): 18, 1996 a.
239. **Lorimer, G. H.:** A quantitative assessment of the role of the chaperonin proteins in protein folding in vivo. *FASEB J.* 10: 5-9, 1996 b.
240. **Losick, R., and Kaiser, D.:** Why and how bacteria communicate. *Sci. Amer.* 276 (2): 52-57, 1997.
241. **Lubben, T. H., Donaldson, G. K., Vitanen, P. V., and Gatenby, A. A.:** Several proteins imported into chloroplast form stable complexes with the GroEL-related chloroplast molecular chaperone. *Plant Cell* 1: 1223-1230, 1989.
242. **Lyon, M. F., Zenthon, J., Evans, E. P., Burtenshaw, M. D., and Willison, K. R.:** Extent of the mouse *t* complex and its inversions shown by *in situ* hybridization. *Immunogenetics* 27: 375-382, 1988.
243. **Maddox, J.:** Does folding determine protein configuration? *Nature* 370: 13, 1994 a.
244. **Maddox, J.:** Towards more measurement in biology. *Nature* 368: 95, 1994 b.
245. **Maldonado, R., Jiménez, I., and Cassadeus, J.:** Changes in ploidy during the *Azotobacter vinelandii* growth cycle. *J. Bacteriol.* 176: 3911-3919, 1994.
246. **Mande, S. C., Mehra, V., Bloom, B. R., and Hol, W. G. J.:** Structure of the heat shock protein chaperonin-10 of *Mycobacterium leprae*. *Science* 271: 203-207, 1996.
247. **Mandel, M., and Higa, A.:** Calcium dependent bacteriophage DNA infection. *J. Mol. Biol.* 53: 159-162, 1970.
248. **Manna, A. C., and Das, H. K.:** Determination of the size of the *Azotobacter vinelandii* chromosome. *Mol. Gen. Genet.* 241: 719-722, 1993.
249. **Manning-Krieg, U. C., Scherer, P. E., and Schatz, G.:** Sequential action of mitochondrial chaperones in protein import into the matrix. *EMBO J.* 10: 3273-3280, 1991.
250. **Marco, S., Carrascosa, J. L., and Valpuesta, J. M.:** Reversible interaction of  $\beta$ -actin along the channel of the TCP-1 cytoplasmic chaperonin. *Biophys J.* 67: 364-368, 1994.

251. **Margolin, W., Wang, R., and Kumar, M.:** Isolation of an *ftsZ* homolog from the archaebacterium *Halobacterium salinarum*: Implications for the evolution of FtsZ and tubulin. *J. Bacteriol.* 178: 1320-1327, 1966.
252. **Margulis, L.:** Origin of Eukaryotic cells. Yale University Press. New Haven CT, 1970.
253. **Martel, R., Cloney, L. P., Pelcher, L. E., and Hemmingsen, S. M.:** Unique composition of plastid chaperonin-60:  $\alpha$  and  $\beta$  polypeptide-encoding genes are highly divergent. *Gene* 94: 181-187, 1990.
254. **Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L., and Hartl, F.-U.:** Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature* 352: 36-42, 1991.
255. **Martin, J., Horwich, A. L., Hartl, F. U.:** Prevention of protein denaturation under heat stress by the chaperonin hsp60. *Science* 258: 995-998, 1992.
256. **Martin, J., Mayhew, M., Langer, T., and Hartl, F. U.:** The reaction cycle of GroEL and GroES in chaperonin-assisted protein folding. *Nature* 366: 228-233, 1993 a.
257. **Martin J., Geromanos, S., Tempst, P., and Hartl, F. U.:** Identification of nucleotide-binding regions in the chaperonin proteins GroEL and GroES. *Nature* 366: 279-282, 1993 b.
258. **Martínez-Romero, E., Segovia, L., Martins-Mercante, F., Franco, A. A., Graham, P., and Pardo, M. A.:** *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. *Int. J. Syst. Bacteriol.* 41: 131-135, 1991.
259. **Mascagni, P., Tonolo, M., Ball, H., Lim, M., Ellis, R. J., and Coates, A.:** Chemical synthesis of 10 kDa chaperonin. Biological activity suggests chaperonins do not require other molecular chaperones. *FEBS Lett.* 286: 201-203, 1991.
260. **Masison, D. C., and Wickner, R. B.:** Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells. *Science* 270: 93-95, 1995.
261. **May, R. M., and Nee, S.:** The species alias problem. *Nature* 378: 447-448, 1995.
262. **May, R. M.:** How many species inhabit the earth? *Sci. Amer.* 267(4): 18-24, Oct., 1992.
263. **Mayhew, M., and Hartl, F. U.:** Lord of the rings: GroES structure. *Science* 271: 161-162, 1996.
264. **Mayhew, M., da Silva, A. C. R., Martin, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F. U.:** Protein folding in the central cavity of the GroEL-GroES chaperonin complex. *Nature* 379: 420-426, 1996.
265. **McDowell, J. M., Huang, S., McKinney, E. C., An, Y.-Q., and Meagher, R. B.:** Structure and evolution of the actin gene family in *Arabidopsis thaliana*. *Genetics* 142: 587-602, 1996.
266. **McElroy, D., Rothenberg, M., Reece, K. S., and Wu, R.:** Characterization of the rice (*Oryza sativa*) actin gene family. *Plant Mol. Biol.* 15: 257-268, 1990.

267. **McLean, B. G., Huang, S., McKinney, E. C., and Meagher, R. B.:** Plants contain highly divergent actin isovariants. *Cell Motil. Cytoskeleton* 17: 276-290, 1990.
268. **McMullin, T. W., Hallberg, R. L.:** A highly evolutionarily conserved mitochondrial protein is structurally related to the protein encoded by the *Escherichia coli* *groEL* gene. *Mol. Cell Biol.* 8: 371-380, 1988.
269. **Melki, R., Vainberg, I. E., Chow, R. L., and Cowan, N. J.:** Chaperonin mediated folding of vertebrate actin-related protein and  $\gamma$ -tubulin. *J. Cell Biol.* 122: 1301-1310, 1993.
270. **Mendoza, J. A., Rogers, E., Lorimer, G. H., Horowitz, P. M.:** Unassisted refolding of urea unfolded rhodanase. *J. Biol. Chem.* 266: 13587-13591, 1991.
271. **Mestel, R.:** Putting prions to the test. *Science* 273: 184-189, 1996.
272. **Miao, G. -H., Hirel, B., Marsoller, M. C., Ridge, R. W., and Verma, D. P. S.:** Ammonia-regulated expression of a soybean gene encoding cytosolic glutamine synthetase in transgenic *Lotus corniculatus*. *Plant Cell* 3: 11-22, 1991.
273. **Michelson, A. M., and Orkin, S. H.:** The 3' untranslated regions of the duplicated human  $\alpha$ -globin genes are unexpectedly divergent. *Cell* 22: 371-377, 1980.
274. **Mifflin, L. C., and Cohen, R. E.:** Characterization of denatured protein inducers of the heat shock (stress) response in *Xenopus laevis* oocytes. *J. Biol. Chem.* 269: 15710-15717, 1994.
275. **Mirsky, A. E., and Pauling, L.:** On the structure of native, denatured and coagulated proteins. *Proc. Natl. Acad. Sci. USA* 22: 439-447, 1936.
276. **Miziorko, H. M., and Lorimer, G. H.:** Ribulose-1, 5-bisphosphate carboxylase-oxygenase. *Ann. Rev. Biochem.* 52: 507-535, 1983.
277. **Monod, J.:** *Chance and Necessity: An essay on the natural philosophy of modern biology.* Wainhouse, Anstryn (trans.). Vintage Books, Random House, 1971.
278. **Mori, M., Murata, K., Kubota, H., Yamamoto, A., Matsushiro, A., and Morita, T.:** Cloning of a cDNA encoding the *Tcp-1* (t complex polypeptide 1) homologue of *Arabidopsis thaliana*. *Gene* 122: 381-382, 1992).
279. **Morimoto, R. I., Tissières, A., and Georgopoulos, C.:** The stress response, function of the proteins, and perspectives. In *Stress Proteins in Biology and Medicine*. R. I. Morimoto, A. Tissières, & C. Georgopoulos Eds. Cold Spring Harbor Laboratory Press. New York. pp. 1- 36, 1990.
280. **Morshauer, R. C., Wang, H., Flynn, G. C., and Zulderweg, E. R. P.:** The peptide-binding domain of the chaperone protein Hsc70 has an unusual secondary structure topology. *Biochemistry* 34: 6261-6266, 1995.
281. **Moulder, J. W.:** Intracellular parasitism: Life in an extreme environment. *J. Infect. Dis.* 130: 300-306, 1974.

282. **Moulder, J. W.:** Comparative biology of intracellular parasitism. *Microbiol. Rev.* 49: 298-237, 1985.
283. **Mummerl, E., Grimm, R., Speth, V., Eckerhorn, C., Schiltz, E., Gatenby, A. A., and Schäfer, E.:** A TCP1 related molecular chaperone from plants refolds phytochrome to its photoreversible form. *Nature* 363: 644-648, 1993.
284. **Mummerl, E., Grimm, R., Speth, V., Eckerhorn, C., Schiltz, E., Gatenby, A. A., and Schäfer, E.:** *Erratum.* *Nature* 372: 709, 1994.
285. **Munson, D., Obar, R., Tzertzinis, G., and Margulis, L.:** The 'tubulin-like' S1 protein of *Spirochaeta* is a member of the hsp65 stress protein family. *Biosystems* 31: 161-167, 1993.
286. **Musgrove, J. E., Johnson, R. A., and Ellis, R. J.:** Dissociation of the ribulosebisphosphate-carboxylase large-subunit binding protein into dissimilar subunits. *Eur. J. Biochem.* 163: 529-534, 1987.
287. **Myolna, P., Pawloski, K., and Bisseling, T.:** Symbiotic nitrogen fixation. *Plant Cell* 7: 869-885, 1995.
288. **Nagata, K.:** Hsp47: a collagen-specific molecular chaperone. *Trends Biochem. Sci.* 21: 23-26, 1996.
289. **Nagpal, P., Jafri, S., Reddy, M. A., and Das, H. K.:** Multiple chromosomes of *Azotobacter vinelandii*. *J. Bacteriol.* 171: 3133-3138, 1989.
290. **Nap, J. -P., and Bisseling, T.:** Developmental biology of a plant-prokaryote symbiosis: The legume root nodule. *Science* 250: 948-954, 1990.
291. **Netzer, W. J., and Hartl, F. U.:** Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388: 343-349, 1997.
292. **Nguyen, T., Zelechowska, M., Foster, V., Bergmann, H., and Verma, D. P. S.:** Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. *Proc. Natl. Acad. Sci. USA* 82: 5040-5044, 1985.
293. **Nilsson, B., Anderson, S.:** Proper and improper folding of proteins in the cellular environment. *Annu. Rev. Microbiol.* 45: 607-635, 1991.
294. **North, G.:** A cytoplasmic chaperonin? *Nature* 354: 434-435, 1991.
295. **Nowak, R.:** Entering the postgenome era. *Science* 270: 368-371, 1995.
296. **Oesch, B., Westaway, D., Wälchli, M., McKinley, M. P., Kent, S. B. H., et al.:** A cellular gene encodes scrapie PrP 27-30 protein. *Cell* 40: 735-746, 1985.
297. **Ogawa, J., and Long, S. R.:** The *Rhizobium meliloti* *groELc* locus is required for regulation of early *nod* genes by the transcription activator NodD. *Genes Devel.* 9: 714-729, 1995.

298. **Ohno, S., and Ohno, M.:** The all pervasive principle of repetitious recurrence governs not only coding sequence construction but also human endeavor in musical composition. *Immunogenetics* 24: 71-78, 1986.
299. **Ohtaka, C., Nakamura, H., and Ishikawa, H.:** Structures of chaperonins from an intracellular symbiont and their functional expression in *Escherichia coli* groE mutants. *J. Bacteriol.* 174: 1869-1874, 1992.
300. **Oishi, M., and Cosloy, S. D.:** The genetic and biochemical basis of the transformability of *Escherichia coli* K12. *Biochem. Biophys. Res. Commun.* 49: 1568-1572, 1972.
301. **Okazaki, A., Ikura, T., Nikaido, K., and Kuwajima, K.:** The chaperonin GroEL does not recognize apo- $\alpha$ -lactalbumin in the molten globule state. *Nature Struc. Biol.* 1: 439-446, 1994.
302. **Olson, M. V.:** A time to sequence. *Science* 270: 394-396, 1995.
303. **Østergaard Jensen, E., Paludan, K., Hyldig-Nielsen, J. J., Jørgensen, P., and Marcker, K. A.:** The structure of a chromosomal leghaemoglobin gene from soybean. *Nature* 291: 677-679, 1981.
304. **Ostermann, J., Horwich, A. L., Neupert, W., and Harl, F. -U.:** Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. *Nature* 341: 125-130, 1989.
305. **Pace, C. N.:** Conformational stability of globular proteins. *Trends Biochem. Sci.* 15: 14-17, 1990.
306. **Pace, N. R.:** A molecular view of microbial diversity and the biosphere. *Science* 276: 734-740, 1997.
307. **Palmer, J. D.:** Organelle genomes: Going, going gone! *Science* 275: 790-791, 1997.
308. **Palleros, D. R., Reid, K. L., Shi, L., Welch, W. J., and Fink, A. L.:** ATP-induced protein-Hsp70 complex dissociation requires K<sup>+</sup> but not ATP hydrolysis. *Nature* 365: 664-666, 1993.
309. **Pan, K-M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., et al.:** Conversion of  $\alpha$ -helices into  $\beta$ -sheets features in the formation of the scrapie prion proteins. *Proc. Natl. Acad. Sci. USA* 90: 10962-10966, 1993.
310. **Parsell, D. A., and Lindquist, S.:** The function of heat-shock proteins in stress tolerance: Degradation and reactivation of damaged proteins. *Annu. Rev. Genet.* 27: 437-496, 1993.
311. **Parsell, D. A., Kowal, A. S., Singer, M. A., and Lindquist, S.:** Protein disaggregation mediated by heat-shock protein Hsp104. *Nature* 372: 475-478, 1994.
312. **Patino, M. M., Liu, J-J., Glover, J.R., and Lindquist, S.:** Support for the prion hypothesis for inheritance of a phenotypic trait in yeast. *Science* 273: 622-626, 1996.
313. **Pelham, H. R. B.:** Speculations on the functions of the major heat shock and glucose regulated proteins. *Cell* 46: 959-961, 1986.
314. **Pelham, H. R. B.:** Functions of the hsp70 protein family: An overview. In *Stress Proteins in Biology and Medicine*. R. I. Morimoto, A. Tissières, & C. Georgopoulos Eds. Cold Spring Harbor Laboratory Press. New York. pp. 287-299, 1990.

315. **Penmetsa, R. V., and Cook, D. R.:** A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. *Science* 275: 527-530, 1997.
316. **Pérez, H. E., Sánchez, N., Vidali, L., Hernández, J. M., Lara, M., and Sánchez, F.:** Actin isoforms in non-infected roots and symbiotic root nodules of *Phaseolus vulgaris* L. *Planta* 193: 51-56, 1994.
317. **Phipps, B. M., Typke, D., Hegerl, R., Volker, S., Hoffmann, A., Stetter, K. O., and Baumeister, W.:** Structure of a molecular chaperone from a thermophilic archaebacterium. *Nature* 361: 475-477, 1993.
318. **Prasad, T. K., and Hallberg, R. L.:** Identification and metabolic characterization of the *Zea mays* mitochondrial homolog of the *Escherichia coli* groEL protein. *Plant Mol. Biol.* 12: 609-618, 1989.
319. **Prasad, T. K., and Steward, C. R.:** cDNA clones encoding *Arabidopsis thaliana* and *Zea mays* mitochondrial chaperonin HSP60 and gene expression during seed germination and heat shock. *Plant Mol. Biol.* 18: 873-885, 1992.
320. **Prusiner, S. B.:** Novel proteinaceous infectious particles cause scrapie. *Science* 216: 136-144, 1982.
321. **Prusiner, S. B.:** Molecular biology of prion diseases. *Science* 252: 1515-1521, 1991.
322. **Prusiner, S. B., Groth, D., Serban, A., Koehler, R., Foster, D., Torchia, M., Burton, D., Yang, S-L., and DeArmond, S. J.:** Ablation of the prion protein (PrP) gene in mice prevents scrapie and facilitates production of anti-PrP antibodies. *Proc. Natl. Acad. Sci. USA* 90: 10608-10612, 1993.
323. **Prusiner, S. B.:** The prion diseases. *Sci. Amer.* 272(1): 30-37, Jan., 1995.
324. **Pititsyn, O. B.:** How the molten globule became. *Trends Biochem. Sci.* 20: 376-379, 1995.
325. **Pititsyn, O. B.:** A determinable but unresolved problem. *FASEB J.* 10: 3-4, 1996.
326. **Putney, S.D., Benkovic, S. J., and Schimmel, P. R.:** A DNA fragment with an  $\alpha$ -phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated *in vivo*. *Proc. Natl. Acad. Sci. USA* 78: 7350-7354, 1981.
327. **Quinto, C., de la Vega, H., Flores, M., Leemans, J., Ceballos, M. A., Pardo, M. A., Azpiroz, R., Girard, M. L., Calva, E., and Palacios, R.:** Nitrogenase reductase: A functional multigene family in *Rhizobium phaseoli*. *Proc. Natl. Acad. Sci. USA* 82: 1170-1174, 1985.
328. **Ranson, N. A., Dunster, N. J., Burston, S. G., and Clarke, A. R.:** Chaperonins can catalyze the reversal of early aggregation steps when a protein misfolds. *J. Mol. Biol.* 250: 581-586, 1995.
329. **Rao, P. V., Horwitz, J., and Zigler, S.:** Chaperone-like activity of  $\alpha$ -crystallin. *J. Biol. Chem.* 269: 13266-13272, 1994.
330. **Reading, D. S., Hallberg, R. L., and Myers, A. M.:** Characterization of the yeast HSP60 gene coding for a mitochondrial assembly factor. *Nature* 337: 655-659, 1989.

331. **Rensing, S. A., and Maier, U.-G.:** Phylogenetic analysis of the stress-70 protein family. *J. Mol. Evol.* 39: 80-86, 1994.
332. **Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R., and Wüthrich, K.:** NMR structure of the mouse prion protein domain PrP(121-231). *Nature* 382: 180-182, 1996.
333. **Rippman, F., Taylor, W.R., Rothbard, J. B., and Green, N. M.:** A Hypothetical model for the peptide binding domain of hsp70 based on the peptide-binding domain of HLA. *EMBO J* 10: 1053-1059, 1991.
334. **Ritossa, F. M.:** A new puffing pattern induced by a temperature shock and DNP in *Drosophila*. *Experientia* 18: 571-573, 1962.
335. **Robinson, C. V., Groß, M., Eyles, S. J., Ewbank, J. J., Mayhew, M., Hartl, F. U., Dobson, C. M., and Radford, S. E.:** Conformation of GroEL-bound  $\alpha$ -lactalbumin probed by mass spectrometry. *Nature* 372: 646-651, 1994.
336. **Rose, G. D., and Wolfenden, R.:** Hydrogen bonding, hydrophobicity, packing, and protein folding. *Annu. Rev. Biophys. Biol.* 22: 381-415, 1993.
337. **Rospert, S., Looser, R., Dubaule, Y., Matouschek, A., Glick, B. S., and Schatz, G.:** Hsp60-independent protein folding in the matrix of yeast mitochondria. *EMBO J.* 15: 764-774, 1996.
338. **Roy, H., Bloom, M., Milos, P., and Monroe, M.:** Studies on the assembly of large subunits of ribulose bisphosphate carboxylase in isolated pea chloroplasts. *J. Cell Biol.* 94: 20-27, 1982.
339. **Rusanganwa, E., and Gupta, R. S.:** Cloning and characterization of multiple groEL chaperonin-encoding genes in *Rhizobium meliloti*. *Gene* 126: 67-75, 1993.
340. **Salbil, H., and Wood, S.:** Chaperonins. *Curr. Op. Struc. Biol.* 3: 207-213, 1993.
341. **Salbil, H. R., Zheng, D., Roseman, A. M., Hunter, A. S., Watson, G. M. F., Chen, S., auf der Mauer, A., O'Hara, B. P., Wood, S. P., Mann, N. H., Barnett, L. K., and Ellis, R. J.:** ATP induces large quaternary rearrangements in a cage-like chaperonin structure. *Curr. Biol.* 3: 265-273, 1993.
342. **Sakaguchi, S., Katamine, S., Nishida, N., Moriuchi, R., Shigematsu, K., et al.:** Loss of cerebellar Purkinje cells in aged mice homozygous for a disrupted PrP gene. *Nature* 380: 528-531, 1996.
343. **Sali, A., Shakhnovich, E., and Karplus, M.:** How does a protein fold? *Nature* 369: 248-251, 1994.
344. **Sambrook, J., Fritsch, E. F., and Maniatis, T.:** *Molecular Cloning. A Laboratory Manual.* Second Edition. Cold Spring Harbor Laboratory Press. New York. 1989.
345. **Sánchez, F., Tobin, S. L., Rdest, U., Zulauf, E., and McCarthy, B. J.:** Two *Drosophila* actin genes in detail. Gene structure, protein structure and transcription during development. *J. Mol. Biol.* 163: 533-551, 1983
346. **Sánchez, F., Campos, F., Padilla, J. E., Bonneville, J. M., Enriquez, C., and Caput, D.:** Purification, cDNA cloning and developmental expression of the nodule-specific uricase from *Phaseolus vulgaris* L. *Plant Physiol.* 84: 1143-1147, 1987.

347. **Sánchez, F., Padilla, J.E., Pérez, H., and Lara, M.:** Control of nodulin genes in root-nodule development and metabolism. *Annu. Rev. Plant Physiol Plant Mol. Biol.* 42: 507-528, 1991.
348. **Sánchez, M., Valencia, A., Ferrándiz, M. - J., Sander, C., and Vicente, M.:** Correlation between the structure and biochemical activities of FtsA, an essential cell division protein of the actin family. *EMBO J.* 13: 4919-4925, 1994.
349. **Sánchez, F., Cárdenas, L., van Spronsen, P., Villanueva, M. A., Valdés, V., Guillen, G., Rongcai, M. Noguez, R., Dantan, E., Vidali, L., Pérez, H., Hepler, P., and Quinto, C.:** Actin cytoskeleton expression and organization during root-nodule development. (Abstract No. 10002). *Plant Physiol.* 111(Supp.): 3, 1996.
350. **Sanger, F., Nicklen, S., and Coulson, A. R.:** DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463-5467, 1977.
351. **Sano, S., and Granick, S.:** Mitochondrial coproporphyrinogen oxidase and protoporphyrin formation. *J. Biol. Chem.* 236: 1173, 1961.
352. **Sato, S., and Ishikawa, H.:** Expression and control of an operon from an intracellular symbiont which is homologous to the *groE* operon. *J. Bacteriol.* 179: 2300-2304, 1997.
353. **Sauk, J. J., Smith, T., Norris, K., and Ferreira, L.:** Hsp47 and the translation-translocation machinery cooperate in the production of  $\alpha 1$  (I) chains of type I procollagen. *J. Biol. Chem.* 269: 3941-3946, 1994.
354. **Savageau, M. A.:** Proteins of *Escherichia coli* come in sizes that are multiples of 14 kDa: Domain concepts and evolutionary implications. *Proc. Natl. Acad. Sci. USA* 83: 1198-1202, 1986.
355. **Schlicher, T., and Soll, J.:** Molecular chaperones are present in the thylakoid lumen of pea chloroplasts. *FEBS Lett.* 379: 302-304, 1996.
356. **Schmid, D., Balci, A., Gehring, H., and Christen, P.:** Kinetics of molecular chaperone action. *Science* 263: 971-973, 1994.
357. **Schmidt, M., and Buchner, J.:** Interaction of GroE with an all- $\beta$  protein. *J. Biol. Chem.* 267: 16829-16833, 1992.
358. **Schmidt, M., Rutkat, K., Rachel, R., Pfeifer, G., Jaenicke, R., Viitanen, P., Lorimer, G., and Buchner, J.:** Symmetric complexes of GroE chaperonins as part of the functional cycle. *Science* 265: 656-659, 1994.
359. **Schneider, G., Lindqvist, Y., Bräden, C.-I., and Lorimer, G.:** Three-dimensional structure of ribulose-1, 5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* at 2.9 Å resolution. *EMBO J.* 5: 3409-3415, 1986.
360. **Schwartz, R. M., and Dayhoff, M. O.:** Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* 199: 395-403, 1978.

361. **Schweltzer, M. H., Marshall, M., Carron, K., Bohle, D. S., Busse, S. C., Arnold, E. V., Barnard, D., Horner, J. R., and Starkey, J. R.:** Heme compounds in dinosaur trabecular bone. *Proc. Natl. Acad. Sci. USA* 94: 6291-6296, 1997.
362. **Searle, J. W., Gorovits, B. M., Ybarra, J., and Horowitz, P. M.:** Reversible oligomerization and denaturation of the chaperonin GroES. *Biochemistry* 35: 4079-4083, 1996.
363. **Selker, J. M. L.:** Three-dimensional organization of uninfected tissue in soybean root nodules and its relation to cell specialization in the central region. *Protoplasma* 147: 178-190, 1988.
364. **Shepherd, J. C. W.:** Method to determine the reading frame of a protein from the purine/pyrimidine genome sequence and its possible evolutionary justification. *Proc. Natl. Acad. Sci. USA* 78: 1596-1600, 1981.
365. **Shepherd, J. C. W.:** Ancient patterns in nucleic acid sequences. In: *Methods in Enzymology* 183. *Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*. Russell F. Doolittle Ed. Academic Press, Inc. San Diego. pp 180-192, 1990.
366. **Sheterline, P., Sparrow, J. C.:** Actin. *Prot. Prof.* 1: 1-63, 1994.
367. **Shinnick, T. M., Vodkin, M. H., and Williams, J. C.:** The *Mycobacterium tuberculosis* 65-kilodalton antigen is a heat shock protein which corresponds to common antigen and to the *Escherichia coli* GroEL protein. *Infect. Immun.* 56: 446-451, 1988.
368. **Short, N.:** On first looking into de human genome. *Nature* 377 Supp.: 1, 1995.
369. **Singh, B., and Gupta, R. S.:** Expression of human 60-kD heat shock protein (HSP60 or P1) in *Escherichia coli* and the development and characterization of corresponding monoclonal antibodies. *DNA Cell Biol.* 11: 489-496, 1992.
370. **Singh, L., and Jones, K. W.:** The use of heparin as a simple cost-effective means of controlling background in nucleic acid hybridization procedures. *Nucl. Acids Res.* 12: 5627-5638, 1984.
371. **Solioz, M., and Bienz, D.:** Bacterial genetics by electric shock. *Trends Biochem.* 15: 175-177, 1990.
372. **Solomon, F.:** Analyses of the cytoskeleton in *Saccharomyces cerevisiae*. *Ann. Rev. Cell Biol.* 7: 633-662, 1991.
373. **Soltys, B. J., and Gupta, R. S.:** Immunoelectron microscopic localization of the 60-kDa heat shock chaperonin protein (Hsp60) in mammalian cells. *Exp. Cell Res.* 222: 16-27, 1996.
374. **Southern, E.:** Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98: 503-517, 1975.
375. **Staden, R.:** Finding protein coding regions in genomic sequences. In: *Methods in Enzymology* 183. *Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences*. Russell F. Doolittle Ed. Academic Press, Inc. San Diego. pp 163-180, 1990.

376. **Stephen, D., Jones, C., and Schofield, J. P.:** A rapid method for isolating high quality plasmid DNA suitable for DNA sequencing. *Nucl. Acids Res.* 18: 7463-7464, 1990.
377. **Sternberg, N.:** Properties of a mutant of *Escherichia coli* defective in bacteriophage  $\lambda$  head formation (*groE*). I. Initial characterization. *J. Mol. Biol.* 67: 1-24, 1973 a.
378. **Sternberg, N.:** Properties of a mutant of *Escherichia coli* defective in bacteriophage  $\lambda$  head formation (*groE*). II. The propagation of phage  $\lambda$ . *J. Mol. Biol.* 76: 25-44, 1973 b.
379. **Sternlicht, H., Farr, G. W., Sternlicht, M. L., Driscoll, J. K., Willison, K., and Yaffe, M. B.:** The *t*-complex polypeptide 1 complex is a chaperonin for tubulin and actin *in vivo*. *Proc. Natl. Acad. Sci. USA* 90: 9422-9426, 1993.
380. **Strauss, E. J., and Falkow, S.:** Microbial pathogenesis: Genomics and beyond. *Science* 276: 707-712, 1997.
381. **Suganuma, N., Tamaoki, M., and Kouchi, H.:** Expression of nodulin genes in plant-determined ineffective nodules of pea. *Plant Mol. Biol.* 28: 1027-1038, 1995.
382. **Sunshine, M., Feiss, M., Stuart, J., and Yochem, J.:** A new host gene (*groPC*) necessary for lambda DNA replication. *Mol. Gen. Genet.* 151: 27-34, 1977.
383. **Tabita, F. R.:** Molecular and cellular regulation and autotrophic carbon dioxide fixation in micro-organisms. *Microbiol. Rev.* 52: 155-189, 1988.
384. **Taguchi, H., and Yoshida, M.:** Chaperonin releases the substrate protein in a form with tendency to aggregate and ability to rebind to chaperonin. *FEBS Lett.* 359: 195-198, 1995.
385. **Takano, T., and Kakefuda, T.:** Involvement of a bacterial factor in morphogenesis of bacteriophage capsid. *Nature New Biol.* 239: 34-37, 1972.
386. **Taubes, G.:** Misfolding the way to disease. *Science* 271: 1493-1495, 1996.
387. **Temple, S. J., Heard, J., Ganter, G., Dunn, K., and Sengupta-Gopalan, C.:** Characterization of a nodule-enhanced glutamine synthetase from alfalfa: Nucleotide sequence, *in situ* localization, and transcript analysis. *Mol. Plant-Microbe Interact.* 8: 218-227, 1995.
388. **Thomas, P. J., Qu, B-H., and Pedersen, P. L.:** Defective protein folding as a basis of human disease. *Trends Biochem. Sci.* 20: 456-459, 1995.
389. **Thomas, P. S.:** Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* 77: 5201-5205, 1980.
390. **Thomson, M. D., Paavola, C. D., Lenvik, T. R., and Gantt, J. S.:** *Chlamydomonas* transcripts encoding three divergent plastid chaperonins are heat-inducible. *Plant Mol. Biol.* 27: 1031-1035, 1995.
391. **Tian, G., Vainberg, I. E., Tap, W. D., Lewis, S. A., and Cowan, N. J.:** Quasi-native chaperonin bound intermediates in facilitated protein folding. *J. Biol. Chem.* 270: 23910-23913, 1995 a.

392. **Tian, G., Vainberg, Y. E., Tap, W. D., Lewis, S. A., and Cowan, N. J.:** Specificity in chaperonin-mediated protein folding. *Nature* 375: 250-253, 1995 b.
393. **Tilney, L. G., and Portnoy, D. A.:** Actin filaments and the growth, movement, and spread of intracellular bacterial parasite *Listeria monocytogenes*. *J. Cell Biol.* 109: 1597-1608, 1989.
394. **Tilly, K., Murialdo, H., and Georgopoulos, C.:** Identification of a second *Escherichia coli* *groE* gene whose product is necessary for bacteriophage morphogenesis. *Proc. Natl. Acad. Sci. USA* 78: 1629-1633, 1981.
395. **Tissières, A., Mitchell, H. K., and Tracy, V. M.:** Protein synthesis in salivary glands of *Drosophila melanogaster*: Relation to chromosome puffs. *J. Mol. Biol.* 84: 389-398, 1974.
396. **Tober, Y., Gaus, S. E., Deboer, T., Achermann, P., Fischer, M., Rüllicke, T., Moser, M., Oesch, B., McBride, P. A., and Manson, J. C.:** Altered circadian activity rhythms and sleep in mice devoid of prion protein. *Nature* 380: 639-642, 1996.
397. **Todd, M. J., Vitanen, P. V., and Lorimer, G. H.:** Dynamics of the chaperonin ATPase cycle: Implications for facilitated protein folding. *Science* 265: 659-666, 1994.
398. **Todd, M. J., Lorimer, G. H., and Thirumalai, D.:** Chaperonin-facilitated protein folding: Optimization of rate and yield by an iterative annealing mechanism. *Proc. Natl. Acad. Sci. USA* 93: 4030-4035, 1996.
399. **Trent, J. D., Nirmesgern, E., Wall, J. S., Harli, F. -U., and Horwich, A. L.:** A molecular chaperone from a thermophilic archaeobacterium is related to the eukaryotic protein t-complex polypeptide-1. *Nature*. 354: 490-493, 1991.
400. **Trent, J. D., Kagawa, H. K., Yaoi, T., Olle, E., and Zaluzec, N. J.:** Chaperonin filaments: The archaeal cytoskeleton? *Proc. Natl. Acad. Sci.* 94: 5383-5388, 1997.
401. **Tuite, M. F.:** Psi no more for yeast prions. *Nature* 370: 327-328, 1994.
402. **Valdés-López, V. M., and Herrera-Estrella, L.:** Isolation and characterization of a non-allelic chaperonin (cpn60-β), partial cDNA clone from *Arabidopsis thaliana*. VII Pan American Association of Biochemical Societies (PAABS) Congress, Ixtapa, México, 1992.
403. **van Brussel, A. A. N., Bakhuizen, R., van Spronsen P. C., Spainik, H. P., Tak, T., Lugtenberg, B. J. J., and Kijne, R. W.:** Induction of pre-infection thread structures in the leguminous host plant by mitogenic lipo-oligosaccharides of *Rhizobium*. *Science* 257: 70-72, 1992.
404. **van der Vlies, S. M., Gatenby, A. I., and Georgopoulos, C.:** Bacteriophage T4 encodes a co-chaperonin that can substitute for *Escherichia coli* GroES in protein folding. *Nature* 368: 654-656, 1994.
405. **Van Dyk, T. K., Gatenby, A. A., and LaRosa, R. A.:** Demonstration by genetic suppression of the interaction of GroE products with many proteins. *Nature* 342: 451-453, 1989.

406. **van Kammen, A.:** Suggested nomenclature for plant genes involved in nodulation and symbiosis. *Plant Mol. Biol. Rep.* 2: 43-45, 1984.
407. **van Spronsen, P. C., Bakhulzen, R., van Brussel, A. A. N., and Kilne, J. W.:** Cell wall degradation during infection thread formation by the root nodule bacterium *Rhizobium leguminosarum* is a two step process. *Eur. J. Cell Biol.* 64: 88-94, 1994.
408. **VanBogelen, R. A., and Neldhardt, F. C.:** Ribosomes as sensors on heat and cold shock in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA.* 87: 5589-5593, 1990.
409. **Varma, M., Aebi, U., Fleming, J., and Leavitt, J.:** A 60-kDa polypeptide in mammalian cells with epitopes related to actin. *Exp Cell Res* 173: 163-173, 1987.
410. **Vasse, J., de Billy, F., Camut, S., and Truchet, G.:** Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J. Bacteriol.* 172: 4295-4306, 1990.
411. **Vasse, J., de Billy, F., and Truchet, G.:** Abortion of infection during *Rhizobium meliloti*-alfalfa symbiotic interaction is accompanied by a hypersensitive reaction. *Plant J.* 4: 555-566, 1993.
412. **Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W.:** Serial analysis of gene expression. *Science* 270: 484-487, 1995.
413. **Verma, D. P. S.:** Signals in root nodule organogenesis and endocytosis of *Rhizobium*. *Plant Cell* 4: 373-382, 1992.
414. **Verma, D. P., and Hong, Z.:** Biogenesis of the peribacteroid membrane in root nodules. *Trends Microbiol* 4: 364-368, 1996.
415. **Viale, A. M., and Arakaki, A. K.:** The chaperone connection to the origins of the eukaryotic organelles. *FEBS Lett.* 341: 146-151, 1994.
416. **Vidali, L., Pérez, H. E., Valdés-López, V., Noguez, R., Zamudio, F., and Sánchez, F.:** Purification, characterization, and cDNA cloning of profilin from *Phaseolus vulgaris*. *Plant Physiol.* 108: 115-123, 1995.
417. **Vierling, E.:** The roles of heat shock proteins in plants. *Annu. Rev. Plant Mol. Biol.* 42: 579-620, 1991.
418. **Vitnanen, P. V., Lorimer, G. H., Seetharam, R., Gupta, R. S., Oppenheim, J., Thomas, J. O., and Cowan N. J.:** Mammalian mitochondrial chaperonin 60 functions as a single toroidal ring. *J. Biol. Chem.* 267: 695-698, 1992 a.
419. **Vitnanen, P. V., Gatenby, A. A., and Lorimer, G. H.:** Purified chaperonin 60 (GroEL) interacts with the non-native states of a multitude of *Escherichia coli* proteins. *Protein Sci.* 1: 363-369, 1992 b.
420. **Vitnanen, P. V., Schmidt, M., Buchner, J., Suzuki, T., Vierling, E., Dickson, R., Lorimer, G. H., Gatenby, A., and Soll, J.:** Functional characterization of the higher plant chloroplast chaperonins. *J. Biol. Chem.* 270: 18158-18164, 1995.

421. **Vijn, I., das Neves, L., van Kammen, A., Franssen, H., and Bisseling, T.:** Nod factors and nodulation in plants. *Science* 260: 1764-1765, 1993.
422. **Vogel, G.:** Parasites shed light on cellular evolution. *Science* 275: 1422, 1997.
423. **von Heijne, G., Steppuhn, J., and Herrmann, R. G.:** Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* 180: 535-545, 1989.
424. **Wakabayashi, S., Marsubara, H., and Webster, D. A.:** Primary sequence of a dimeric bacterial hemoglobin from *Vitreoscilla*. *Nature* 322: 481-483, 1986.
425. **Walter, S., Lorimer, G. H., and Schmid, F. X.:** A thermodynamic coupling mechanism for GroEL-mediated unfolding. *Proc. Natl. Acad. Sci. USA* 93: 9425-9430, 1996.
426. **Wallington, E. J., and Lund, P. A.:** *Rhizobium leguminosarum* contains multiple chaperonin (*cpn60*) genes. *Microbiology* 140: 113-122, 1994.
427. **Wang, C-C, and Tsou, C-L.:** The insulin A and B chains contain sufficient structural information to form the native molecule. *Trends Biochem. Sci.* 16: 279-281, 1991.
428. **Wang, T.-F., Chang, J., and Wang, C.:** Identification of the peptide binding domain of hsc70. 18-kilodalton fragment located immediately after ATPase domain is sufficient for high affinity binding. *J. Biol. Chem.* 268: 26049-26051, 1993.
429. **Weissman, C.:** The prion connection: Now in yeast? *Science* 264: 528-530, 1994.
430. **Weissman, J. S., Kashi, Y., Fenton, W. A., and Horwich, A. L.:** GroEL-mediated protein folding proceeds by multiple rounds of binding and release of nonnative forms. *Cell* 78: 693-702, 1994.
431. **Weissman J. S., Rye, H. S., Fenton, W. A., Beechem, J. M., and Horwich, A. L.:** Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* 84: 481-490, 1996.
432. **Welch, W. J.:** The mammalian stress response: Cell physiology and biochemistry of stress proteins. In *Stress Proteins in Biology and Medicine*. R.I. Morimoto, A. Tissières, & C. Georgopoulos Eds. Cold Spring Harbor Laboratory Press. New York. pp. 223-278, 1990.
433. **Welch, J.:** How cells respond to stress. *Sci. Amer.* 268(5): 34-41, May, 1993.
434. **Wickner, R. B.:** (URE3) as an altered URE2 protein: Evidence for a prion analog in *Saccharomyces cerevisiae*. *Science* 264: 566-569, 1994.
435. **Wiech, H., Buchner, J., Zimmermann, R., and Jakob, U.:** Hsp90 chaperones protein folding *in vitro*. *Nature* 358: 169-170, 1992.
436. **Willison, K., Lewis, V., Zuckerman, K. S., Cordell, J., Dean, C., Miller, K., Lyon, M. F., and Marsh, M.:** The *t* complex polypeptide 1 (TCP-1) is associated with the cytoplasmic aspect of Golgi membranes. *Cell* 57: 621-632, 1989.
437. **Woese, C. R., and Fox, G. E.:** Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. USA* 74: 5088-5090, 1977.

438. **Woese, C. R.:** Archaeobacteria. *Sci. Amer.* 244(6): 94-106, 1981.
439. **Woese, C. R., Kandler, O., and Wheel, M. L.:** Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Nat. Acad. Sci. USA* 87: 4576-4579, 1990.
440. **Wolynes, P. G., Onuchic, J. N., and Thirumalai, D.:** Navigating the folding routes. *Science* 267: 1619-1620, 1995.
441. **Wu, H. B., Feist, G. L., and Hemmingsen, S. M.:** A modified *Escherichia coli* chaperonin (GroEL) polypeptide synthesized in tobacco and targeted to the chloroplasts. *Plant Mol. Biol.* 22: 1087-1100, 1993.
442. **Wülfing, C., and Plückthun, A.:** Protein folding in the periplasm of *Escherichia coli*. *Mol. Microbiol.* 12: 685-692, 1994.
443. **Yaffe, M. B., Farr, G. W., Mikos, D., Horwich, A. L., Sternlicht, M. L., and Sternlicht, H.:** TCP1 complex is a molecular chaperone in tubulin biogenesis. *Nature* 358: 245-248, 1992.
444. **Ybarra, J., and Horowitz, P. M.:** Refolding and reassembly of active chaperonin GroEL after denaturation. *J Biol. Chem.* 270: 22113-22115, 1995.
445. **Zabaleta, E., Oropeza, A., Jiménez, B., Salerno, G., Crespi, M., and Herrera-Estrella, L.:** Isolation and characterization of genes encoding chaperonin 60 $\beta$  from *Arabidopsis thaliana*. *Gene* 111: 175-181, 1992.
446. **Zabaleta, E., Assad, N., Oropeza, A., Salerno, G., and Herrera-Estrella, L.:** Expression of one of the members of the *Arabidopsis* chaperonin 60 $\beta$  gene family is developmentally regulated and wound-repressible. *Plant Mol. Biol.* 24: 195-202, 1994 a.
447. **Zabaleta, E., Oropeza, A., Assad, N., Mandel, A., Salerno, G., and Herrera-Estrella, L.:** Antisense expression of chaperonin 60 $\beta$  in transgenic tobacco plants leads to abnormal phenotypes and altered distribution of photoassimilates. *Plant J.* 6: 425-432, 1994 b.
448. **Zahn, R., Spitzfaden, C., Otfinger, M., Wüthrich, K., and Plückthun, A.:** Destabilization of the complete protein secondary structure on binding to the chaperone GroEL. *Nature* 368: 261-265, 1994.
449. **Zahn, R., Perrett, S., Stenberg, G., and Fersht, A. R.:** Catalysis of amide proton exchange by the molecular chaperones GroEL and SecB. *Science* 271: 642-645, 1996.
450. **Zal, F., Lallier, F. H., Wall, J. S., Vinogradov, S. N., and Toulmond A.:** The multi-hemoglobin system of the hydrothermal vent tube worm *Riftia pachyptila*. I. Reexamination of the number and masses of its constituents. *J. Biol. Chem.* 271: 8869-8874, 1996.
451. **Zardeneta, G., and Horowitz, P. M.:** Micelle-assisted protein folding. *J. Biol Chem.* 267: 5811-5816, 1992.

452. **Zeilstra-Ryalls, J., Fayet, O., and Georgopoulos, C.:** The universally conserved GroE (Hsp60) chaperonins. *Annu. Rev. Microbiol.* 45: 301-325, 1991.
453. **Zhang, L., Zhou, W., Veculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W.:** Gene expression profiles in normal and cancer cells. *Science* 276: 1268-1272, 1997.
454. **Zhou, C., Yang, Y., and Jong, A. Y.:** Mini-prep in ten minutes. *BioTechniques* 8: 172-173, 1990.
455. **Zhu, H., Ownby, D. W., Riggs, C. K. Nolasco N. J., James K. Stoops, and Riggs, A. F.:** Assembly of the gigantic hemoglobin of the earthworm *Lumbricus terrestris*. Roles of subunit equilibria, non-globin linker chains, and valence of the heme iron. *J. Biol. Chem.* 271: 30007-30021, 1996
456. **Zimmerman, S. B., and Minton, A. P.:** Macromolecular crowding: Biochemical, biophysical and physiological consequences. *Annu. Rev. Biophys Biomol. Struct.* 22: 27-65, 1993.