

UCH - FC
DOG - B
A 664
C.1

**CAMBIOS CONFORMACIONALES CONSERVADOS EN LA UNIÓN DE
LIGANDOS A FAMILIAS DE PROTEÍNAS**

Tesis
Entregada a la
Universidad de Chile
En cumplimiento parcial de los requisitos
para optar al grado de
Doctor en Ciencias con mención en Biología Molecular, Celular y Neurociencias



Facultad de Ciencias

por

JOSÉ JAIME ARBILDUA SEPÚLVEDA

Marzo, 2007

Director de Tesis
Dr. Octavio Monasterio O.

Co-director de Tesis
Dr. Roberto Sánchez K.

**FACULTAD DE CIENCIAS
UNIVERSIDAD DE CHILE**

**INFORME DE APROBACION
TESIS DE DOCTORADO**

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por el candidato.

JOSÉ JAIME ARBILDUA SEPÚLVEDA

Ha sido aprobada por la comisión de Evaluación de la tesis como requisito para optar al grado de Doctor en Ciencias con mención en Biología Molecular, Celular y Neurociencias, en el examen de Defensa de Tesis rendido el día 20 de Diciembre de 2006.

Director de Tesis :
Dr. Octavio Monasterio



Co-director de Tesis :
Dr. Roberto Sánchez



Comisión de Evaluación de la Tesis :
Dr. Osvaldo Álvarez (Presidente)



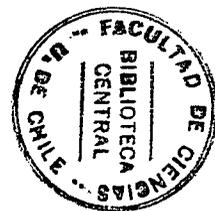
Dr. Víctor Cifuentes



Dr. Danilo Gonzalez



A mi padre.





Biografía

Nací en Vallenar, tercera región, el 23 de agosto de 1968, debido al trabajo de mi padre, mi familia se mudó a Coquimbo, posteriormente a La Serena, Santiago y finalmente a Concepción, donde realicé mi educación media en el Colegio Saint John. Estudié Bioquímica en la Universidad de Concepción, entre los años 1986 a 1990, fue este último año donde junto a dos amigos, Esteban Nova y Roberto Sánchez, decidimos realizar nuestras tesis de pre grado en el área de estructura de proteínas. Los tres empezamos en 1991 con nuestras respectivas tesis, dirigidos por el Dr. Octavio Monasterio, en el laboratorio de Bioquímica de la Facultad de Ciencias de la Chile. En 1992 defendí mi tesis y comencé a trabajar en el Centro de Equipo Mayor de la Facultad de Ciencias, donde he estado involucrado en el desarrollo de técnicas de análisis químico y biológico y en actividades docentes, especialmente relacionadas a instrumentación y bioinformática. Ingresé al programa de Doctorado el año 2001.

AGRADECIMIENTOS

Deseo agradecer en forma muy especial al Dr. Octavio Monasterio y al Dr. Roberto Sánchez, por el apoyo y la confianza que me entregaron durante el desarrollo de esta tesis, como también por el tiempo y la paciencia que dedicaron a discutir los resultados de la misma.

Deseo agradecer a mi familia por todo su apoyo, sobretodo durante mi estadía en Estados Unidos, en el laboratorio del Dr. Roberto Sánchez, que significó que no pudieran contar conmigo durante 3 meses.

En el trabajo día a día, agradezco el aporte y amistad de estudiantes y profesores del laboratorio de Biología Estructural y Molecular, en las diferentes etapas de esta tesis. Especialmente de Esteban Nova, Andrea Garcés, Rodrigo Díaz, Jae Yen Shin y Felipe Montecinos, con quienes trabajé de manera más cercana debido a que nuestras tesis comprendían temas relacionados. También agradezco la ayuda y amistad de Carlos Madrid, biólogo de la Facultad de Ciencias, que actualmente trabaja en Nueva York, USA, sin la cual la configuración del "cluster" de servidores utilizados para realizar esta tesis no habría sido posible. Por último quiero agradecer al Dr. Suvobrata Chakravarty (Suvo), quien realiza su post doctorado en el laboratorio del Dr. Sánchez, por su prestancia para ayudar y orientar a las personas que trabajamos con el.



INDICE DE MATERIAS

	Pág.
Dedicatoria	ii
Agradecimientos	iii
Índice de materias	iv
Lista de tablas	vii
Lista de figuras	viii
Lista de abreviaturas	x
RESUMEN	xi
ABSTRACT	xv
INTRODUCCIÓN	1
Cambios conformacionales	1
La FtsZ	4
Similitudes estructurales y funcionales en la familia de las tubulinas	8
Objetivos	11
MATERIALES Y METODOS	12
Construcción de una base de datos de pares de proteínas de estructura conocida en su estado libre y acomplejadas con ligando	12
Determinación de los contactos de los ligandos en la proteína	13
Comparación de las coordenadas de los carbono alfa ($C\alpha$) y los carbono beta ($C\beta$) en la apo-proteína y la proteína-ligando	14
Alineamiento múltiple para cada familia	14
Cambios conformacionales relacionados a la unión de ligando (LRCC)	15
Factor B normalizado	15
Análisis de χ^2 para determinar la independencia de dos valores	16
Análisis de χ^2 para determinar si dos conjuntos de datos presentan distribuciones diferentes	16



Número de capas de aminoácidos en la estructura de la proteína respecto a los residuos que constituyen el sitio activo (SA)	17
Análisis de componentes principales (PCA)	17
Determinación de estructura secundaria	18
Grado de exposición de los aminoácidos	18
Diseño de mutantes	19
Modelo 3D de EcFtsZ	19
Análisis y representación de resultados	21
Programas y bases de datos	21
Mutagénesis sitio-dirigida	22
Sobre expresión y purificación de EcFtsZ silvestre y mutantes	22
Ensayo de complementación <i>in vivo</i>	23
Dispersión de la luz asociada a la polimerización de EcFtsZ	24
Actividad GTPásica	24
Dicroísmo circular (DC) y espectroscopia de fluorescencia	25
RESULTADOS	26
Caracterización de los cambios conformacionales relacionados a la unión de ligandos pequeños. Base de datos de pares de proteínas en estado libre y como complejo proteína-ligando	26
Proporción de residuos que participan en LRCC, y su relación espacial respecto al sitio activo (SA)	28
Composición aminoacídica de los LRCC y su distribución en estructura secundaria	30
Grado de exposición y flexibilidad de los LRCC y SA	34
Conservación de los LRCC en familias de proteínas Conservación estructural de los LRCC en familias de proteínas con un solo sitio activo	38
Conservación estructural de los LRCC en familias de proteínas con dos sitios activos	42
Efecto de los contactos cristalográficos sobre los LRCC	45

Determinación de los LRCC en FtsZ	
Comparación de FtsZ de distintos organismos en conformación GTP y GDP	47
Comprobación experimental del efecto esperado para la introducción de mutaciones en zonas LRCC y no LRCC de FtsZ	54
Fenotipo de mutantes puntuales en regiones involucradas en LRCC de EcFtsZ	
Caracterización estructural de las mutantes puntuales de EcFtsZ, F40W, F135W e I294W	55
Caracterización funcional de las mutantes puntuales de EcFtsZ, F40W, F135W e I294W	56
Conformación de la mutante I294W	61
DISCUSION	65
Análisis de la base de datos de proteínas en estado libre y como complejo proteína-ligando	65
Conservación de los LRCC en algunas familias de proteínas	67
Elementos estructurales que participan en los LRCC	71
Influencia de la flexibilidad sobre la función de FtsZ	73
CONCLUSIONES	80
BIBLIOGRAFIA	82



LISTA DE TABLAS

		Pág.
Tabla 1	Principales familias de la base de datos de proteínas en estado libre y como complejo proteína-ligando	27
Tabla 2	Lista de los archivos PDB, de la familia de las serino proteasas, en una conformación con y sin ligando, analizados en la figura 12	48
Tabla 3	Características de la comparación de estructuras en estado GTP y GDP de la figura 13A, 13B y figura 14A	52



LISTA DE FIGURAS

		Pág.
Figura 1	Modelo por homología de EcFtsZ, basado en la estructura cristalográfica de MjFtsZ (1fsz.pdb)	5
Figura 2	Esquema de la regulación funcional de FtsZ mediante la hidrólisis e intercambio de GDP – GTP	7
Figura 3	Mutantes puntuales en EcFtsZ	20
Figura 4	Porcentaje de distribución de los residuos de la proteína involucrados en el SA y en los LRCC	26
Figura 5	Porcentaje de distribución de residuos de los LRCC y del total de residuos de la proteína en las diferentes capas alrededor del SA	31
Figura 6	Composición aminoacídica en los LRCC y los SA y frecuencia de los LRCC en los elementos de estructura secundaria	33
Figura 7	Grado de exposición de los residuos que participan en LRCC, en SA y el total de residuos en las proteínas de la base de datos	35
Figura 8	Distribución del factor térmico normalizado para todos los residuos de las proteínas en la base de datos, los LRCC y los SA	37
Figura 9	Conservación de los cambios conformacionales asociados a un mismo SA, en la familia de las fosfotransferasas	39
Figura 10	Comparación de LRCC generados por la unión de diferentes ligandos a una misma proteína (A y B) y a otras proteínas de la familia fosfotransferasas (C)	41
Figura 11	LRCC para el sitio de unión de dihidrofolato y el sitio de unión de NADPH ₂ en la familia de la DHF	44
Figura 12	Caracterización de los LRCC para la familia de las serino proteasas	46
Figura 13	LRCC y flexibilidad intrínseca en FtsZ de distintos microorganismos	50
Figura 14	Gráfico de barras de los desplazamientos de los C α para MtbFtsZ y MjFtsZ entre la conformación GDP y GTP, y variaciones del factor térmico normalizado para ambas proteínas en la conformación GDP	51
Figura 15	Espectros de emisión de fluorescencia y dicroísmo circular	57

Figura 16	Caracterización funcional de las mutantes de EcFtsZ, F40W, F135W e I294W	59
Figura 17	Microfotografías electrónicas de EcFtsZ wt y la mutante I294W	60
Figura 18	Mutaciones correlacionadas en las posiciones 275 y 294 de EcFtsZ	63
Figura 19	Regiones alejadas del SA que participan en LRCC en la familia DHFR	72
Figura 20	Modelo de interacción longitudinal y lateral en los doble filamentos de FtsZ	79



LISTA DE ABREVIATURAS

CATH	: Base de datos que clasifica las proteínas de estructura conocida sobre la base de <u>C</u> lase, <u>A</u> rquitectura, <u>T</u> opología y <u>H</u> omología (CATH)
CD	: Dicroísmo Circular
FtsZ	: Proteína "Filamentous temperature sensitive Z"
GTP	: Nucleótido trifosfato de guanosina
GdmCl	: Cloruro de guanidinio
NMR	: Resonancia Magnética Nuclear
PDB	: "Protein Data Bank"
X-ray	: Difracción de Rayos X
RMSD	: "Root Mean Square Deviation"
SA	: Sitio activo
LRCC	: Cambios conformacionales relacionados a la unión de Ligando
MjFtsZ	: FtsZ de <i>Methanococcus jannaschii</i>
MtbFtsZ	: FtsZ de <i>Mycobacterium tuberculosis</i>
EcFtsZ	: FtsZ de <i>Escherichia coli</i>
PaFtsZ	: FtsZ de <i>Pseudomona aeuroginosa</i>

Otras abreviaturas se definen en el texto.

RESUMEN

Se caracterizaron los cambios conformacionales inducidos por ligandos en familias de proteínas, al comparar la estructura tridimensional, 3D, de la apo-proteína con la de la proteína acomplejada. La comparación de las estructuras 3D mostró cambios conformacionales inducidos por ligando, en todas las proteínas examinadas. En promedio, un 8% de los residuos de la proteína participan en los cambios conformacionales. Los residuos del sitio activo comprenden, en promedio un 3% de los residuos totales de la proteína, y participan con uno de cada cinco aminoácidos en los cambios conformacionales. Los sitios de unión a ligando se caracterizaron por una mayor rigidez en la estructura de la proteína, determinada a partir del factor térmico normalizado, con una exposición media de sus residuos al solvente (grado de exposición entre el 5 – 40%), alta conservación de identidad de secuencia y por una composición aminoacídica característica, rica en aminoácidos aromáticos e histidina. En cambio, las secciones de la proteína que participan en los cambios conformacionales se caracterizaron por ser más flexibles pues predominan estructuras tipo lazo, con una exposición al solvente mayor al 40%, por una baja conservación de su identidad de secuencia, y por una composición aminoacídica rica en glicina, serina, ácido aspártico y asparragina.

Con el objeto de determinar la conservación de los cambios conformacionales en proteínas relacionadas evolutivamente, se agrupó las proteínas en familias de acuerdo a la clasificación de estructuras definida por la base de datos CATH, que usa como criterios su Clase, Arquitectura, Topología y Homología. Para cada familia se construyó alineamientos estructurales múltiples donde se codificó la información de los residuos que participan en los cambios conformacionales. El análisis de los alineamientos,

mostró cambios conformacionales similares cuando distintos ligandos se unen a un mismo sitio activo en una misma proteína. Además se observó cambios conformacionales similares con la unión de distintos ligandos a diferentes proteínas de la familia.

Los resultados del análisis de la estructura tridimensional de las proteínas, en su conjunto, muestran un patrón común para los cambios conformacionales, que se conserva en la estructura 3D de una misma familia. Sin embargo, en general no se observa conservación en la secuencia aminoacídica de estos patrones.

El valor predictivo de la información obtenida en la caracterización de los cambios conformacionales, se corroboró con la determinación de las regiones en la estructura de FtsZ que participan en cambios conformacionales asociados a la hidrólisis del nucleótido GTP. Con este objetivo se comparó las estructuras cristalográficas de la proteína FtsZ, 1rlu y 1rq7 de *Mycobacterium tuberculosis* y 1w58 y 1fsz de *Methanococcus jannaschii*, unidas a GTP y GDP, respectivamente. Las regiones determinadas que participan en cambios conformacionales incluyen la hélice H2 y parte de la hélice H1, los lazos T3, T4, el lazo entre la hebra β S5 y hélice H5, y el lazo entre las hélices H6 y H7, todos ellos elementos estructurales ubicados alrededor del sitio de unión de GTP. También participan en los cambios conformacionales regiones alejadas del sitio de unión del nucleótido, tales como las hebras β S7 y S9, lazos T7, S8-H10 y H10-S9 (nomenclatura de Oliva 2004). Para evaluar el movimiento de estas regiones se construyó tres mutantes de FtsZ de *Escherichia coli*: F40W ubicada en una región que no participa en cambios conformacionales (hebra β S2) , F135W en el dominio amino terminal en una región cercana al sitio de unión del nucleótido que participa en los cambios conformacionales (lazo S5-H5), e I294W ubicada en el dominio carboxilo

terminal alejada del sitio de GTP (hebra β S9) y que participa en los cambios conformacionales. Los espectros de CD de las tres mutantes no mostraron cambios significativos en la estructura secundaria de las proteínas, lo que indica que su estructura es similar a FtsZ silvestre. La caracterización funcional de la mutante F40W mostró complementación *in vivo*, es decir pudo sustituir la función de FtsZ silvestre silenciada, mostró un 80% de la velocidad inicial de hidrólisis de GTP respecto a FtsZ silvestre, y un nivel de polimerización de 70% inducida por 1 mM de GTP y evaluada por turbidez a 350 nm. La segunda mutante, F135W, complementó *in vivo*, mostró un 71% de la velocidad inicial de hidrólisis del nucleótido y un 32% del nivel de polimerización respecto a FtsZ silvestre. I294W por su parte, no complementó *in vivo* y GTP no indujo su polimerización, sin embargo presenta turbidez basal en presencia de GDP y su velocidad inicial de hidrólisis de GTP es de un 13% respecto a FtsZ silvestre. Los resultados de las mutaciones puntuales en FtsZ, no mostraron cambios significativos en la estructura 3D de la proteína. Como se esperaba, la introducción en regiones importantes para la función de la proteína del residuo triptófano, posiciones F135 y I294 que participan en cambios conformacionales, disminuyeron de manera significativa su función. La mutante F40W ubicada en una región rígida de la proteína que no participa en cambios conformacionales no alteró de manera significativa la función de FtsZ, sin embargo, mostró una variación del 2% en su contenido de estructura secundaria respecto a FtsZ silvestre, por la perturbación que provocó la introducción del residuo triptofano.

Los resultados en su conjunto, muestran que regiones de la estructura tridimensional de las proteínas que participan en los cambios conformacionales, son regiones superficiales y flexibles que no están involucradas en la estabilidad de la proteína, por lo

tanto son las regiones a considerar en ingeniería de proteínas, si se desea obtener proteínas con determinadas características funcionales, o se pretende bloquear una determinada función.

ABSTRACT

Conformational changes induced by ligands in protein families were studied by structural comparison between the apo-protein and the protein-ligand complex structures. Comparison of single chain proteins in the PDB database in the free and ligand bound state, showed that conformational changes induced by ligand are a common behavior in proteins and involve around 8% of the total residues. Active site residues are approximately 3% of the total protein amino acids and only 20% of them are involved in conformational changes. Binding sites were characterized by their rigidity (analyzed from B-factor normalized data), a solvent exposure between 5 – 40%, a high sequence conservation and a composition rich in aromatics residues and histidine. On the other hand, sections of the protein involved in conformational changes were characterized by their flexibility, where loops are the most common secondary structure elements, the solvent exposure is above 40% and the sequence conservation is below the protein average. The amino acid composition of the regions involved in conformational changes presented high frequency of glycine, serine, aspartic acid and asparagine.

With the aim to determine if there is a common conformational change pattern in each protein family, proteins selected from the PDB database were divided into families in agreement with the structural classification found on CATH database, that use the criteria of Class, Architecture, Topology and Homology for the classification. For each protein in the family, residues participating in conformational changes were determined by structural superposition of the apo-protein and the protein-ligand complex. To compare protein conformational changes within a family, a multiple structural alignment was performed. The alignment analysis showed similar conformational changes when

different ligands bind to the same active site. In the same way, different proteins within a family showed similar conformational changes upon ligand bound.

The results of the three-dimensional structure analysis of the proteins showed a common pattern of conformational changes conserved within a family. This was not seen for the sequence amino acid conservation.

The predictive value of the information obtained from conformational changes, was assayed on the FtsZ protein. Regions of the structure important in conformational changes, related to GTP hydrolysis, were found by structural comparison of FtsZ in the GTP and GDP conformations, PDB code 1rlu and 1rq7 from *Mycobacterium tuberculosis*, and 1w58 and 1fsz from *Methanococcus jannaschii*. These regions include H2 helix, H1 helix, T3, T4 and H5-H6 loop near the GTP binding site, and the T7, S8-H10, S9-H10 loops, S7, S9 β strand (definition of FtsZ secondary structure elements by Oliva et al. 2004) in the C-terminal domain. To evaluate these regions in the laboratory, three mutants in FtsZ of *Escherichia coli* were analyzed, F40W located in S2 β strand a region in the amino terminal domain away from the conformational changes associated to GTP hydrolysis, F135W near the nucleotide binding site in the H5-S5 loop position that belong to conformational changes, an I294W in the carboxy terminal domain away from the GTP binding site in the S9 β strand, also a position involved in conformational changes. Circular dichroism spectra from the three mutants showed a very similar behavior respect to FtsZ wild type, suggesting no alterations in the secondary structure content of FtsZ with the point mutations. The functional characterization of F40W showed *in vivo* complementation, determined as the capacity of the mutant to replace the silenced FtsZ wt in the bacteria, 80% of the initial hydrolysis velocity respect to wild type and a GTP induced polymerization level of 70% respect to

wild type (evaluated by turbidity at 350 nm). The second mutant F135W, showed *in vivo* complementation, a 71% of the nucleotide initial hydrolysis velocity and a 32% of the polymerization level of the FtsZ wild type. I294W on the other hand, did not showed *in vivo* complementation or polymerization induced by GTP, however the mutant has a basal polymerization and an initial GTP hydrolysis velocity of 13% respect to wild type. The results from the point mutations in FtsZ did not showed alterations on the 3D structure, but showed changes in functionality. Mutations on flexible regions involved in conformational changes, F135W and I294W, showed drastic alterations on functionality even when none of these amino acid positions belong to the nucleotide binding site, more surprisingly I294W is in the C-terminal domain. Mutant F40W showed only minor alterations in functionality but changes of 2% in the content of secondary structure suggest a role in the stability of the protein.

The results as a whole showed that regions of the structure involved in conformational changes presented high flexibility, above 40% solvent accessibility and did not altered protein fold. These characteristics made the regions belonging to conformational changes the ideal substrate to look for a protein with certain functional parameters, by protein engineering, or to block a specific activity.

INTRODUCCIÓN

Cambios conformacionales

La propiedad que permite a las proteínas cumplir sus diversas funciones biológicas, es la capacidad para interactuar con diferentes tipos de ligandos, tales como; otras proteínas, DNA, RNA, nucleótidos, carbohidratos, etc. La interacción proteína-ligando cumple diversas funciones: 1) estructurales en el citoesqueleto, (Walker, Burgess et al. 2000; Schuyler and Pellman 2001) 2) en el metabolismo, por medio de complejos enzimáticos (Smith 1994; Sprang 1997) 3) en la transducción de señales, mediante la acción de integrinas, proteína G, canales de iones, etc. (Sprang 1997; Kleanthous 2000). Al producirse la interacción proteína-ligando se generan cambios conformacionales que permiten a la proteína adoptar su conformación activa, (o en el caso de inhibidores alostéricos salir de su conformación activa) estos pueden ser pequeños, como en el caso de algunos rotámeros en las cadenas laterales de algunos residuos aminoacídicos (Gerstein and Echols 2004), o grandes como en el caso de cambios en la disposición tridimensional de uno o más dominios de la proteína (Yon, Perahia et al. 1998; Gerstein, Echols et al. 2003). Los cambios conformacionales de los complejos proteína-ligando son dependientes de la temperatura, así estos no ocurren a bajas temperaturas, pues a pesar de que la proteína posee su estructura nativa no es capaz de adoptar una conformación activa y pierde su funcionalidad. A medida que la temperatura aumenta la proteína recupera su actividad. La mínima temperatura a la cual la proteína puede adoptar su conformación activa se denomina "temperatura de transición dinámica (T_c)" (Parak 2003). El mismo comportamiento se ha observado en

proteínas de organismos termófilos las que no son activas a temperatura ambiente, sin embargo son totalmente funcionales a las temperaturas óptimas de desarrollo del microorganismo (Vihinen 1987; Závodszy, Kardos et al. 1998). Estos resultados sugieren que debe existir un cierto grado de flexibilidad en la estructura de la proteína que le permita adoptar la conformación activa una vez unido el ligando. Otra característica de los cambios conformacionales es su reversibilidad, es decir una vez liberado el ligando la apo-proteína recupera su conformación, por lo tanto los movimientos en la estructura 3D de la proteína al formar el complejo proteína-ligando, no deben afectar mayormente su estabilidad. Estudios de intercambio de deuterio seguidos por NMR en proteínas con y sin ligando y simulaciones de dinámica molecular, mostraron que las regiones en la estructura 3D de las proteínas que participan en cambios conformacionales son zonas que presentan cooperatividad local en la estructura y no cooperatividad global, es decir son zonas que se pueden mover de manera independiente del resto de la proteína sin afectar mayormente la estabilidad global de la estructura (Hilser, Dowdy et al. 1998; Freire 1999; Kern and Zuiderweg 2003).

Debido a la importancia que tienen los cambios conformacionales en la función de las proteínas y en la traducción de la unión del ligando a una función, poder predecir los cambios conformacionales es un área de gran interés en bioinformática estructural y en ingeniería de proteínas (Mizoue and Chazin 2002). Ejemplos en la literatura sugieren que proteínas de una misma familia suelen tener cambios conformacionales similares, es el caso de las integrinas, receptores de la membrana externa de las células que al unir ligandos como fibronectina generan una respuesta celular mediada por el desplazamiento de más de 10 Å de una hélice del dominio MIDAS (Xiao, Takagi et al.

2004). Otro ejemplo, son las proteínas G, que muestran cambios conformacionales en las regiones "switch I" y "switch II" tras la hidrólisis de GTP (Sprang 1997). En proteínas solubles, cambios conformacionales similares se han observado por NMR en lisozima humana y de *Gallus domesticus* (Mine, Ueda et al. 2000). Estos resultados en su conjunto, sugieren que una manera de predecir los cambios conformacionales dentro de una familia de proteínas, podría hacerse a partir de información de otras proteínas de la familia. Esta hipótesis abre otras interrogantes, como: 1) ¿existe alguna relación entre la identidad de secuencia y la conservación de los cambios conformacionales en una familia de proteínas?, resultados de Todd et al, y Aloy y Russell (Todd, Orengo et al. 2001; Aloy and Russell 2002) muestran que proteínas de la misma familia con identidades de secuencias menores al 30% pueden presentar funciones enzimáticas distintas e incluso no ser enzimas; 2) ¿presentan distintos sitios de unión los mismos cambios conformacionales en una familia de proteínas?, si suponemos que los cambios conformacionales dependen de la unión del ligando a la proteína, distintos sitios de unión o diferentes funciones en una familia podrían tener cambios conformacionales disímiles.

Una manera de conocer las regiones que participan en los cambios conformacionales, es comparar la estructura tridimensional de la proteína libre y acomplejada (Domingues, Rahnenfuhrer et al. 2004). Esta aproximación se ha usado en proteínas donde los cambios conformacionales son bastante evidentes (Gerstein, Echols et al. 2003), sin embargo no está claro si es factible usar esta aproximación con cambios estructurales pequeños, donde los contactos cristalográficos y la distinta resolución de los cristales de proteínas pueden interferir en la comparación del estado libre y acomplejado (Luque and Freire 2002; Andre, Nonato et al. 2005).

La FtsZ

FtsZ es una proteína de dos dominios, uno amino terminal con un plegamiento tipo "Rossmann" y un sitio de unión a GTP, y uno carboxilo terminal con un plegamiento similar a la Corismato Mutasa (Oliva, Cordell et al. 2004; Osawa and Erickson 2005) (ver fig. 1). FtsZ forma el anillo Z, principal componente del divisoma bacteriano (Lutkenhaus and Addinall 1997; Weiss 2004). La formación del anillo se confirmó por microscopía de fluorescencia, pues FtsZ se ubica en el ecuador de la bacteria durante la división celular y forma una estructura parecida a un anillo asociado a la membrana interna. Este anillo se contrae para producir la división celular y posteriormente vuelve a formarse en cada una de las células hijas, para comenzar con un nuevo ciclo de división (Weiss 2004; Goehring and Beckwith 2005). Para formar el anillo, FtsZ debe interactuar consigo misma de manera longitudinal dando origen a protofilamentos, visibles por microscopía electrónica (Oliva, Huecas et al. 2003; Oliva, Cordell et al. 2004). *In vitro* los protofilamentos interactúan lateralmente y dan origen a dobles filamentos que se asocian para formar sábanas. Una vez formado el anillo se produce la interacción de ZipA y FtsA y de otras dos proteínas reguladoras ZapA y SulA, donde las dos primeras fijarían el anillo a la membrana, además FtsA reclutaría a las proteínas que participan en etapas más tardías de la formación del divisoma. La proteína ZapA actuaría como un activador y SulA actuaría como un inhibidor de la formación del anillo (Mosyak, Zhang et al. 2000; Pichoff and Lutkenhaus 2002; Cordell, Robinson et al. 2003; Goehring and Beckwith 2005). La función de FtsZ se regula por la hidrólisis de GTP, así FtsZ-GTP polimeriza para formar protofilamentos que *in vitro* se asocian en dobles filamentos que dan origen a sábanas, la hidrólisis del fosfato gama del GTP

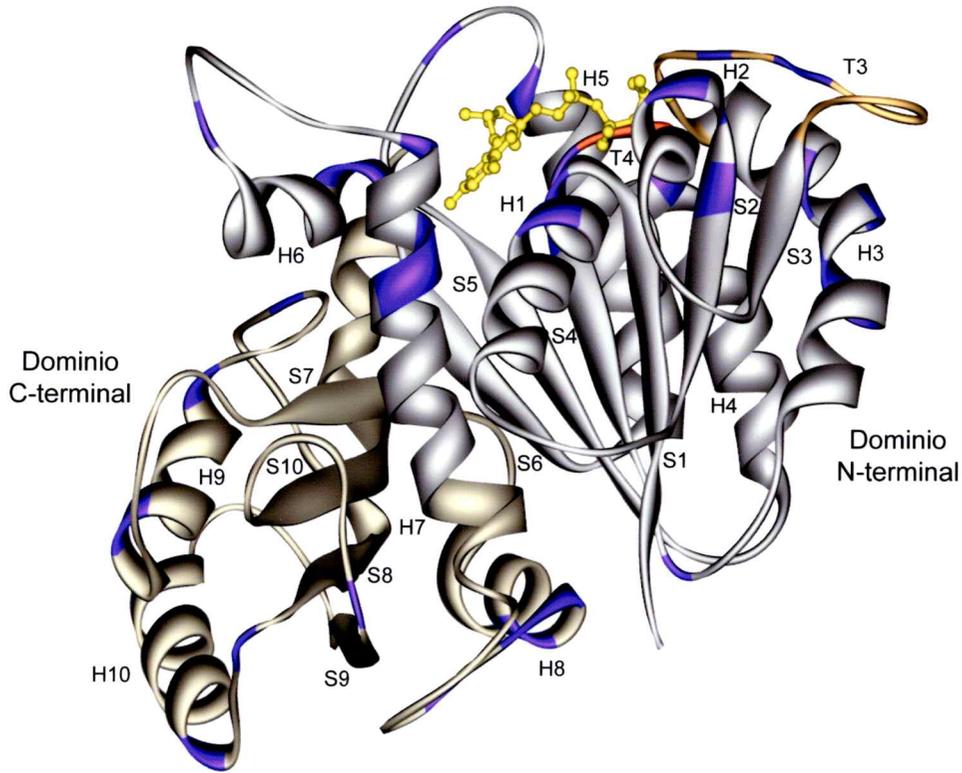


Figura 1. Modelo por homología de EcFtsZ, basado en la estructura cristalográfica de MjFtsZ (1fsz.pdb). La estructura 3D de un modelo de EcFtsZ está representada en esquema de cintas, se indican los dominios N-terminal y C-terminal, en gris y marrón, respectivamente; el nucléotido está representado en amarillo. Se muestran los lazos T3 y T4 en anaranjado y rojo, respectivamente. La ubicación de algunas mutaciones puntuales en la superficie de EcFtsZ, que alteran la función de la proteína se muestran en morado (Bi and Lutkenhaus 1990; RayChaudhuri and Park 1992; Bramhill and Thompson 1994; Dai, Mukherjee et al. 1994; Wang, Huang et al. 1997; Trusca, Scott et al. 1998; Lu, Stricker et al. 2001; Mukherjee, Saez et al. 2001; Scheffers, de Wit et al. 2001). La nomenclatura usada para indicar los elementos de estructura secundaria en FtsZ muestra las hélices α como "H", las hebras β como "S" y los lazos como "T" (Nogales, Downing et al. 1998).

induce la despolimerización (ver figura 2). Por lo tanto existirían a lo menos dos conformaciones funcionales de FtsZ, la conformación GTP y la conformación GDP (Lu, Reedy et al. 2000; Mingorance, Rueda et al. 2001; Oliva, Cordell et al. 2004).

Las regiones en la estructura de FtsZ que participan en los cambios conformacionales asociados al paso de la conformación GTP a GDP, se han obtenido a partir de la comparación de las coordenadas 3D de FtsZ de *Methanococcus jannaschii* (MjFtsZ) o *Mycobacterium tuberculosis* (MtbFtsZ), en estado GTP y en estado GDP. Los cristales de ambas estructuras son muy similares, y solo muestran un efecto del estado de fosforilación del nucleótido sobre regiones cercanas al sitio activo, es el caso del lazo T3 y la hélice H2 (ver figura 1). El lazo T3, no está unido directamente al nucleótido y muestra, en MtbFtsZ, una transición a una conformación ordenada solo en estado GTP (Leung, White et al. 2004; Oliva, Cordell et al. 2004). Experimentos de fluorescencia, donde se mutó una treonina en el lazo T3 por triptófano, y simulaciones de dinámica molecular, también sugieren un cambio conformacional del lazo T3 asociado a la hidrólisis de GTP (Díaz, Kralicek et al. 2001). Estos resultados solo muestran parcialmente los cambios conformacionales que ocurren en FtsZ con la hidrólisis de GTP, pues no explican el efecto de un número importante de mutantes puntuales, superficiales y alejadas del sitio de unión del nucleótido, sobre la función de FtsZ de *Escherichia coli* (EcFtsZ, figura 1) (Lu, Stricker et al. 2001; Osawa and Erickson 2005; Redick, Stricker et al. 2005).

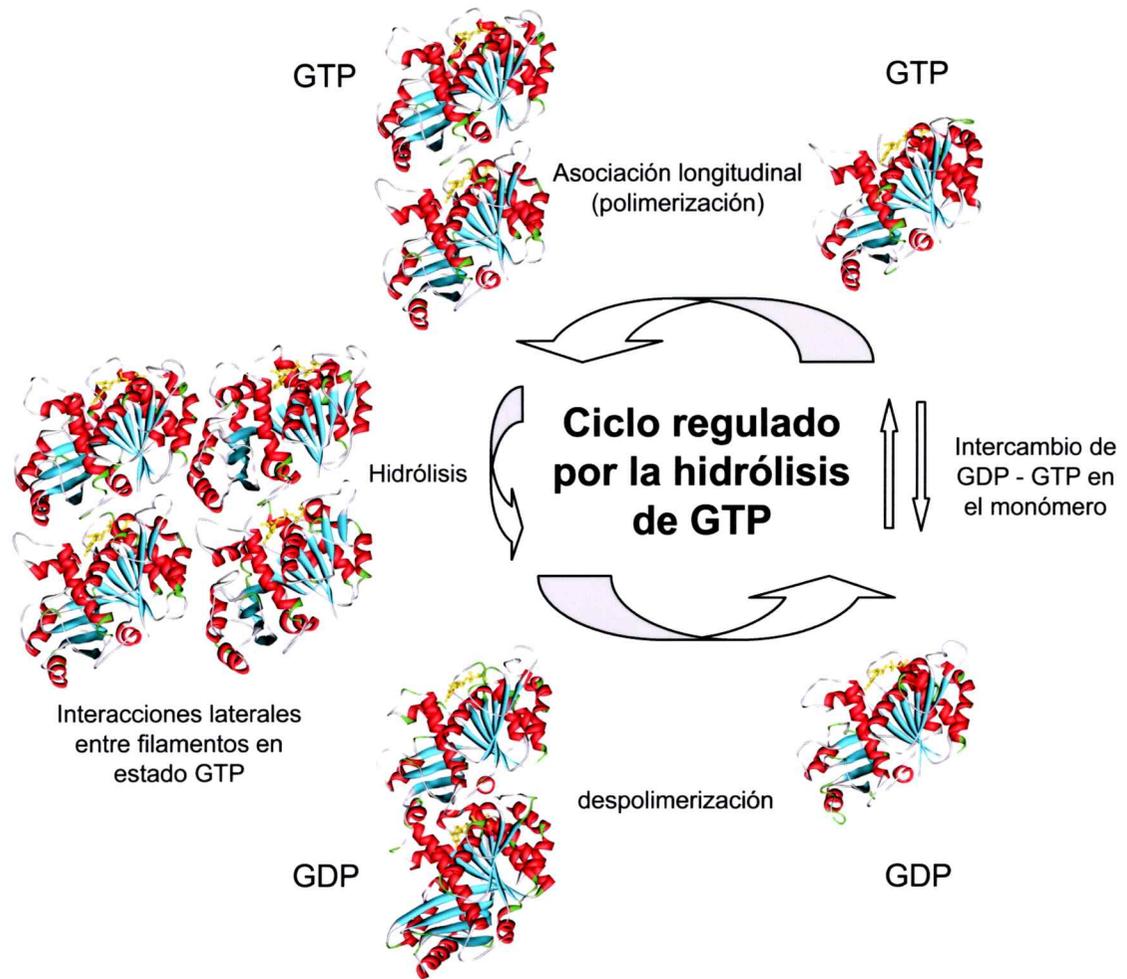


Figura 2. Esquema de la regulación funcional de FtsZ mediante la hidrólisis e intercambio de GDP – GTP. Las estructuras de FtsZ en sus distintos grados de oligomerización están representadas en cintas, donde hélices, hebras beta, vueltas y lazos se muestran en rojo, celeste, verde y gris, respectivamente.

Similitudes estructurales y funcionales en la familia de las tubulinas

Las proteínas α , β , γ , δ tubulinas, las FtsZ y recientemente la BtubA y la BtubB forman una familia de GTPasas (Erickson 1995; Sontag, Staley et al. 2005) caracterizadas por el motivo de unión a nucleótido GGTG (lazo T4 de FtsZ). Las proteínas de esta familia con estructura conocida son: el heterodímero $\alpha\beta$ -tubulina determinado mediante difracción de electrones con una resolución de 3,7 Å (Nogales, Wolf et al. 1998); FtsZ de *Methanococcus jannaschii* (MjFtsZ, 1fsz.pdb) determinada por difracción de rayos X a una resolución de 2,6 Å (Lowe 1998; Lowe and Amos 1998); FtsZ de *Pseudomona aeuroginosa* (PaFtsZ, 1ofu.pdb), con una resolución de 2,1 Å (Cordell, Robinson et al. 2003); FtsZ de *Mycobacterium tuberculosis* (MtbFtsZ, 1rq7.pdb, 1rlu.pdb) (Leung, White et al. 2004), también obtenida por difracción de rayos X, con una resolución de 2,6 Å y Btub AB (2btq.pdb) con una resolución de 3,2 Å (Schlieper, Oliva et al. 2005). La similitud estructural de las FtsZ con el heterodímero de tubulina y el heterodímero Btub AB, comprende los dominios amino terminal y carboxilo terminal de las FtsZ y los dominios amino terminal e intermedio de α y β tubulina y Btub A y B. La superposición de los dominios amino terminal de FtsZ y tubulina o FtsZ y Btub, muestran valores de RMSD menores a 2,5 Å para 179 residuos, esta similitud estructural disminuye a 2,9 Å para 79 residuos en el dominio carboxilo terminal. La similitud estructural aumenta si se compara tubulina y Btub, con un RMSD de 1,7 Å para 231 residuos del dominio amino terminal y 1,8 Å en 86 residuos del dominio intermedio. Además, tubulina y Btub AB presentan un tercer dominio, carboxilo terminal, constituido por dos extensas α -hélices, no presentes en FtsZ. No obstante su similitud estructural las identidades de secuencia entre estas proteínas es baja, siendo las tubulinas y Btub AB las más similares con alrededor de un 35% de identidad de secuencia y las FtsZ las más disímiles respecto a

las anteriores, con alrededor de un 13% y un 10% de identidad de secuencia, respectivamente.

Funcionalmente, $\alpha\beta$ -tubulina y FtsZ participan en la división celular en eucariontes y procariontes, respectivamente. Más aun ambas proteínas polimerizan longitudinalmente, en presencia de GTP, formando protofilamentos que luego interactúan lateralmente para formar túbulos (microtúbulos), en el caso de la tubulina y sábanas o pares de filamentos en el caso de FtsZ (Erickson, Taylor et al. 1996). La unidad funcional de tubulina es el heterodímero $\alpha\beta$, en cambio en FtsZ es un monómero (Joshi, Palacios et al. 1992; Erickson 2000; Margolin 2000). La polimerización longitudinal de tubulina o FtsZ, en conformación GTP permite una activación "trans" de la actividad GTPásica, al interactuar la cadena lateral del ácido glutámico 254 de la α -tubulina con el sitio de unión de nucleótido en la β -tubulina del heterodímero vecino en el protofilamento, o de forma análoga, el ácido aspártico 238 de MjFtsZ activaría en "trans" la hidrólisis de nucleótido en el sitio de unión de GTP del monómero vecino en el polímero (Nogales, Whittaker et al. 1999). Si los monómeros que se adicionan al extremo del polímero están en conformación GDP, un cambio conformacional altera los contactos laterales entre los protofilamentos, los que se separan y se curvan para finalmente despolimerizarse (Mandelkow, Mandelkow et al. 1991).

La comparación de polímeros de tubulina, mediante estudios de criomicroscopía y "docking", en conformación recta (estado GTP) y en conformación curva (estado GDP), mostró cambios en las interfases intra e interdímero (Wang and Nogales 2005), los cuales serían responsables de la pérdida de los contactos laterales y posterior despolimerización de los microtúbulos. Para que la despolimerización de tubulina se inicie, los autores proponen 3 efectos de la hidrólisis de GTP: 1) induce cambios

alrededor del sitio de unión de nucleótido que afecta la interfase entre heterodímeros; 2) induce cambios conformacionales a larga distancia que relacionan las interfases intra e interdímero; 3) altera la conformación de la proteína para debilitar los contactos laterales. Debido a las similitudes entre tubulina y FtsZ, los cambios conformacionales sugeridos que comunicarían el sitio de unión de GTP y los contactos laterales en tubulina, también pueden estar presentes en FtsZ.

Las similitudes estructurales y funcionales que presentan proteínas de una misma familia, como es el caso de las tubulinas, y la conservación de regiones en la estructura involucradas en los cambios conformacionales en proteínas de las familias de las integrinas, proteína G y lisozima, nos han llevado a postular la hipótesis: "Los cambios conformacionales inducidos por ligandos en familias de proteínas, son conservados para un mismo sitio activo". Si este es el caso, la información cristalográfica que existe de FtsZ en estado GTP y GDP, en distintos organismos, nos permitiría predecir los cambios conformacionales en la familia de las tubulinas, compuesta por: FtsZ, α , β y γ tubulinas, Btub A y Btub B.

Objetivo General:

Determinar la similitud de los cambios conformacionales en familias de proteínas.

Objetivos Específicos:

- Obtener una base de datos de proteínas de estructura conocida en estado libre y acomplejadas.
- Determinar por métodos computacionales en complejos proteína – ligando pertenecientes a una misma familia, los cambios conformacionales relacionados a la unión de ligando y su grado de conservación a nivel de estructura tridimensional y a nivel de secuencia aminoacídica.
- Determinar el efecto de diferentes ligandos sobre los cambios conformacionales, en un mismo sitio activo y en distintos sitios activos.
- Caracterizar los residuos que participan en los cambios conformacionales.
- Determinar las regiones en la estructura de EcFtsZ involucradas en los cambios conformacionales inducidos por la hidrólisis de GTP.

MATERIALES Y METODOS

Construcción de una base de datos de pares de proteínas de estructura conocida en su estado libre y acomplejadas con ligando. Se seleccionó de la base de datos PDB (Berman, Westbrook et al. 2000), actualizada al 23 de mayo del 2004, aquellos registros que contienen más de 30 residuos y son estructuras determinadas por difracción de rayos X o NMR. Para seleccionar una proteína que aparece varias veces en la base de datos PDB, como por ejemplo: con y sin ligando, o con una estructura determinada por diferentes métodos, etc., se comparó la secuencia de la proteína con un banco de datos de secuencias elaborado a partir de la base de datos PDB, de acuerdo a la última información de estandarización de los registros PDB (Westbrook, Feng et al. 2002). Con este propósito se decidió utilizar el campo SEQRES, donde está la información de la secuencia de cada proteína con su número de cadenas y luego, se comparó con *blast* (Altschul, Madden et al. 1997) cada una de las cadenas contra la base de datos generada y se seleccionaron aquellos casos que mostraron un 100% de identidad de secuencia, y por ende todos los registros del PDB que involucraban a una misma proteína. Se excluyó aquellos casos redundantes, como por ejemplo una misma estructura determinada a distintas resoluciones, o por metodologías diferentes. Como criterio de selección se utilizaron las estructuras de mayor resolución y las determinadas por difracción de rayos X en lugar de las determinadas por NMR, cuando las coordenadas tridimensionales estaban disponibles por ambas técnicas (Fan and Mark 2003). El conjunto de datos obtenidos se separó en tres subconjuntos de interés: 1) Los

casos que tenían el mismo número e identidad de cadenas, se seleccionaron de acuerdo al campo HETATM buscando pares de proteínas con y sin un ligando, estos casos se agruparon en la base de datos "ligando n-1". 2) Otro subconjunto se generó con los casos que tenían el mismo número e identidad de cadenas y el mismo número de ligandos donde uno de ellos era un subconjunto del otro, por ejemplo : la proteína A cristalizada con GTP y la proteína A cristalizada con GDP. Estos casos se agruparon en la base de datos "ligando parcial". 3) El tercer subconjunto agrupó interacciones proteína-proteína (proteína RNA y DNA) y se seleccionó a partir de aquellos casos donde todas las cadenas en un archivo PDB y todos sus ligandos estaban presentes en el segundo archivo, sin embargo el segundo archivo presentaba otras cadenas adicionales, estas proteínas se agruparon en la base de datos "proteína – proteína".

Determinación de los contactos de los ligandos en la proteína. Los contactos que presenta el ligando en la proteína se obtuvieron con el programa complex-info de Ernesto Moreno (Sanchez 2003). Este programa discrimina por distancia si el ligando está unido covalentemente o no a la proteína. Si la distancia entre cualquier átomo del ligando y cualquier átomo de la proteína es menor que 1,9 Å supone una unión covalente, como por ejemplo en el caso de una metilación. Esta distancia se obtuvo del mejor compromiso que se encontró al analizar varias familias de la base de datos "ligando n-1". Además este programa entrega una lista de todos los residuos de aminoácidos que están en contacto con el ligando (menos de 6 Å), definido en esta tesis como el sitio activo (SA).

Comparación de las coordenadas de los carbono alfa ($C\alpha$) y los carbono beta ($C\beta$) en la apo-proteína y la proteína-ligando. Para comparar las proteínas con y sin ligando, se extrajo la información de su secuencia de los campos ATOM de los respectivos archivos PDB. Luego se alinearon ambas secuencias mediante el alineamiento de pares de secuencias de Blast, y se procesó el alineamiento con un programa en perl que usa la rutina align3d de Modeller (Sali and Blundell 1993). Esta rutina calcula las distancias entre los $C\alpha$ equivalentes de la apo-proteína y el complejo proteína-ligando $(C\alpha_{\text{sin ligando}} - C\alpha_{\text{con ligando}})/2$ y entrega una tabla con los resultados. Con esta tabla se construyeron los perfiles de distancia para cada par de proteínas, donde se compararon las coordenadas de la proteína libre y acomplejada. De la misma manera, se determinó los perfiles de distancia para los $C\beta$.

Alineamiento múltiple para cada familia. Se ensayó varios métodos de alineamiento múltiple automático para cada familia, y los mejores resultados se obtuvieron con un protocolo que seleccionó las estructuras de las proteínas sin ligando, desde los pares de proteína libre y con ligando. Con ellas se generó un alineamiento múltiple de estructuras con la rutina "malign3d" del programa Modeller. Luego se usó este alineamiento como un bloque (no se altera el alineamiento interno entre estas secuencias) y se alineó iterativamente cada una de las secuencias de las proteínas restantes contra el bloque, mediante la rutina "align". Como restricción impuesta a este protocolo, se descartaron aquellas secuencias que tuvieran un largo mayor que 1,5 veces la secuencia más larga en el alineamiento de bloque. Para los casos de familias en que no habían estructuras sin ligando para generar el primer alineamiento de bloque, se usó las estructuras con un solo ligando para producirlo, y posteriormente al igual que en el caso anterior se agregó una a una las secuencias de las proteínas restantes. Si este tampoco era el caso,

simplemente se usó la rutina "malign" de Modeller para generar un alineamiento múltiple con todas las secuencias de la familia de proteínas.

Cambios conformacionales relacionados a la unión de ligando (LRCC). Los aminoácidos que participan en los cambios conformacionales se obtuvieron a partir de las tablas de distancias de $C\alpha$ o $C\beta$ para cada proteína seleccionada con y sin ligando. A cada uno de estos perfiles de distancias se les determinó de forma independiente el ruido. Para determinar el ruido, se calculó el promedio con todas las distancias entre aminoácidos estructuralmente equivalentes de la tabla, luego se eliminó todos los valores de distancias mayores a dos veces el promedio, y se volvió a calcular el promedio, esto se hizo en forma iterativa hasta que el número de distancias usadas para el cálculo del promedio no varió. El ruido se definió como dos veces el valor final del promedio. Se definió un residuo de amino ácido como perteneciente a un LRCC (Ligand Related Conformational Change) si la distancia $C\alpha_{\text{sin ligando}} - C\alpha_{\text{con ligando}}$ o $C\beta_{\text{sin ligando}} - C\beta_{\text{con ligando}}$ es mayor que dos veces el ruido.

Factor B normalizado. El factor B o factor térmico, se relaciona con el promedio al cuadrado del desplazamiento en torno a un punto medio de un átomo e informa sobre la flexibilidad en la estructura determinada por difracción de rayos X. Este factor depende de las medidas experimentales y del refinamiento posterior de los datos. En algunas estructuras se utiliza otro factor denominado U que esta directamente relacionado con B por la expresión:

$$B = 8\pi^2 U^2$$

Para poder comparar los valores del factor B entre distintas proteínas y obviar aquellos casos en que se informa U, se normalizó el factor B, para cada proteína. La normalización consistió en considerar solamente los carbono alfa de los aminoácidos de

la proteína, y a cada valor de factor térmico restarle el valor promedio de factor B y dividirlo por la desviación estándar (Vihinen 1987; Parthasarathy and Murthy 1997; Lovell, Word et al. 1999; Yuan 2003).

Análisis de χ^2 para determinar la independencia de dos valores. El análisis de independencia de dos valores mediante la prueba de χ^2 , permitió evaluar si el valor encontrado para un determinado parámetro dentro de la base de datos de cambios conformacionales, por ejemplo: la abundancia del aminoácido glicina en regiones de la estructura que participan en cambios conformacionales (valor observado) es significativamente diferente al valor esperado para ese parámetro, por ejemplo: la abundancia natural del aminoácido glicina en las proteínas (valor esperado). La probabilidad de χ^2 se calculó de acuerdo a la ecuación:

$$\chi^2 = \frac{(\text{valor}_{\text{observado}} - \text{valor}_{\text{esperado}})^2}{\text{valor}_{\text{esperado}}}$$

El valor obtenido de la ecuación anterior se comparó con el valor de χ^2 encontrado en tablas para un 95% de confianza y un grado de libertad. Si el valor calculado es mayor que el encontrado en las tablas la diferencia es significativa (Spiegel 1988).

Análisis de χ^2 para determinar si dos conjuntos de datos presentan distribuciones diferentes. El análisis de χ^2 permitió comparar estadísticamente dos conjuntos de datos, independiente de su distribución (análisis no paramétrico) (Press; Carugo and Pongor 2002). Debido a la imposibilidad de determinar con un 100% de certeza que dos conjuntos de datos no pertenecían a una misma distribución, se entregó como criterio de comparación una probabilidad de χ^2 . Una probabilidad de 1, se interpretó como que ambos conjuntos de datos provienen de una misma distribución, y a medida que el valor

de probabilidad se aproximó a cero, la probabilidad de que fueran distintos aumentó. El programa perl que se usó para calcular esta probabilidad, separó la distribución de datos en intervalos (bin), donde cada intervalo tuvo a lo menos un 5% de los datos, para posteriormente comparar los intervalos correspondientes entre los dos conjuntos de datos.

Número de capas de aminoácidos en la estructura de la proteína respecto a los residuos que constituyen el sitio activo (SA). La vecindad de los residuos de la proteína o de los residuos que participan en los LRCC respecto a aminoácidos del sitio activo, se determinó usando la rutina "write_data" del programa Modeller. Esta rutina define como aminoácidos de una capa contigua a todos aquellos residuos que tengan uno o más átomos a menos de 6 Å de distancia de los aminoácidos de referencia. Capa 0, en el análisis del total de residuos de la proteína, es el número total de residuos que forman el SA, en cambio, capa 0 en el análisis de LRCC se refiere al número de aminoácidos del SA que también forman parte de los LRCC. Capa 1, se define como todos aquellos residuos que están en directo contacto (menos de 6 Å) de cualquier aminoácido del SA. Capa 2 abarca a todos los residuos en contacto con la primera capa de vecinos (capa 1) y así sucesivamente.

Análisis de componentes principales (PCA). El análisis de PCA nos permitió comparar los residuos estructuralmente equivalentes que participan en los cambios conformacionales dentro de una familia de proteínas. La información de cambios conformacionales de cada proteína se codificó en su secuencia aminoacídica, donde los residuos que participan en los LRCC se anotan como una alanina (A) en la secuencia y los demás residuos se anotan como una leucina (L). Las secuencias codificadas se reemplazan en el alineamiento múltiple de la familia, agregándose las inserciones

correspondientes, y luego se comparan entre sí mediante la rutina de Modeller "ID_TABLE", esta rutina entrega una matriz donde están ordenadas todas las comparaciones de identidad de secuencia entre cada par de proteínas. Estos datos se analizan mediante la rutina "PRINCIPAL_COMPONENTS", que determina las correlaciones entre los valores de identidad de secuencia de las proteínas. El gráfico de componentes principales permite discriminar si los LRCC son similares para todas las proteínas de la familia o si existen LRCCs distintos para subgrupos de proteínas. Esta estrategia, para comparar los LRCC, tiene la ventaja de evaluar las zonas equivalentes que se mueven en la estructura de las proteínas, y a su vez evaluar la concordancia de regiones que no se mueven.

Determinación de estructura secundaria. La asignación de estructura secundaria por aminoácido, a partir de las coordenadas cristalográficas de las proteínas acomplejadas, se realizó mediante el programa en Fortran de Suvobrata Chakravarty, posdoc en el laboratorio del Dr. Roberto Sánchez, Departamento de Fisiología y Biofísica de la Escuela de Medicina de Mount Sinai, New York, USA. El programa asigna la estructura secundaria sobre la base de los ángulos diedros ϕ y ψ de la cadena principal, no asigna los residuos amino y carboxilo terminal, y se requieren a lo menos 4 residuos consecutivos para identificar estructuras secundarias de hélice α o hebra β .

Grado de exposición de los aminoácidos. El grado de exposición de cada aminoácido se determinó mediante el programa NACCESS (Hubbard and Thornton 1993) que usa el método de Lee y Richards 1971 (Lee and Richards 1971). El método consiste en calcular el ASA (accessible surface area) de cada aminoácido en la proteína respecto al ASA del mismo aminoácido en un tripéptido glicina-X-glicina. Si se obtiene una razón $\geq 0,4$, el

residuo se considera expuesto; si la razón es $< 0,05$, el aminoácido se considera oculto, y valores entre ambos límites se definen como accesibilidad intermedia.

Diseño de mutantes. A partir de un alineamiento múltiple de 474 FtsZ, mediante un análisis de conservación, se diseñó 3 mutaciones puntuales en la estructura de EcFtsZ (fig.3). El criterio fue buscar posiciones en el alineamiento múltiple, con alta frecuencia de aminoácidos aromáticos y expuestas al solvente, según un modelo 3D de la estructura de EcFtsZ. La elección de una sustitución puntual por triptófano, se debió a que EcFtsZ no tiene triptófanos en su estructura primaria y con estas mutantes se podía caracterizar el entorno por espectroscopia de fluorescencia y establecer relaciones entre la estructura y función de EcFtsZ. Todas las mutantes de EcFtsZ presentan algún grado de exposición al solvente, la primera F135W, está cercana al SA en el dominio amino terminal, y forma parte de los aminoácidos involucrados en los cambios conformacionales asociados a la hidrólisis de GTP (ver figura 12). F135 mostró en el alineamiento múltiple, más de un 99% de conservación de residuos aromáticos. En el dominio carboxilo terminal, se sustituyó por triptófano el residuo I294, posición que presenta un 50% de carácter aromático en el alineamiento de FtsZ y se sugiere que participa en los cambios conformacionales (ver resultados). La tercera mutación F40W, se diseñó como control, pues está localizada en el dominio amino terminal en una región más rígida de la estructura de FtsZ, que no participaría en los cambios conformacionales asociados a la hidrólisis de GTP.

Modelo 3D de EcFtsZ. Por modelado comparativo de proteínas se obtuvo un modelo de la estructura 3D de EcFtsZ a partir de el molde MjFtsZ (1fsz), proteína que presenta un 40% de identidad de secuencia con la secuencia a modelar. El modelo de EcFtsZ comprende desde la posición 9 de la secuencia aminoacídica a la posición 324,

con su sitio activo ocupado por el nucleótido GDP. La confiabilidad del modelo de EcFtsZ se evaluó mediante el programa PROSAIL (Sippl 1993).

La mutante I294W se modeló de la misma manera que EcFtsZ, con la ventaja adicional que el molde usado, MjFtsZ, tiene un triptófano en la posición que alinea con I294, esto permitió que el programa Modeller copiara el rotámero y el ambiente del triptófano de MjFtsZ en el modelo de EcFtsZ.

Análisis y representación de resultados. Los gráficos de densidad y componentes principales se realizaron con el programa ASGL (Sanchez 2003), para los demás gráficos se usó el programa Sigmaplot 2001, de Systat software Inc.

Las representaciones de proteínas que incluyen superficies de van der Waals se realizaron con el programa MolMol 2K.2 (Koradi, Billeter et al. 1996), para el resto de las representaciones se usó el programa DS Visualizer 1.5, de Accelrys Software Inc, seguido de una optimización de la imagen con el programa POV-Ray 3.6, de Persistente Vision.

Programas y bases de datos. La base de datos de proteínas de estructura conocida (PDB) se actualizó semanalmente a través del sitio web del Protein Data Bank (<ftp.rcsb.org>) y fue el punto de partida para obtener la base de datos de cambios conformacionales (proteínas en conformación libre y acomplejada), utilizada en el desarrollo de esta tesis. Además de la base de datos PDB, se utilizó las bases de datos de secuencias no redundantes nr y pdbaa, actualizadas semanalmente en los servidores del laboratorio, a partir del sitio Web del NIH (<ftp.ncbi.nih.gov>). La automatización de los análisis efectuados sobre la base de datos de cambios conformacionales, ya sea para análisis sencillos que solo requerían el uso de un servidor, o para análisis más

complejos que requerían un procesamiento de datos distribuido, se utilizó un “cluster” de diez servidores dual PIII de 1 GHz, con programas escritos en lenguaje Perl.

Mutagénesis sitio-dirigida. Para realizar la mutagénesis sitio-dirigida se utilizó el plasmidio pMFV57, construido a partir del plasmidio pMFV56 (Rivas, Lopez et al. 2000) donde el gen de la beta-lactamasa se clonó en el sitio *EcoR* V interrumpiendo el gen de kanamicina. Las mutantes en EcFtsZ se construyeron mediante mutagénesis sitio-dirigida con el “kit” comercial *Quick Change® XL Site Directed Mutagenesis* (Stratagene). Todos los plasmidios fueron purificados con el “kit” *E.Z.N.A® Plasmid Miniprep I* (Omega Bio-tek). Cada gen de *ftsZ* con su respectiva mutación puntual fue totalmente secuenciado.

Sobre-expresión y purificación de EcFtsZ silvestre y mutantes. La EcFtsZ recombinante y las mutantes puntuales se sobre expresaron en cepas de *E.coli* BL21(DE3) por inducción con IPTG. Las células se crecieron en 2 l de caldo LB con ampicilina 100 µg/ml a 37 °C y con agitación, hasta un valor de DO₆₀₀ de 0,7, luego se agregó al cultivo 0,4 mM de IPTG y las células se continuaron creciendo durante 4 horas más. Para la purificación de cada proteína mutante y de EcFtsZ silvestre se utilizó el método de glutamato (Beuria, Krishnakumar et al. 2003), adaptado para el laboratorio. Este sistema se basa en la capacidad de EcFtsZ de polimerizar *in vitro* en presencia de GTP. El efecto del glutamato sobre la polimerización es estimular la formación de manojos de filamentos de FtsZ. Esto facilita su precipitación y posterior purificación, además de disminuir la pérdida de proteína en el proceso. De manera resumida, el método de purificación consistió en centrifugar las células a 5000 rpm en un rotor SORVALL H6000A por 45 min, el precipitado se suspendió en amortiguador A (Tris-HCl 50 mM pH 8, KCl 50mM, EDTA 1mM y glicerol 10%) y se lisó las bacterias por

sonicación mediante 5 pulsos de 45 watts y 2 pulsos de 75 watts. Las células lisadas se centrifugaron a 92000 g por 90 min, y con el sobrenadante se efectuó un corte con sulfato de amonio al 25% de saturación. La solución se centrifugó a 9391 g a 4 °C, el precipitado se disolvió en amortiguador A y se dializó toda la noche contra el mismo amortiguador para eliminar el exceso de sal. Se purificó la muestra por cromatografía de intercambio iónico mediante un equipo de HPLC (Beckman System Gold), utilizando un gradiente de KCl para soltar la proteína retenida en la columna (amortiguador B: Tris-HCl 50 mM pH 8, KCl 50 mM, EDTA 1 mM, KCl 1 M y glicerol 10%). Con la FtsZ obtenida, se realizó un ciclo de polimerización y despolimerización en una solución de Pipes 50 mM, glutamato 1 M, CaCl₂ 10 mM, MgCl₂ 10 mM, GTP 2 mM a 37 °C durante 30 min y se centrifugó los filamentos resultantes a 9391 g durante 30 min a 37 °C. La proteína se suspendió en 5ml de amortiguador A, y se lavó 2 veces en el mismo amortiguador a 4 °C. A esta temperatura, los polímeros se desarman y la proteína queda principalmente en estado monomérico. La fase soluble que contenía la proteína se concentró con el sistema Centriprep y almacenó a -80 °C. La concentración de la proteína purificada y el contenido de nucleótido se determinó por absorbancia a 280 nm y a 254 nm en 6 M de GdmCl con amortiguador fosfato pH 6,5 (Rivas, Lopez et al. 2000). Además, se utilizó el método de Bradford adaptado para EcFtsZ como sistema secundario de medición de la concentración. El contenido de nucleótido se determinó por HPLC en una columna SUPELCOSIL™ LC-18-D13.

Ensayo de complementación *in vivo*. Para los ensayos de complementación, se utilizó la cepa de *E. coli* VIP2- λ DE3. En esta cepa la FtsZ cromosomal se interrumpió mediante una inserción, y se complementó con una copia del gen de la proteína en un plasmidio con un origen de replicación termo-sensible pLAR9, lo que impide la

sobrevivencia del vector a temperaturas mayores a 37 °C. El vector posee resistencia al antibiótico kanamicina, que se usa como medio de selección. Cultivos de VIP2(DE3)/pLAR9 crecidos durante la noche a temperatura permisiva (30 °C), se transformaron por electroporación con plasmidios que contenían EcFtsZ silvestre o los genes de *ftsZ* mutantes descritos anteriormente. Se cultivó las células transformadas en 5 ml de LB-Ampicilina-Kanamicina-Cloranfenicol, y se creció a 30 °C con agitación hasta alcanzar una O.D. de 0,6. Diluciones seriadas de cada cultivo se plaquearon a ambas temperaturas, la permisiva (30 °C) y la no-permisiva (42 °C) en LB-Ampicilina-Kanamicina-Cloranfenicol y luego se contó el número de colonias.

Dispersión de la luz asociada a la polimerización de EcFtsZ. Se utilizó un espectrofluorímetro Perkin-Elmer LS-50 para determinar la variación de la dispersión de luz asociada a la formación de los polímeros. Para las mediciones se usó una longitud de onda de excitación de 350 nm y atenuadores al 1%. Cada proteína EcFtsZ mutante y silvestre se pre-incubó en una solución de polimerización que contenía MES 50 mM pH 6,5, KCl 50 mM y MgCl₂ 10 mM, a 30 °C. La reacción se inició agregando GTP 1 mM.

Actividad GTPásica. Se incubó la proteína (12 µM) a 30 °C por 3 min en una solución que contenía MES 50 mM pH 6,5, KCl 50 mM, MgCl₂ 10 mM. La polimerización se inició agregando GTP 1 mM a la solución y posteriormente se detuvo a diferentes tiempos agregando 10% de ácido perclórico, se neutralizó la reacción con KHCO₃ 1M y se removió las burbujas de CO₂ con 5 min de centrifugación en Speed-Vac. La proteína denaturada y la sal insoluble de KClO₄ se removió por centrifugación a 17530 g por 10 min, y el sobrenadante se analizó por HPLC en una columna SUPELCOSIL™ LC-18-D13. La separación de GDP y GTP se realizó usando una fase móvil que contenía 4 mM TBAP (Tetrabutyl fosfato hidrógeno de amonio), 0,2 M de K₂HPO₄ y 0,1 M de ácido

acético glacial. A partir de las concentraciones de GDP calculadas y normalizadas por la concentración total de nucleótido, se construyó las curvas de progreso para cada mutante, y se determinó la velocidad inicial de hidrólisis.

Dicroísmo circular (DC) y espectroscopia de fluorescencia. El contenido de estructura secundaria para EcFtsZ silvestre y las mutantes se determinó mediante DC en un espectropolarímetro JASCO J-600 usando celdas de 1 mm de paso óptico. El espectro de DC en el UV-lejano (190 nm hasta 260 nm) de cada muestra, se obtuvo con 4 barridos, usando una velocidad de 20 nm/min, con una sensibilidad de 1 nm y una constante de tiempo de 2 s. Los espectros de emisión de fluorescencia se obtuvieron con un espectrofluorímetro Perkin Elmer LS-50. La metodología consistió en obtener 3 barridos por muestra, usando una excitación de 295 nm y una velocidad de barrido de 140 nm/min. El análisis de los espectros de DC y emisión de fluorescencia se realizó con el programa GRAMS™.

RESULTADOS

Caracterización de los cambios conformacionales relacionados a la unión de ligandos pequeños.

Base de datos de pares de proteínas en estado libre y como complejo proteína-ligando. El estudio teórico de los cambios conformacionales relacionados a la unión de ligando (LRCC), es decir la identificación de los residuos que cambian las coordenadas de su $C\alpha$ al pasar la proteína del estado libre a complejo, requirió encontrar el mayor número de proteínas con su estructura conocida en ambos estados, en la base de datos PDB, y por medio del sistema de clasificación CATH (Orengo, Michie et al. 1997) agruparlas de acuerdo a su tipo de plegamiento en familias de proteínas.

Una vez obtenidos los conjuntos de proteínas del tipo "ligando n-1", "ligando parcial" y "proteína-proteína" (ver métodos) se trabajó con el sistema más simple posible, con este propósito se seleccionaron desde el banco PDB las proteínas con "ligando n-1" que tenían una sola cadena peptídica (ver tabla 1). No se consideraron los casos de "ligando parcial" porque se encontró solo un tercio de la población de "ligando n-1" y el número de proteínas en cada familia no permitió analizar los LRCC en un rango de distintas identidades de secuencia. Tampoco se estudiaron los complejos proteína-proteína debido a la dificultad de determinar si realmente existe una interacción proteína-proteína y no se debe a un problema de contactos cristalográficos (Jones and Thornton 1996; Henrick and Thornton 1998; Kleanthous 2000), lo que implica un análisis más exhaustivo en cada caso para establecer si la interacción existe o no.

Tabla 1. Principales familias de la base de datos de proteínas en estado libre y como complejo proteína-ligando.

CATH	Proteínas	Topología	Homología
1.10.238.10	12	Dominio 1, Recoverina	Mano EF
1.10.246.10	14	Dominio 1, cadena A, Seroalbumina	ND
1.10.420.10	34	Dominio 2, Peroxidasa	ND
1.10.490.10	64	Tipo Globina	Globinas
1.10.510.10	40	Dominio 1, fosfotransferasa	ND
1.10.530.10	23	Lisozima	ND
1.10.530.40	33	Lisozima	ND
1.10.630.10	23	Citocromo P450	ND
1.20.90.10	8	Fosfolipasa A2	ND
2.120.10.10	21	Neuraminidasa	ND
2.20.25.50	16	Dominio N-terminal de Tfilb	ND
2.40.10.10	48	Subunidad H, Trombina	Serino proteasas tipo Tripsina
2.40.128.20	25	Subunidad I, Serratia Metalo proteinasa	ND
2.60.120.200	25	Jelly Rolls	ND
3.10.130.10	10	Protelna P-30	Protelna P-30
3.10.200.10	44	Anhidrasa Carbonica II	Anhidrasa Carbonica II
3.10.50.40	9	Dominio 3, Quitinasa A	ND
3.20.20.80	23	Tim Barrel	Glicosidasas
3.20.20.90	14	Tim Barrel	Triosa Fosfato Isomerasas *
3.30.390.30	8	Dominio 1, tipo Enolasa	ND
3.30.420.10	9	Dominio 5, Nucleotidiltransferasa	ND
3.30.450.20	8	Beta-Lactamasa	ND
3.30.572.10	10	Cadena A, Timidilato Sintasa	Cadena A, Timidilato Sintasa
3.40.390.10	10	Dominio catalítico, Colagenasa	Dominio catalítico, Colagenasa
3.40.420.10	12	Dominio 1, subunidad A, Ricina	Dominio 1, subunidad A, Ricina
3.40.430.10	19	Subunidad A, Dihidrofolato Reductasa	Dihidrofolato Reductasa
3.40.50.1820	27	Plegamiento Rossmann	ND
3.40.50.200	10	Plegamiento Rossmann	ND
3.40.50.300	17	Plegamiento Rossmann	"P-loop" triphosphate hydrolase
3.40.50.720	27	Plegamiento Rossmann	Dominio de union a NAD (P)
3.40.630.10	14	Aminopeptidasa	Zn peptidasas
3.90.190.10	16	Subunidad A, Tirosina Fosfatasa	Subunidad A, Tirosina Fosfatasa

*Otras funciones enzimáticas conocidas para las proteínas que constituyen la familia de triosa fosfato isomerasas, son: FMN – oxidoreductasa, enzimas que unen fosfato, enzimas de la biosíntesis de triptófano.

Tabla 1: De izquierda a derecha las columnas de la tabla muestran: el código CATH de la familia, el número de proteínas con estructura conocida en conformación libre y acomplejada en cada familia, información del tipo de topología y homología que presentan en la base de datos CATH.

Proporción de residuos que participan en LRCC, y su relación espacial respecto al sitio activo (SA). Para determinar la proporción de residuos de la proteína que participan en el SA y en los LRCC, se analizó la base de datos de proteínas con y sin ligando. El gráfico de la figura 4, muestra una frecuencia máxima de 3% para el porcentaje de residuos de la proteína que participan en el SA, con una distribución que se ajusta bien a una curva de Gauss del tipo: $y = y_0 + ae^{[-0.5(x - \bar{x})^2/b]}$. Los residuos de los LRCC muestran una frecuencia máxima de 8%, que también se ajusta a una curva de Gauss, pero que presenta una distribución más amplia que la observada para el SA. Este resultado sugiere que los cambios conformacionales, relacionados a la unión de ligandos pequeños, en general no se limitan al sitio activo de las proteínas, sino que involucran otras regiones de la estructura.

La distribución de los LRCC en la estructura de la proteína alrededor del SA, se determinó como el mínimo número de capas entre los aminoácidos pertenecientes a los LRCC y aquellos en contacto con el ligando, que se denominó capa cero (ver métodos). En la figura 5, se muestra el porcentaje de los aminoácidos que está presente en cada capa tomando como punto de partida el SA. Las barras negras representan el porcentaje sobre el total de los residuos y las grises el porcentaje sobre el total de residuos que pertenecen a LRCC. En el gráfico se observa, que alrededor de un 10% de los residuos que participan en los LRCC son residuos en contacto con el ligando, capa 0, estos residuos luego propagan el movimiento a la capa contigua de aminoácidos, capa 1, que corresponde al 15% de los residuos que participan en LRCC, las tres siguientes capas de aminoácidos, vecindades 2 a 4 mantienen una proporción de residuos involucrados en los cambios conformacionales constante de alrededor del 12%, disminuyendo

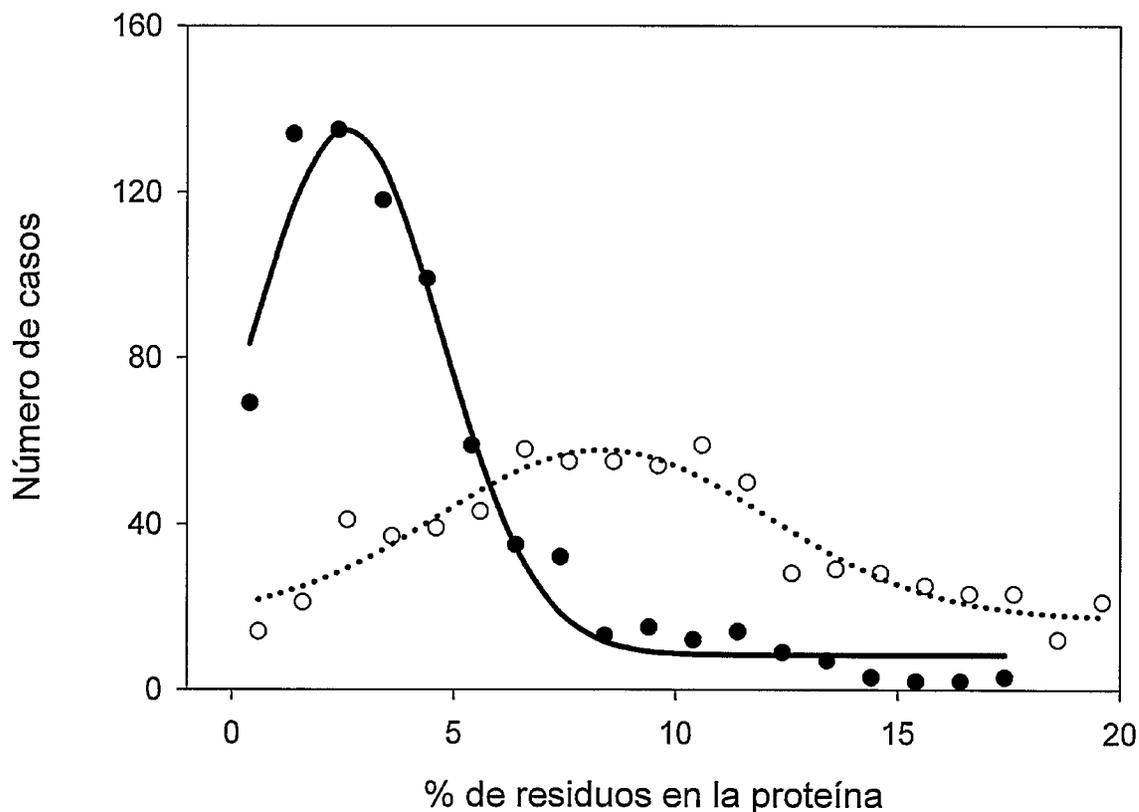


Figura 4. Porcentaje de distribución de los residuos de la proteína involucrados en el SA y en los LRCC. En puntos negros se muestra la distribución del porcentaje de residuos de las proteínas que participan en el SA. En círculos se muestra la distribución del porcentaje de residuos de las proteínas que participan en los LRCC. La curva continua y la punteada son ajustes de los datos a curvas de Gauss del tipo: $y = y_0 + ae^{[-0.5(x - \bar{x})^2/b]}$. La base de datos analizada contenía 786 proteínas con y sin ligando.

rápidamente su participación en las siguientes capas más alejadas del SA. Estos resultados muestran que los cambios conformacionales, considerando toda la base de datos, se extienden a lo menos a 4 residuos de distancia del sitio activo. Esto no descarta que existan residuos más alejados del sitio activo que tengan movimiento. Como control se graficó la distribución del total de residuos de las proteínas en cada capa (barras negras), en la capa 0 se muestra que un 4% de los residuos corresponden a el SA. Este valor es similar al observado en la figura 4, para la frecuencia máxima de residuos que pertenecen al SA en cada proteína. Este control muestra un aumento sostenido en la proporción de residuos en las capas subsiguientes hasta la capa 5, donde empieza a decaer la proporción de vecinos debido a la geometría globular de las proteínas. Se hace evidente al comparar los datos de distribución de los aminoácidos de los LRCC y del control, que los cambios conformacionales ocurren principalmente alrededor del SA. El análisis de los aminoácidos del SA que forman parte de los LRCC, mostró que en promedio solo un 18% de los residuos que forman parte de los sitios activos se mueven, es decir solo uno de cada cinco aminoácidos del SA pertenece a los LRCC. Estos resultados sugieren que en el SA habrían residuos inmóviles que no cambian las coordenadas de su $C\alpha$ al unir el ligando, y otros que son sensibles a la unión del ligando y comunicarían los cambios conformacionales a otras regiones de la estructura.

Composición aminoacídica de los LRCC y su distribución en estructura secundaria. Para determinar si existe una composición aminoacídica que sea característica de los cambios conformacionales, se cuantificó la frecuencia de aparición de cada residuo en los LRCC y se corrigió el valor de acuerdo a la abundancia natural de cada aminoácido. De la misma manera, se determinó la composición aminoacídica de

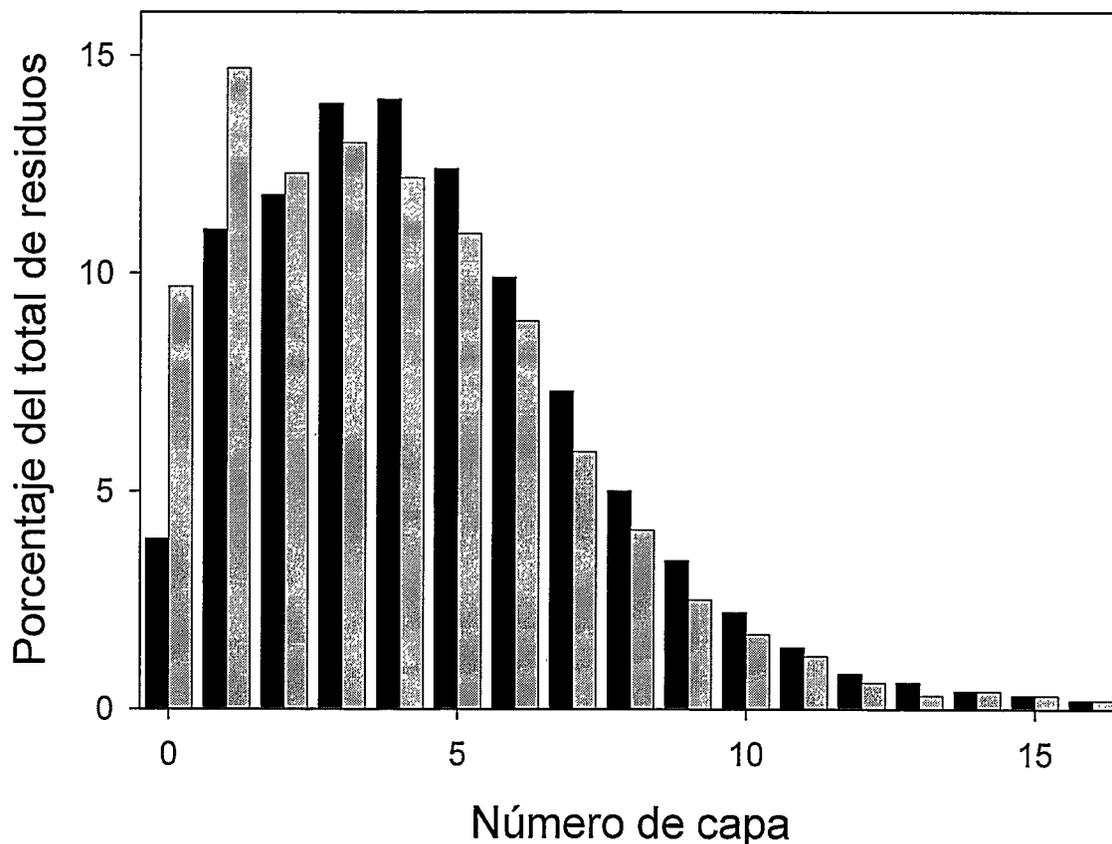


Figura 5. Porcentaje de distribución de residuos de los LRCC y del total de residuos de la proteína en las diferentes capas alrededor del SA. Las barras grises muestran las capas de los LRCC respecto a los residuos del SA, capa 0 corresponde a los aminoácidos que forman parte de los LRCC y del SA; capa 1, son aquellos residuos en contacto directo con algún residuo del SA; capa 2, muestra a los residuos en contacto directo con cualquier residuo de la capa 1, y así sucesivamente. Las barras negras muestran las capas de todos los residuos de la proteína respecto al sitio activo, siendo la capa 0, los residuos de la proteína que forman parte del SA.

los SA, para evaluar el grado de especificidad en la composición de los LRCC. Para ponderar en igual medida la composición aminoácidica de las diferentes familias de proteínas, debido al alto grado de conservación que presentan los residuos del SA, se analizó sólo una proteína de cada familia de la base de datos de cambios conformacionales. La figura 6A, muestra un gráfico de barras de la frecuencia de cada aminoácido en los LRCC (barras grises), aquellos residuos con una frecuencia mayor a 1 muestran una presencia en los LRCC mayor a su abundancia natural, por el contrario valores menores a 1 indican que los aminoácidos son menos frecuentes. La prueba de chi cuadrado (χ^2) para determinar los valores que se diferencian significativamente de la abundancia natural (con un 95% de confianza), mostró que los residuos glicina, serina, asparragina y ácido aspártico, muestran una mayor abundancia en los LRCC que en el resto de la proteína, en cambio fenilalanina, tirosina, triptófano e histidina son más abundantes en el SA (barras negras). Si se comparan las abundancias relativas de los aminoácidos más frecuentes en LRCC y SA, se observa una mayor selectividad en los SA que en los LRCC. Esta mayor selectividad de los aminoácidos del SA, respecto al resto de la proteína se acentúa por la presencia de aminoácidos poco frecuentes en las proteínas, es el caso de tirosina, triptófano e histidina.

En la figura 6B, se muestra la distribución de los LRCC (barras grises) en los distintos tipos de estructura secundaria, como también la distribución promedio de los elementos de estructura secundaria en las proteínas que componen la base de datos (barras negras). Se observa que la abundancia de los LRCC en la estructura secundaria sigue el siguiente orden: lazos >> hélice α > hebra β > vuelta. Si analizamos la composición aminoácidica de los LRCC, donde se favorece la presencia de aminoácidos pequeños y

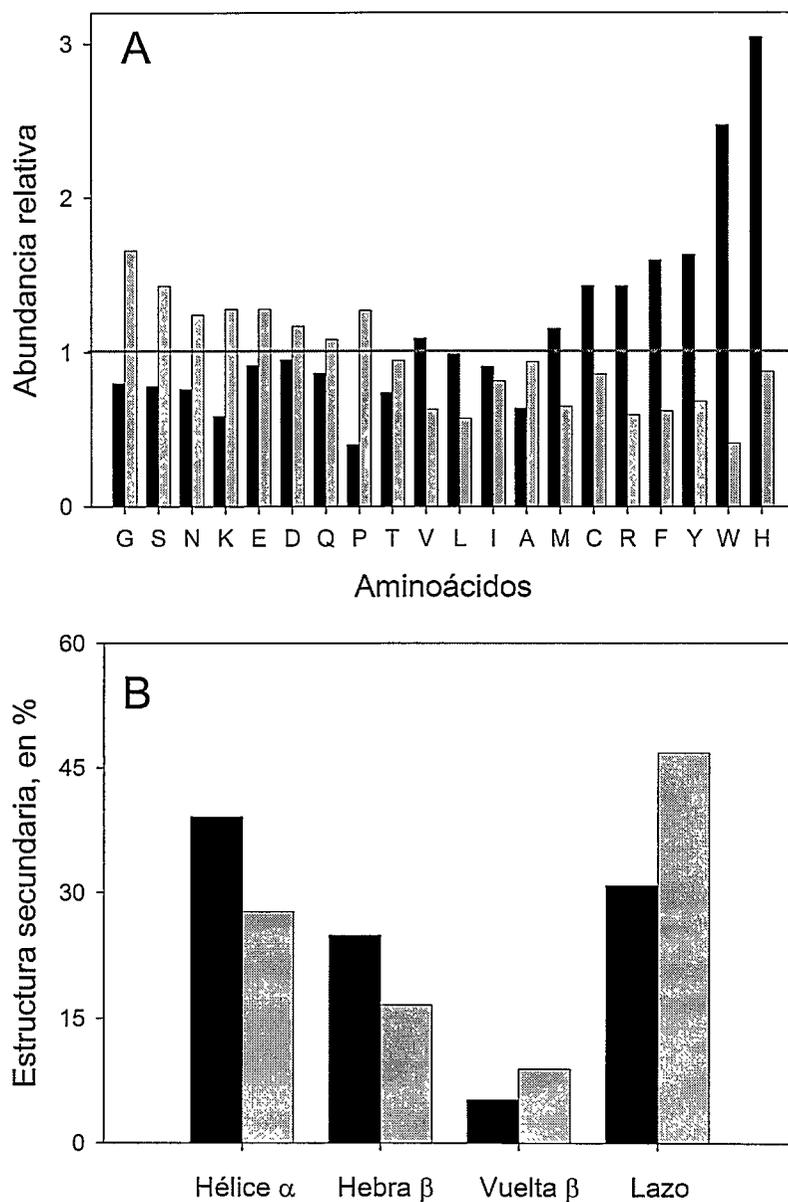


Figura 6. Composición aminoacídica en los LRCC y los SA y frecuencia de los LRCC en los elementos de estructura secundaria. En (A) se muestra la composición aminoacídica de los LRCC (barras grises) y de los SA (barras negras) respecto a la abundancia natural, valor igual a 1, de cada residuo en la base de datos de proteínas libres y como complejos. Debido a que los residuos del SA son muy conservados en familias de proteínas, se seleccionó solo una proteína de cada familia para determinar la composición aminoacídica de los SA y LRCC. En (B), en barras grises, se muestra la frecuencia de los LRCC en distintos tipos de estructura secundaria; en barras negras, se muestra el porcentaje promedio de estructura secundaria en las proteínas de la base de datos.

polares o cargados, y su predilección por elementos de estructura secundaria, es evidente que los cambios conformacionales relacionados a la unión de ligandos pequeños, son fenómenos que ocurren principalmente en la superficie de las proteínas.

Grado de exposición y flexibilidad de los LRCC y SA. Dado el carácter polar de los residuos que participan en LRCC se consideró pertinente caracterizar el grado de exposición y la flexibilidad de los LRCC respecto a los residuos expuestos al solvente de la proteína como los del SA, pues permitirían discriminar de mejor manera las propiedades de los LRCC, que comparándolos contra el total de los residuos de la proteína, donde se promedian propiedades para aminoácidos expuestos y no expuestos. Para determinar el grado de exposición de los residuos se utilizó el programa NACCESS y la clasificación de exposición relativa de Lee y Richards 1971. La figura 7, muestra en la primera columna de la izquierda, el grado de exposición de todos los residuos de las proteínas incluidas en la base de datos de cambios conformacionales, se observa que aproximadamente un tercio de los residuos están ocultos (color gris oscuro), un tercio de los residuos presentan una exposición intermedia (color gris) y un tercio de los residuos están expuestos (gris claro). La segunda barra del gráfico de la figura 7, muestra que solo un 10% de los residuos que participan en los LRCC están ocultos, o tienen menos de un 5% de accesibilidad al solvente, la mayoría de los residuos, alrededor de un 55%, están expuestos al solvente y aproximadamente un 35% están medianamente expuestos, confirmándose los resultados anteriores de composición aminoacídica y preferencia por elementos de estructura secundaria, que sugerían que los LRCC están mayormente en la superficie. Un comportamiento parecido en términos del total de residuos expuestos, más los no expuestos se observa en la barra

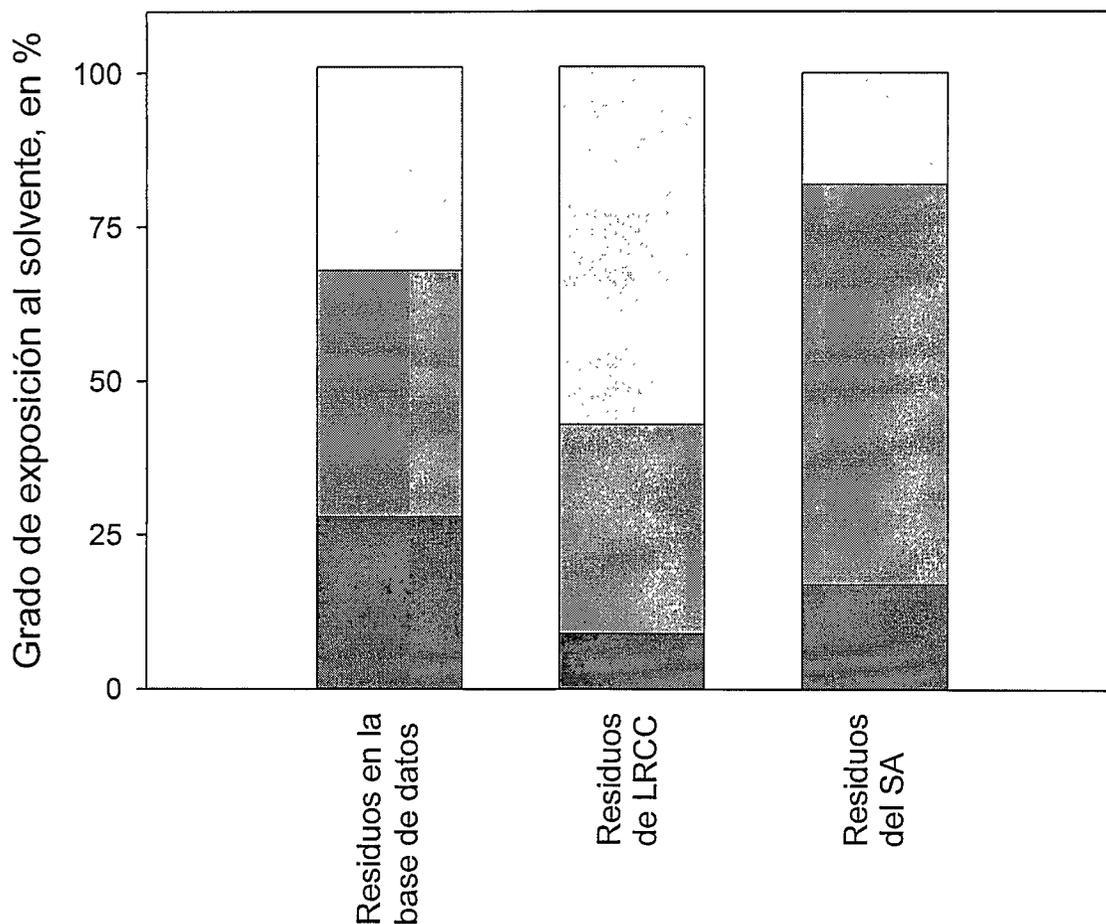


Figura 7. Grado de exposición de los residuos que participan en LRCC, en SA y el total de residuos en las proteínas de la base de datos. El gráfico de barras muestra de izquierda a derecha, el grado de exposición de los residuos de una proteína promedio de la base de datos, de los residuos que participan en LRCC y de los residuos que interactúan con el ligando (SA). Cada barra muestra en gris oscuro la fracción de residuos ocultos, menos de un 5% de accesibilidad al solvente; en gris la fracción de residuos con exposición intermedia, entre un 5 y un 40% de accesibilidad y en gris claro la fracción de residuos expuestos al solvente, más de un 40% de exposición.

de la derecha para los residuos que participan en el SA. Sin embargo, este perfil de exposición se diferencia de los residuos en LRCC debido a que el 70% de ellos está solo medianamente expuesto al solvente a diferencia de solo un 35% para los LRCC. Estos resultados sugieren que es posible discriminar a los LRCC incluso de otros residuos presentes en la superficie de la proteína, como son los aminoácidos presentes en el SA.

La información de factor térmico, determinada a partir de los desplazamientos isotrópicos promedio de cada átomo en los datos cristalográficos, se incluye en los archivos de coordenadas tridimensionales de las proteínas, denominados archivos PDB. Esta información permite evaluar la flexibilidad de las distintas regiones de una proteína. En la figura 8, se muestra la distribución del factor térmico normalizado para todos los aminoácidos de las proteínas en estado libre en la base de datos (línea discontinua), en los residuos del SA (línea continua) y en los residuos de los LRCC (línea punteada). Las distribuciones muestran claramente que los residuos del SA, en su conformación sin el ligando, presentan una mayor rigidez, corrimiento a la izquierda, que los residuos totales de la proteína, en cambio los residuos de los LRCC muestran una distribución más amplia, con un corrimiento hacia la derecha, valores más flexibles. Con el fin de obtener un parámetro cuantitativo para la comparación de las distribuciones de factor térmico normalizado, se determinó la probabilidad de χ^2 para el análisis de las curvas de los residuos del SA y de los LRCC, respecto a la curva del movimiento o factor térmico de todos los residuos de la proteína. El valor 0,91 para χ^2 de la distribución del movimiento en el SA indica un 91 % de probabilidad que la distribución del movimiento en la curva del SA corresponda con la distribución de los residuos totales de

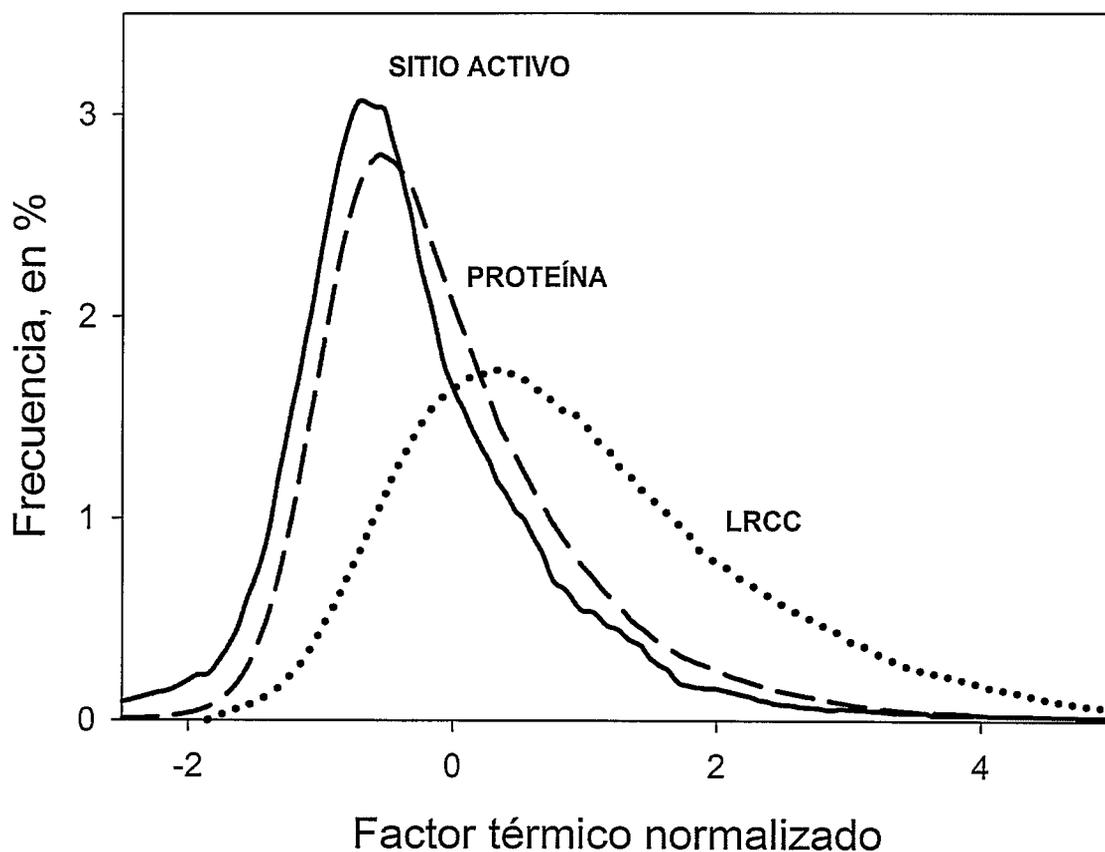


Figura 8. Distribución del factor térmico normalizado para todos los residuos de las proteínas en la base de datos, los LRCC y los SA. Las curvas del gráfico muestran, en línea continua la distribución del factor térmico normalizado para los residuos que pertenecen al SA, en línea discontinua la distribución para todos los residuos de las proteínas y en línea punteada la distribución de los LRCC. Los valores de factor térmico analizados corresponden exclusivamente a las proteínas de la base de datos en conformación libre.

la proteína en un muestreo al azar de un mismo conjunto de datos. La prueba de χ^2 para los LRCC mostró un valor de probabilidad de 0 que indica dos poblaciones diferentes. El análisis del efecto del ligando unido sobre los valores de factor térmico normalizado solo mostró una pequeña diferencia en rigidez cuando se compararon las distribuciones para los residuos del SA con y sin ligando, observándose un pequeño corrimiento a la izquierda, mayor rigidez, con la unión de ligando (no se muestran los resultados), sin embargo la prueba de χ^2 mostró una probabilidad de 1, indicando que las diferencias de rigidez no son significativas, en presencia y ausencia del ligando.

Los resultados de los análisis del grado de exposición en conjunto con los de flexibilidad, para los residuos de los SA y los LRCC, muestran una correlación entre un mayor grado de exposición y una mayor flexibilidad en regiones superficiales de la estructura.

Conservación de los LRCC en familias de proteínas.

Conservación estructural de los LRCC en familias de proteínas con un solo sitio

activo. Para estudiar el grado de conservación de los LRCC en cada familia de proteínas, se seleccionó, de la base de datos de cambios conformacionales, las familias que presentaban un solo sitio de unión a ligando, conservado para todas las proteínas analizadas, independiente de la identidad del ligando unido. De esta manera eliminamos la posibilidad de mezclar efectos debidos a la conservación de los LRCC con la posibilidad de estar analizando distintos LRCC producidos por distintos sitios de unión.

La figura 9A, muestra un esquema de la metodología usada para comparar los LRCC, donde primero se superpusieron las coordenadas tridimensionales de la proteína libre con las coordenadas de la proteína con ligando, se graficaron las distancias entre los $C\alpha$ equivalentes y se determinó el valor del ruido para esta comparación de estructuras en

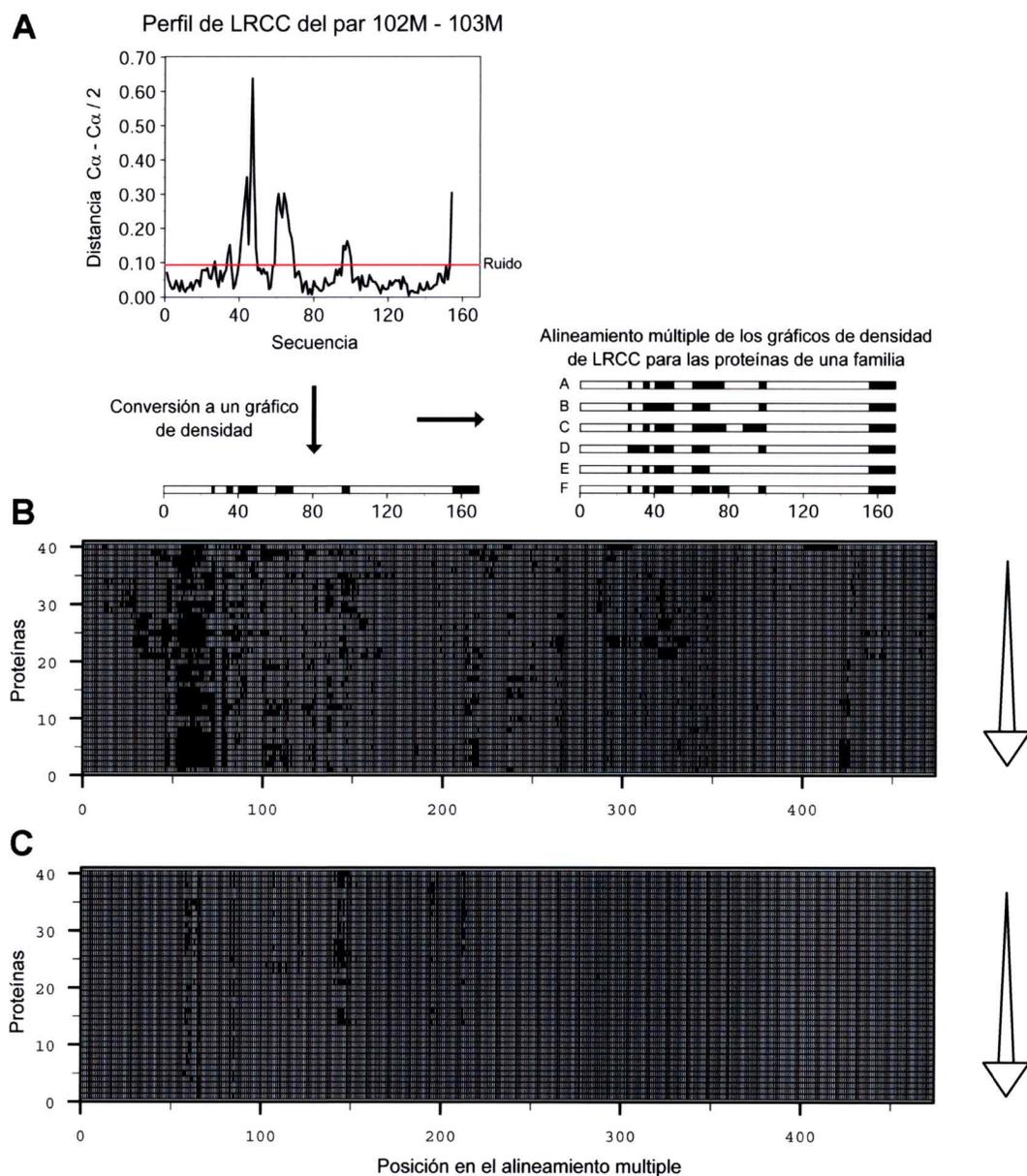


Figura 9. Conservación de los cambios conformacionales asociados a un mismo SA, en la familia de las fosfotransferasas. En A, esquema de la obtención de cambios conformacionales a partir de las diferencias en la posición de los $C\alpha$ al comparar la proteína libre y con ligando. Las flechas indican los diferentes pasos seguidos para la comparación de los perfiles de densidad de los LRCC de diferentes proteínas (A,B,C,etc) En B, gráfico de densidad de los LRCC para las proteínas de la familia de las fosfotransferasas. Los píxeles negros indican las regiones que sufren cambios conformacionales. En C, gráfico de densidad de los SA, definidos como aminoácidos de la proteína en contacto con cualquier átomo del ligando. El orden en que se muestran los patrones de LRCC y SA para cada proteína es el mismo. Los gráficos de densidad están ordenados de acuerdo a la identidad de secuencia de las proteínas como lo indica el grosor de la flecha a la derecha.

particular (ver Materiales y Métodos), el que se muestra como una línea horizontal roja en el gráfico. Para convertir el gráfico 9A a un gráfico de densidad, se marcaron las posiciones de los C α que están sobre el ruido como píxeles negros, que corresponden a los LRCC, y las posiciones bajo el ruido se dejaron en blanco. La información obtenida se comparó con las otras proteínas de la familia en un alineamiento múltiple, figura 9B. El gráfico de densidad de la figura 9C muestra el mismo tipo de representación que 9B, pero esta vez para los residuos en contacto con el ligando (SA). En la figura 9B, la información de los LRCC para la familia de las fosfotransferasas, (código CATH 1.10.510.10) muestra que los cambios conformacionales no se distribuyen homogéneamente en todo el alineamiento y se observan regiones discretas. Una región muy conservada de cambios conformacionales se ubica entre las posiciones 50 a 75 del alineamiento múltiple, y otras regiones menos conservadas alrededor de la posición 140, 210, 270, 345 y 420. Este resultado indica un patrón de cambios conformacionales conservado para la familia de las fosfotransferasas.

Al comparar la figura 9B y 9C, se observa un mismo patrón de contactos para los ligandos en las proteínas de la familia de las fosfotransferasas, confirmando que los LRCC mostrados en la figura 9B, se derivan de un mismo sitio de unión a ligando. La figura 9C, también muestra diferente número de contactos para los distintos ligandos unidos al SA.

Estos resultados en su conjunto muestran que los LRCC generados por ligandos unidos a un mismo sitio activo son similares, pues presentan patrones comunes para cada uno de los ligandos unidos a diferentes proteínas. Con el objeto de separar estas conductas, en la figura 10 se muestra en las dos primeras filas (10A y 10B), el comportamiento de una misma proteína de la familia fosfotransferasa que une diferentes ligandos, donde la

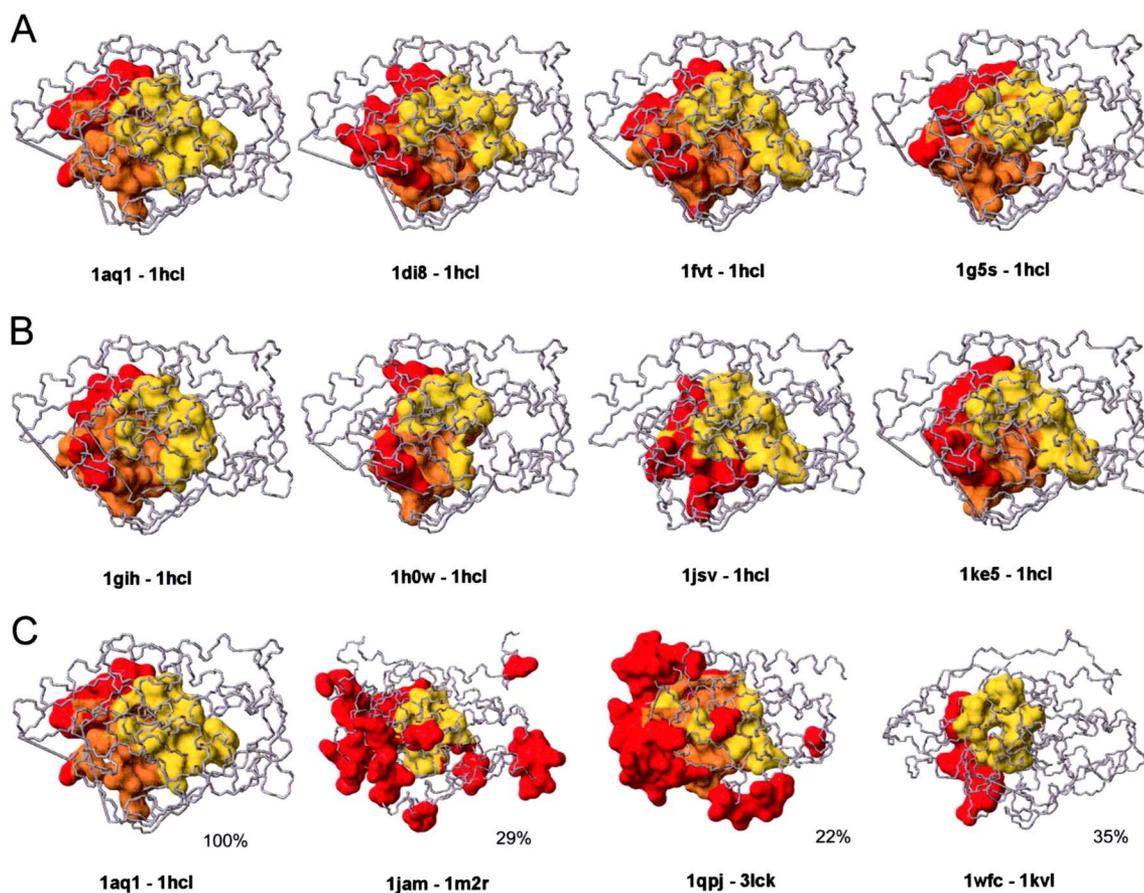


Figura 10. Comparación de LRCC generados por la unión de diferentes ligandos a una misma proteína (A y B) y a otras proteínas de la familia fosfotransferasa (C). La primera y segunda fila (A y B) muestran la estructura con diferentes ligandos (indicado por los primeros 4 caracteres que aparecen bajo cada estructura) en un mismo SA de una misma proteína (1hcl), donde se representa el esqueleto de la proteína como un cable gris, los residuos del SA como superficie amarilla de van der Waals, los residuos que forman parte del SA y LRCC como superficie naranja de van der Waals, y los LRCC como superficie roja de van der Waals. La última fila (C) muestra la estructura con diferentes ligandos de distintas proteínas de la familia, con identidades de secuencia de: 100%, 35%, 29% y 22%, donde los códigos PDB de 4 letras tienen el mismo significado que en (A y B).

proteína se representa como cable gris; los contactos del ligando como superficies de van der Waals, de color amarillo; los residuos que forman parte del SA y de LRCC como superficies de van der Waals de color naranja; y los LRCC como superficies de van der Waals de color rojo. En cambio, en la última fila de la figura 10 se comparan distintas proteínas de la familia con identidades de secuencia de 100%, 35%, 29% y 22%. La misma proteína con distintos ligandos (fig. 10A y B) mostró un patrón de LRCC más conservado que para distintas proteínas de la familia (fig. 10C), sin embargo el patrón de LRCC continúa siendo similar con diferencias en la amplitud de los cambios conformacionales.

Otros ejemplos de familias con un solo sitio de unión a ligando, que mostraron un mismo patrón de LRCC son: la familia de las globinas con código CATH 1.10.490.10, la familia del citocromo P450 con código CATH 1.10.630.10, la familia de las serino proteasas con código CATH 2.40.10.10, la familia de la anhidrasa carbónica II con código CATH 3.10.200.10 y la familia de la timidilato sintetasa código CATH 3.30.572.10 (no se muestran los resultados).

Conservación estructural de los LRCC en familias de proteínas con dos sitios activos. Se demostró que los LRCC son conservados en familias de proteínas que unen distintos ligandos en una misma región de la estructura, pero ¿Qué ocurre si la proteína tiene más de un sitio activo? En la base de datos de cambios conformacionales se encontró solo una familia de proteínas con dos sitios de unión suficientemente poblados para permitir un análisis de cambios conformacionales, que es la de la dihidrofolato reductasa (DHF) con código CATH 3.40.430.10. La DHF cataliza la transformación de dihidrofolato (DDF) en tetrahidrofolato mediante la transferencia de un hidrógeno desde la coenzima NADPH_2 al N5 del sustrato (Agarwal, Billeter et al. 2002; Rod and Brooks

2003; Rod, Radkiewicz et al. 2003). DHF tiene un sitio de unión para la coenzima NADPH_2 y otro para el sustrato DDF. En la base de datos de cambios conformacionales están representadas DHF de distintos organismos acomplejadas con análogos de NADPH_2 , de DDF, o con ambos ligandos. En la figura 11, se muestran los LRCC resultantes de la comparación de la estructura de DHF acomplejada con ambos ligandos (estado unido) y DHF acomplejada con un solo ligando (estado libre), el ligando puede ser un análogo de DDF o de NADPH_2 . La figura 11A, muestra en esquema de cables el esqueleto de las proteínas acomplejadas, y en superficie de van der Waals los residuos en contacto con los distintos análogos del ligando DDF, es decir en el sitio activo del dihidrofolato. En el gráfico de densidad de la figura 11A, se observan como píxeles negros los LRCC producidos por la unión de DDF, donde se destacan principalmente las regiones entre los residuos 25 - 40, 65 - 75 y 190 - 195 del alineamiento múltiple, se observa que en menor grado participan en LRCC residuos alrededor de la posición 155 y 165 del alineamiento. La unión de análogos de NADPH_2 en el segundo sitio activo, figura 11B, muestra un patrón de LRCC más conservado donde destacan las regiones entre los residuos 20 - 30, 80 - 90, además de residuos alrededor de 155 y 165 y de menor importancia son los residuos de las posiciones 75 y 130 del alineamiento. Los LRCC producidos por la unión de análogos de DDF y los LRCC producidos por la unión de análogos de NADPH_2 , muestran patrones de cambios conformacionales conservados, que son diferentes, aun cuando tienen algunas regiones comunes alrededor de la posición 75, 155 y 165.

Los distintos patrones de LRCC para dos sitios activos en la familia DHF, sugieren que distintos sitios activos en proteínas producen cambios conformacionales diferentes, independientes de la naturaleza del ligando.

Efecto de los contactos cristalográficos sobre los LRCC. Los resultados anteriores del análisis de la base de datos de cambios conformacionales indican que los LRCC se producen principalmente en la superficie de las proteínas, y que son conservados para un mismo sitio activo dentro de una familia de proteínas. La metodología para determinar los LRCC consistió en comparar dos estructuras cristalográficas con y sin ligando, sin embargo, dependiendo de la celda cristalográfica con que se obtuvo la estructura, podrían existir distintos patrones de contactos cristalográficos para la proteína, estos contactos cristalográficos podrían alterar la posición de los aminoácidos superficiales y producir alteraciones en los LRCC. Por lo tanto una pregunta válida es ¿Influye en la comparación de estructuras el uso de distintas celdas cristalográficas sobre el patrón de LRCC? La figura 12, muestra el alineamiento múltiple de LRCC y del SA para la familia de las serino proteasas, donde la tripsina de bovino que ha sido cristalizada en presencia de distintos inhibidores está representada en 44 de los 48 perfiles mostrados. El gráfico de densidad de la figura 12A muestra, en píxeles negros, los perfiles de LRCC para las proteínas de la familia, donde se observa un patrón conservado de cambios conformacionales que incluye las posiciones 70, 95, 105, 135, 155, 170 - 180, 220 - 225 y en menor grado la posición 285 del alineamiento múltiple. La figura 12B, muestra los contactos del ligando en el SA de cada uno de los complejos evaluados, donde la conservación de estos contactos indica la presencia de un solo sitio activo en las proteínas analizadas. La diferencia en el número de contactos del ligando en cada proteína, muestra la heterogeneidad de los inhibidores presentes en la base de datos, para las serino proteasas. La figura 12C, muestra una representación gráfica del análisis de componentes principales (PCA) realizados a los perfiles de LRCC (ver Materiales y Métodos), la abscisa y la ordenada del gráfico (p y q) corresponden al

primer y segundo componente principal, respectivamente. El resultado de PCA mostró dos poblaciones de datos, la primera que incluye al 80% de las proteínas esta centrada en $p = -4$ y $q = 0$, delimitada por un óvalo azul, y la segunda centrada en $p = 10$ y $q = 2$, esta demarcada por un óvalo rojo. Las dos poblaciones corresponden a dos patrones de cambios conformacionales que son posibles de diferenciar en esta familia. El estudio de la información anexa, en los archivos PDB, mostró que la mayoría de los LRCC agrupados en el óvalo rojo, corresponden mayoritariamente a comparaciones entre estructuras con celdas cristalográficas distintas, grupos espaciales $P 2_1 2_1 2_1$ y $P 3_1 2 1$, (ver tabla 2) en cambio en el óvalo azul están presentes los LRCC obtenidos al comparar estructuras con las mismas celdas cristalográficas, ya sea de grupo espacial $P 2_1 2_1 2_1$ entre sí, o de simetría $P 3_1 2 1$, las únicas excepciones son las comparaciones de los LRCC producidos por las proteínas con código PDB 1az8 y 1tgn, con una celda $P 2_1 2_1 2_1$ y $P 3_1 2 1$, respectivamente, y las proteínas con código PDB 1aq7 y 1tgn, con celdas cristalográficas $P 6_1$ y $P 3_1 2 1$, respectivamente. Este resultado sugiere que para determinar los LRCC entre dos conformaciones de una proteína, se deben comparar estructuras con las mismas celdas unitarias, pues comparar distintas celdas puede producir un artefacto debido a los distintos contactos cristalográficos en una y otra conformación.

Determinación de los LRCC en FtsZ.

Comparación de FtsZ de distintos organismos en conformación GTP y GDP

El gran número de estructuras determinadas experimentalmente para FtsZ, en distintos microorganismos (mesófilos y termófilos), y bajo distintas condiciones de cristalización, permitió seleccionar estructuras en estado GDP y GTP de *Mycobacterium tuberculosis*, archivos 1RQ7 y 1RLU, y de *Methanococcus jannaschii*, archivos 1W58 y 1FSZ,

Tabla 2. Lista de los archivos PDB, de la familia de las serino proteasas, en una conformación con y sin ligando, analizados en la figura 12.

PDB unido	Celda	PDB libre	Celda	Ligando	Grupo
1uto	P 3 ₁ 2 1	1tnj	P 2 ₁ 2 ₁ 2 ₁	GOL	Rojo
1utp	P 3 ₁ 2 1	1tni	P 2 ₁ 2 ₁ 2 ₁	GOL	Rojo
1xuj	P 3 ₁ 2 1	1tld	P 2 ₁ 2 ₁ 2 ₁	BOZ	Rojo
1xuk	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	BAB	Rojo
1gj2	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	122	Rojo
1gi1	P 2 ₁ 2 ₁ 2 ₁	1gi3	P 2 ₁ 2 ₁ 2 ₁	SO4	Rojo
1gi1	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	BMZ	Rojo
1gi0	P 3 ₁ 2 1	1gi1	P 2 ₁ 2 ₁ 2 ₁	MG	Rojo
1c5s	P 3 ₁ 2 1	1tld	P 2 ₁ 2 ₁ 2 ₁	ESX	Rojo
1c1t	P 3 ₁ 2 1	1xuk	P 2 ₁ 2 ₁ 2 ₁	MG	Rojo
1g3b	P 2 ₁ 2 ₁ 2 ₁	1c1o	P 3 ₁ 2 1	108	Rojo
1c1o	P 3 ₁ 2 1	1tld	P 2 ₁ 2 ₁ 2 ₁	MG	Rojo
1jir	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	AML	Azul
1k1j	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	FD2	Azul
1k1l	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	FD3	Azul
1k1m	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	FD4	Azul
1lqe	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	IMA	Azul
1ql7	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	ZEN	Azul
1gi5	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	123	Azul
1g36	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	R11	Azul
1f0u	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	RPR	Azul
1f0t	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	PR1	Azul
1eb2	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	BPO	Azul
1oyq	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	T87	Azul
1o38	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	653	Azul
1o35	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	802	Azul
1o2q	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	991	Azul
1o2o	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	950	Azul
1o2n	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	762	Azul
1o2i	P 2 ₁ 2 ₁ 2 ₁	1tld	P 2 ₁ 2 ₁ 2 ₁	655	Azul
1c2k	P 3 ₁ 2 1	1c2l	P 3 ₁ 2 1	ABI	Azul
1c2j	P 3 ₁ 2 1	1c2m	P 3 ₁ 2 1	BAK	Azul
1c1p	P 3 ₁ 2 1	1c1q	P 3 ₁ 2 1	DMS	Azul
1c1r	P 3 ₁ 2 1	1c1p	P 3 ₁ 2 1	ZN	Azul
1c1q	P 3 ₁ 2 1	1c1o	P 3 ₁ 2 1	BAI	Azul
1c1t	P 3 ₁ 2 1	1c1o	P 3 ₁ 2 1	BAB	Azul
1c2m	P 3 ₁ 2 1	1c1o	P 3 ₁ 2 1	ZN	Azul
1c5p	P 3 ₁ 2 1	1c1o	P 3 ₁ 2 1	BAM	Azul
1c5u	P 3 ₁ 2 1	1c1o	P 3 ₁ 2 1	ESP	Azul
1c1o	P 3 ₁ 2 1	1c5v	P 3 ₁ 2 1	SO4	Azul
1ghz	P 3 ₁ 2 1	1c1o	P 3 ₁ 2 1	120	Azul
1gi0	P 3 ₁ 2 1	1c1o	P 3 ₁ 2 1	BMZ	Azul
1az8	P 2 ₁ 2 ₁ 2 ₁	1tgn	P 3 ₁ 2 1	IN4	Azul
1aq7	P 6 ₁	1tgn	P 3 ₁ 2 1	AEB	Azul
1fn6	P 2 ₁ 2 ₁ 2 ₁	1fni	P 2 ₁ 2 ₁ 2 ₁	MOH	Azul
1lkb	P 2 ₁ 2 ₁ 2 ₁	1qnj	P 2 ₁ 2 ₁ 2 ₁	CL	Azul
1ssx	P 3 ₂ 2 1	2alp	P 3 ₂ 2 1	GOL	--
1tai	P 3 ₂ 2 1	2alp	P 3 ₂ 2 1	TAM	--

Tabla 2. La primera columna de la izquierda muestra el código PDB de la proteína acomplejada, la segunda columna indica la celda en que fue cristalizada, la tercera columna muestra el código PDB de la proteína libre, luego se indica la celda cristalográfica de la proteína libre, el código PDB para el ligando unido y la sexta columna indica el grupo al que pertenece, de acuerdo al análisis de PCA de la figura 12C.

respectivamente. Solo se condicionó la selección a proteínas que presentaran la misma celda cristalográfica, en ambas conformaciones.

La figura 13A, muestra de izquierda a derecha, en esquema de cintas, la estructura 3D de MtbFtsZ y MjFtsZ, respectivamente, estas FtsZ de diferentes organismos, presentan una identidad de secuencia de 48% y un RMS de 1,45 Å, para todos los C α de sus estructuras cristalográficas. El esquema de cintas de la figura 13A, muestra en rojo los residuos del esqueleto de la proteína que participan en los LRCC, la disposición de estos grupos en la estructura es similar para las dos proteínas, observándose una distribución de estos grupos alrededor del sitio activo, en elementos de estructura secundaria que pertenecen al dominio amino terminal, como el lazo T3, lazo H6-H7 y hélice H2. Otros residuos se distribuyen en regiones alejadas del SA como el lazo H8-S7, el lazo entre la hebra β S8 y hélice H10 y el lazo S9 – S10, en el dominio carboxilo terminal de FtsZ. La magnitud de los cambios conformacionales no es igual para las dos proteínas, MtbFtsZ (fig. 13A), muestra una mayor población de residuos que participan en los LRCC, respecto a MjFtsZ (fig. 13B). Un análisis más detallado de la magnitud de los cambios conformacionales en los distintos elementos de estructura secundaria se observa en la figura 14A, donde MjFtsZ (barras rojas) muestra movimiento de los carbono α de los lazos H1 – S1, T3, T4, H4 – S5 y de la hélice H3 en el dominio amino terminal. En general, la magnitud de los cambios conformacionales es mayor para elementos de estructura secundaria del dominio carboxilo terminal, donde destacan los lazos descritos en el análisis de la figura 13A. La mayor movilidad del dominio carboxilo terminal con la hidrólisis de GTP se comprueba analizando los valores de RMS para cada dominio de MjFtsZ por separado, con un valor de 0,3 Å

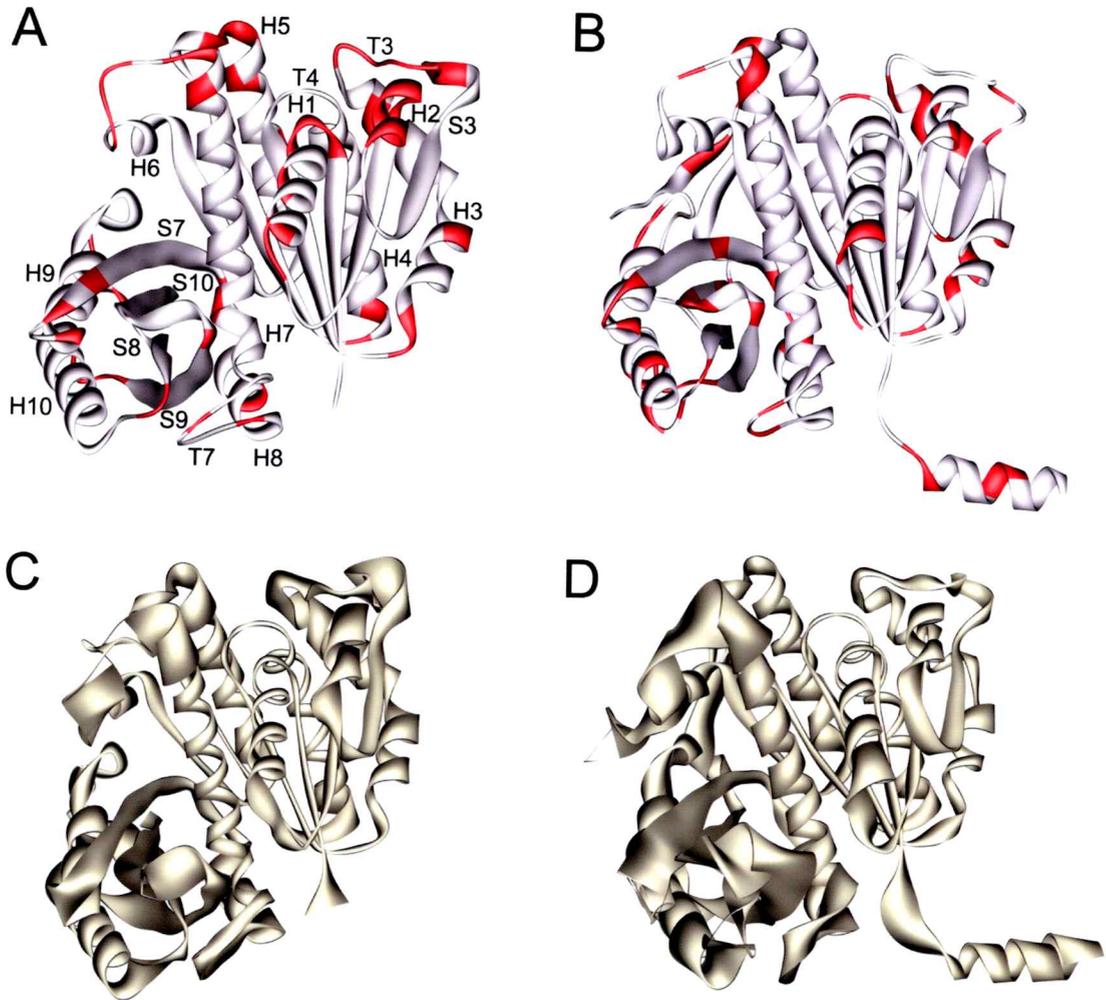


Figura 13. LRCC y flexibilidad intrínseca en FtsZ de distintos microorganismos. Las estructuras A y B de FtsZ se representan con el esqueleto de las proteínas en cinta y los aminoácidos que participan en los cambios conformacionales coloreados en rojo. Las estructuras C, D muestran el esqueleto de las proteínas como cintas de ancho variable dependiendo de los valores del factor térmico de los respectivos archivos PDB, (más anchas si la región es más flexible). La orientación de las proteínas muestra el bolsillo de unión a nucleótido en el centro de la cara superior. En A, se observan los LRCC para MtbFtsZ, donde se identifican algunos elementos de estructura secundaria relacionados con los cambios conformacionales, y en B se observan los LRCC para MjFtsZ. En C y D, se muestra la flexibilidad del esqueleto de MtbFtsZ y MjFtsZ, respectivamente.

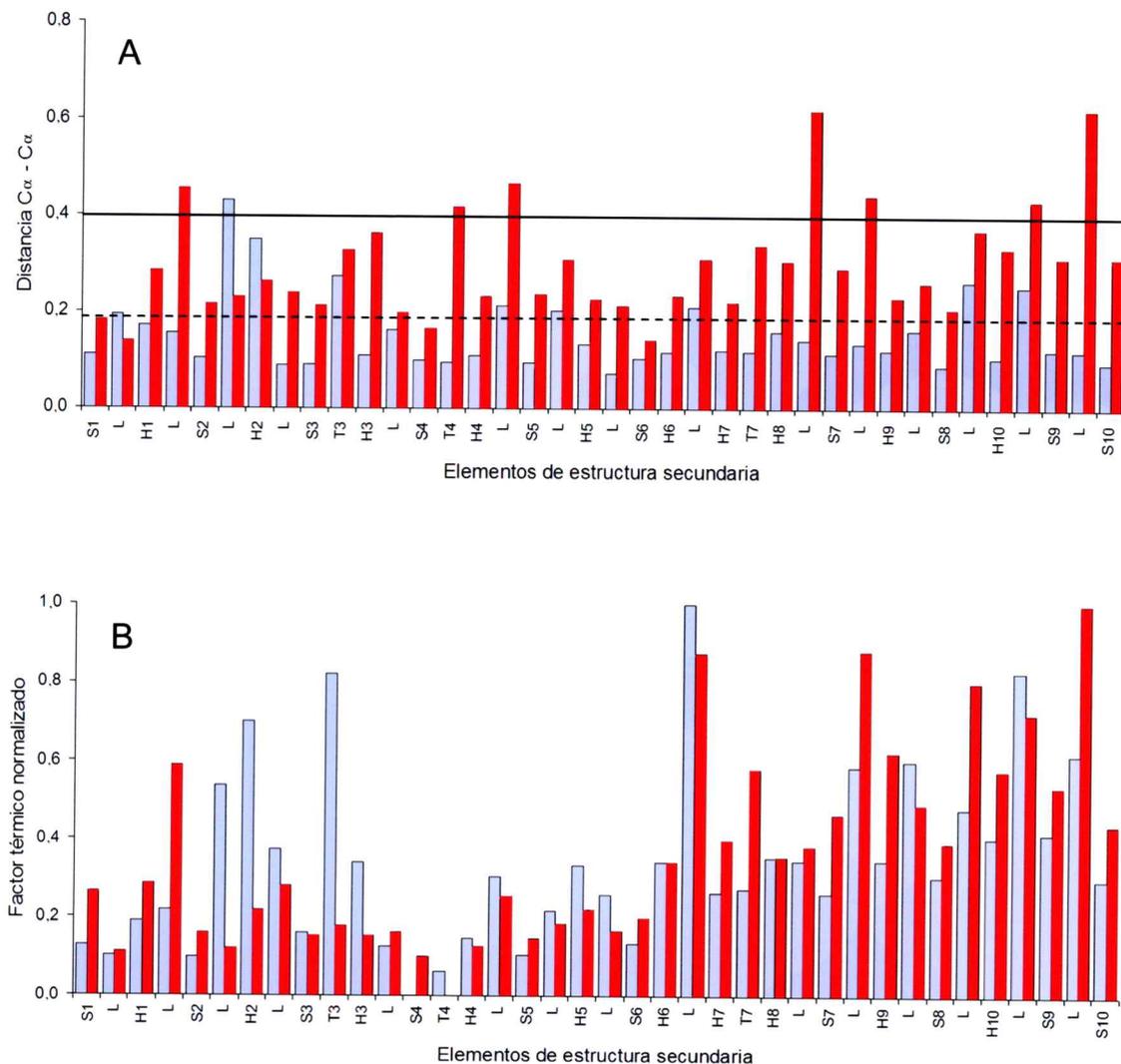


Figura 14. Gráfico de barras de los desplazamientos de los $C\alpha$ para MtbFtsZ y MjfFtsZ entre la conformación GDP y GTP, y variaciones del factor térmico normalizado para ambas proteínas en la conformación GDP. En A, se muestran las magnitudes de los desplazamientos de los $C\alpha$ de MjfFtsZ entre la conformación GDP y GTP (barras rojas), respecto a los elementos de estructura secundaria de la proteína. De la misma manera se muestran en barras celestes las magnitudes de los desplazamientos de los $C\alpha$ para MtbFtsZ. La línea continua y la discontinua muestran el error calculado para los desplazamientos de los $C\alpha$ para MjfFtsZ y MtbFtsZ, respectivamente. En B, se muestran los valores de factor térmico normalizado para los elementos de estructura secundaria de MjfFtsZ y MtbFtsZ en barras moradas y celestes, respectivamente.

Tabla 3. Características de la comparación de estructuras en estado GTP y GDP de la figura 13A, 13B y figura 14A.

PDB GTP/GDP	Organismo	Grupo espacial GTP/GDP ^a	RMS Å GTP/GDP ^b	Conformación GTP/GDP
1W58 / 1FSZ	<i>M. jannaschii</i>	1 2 ₁ 3 / 1 2 ₁ 3	0.3	Sustitución en el cristal
	Dominio N		0.3	
	Dominio C		0.4	
1RLU / 1RQ7	<i>M. tuberculosis</i>	P 6 ₅ / P 6 ₅	0.2	Sustitución en el cristal
	Dominio N		0.2	
	Dominio C		0.2	

Tabla 3. De izquierda a derecha, la primera columna muestra el código PDB de las estructuras cristalinas en estado GTP y GDP para un determinado organismo, la segunda columna indica a que organismo pertenece la FtsZ, y la tercera columna las celdas cristalográficas de las estructuras comparadas en la figura 13, en la cuarta columna se indica el RMS entre los C α de la conformación GTP y GDP, y en la quinta columna se informa de las condiciones en que se obtuvieron las coordenadas cristalográficas de las proteínas en ambos estados.

para el dominio amino terminal y de 0,4 Å para el dominio carboxilo terminal (tabla 3). El análisis del movimiento de los carbono α de MtbFtsZ (fig 14A, barras celestes) muestra que los lazos S2 - H2, T3, H4 - S5, S5 - H5, H6 - H7 y la hélice H2 del dominio amino terminal y los lazos S8 - H10 y H10 - S9 del dominio carboxilo terminal presentan los mayores desplazamientos. La magnitud de los cambios conformacionales en MtbFtsZ inducidos por la hidrólisis de GTP son similares en ambos dominios de la proteína, con un valor de RMS de 0,2 Å. Si se contrastan los resultados de cambios conformacionales con la información de flexibilidad (factor térmico) de MjFtsZ y MtbFtsZ, representada por el ancho de la cinta de la figura 13C y 13D, se observa que el SA formado por el lazo T4 y el segmento amino de las hélices H1 y H4 es más rígido que el resto de la estructura, en cambio en el dominio carboxilo terminal las hélices H9, H10, hebra β S9 y los lazos entre estos elementos de estructura secundaria son flexibles. El análisis de flexibilidad, más detallado en la figura 14B, muestra una mayor rigidez de la estructura en el dominio

amino terminal, que incluye: las hebras β S1, S2, S3, S4, S5 y S6; las hélices α H1, H3, H4 y H5; y los lazos S1 - H1, H3 - S4, T4, H4 - S5, S5 - H5 y H5 - S6. En cambio el dominio carboxilo terminal muestra en general una mayor flexibilidad, principalmente en las hebras β S7, S9, las hélices α H9, H10 y los lazos S7 - H9, H9 - S8, S8 - H10, H10 - S9 y S9 - S10.

Los resultados de los análisis de LRCC y de factor térmico presentan patrones similares en el dominio carboxilo terminal de FtsZ, sin embargo hay regiones de bajo factor térmico normalizado en la estructura que también participan en los LRCC, es el caso de los residuos cercanos al nucleótido, como las hélices H1, H4 y el lazo T4, del dominio amino terminal.

La tabla 3, muestra algunas características relevantes de la comparación de estructuras presentada en la figura 13 y 14, como por ejemplo, en la segunda columna se observa que las estructuras comparadas tenían la misma celda cristalográfica. Otro parámetro importante es el análisis de RMS entre los $C\alpha$ de las conformaciones GTP y GDP, que muestra que tanto MjFtsZ como MtbFtsZ presentan en general pequeñas variaciones entre ambas conformaciones, lo que se explica debido a que la determinación de las conformaciones GTP y GDP para FtsZ, en ambos organismos, se obtuvo mediante el lavado del cristal con el ligando que se deseaba analizar, ver columna 5 de la tabla 3.

Los resultados de FtsZ en su conjunto, le dan un contexto a las características de los LRCC derivadas del análisis de la base de datos de cambios conformacionales, donde los LRCC se distribuyen alrededor del SA que presenta una mayor rigidez y se extienden a lo menos a 4 residuos de distancia o más, probablemente dependiendo de la existencia de regiones más flexibles, que en el caso de la FtsZ incluyen regiones en el dominio carboxilo terminal.

Comprobación experimental del efecto esperado para la introducción de mutaciones puntuales en zonas LRCC y no LRCC de FtsZ

Para evaluar el efecto de las regiones rígidas y flexibles sobre la estabilidad y función de EcFtsZ, se introdujeron mediante mutaciones puntuales, triptófanos en distintas regiones de la proteína. Este aminoácido es ideal para el estudio de EcFtsZ que no tiene triptófanos en su secuencia aminoacídica y además puede ser usado como sonda fluorescente. Los aminoácidos que se sustituyeron por triptófano se seleccionaron mediante un alineamiento múltiple de secuencias de la familia de FtsZ. La mutante F135W se seleccionó por el alto contenido de aminoácidos aromáticos en esa posición (99%). El análisis de cambios conformacionales de MtbFtsZ (fig. 13B y 14A) muestra que F135 localizada en el lazo S5 - H5 se desplaza con la hidrólisis del nucleótido, además se ubica en la misma región que el lazo H6 - H7 en la estructura 3D (fig. 13A y fig. 3), uno de los elementos de estructura secundaria más flexibles de FtsZ. En el dominio carboxilo terminal la posición I294 presenta un 50% de conservación en el contenido de aromático. Este residuo se localiza en la hebra β S9 en la región más flexible del dominio carboxilo terminal y a un par de aminoácidos de distancia del lazo H10 - S9, elemento de estructura secundaria que cambia su posición al pasar de la conformación GTP a la conformación GDP. Para determinar el efecto de una zona rígida en la función de FtsZ, se construyó la mutante F40W que se localiza en la hebra β S2 (fig. 14B). Si los residuos que participan en los LRCC en MjFtsZ y MtbFtsZ, también participan en los LRCC en EcFtsZ (LRCC conservados), se espera que las mutaciones puntuales en estas regiones, F135W e I294W, modifiquen parámetros funcionales de la proteína de *Escherichia coli*, y sin embargo no afecten el plegamiento de la proteína, debido a que estas regiones son flexibles y podrían acomodar a un residuo voluminoso

como triptófano. En cambio, una mutación en una región de bajo factor térmico normalizado y alejada de las zonas que participan en LRCC, F40W, no debería afectar la función de EcFtsZ, pero podría producir problemas estéricos y afectar la estructura tridimensional de la proteína.

Fenotipo de mutantes puntuales en regiones involucradas en LRCC de EcFtsZ.

Caracterización estructural de las mutantes puntuales de EcFtsZ, F40W, F135W e I294W.

Posibles cambios en la estructura 3D de FtsZ producidos por la introducción de un triptófano se analizaron por espectros de emisión de fluorescencia y dicroísmo circular (CD) en el UV-lejano. Estas medidas permiten determinar cambios en el medio que rodea a la sonda fluorescente en la estructura de las mutantes y en el contenido de estructura secundaria, respecto a FtsZ wt.

Fluorescencia. Los espectros de emisión de fluorescencia normalizados de las mutantes F40W, F135W e I294W, se muestran en la figura 15A. El espectro de emisión de fluorescencia de la mutante F135W presentó un máximo de emisión a 343 nm, en cambio I294W presentó un máximo de emisión a 339 nm, a la misma concentración de proteína. La intensidad de fluorescencia de la mutante control, F40W, alcanza un máximo de emisión a 340 nm. Estos resultados sugieren que el triptófano en las tres mutantes se encuentra en un medio ambiente parecido diferente al del solvente, donde el de mayor hidrofobicidad es el de la mutante I294W. Si hubiese uno totalmente expuesto debería presentar un máximo a 353 nm que es el del triptófano en una proteína desplegada (Lakowics 1999).

Dicroísmo Circular. Los espectros de CD en el UV-lejano de EcFtsZ wt y las mutantes F40W, F135W e I294W, se muestran en la figura 15B. Los espectros son típicos de

proteínas con estructura α/β , donde se observan dos mínimos uno a 208 nm y otro a 222 nm, asociado a la absorción del enlace peptídico en una estructura de hélice α , similar para EcFtsZ wt y las mutantes. Las pequeñas diferencias observadas a 208 nm para la mutante I294W, no afectaron el contenido de estructura secundaria de esta mutante, sin embargo el cambio a 222 nm para la mutante F40W, modificó su contenido de estructura secundaria respecto a FtsZ wt en alrededor de un 2%. Este cambio es compatible con la introducción de un aminoácido voluminoso en una región más rígida de la proteína.

Caracterización funcional de las mutantes puntuales de EcFtsZ, F40W, F135W e I294W

La caracterización funcional de las mutantes de EcFtsZ se hizo por medio de la actividad GTPásica, la polimerización y el grado de complementación *in vivo* de las mutantes F40W, F135W e I294W, respecto a EcFtsZ wt.

Actividad GTPásica. La mutante F40W, ubicada en una zona más rígida de la estructura de FtsZ distante de los cambios conformacionales, mostró una disminución en la actividad GTPásica similar a la obtenida para la otra mutante del dominio amino terminal, F135W, 80% y 71%, respectivamente. En cambio I294W, ubicada en el dominio carboxilo terminal, mostró solo un 13% de la actividad GTPásica de EcFtsZ wt (figura 16A).

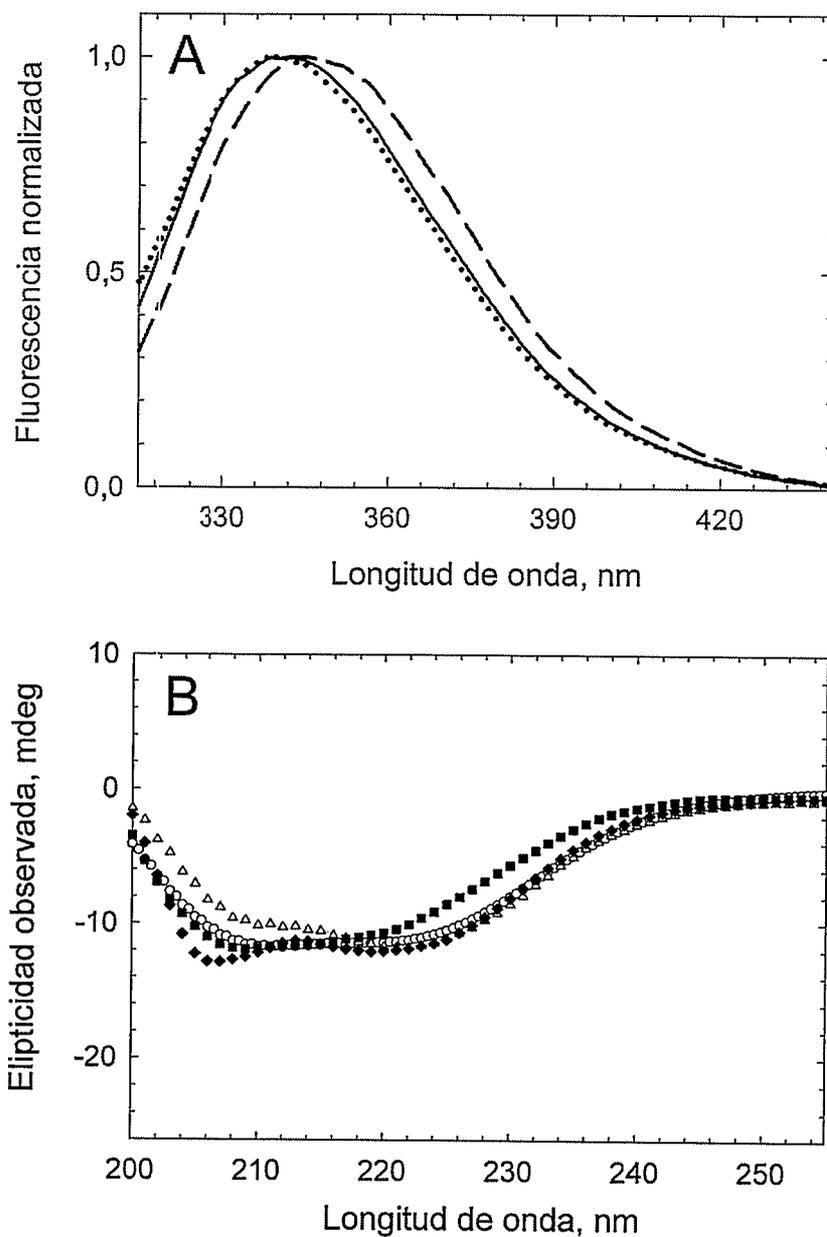


Figura 15. Espectros de emisión de fluorescencia y dicroísmo circular. En A, se muestran los espectros de emisión de fluorescencia normalizados de las tres mutantes puntuales de EcFtsZ, F40W en línea continua, F135W en líneas discontinuas e I294W en líneas punteadas. Longitud de onda de excitación 295 nm. Concentración de las mutantes 12 μ M. En B, se muestran los espectros de CD en el UV-lejano para EcFtsZ *wt* (\blacklozenge); y las mutantes F40W (\blacksquare), F135W (O), e I294W (Δ), a 25°C en amortiguador Tris-HCl 50 mM pH 7,5. (Resultados de R. Díaz y F. Montecinos).

Polimerización. La figura 16C muestra el curso temporal de la turbidez de las soluciones de FtsZ después de agregar GTP. La turbidez es una medida de la polimerización de FtsZ, y depende del largo y la forma de los polímeros. La curva de la mutante F40W, presentó un 70% de la máxima turbidez respecto a EcFtsZ wt, para la misma concentración de proteína (25 μ M), la extensión del máximo de la curva aumentó respecto a EcFtsZ wt, indicando que la proteína permanece más tiempo polimerizada, lo que está de acuerdo con su menor actividad GTPásica, respecto a EcFtsZ wt. La mutante F135W, disminuyó la turbidez asociada a la polimerización a un 32%, comparada con EcFtsZ wt, con un aumento en la extensión del máximo de la curva. La mutante I294W (inserto, figura 16C) mostró una turbidez inicial dos veces mayor que la EcFtsZ wt inducida por GTP y no presentó el incremento típico de la turbidez al agregar GTP, que induce la formación de polímeros. Para analizar el comportamiento anómalo en la polimerización de I294W, se observó el fenotipo de la mutante y de EcFtsZ wt por microscopía electrónica, figura 17. En la figura 17A, se observa un campo representativo de la micrografía electrónica, obtenida en condiciones no polimerizantes para EcFtsZ wt, donde no se observa la presencia de estructuras poliméricas. La figura 17B, corresponde a EcFtsZ wt en presencia de GTP, condiciones polimerizantes, donde se observan los típicos filamentos simples y dobles de FtsZ polimerizada. La figura 17C, muestra los polímeros responsables de la turbidez de la mutante I294W en ausencia de GTP. Los polímeros de I294W son polímeros rectos de varios protofilamentos (hojas) que comparados con los del control, EcFtsZ - GTP, sugieren que la mutación favorecería las interacciones laterales para formar las hojas de 4 o más filamentos. En otros campos de la preparación, se observaron manojos de láminas de FtsZ (no se muestran los resultados).

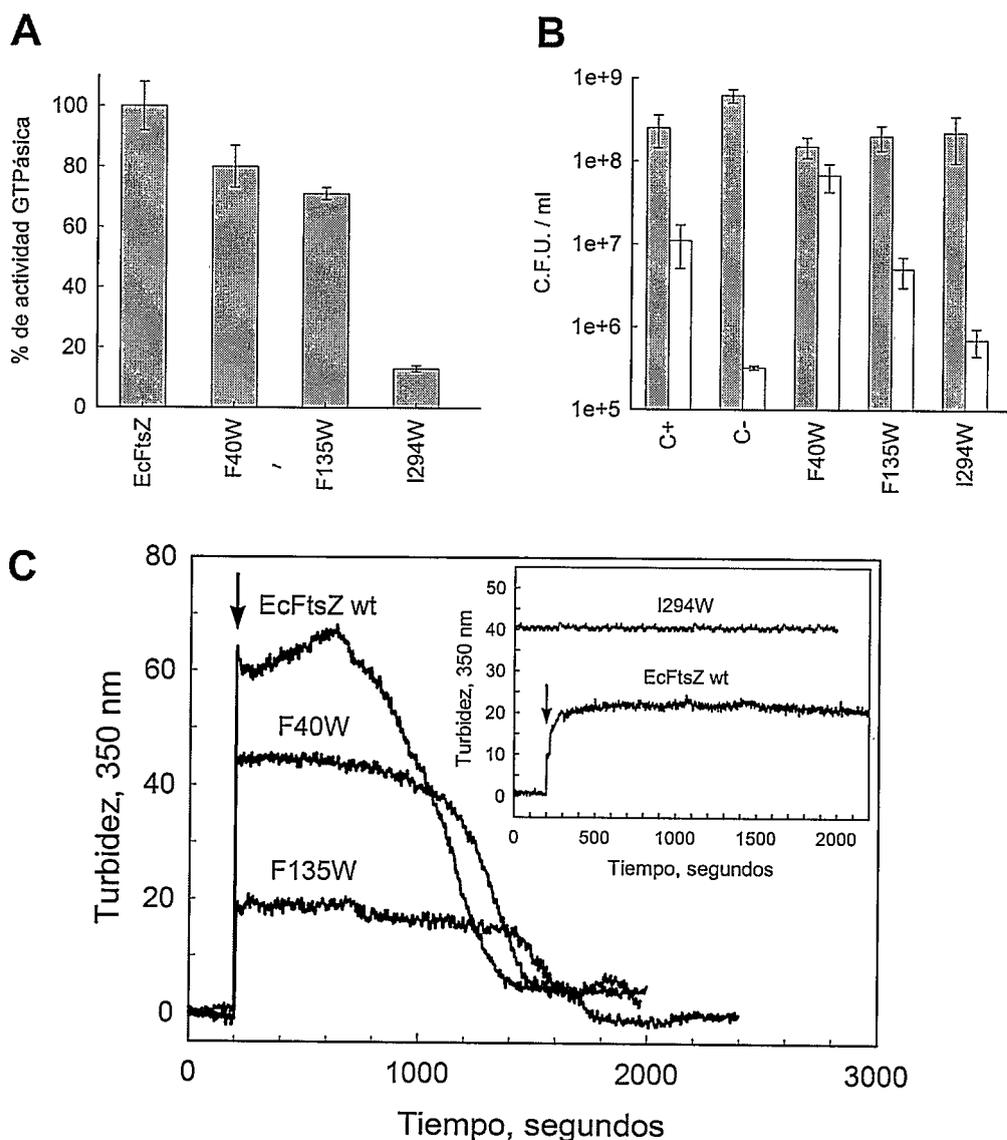


Figura 16. Caracterización funcional de las mutantes de EcFtsZ, F40W, F135W e I294W. En A, la hidrólisis del nucleótido se inicia agregando 1 mM GTP y se detiene con ácido perclórico al 10%. Cada proteína, 12 μ M, se incubó en tampón MES 50 mM pH 6,5 y 10 mM de $MgCl_2$ a 37°C, por 5 minutos. Los porcentajes de actividad son respecto a EcFtsZ wt. Los errores se obtuvieron del análisis de regresión lineal. En B, la complementación in vivo se estimó a partir del número de colonias a la temperatura restrictiva (42°C, barras blancas) comparado con el número de colonias a la temperatura permisiva (30°C, barras grises), para EcFtsZ wt y las tres mutantes transformadas en la cepa VIP-2. C+ corresponde a la cepa VIP-2 complementada con un plasmidio con el gen *ftsZ* silvestre, y C- es la cepa sin complementar. En C, polimerización de cada mutante y EcFtsZ wt, a una concentración de 25 μ M, seguida por dispersión de luz a 90°. En el inserto polimerización de EcFtsZ wt y la mutante I294W, a una concentración de 12 μ M. Las flechas indican el momento en que se agrega el GTP (resultados de R. Díaz y F. Montecinos).

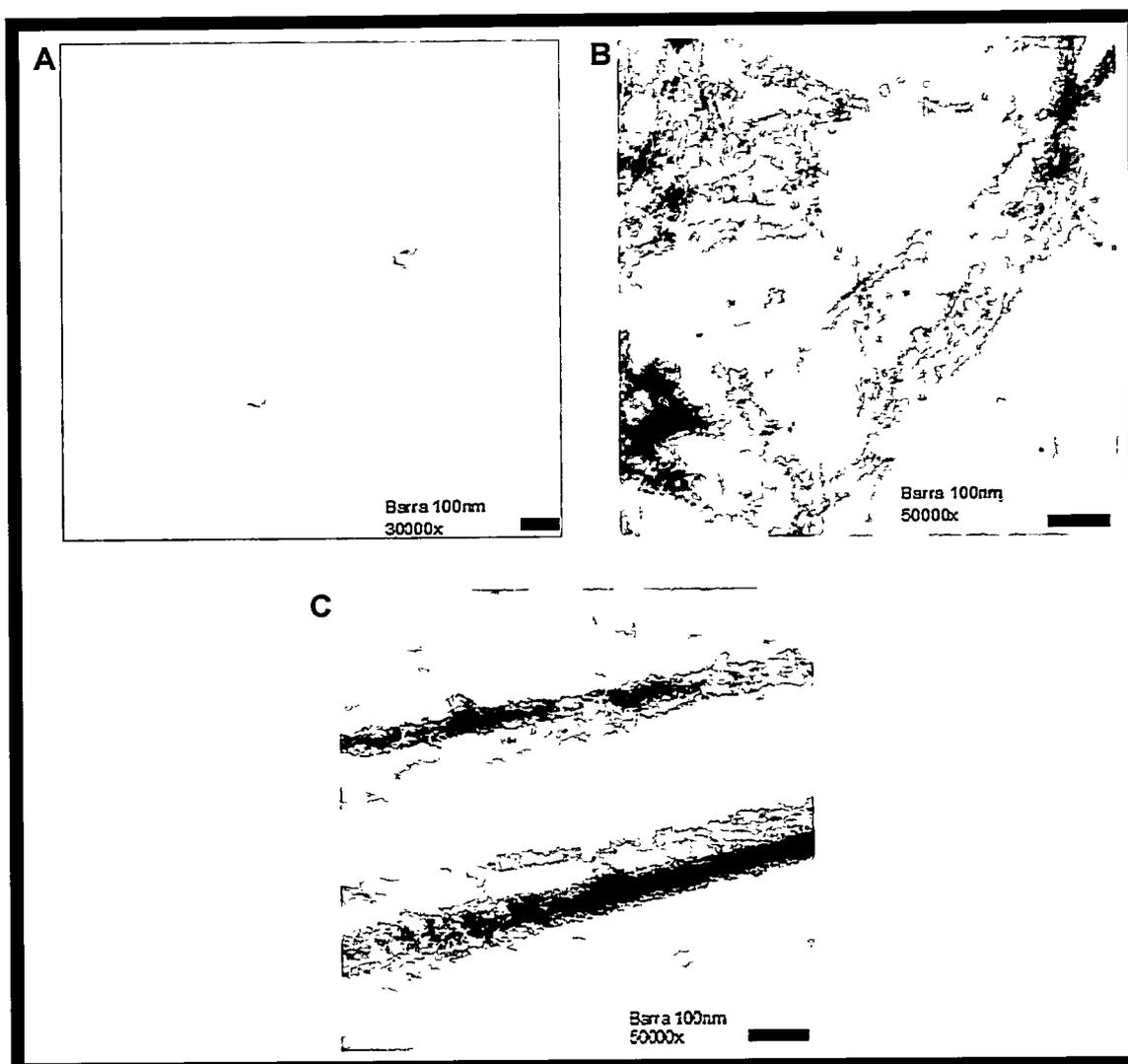


Figura 17. Micrografías electrónicas de EcFtsZ wt y la mutante I294W. En A, micrografía electrónica de EcFtsZ wt en condiciones no polimerizantes, sin GTP. En B, EcFtsZ wt en amortiguador de polimerización con GTP. C, micrografía electrónica de la mutante I294W en condiciones no polimerizantes, sin GTP. Las muestras se prepararon por tinción negativa con acetato de uranilo 1,0% (fotos de F. Montecinos).

Complementación in vivo. La cepa VIP-2/pLAR9 de *Escherichia coli* presenta una mutación en el gen cromosomal de FtsZ, esta mutación letal ha sido complementada con el plasmidio pLAR9 que le permite a la cepa replicarse normalmente a temperaturas menores a 37 °C. En condiciones permisivas (42 °C) pLAR9 no produce FtsZ y la bacteria no se replica (control negativo, C-), sin embargo al complementar la cepa VIP-2/pLAR9 con un segundo plasmidio que no presenta un origen de replicación termosensible y que incluye el gen de la *ftsZ*, las bacterias recuperan su capacidad de dividirse a 42 °C (control positivo, C+). La complementación de la cepa VIP-2/pLAR9 con un segundo plasmidio que contiene las mutantes de EcFtsZ, permite que a temperatura permisiva solo se produzca FtsZ mutante, y que el experimento de complementación permita evaluar si estas mutantes son funcionales *in vivo*. La mutante F40W complementó *in vivo* a la temperatura restrictiva para EcFtsZ wt, 42°C (figura 16B). La mutante F135W, dentro del error experimental también complementó, en cambio la mutante I294W prácticamente no complementó a 42°C.

Conformación de la mutante I294W

La caracterización funcional de la mutante I294W mostró la pérdida de un 87% de la actividad GTPásica y un estado polimerizado permanente. Estas alteraciones funcionales impiden que I294W complemente *in vivo*, a pesar de presentar características estructurales muy similares a EcFtsZ wt y a las mutantes F40W y F135W. Esto no se explica por la introducción de un grupo voluminoso como el triptófano, que podría alterar la estructura tridimensional de la proteína, pues esta es una zona móvil. Para estudiar las bases estructurales de la alteración funcional de I294W, se infirió la estructura de EcFtsZ, por modelado comparativo de proteínas, usando como molde MjFtsZ. El modelo de la mutante mostró que el triptófano 294 estaba

separado solo por 4,5 Å de la fenilalanina 275, sugiriendo que ambos residuos podrían interactuar alterando la flexibilidad del dominio carboxilo terminal y por ende la función de FtsZ, figura 18A. Si este fuera el caso debería existir una relación evolutiva entre ambas posiciones, mutaciones correlacionadas, que impidiera la aparición simultánea de dos aminoácidos aromáticos en estas posiciones en FtsZ. Para evaluar esta hipótesis, se analizaron las mutaciones correlacionadas para la posición I294, a partir de un alineamiento múltiple de FtsZ, figura 18B. La posición equivalente al aminoácido 275 de EcFtsZ, en el alineamiento múltiple mostró una clara alternancia con la posición 294 en el contenido de aminoácidos aromáticos, es decir si la posición 294 correspondía a un aminoácido aromático, la posición 275 presentaba un aminoácido alifático y viceversa, por ejemplo MjFtsZ presenta un triptófano en la posición equivalente a 294 y una alanina en la equivalente a 275 y MtbFtsZ una isoleucina en 275 y una fenilalanina en 294. Estos resultados sugieren que la causa de la inhibición producida por la mutación I294W sería la formación de un puente estabilizado por la interacción entre los anillos aromáticos de la fenilalanina 275 y el triptófano I294, que estabilizaría la conformación GTP de EcFtsZ. Los resultados experimentales obtenidos para las mutantes puntuales de EcFtsZ F135W e I294W mostraron un marcado efecto sobre la funcionalidad sin alterar la estructura secundaria y terciaria de la proteína, como era de esperar si estas mutantes pertenecían a regiones involucradas en cambios conformacionales, sin embargo no se esperaba obtener un fenotipo que sugiere que la proteína permanece en conformación GTP (I294W). Por otra parte, la mutante F40W que no participa en los LRCC presentó parámetros funcionales similares a EcFtsZ silvestre, como se esperaba por estudios bioinformáticos. Las pequeñas diferencias respecto a FtsZ wt podrían deberse a

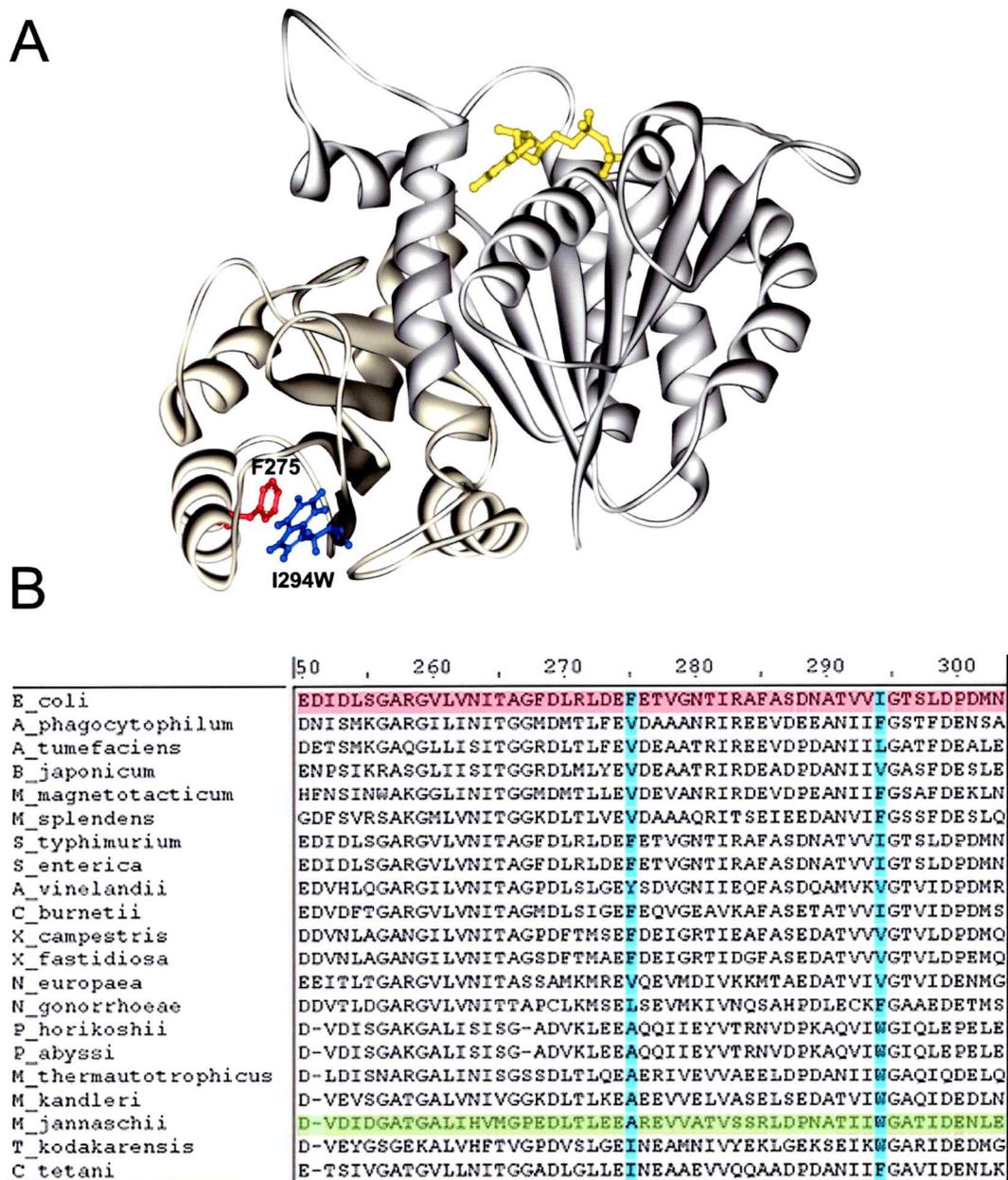


Figura 18. Mutaciones correlacionadas en las posiciones 275 y 294 de EcFtsZ. En A, se muestra un modelo de la estructura de la mutante I294W representada en cintas, donde se indica la posición del residuo triptófano 294 en rojo y fenilalanina 275 en azul. En esferas y cable amarillo se muestra el nucleótido. En B, se observa una parte del alineamiento múltiple de FtsZ, donde la columna de la izquierda indica el organismo y la columna de la derecha su respectiva secuencia aminoacídica. En azul se muestran las posiciones correlacionadas, en rojo la secuencia de EcFtsZ y en verde la secuencia de MjFtsZ (resultados de Andrea Garcés).

diferencias en la estructura tridimensional de la proteína, como lo sugiere la variación de un 2% del contenido de estructura secundaria de la mutante.

DISCUSIÓN

Análisis de la base de datos de proteínas en estado libre y como complejo proteína-ligando. El análisis de la base de datos PDB muestra que existe un número suficiente de estructuras de proteínas disponibles con y sin ligando, para caracterizar los cambios conformacionales en familias de proteínas, más aun con proteínas de una sola cadena polipeptídica. Con este subconjunto de estructuras se determinaron los LRCC sin considerar relaciones más complejas como el grado de oligomerización de la proteína. Esto no significa que todas las proteínas estudiadas tengan como forma biológica activa el estado monomérico, pues muchas veces la notación en los archivos PDB solo incluye la unidad asimétrica (ASU), pero el estado biológico de la proteína es un homodímero, homotrímero, etc (Jones and Thornton 1996; Henrick and Thornton 1998; Ponstingl, Kabir et al. 2003). Además la obtención automática de las proteínas en estado libre y acomplejada presentó varias dificultades principalmente dadas por variaciones en la manera de anotar la información en los archivos PDB, donde aparecen, por ejemplo, modificaciones covalentes en la proteína anotadas como heteroátomos (ligandos) (Westbrook, Feng et al. 2002). En estos casos se usó un filtro de distancia, donde las distancias menores a 1,9 Å, entre cualquier átomo del ligando y cualquier átomo de la proteína, representa una unión covalente, sin embargo como cualquier valor de corte para discriminar, produce falsos positivos y omisiones. Otras proteínas seleccionadas mostraron como ligandos más de un átomo metálico. No se descartaron estas proteínas, pues en algunos casos de proteínas seleccionadas se consideraba a los aniones como Cl^- , Br^- , y cationes como Na^+ , Ca^{2+} , etc., como ligandos. En estos casos es difícil determinar si estos iones son parte del medio de cristalización y su interacción

con la proteína inespecífica, y por lo tanto no serían verdaderos ligandos. Para estas proteínas se consideró que todos los iones correspondían a un solo ligando y de esa manera se analizó. También se observó que algunas proteínas en la base de datos que presentan distintas condiciones de cristalización, como por ejemplo diferentes pH, donde el cambio conformacional de la proteína podría deberse a la condición de cristalización y no a la unión del ligando. A pesar de estos casos, la gran mayoría de las proteínas en estado libre y acomplejada cumplieron con los criterios establecidos para la comparación.

La predicción de cambios conformacionales al comparar las estructuras 3D de una misma proteína en estado libre y acomplejada, mostró que para proteínas cristalizadas en distintos grupos espaciales, se pueden producir errores en la predicción de cambios conformacionales (figura 12) o incluso como lo muestra un estudio, omitir el papel de residuos claves en la función de una proteína (Andre, Nonato et al. 2005). Más aun la correcta predicción de los cambios conformacionales depende directamente de la magnitud de los movimientos en la estructura al unir el ligando, debido a que cambios en la orientación de las cadenas laterales de los residuos en la superficie de la proteína, se observan normalmente en difracción de rayos X al generar varios cristales de la misma proteína (Luque and Freire 2002). Para minimizar este problema, en esta tesis se compararon solo los C α de las proteínas en estado libre y acomplejada.

La caracterización de los LRCC a partir del análisis de la base de datos de cambios conformacionales mostró que alrededor de un 8% de los residuos de la proteína participan en cambios conformacionales, un resultado esperable si se considera que para que una proteína cumpla su función biológica debe ser estable y presentar cierto grado de flexibilidad en la superficie, que le permita unir el ligando y adoptar una

conformación activa (Vihinen 1987; Závodszky, Kardos et al. 1998). La estabilidad estructural se lograría mediante un núcleo rígido de la proteína y la función mediante un sitio de unión del ligando y regiones flexibles en la estructura responsable de los cambios conformacionales (Hilser, Dowdy et al. 1998; Luque and Freire 2000). Los resultados obtenidos en esta tesis, sugieren que el SA tendría un papel dual, por un lado formaría parte del núcleo rígido responsable de la estabilidad de la proteína (figura 8), y por otro lado une al ligando y permite la propagación de los cambios conformacionales mediante una fracción de sus residuos (alrededor del 18%), a través de regiones flexibles (LRCC, figura 8), en la superficie de la proteína (figura 7). La distribución de los residuos que participan en LRCC alrededor del SA indica que los cambios conformacionales están directamente relacionados a la unión del ligando (figura 5) y en promedio se propagan al menos a 4 residuos de distancia del SA, lo que se ajusta a la gran cantidad de evidencia en literatura, y en esta tesis sobre el efecto que presentan mutaciones puntuales alejadas del SA sobre la función de las proteínas.

Conservación de los LRCC en algunas familias de proteínas. Existen proteínas de una misma familia, descritas en la literatura, que muestran conservación de los cambios conformacionales, es el caso de las integrinas, que median procesos de adhesión celular y que para la transducción del estímulo dado por el ligando mueven una hélice del dominio MIDAS ("Metal ion-dependent adhesion site") más de 10 Å hacia el dominio intracelular (Kleanthous 2000; Xiao, Takagi et al. 2004). Otro ejemplo son las proteínas G (Sprang 1997), esta familia une e hidroliza GTP, que produce cambios conformacionales conservados en las regiones "switch I" y "switch II", que le permiten formar complejos proteína - proteína. Los ejemplos de conservación de cambios conformacionales en las proteínas de familias de receptores de membrana, como las

integrinas y la proteína G, se extienden a otras proteínas globulares y solubles de dos especies distintas que presentan conservación de sus cambios conformacionales, como en el caso de proteínas de la familia de las lisozimas, donde se han comparado por NMR y difracción de rayos X los cambios conformacionales entre la lisozima humana y de *Gallus domesticus* y se ha demostrado que son las mismas regiones en la estructura las que se mueven con la unión del ligando (Mine, Ueda et al. 2000; Ohmura, Motoshima et al. 2002). El análisis bioinformático de las proteínas de la familia de las lisozimas presentes en la base de datos de cambios conformacionales (1.10.530.10), mostró patrones de LRCC conservados para 27 proteínas con un rango de identidad de secuencia entre 56% y 100% (no se muestran los resultados), confirmando lo observado por NMR y difracción de rayos X en lisozima humana y de *Gallus domesticus*. Esta tendencia a la conservación de los cambios conformacionales se hace evidente en los resultados del análisis de la base de datos de cambios conformacionales, ver figuras 9, 10, 11 y 12. En la familia de las fosfotransferasas (fig. 9) existe gran variación en la identidad de secuencia de las proteínas que la conforman, con identidades de secuencia de hasta un 22 %, sin embargo los ligandos se unen al mismo sitio activo y los patrones de cambios conformacionales son similares. En la figura 12, se muestra otro ejemplo de la conservación de los LRCC, esta vez para la familia de las serino proteasas (Perona and Craik 1995) que presenta un sitio de unión conservado, S1, donde ocurre la hidrólisis de la unión peptídica. Para las serino proteasas, tipo tripsina existen una gran variedad de inhibidores, estos ligandos emulan péptidos unidos al SA. En presencia de cationes divalentes que se comportan como coinhibidores y se unen al mismo sitio que los inhibidores orgánicos hay un aumento de la afinidad de los péptidos al coordinarse los cationes tanto a la proteína como al inhibidor (Katz, Clark et al. 1998). Estos

coinhibidores muestran el mismo patrón de LRCC que se produce con los inhibidores orgánicos. Los resultados también mostraron patrones conservados de cambios conformacionales para las familias de las globinas (1.10.490.10), citocromo P450 (1.10.630.10), anhidrasa carbónica (3.10.200.10) y timidilato sintetasa (3.30.572.10), resultados que no se muestran.

El hecho que la familia de las fosfotransferasas (1.10.510.10) presente LRCC similares para la unión de diferentes ligandos a un mismo SA, en la misma proteína (fig. 9), se explica porque solo participan en los LRCC en promedio uno de cada cinco aminoácidos del SA, por lo tanto un ligando que contacte solo algunos residuos del SA puede generar un patrón similar de cambios conformacionales que otro donde todos los residuos del SA participan en la unión. Por otro lado, la observación de patrones similares de cambios conformacionales para los residuos de LRCC cercanos al SA en proteínas de la misma familia con distinta identidad de secuencia, puede explicarse, debido a que los aminoácidos que participan en el SA de las proteínas son muy conservados (Aloy, Querol et al. 2001; Gutteridge, Bartlett et al. 2003). En cambio los residuos involucrados en LRCC alejados del SA, en general no son conservados a nivel de identidad de secuencia (resultados no mostrados), no obstante se ha observado que algunas propiedades en la estructura, como el grado de flexibilidad inferidas mediante el factor térmico, clasifican a estas regiones de la estructura como flexibles. Esta propiedad esta directamente relacionada a la composición aminoacídica de estas regiones de la estructura, que son ricas en aminoácidos expuestos y polares: G, S, N y D, figura 6A, resultado sustentado por metodologías de predicción de zonas rígidas y flexibles (Radivojac, Obradovic et al. 2004). Un ejemplo puntual es la proteína FtsZ, donde se

observó un patrón similar de flexibilidad en las estructuras provenientes de dos especies distintas que presentan una identidad de secuencia del 48% (fig. 13 C,D).

El comportamiento de los LRCC inducidos por la unión de ligandos a dos sitios de unión distintos en las proteínas, se analizó con la familia dihidrofolato reductasa (DHFR, código CATH 3.40.430.10), las proteínas de esta familia catalizan el paso de dihidrofolato a tetrahidrofolato, y presentan un sitio de unión para el sustrato y un segundo sitio de unión para la coenzima NADPH_2 que participa en la reacción de oxido-reducción. En estudios espectroscópicos con mutaciones sitio dirigidas (Agarwal, Billeter et al. 2002) se observó que la actividad de la enzima se perturba si se mutan aminoácidos del SA, y también residuos que pertenecen a elementos estructurales alejados del sitio activo, como el "lazo" Met 20 y el "lazo" $\beta\text{F} - \beta\text{G}$. Por simulaciones de dinámica molecular, se determinó que estos lazos modifican su posición hasta permitir que el sustrato y la coenzima queden muy próximos y debidamente orientados para la reacción de transferencia de un protón con su electrón. En esta tesis, la mayor parte de los pares de proteínas libres y acomplejadas de la familia DHFR analizadas, mostraron la unión del segundo ligando al sitio de NADPH_2 o al sitio de dihidrofolato (DDF), es decir a uno de ellos mientras el otro contenía el ligando unido, la figura 11A muestra las estructuras de las proteínas con y sin un ligando en el sitio de unión de DDF, con el sitio de unión para la coenzima NADPH_2 ocupado. De la misma manera en la figura 11B se observan las estructuras con y sin un ligando en el sitio de unión de NADPH_2 , donde DDF esta unido a la proteína. Los patrones de LRCC confirman la importancia de los lazos Met-20 y $\beta\text{F} - \beta\text{G}$ en los cambios conformacionales en la familia DHFR (figura 19). Sin embargo cabe destacar que los resultados demuestran que ambos sitios de unión presentan LRCC diferentes y característicos (fig. 11 A y B). Los resultados en su conjunto, indican que el

plegamiento de la proteína es el determinante principal de los LRCC, pues son conservados dentro de la misma familia de proteínas e independientes de la naturaleza del ligando, siempre y cuando éste se una a una misma región en la superficie de la proteína.

Elementos estructurales que participan en los LRCC. En la figura 6, se muestran las distintas proporciones de elementos de estructura secundaria presentes en los LRCC, donde se observa que los lazos son los elementos estructurales más frecuentes, y en menor grado las vueltas β . No es de sorprender la presencia de lazos en los LRCC, pues debido a sus características estructurales son los elementos ideales para participar en cambios conformacionales. Destacan entre sus características, que no tienen patrones definidos de puentes hidrógeno (Eswar, Ramakrishnan et al. 2003), normalmente se encuentran expuestos en la superficie de las proteínas y presentan alto grado de flexibilidad, determinado por estudios de NMR y difracción de rayos X. Esta flexibilidad de los lazos es el factor determinante para orientar los aminoácidos catalíticos en enzimas como lipasas y enolasas (Gunasekaran, Ma et al. 2003; Gunasekaran and Nussinov 2004), y le proporciona cierta flexibilidad al sitio catalítico. Así lo demuestran estudios en la familia de las serino proteasas, donde se modificaron con mutaciones sitio dirigidas el entorno del sitio activo, con el fin de aumentar su flexibilidad, y como consecuencia la enzima amplió el rango de ligandos que reconoce (Perona and Craik 1995). La caracterización de los cambios conformacionales de las enzimas dihidrofolato reductasa (Agarwal, Billeter et al. 2002) y lisozima (Ohmura, Motoshima et al. 2002) mostraron como los lazos fueron elementos estructurales claves para cumplir su función.

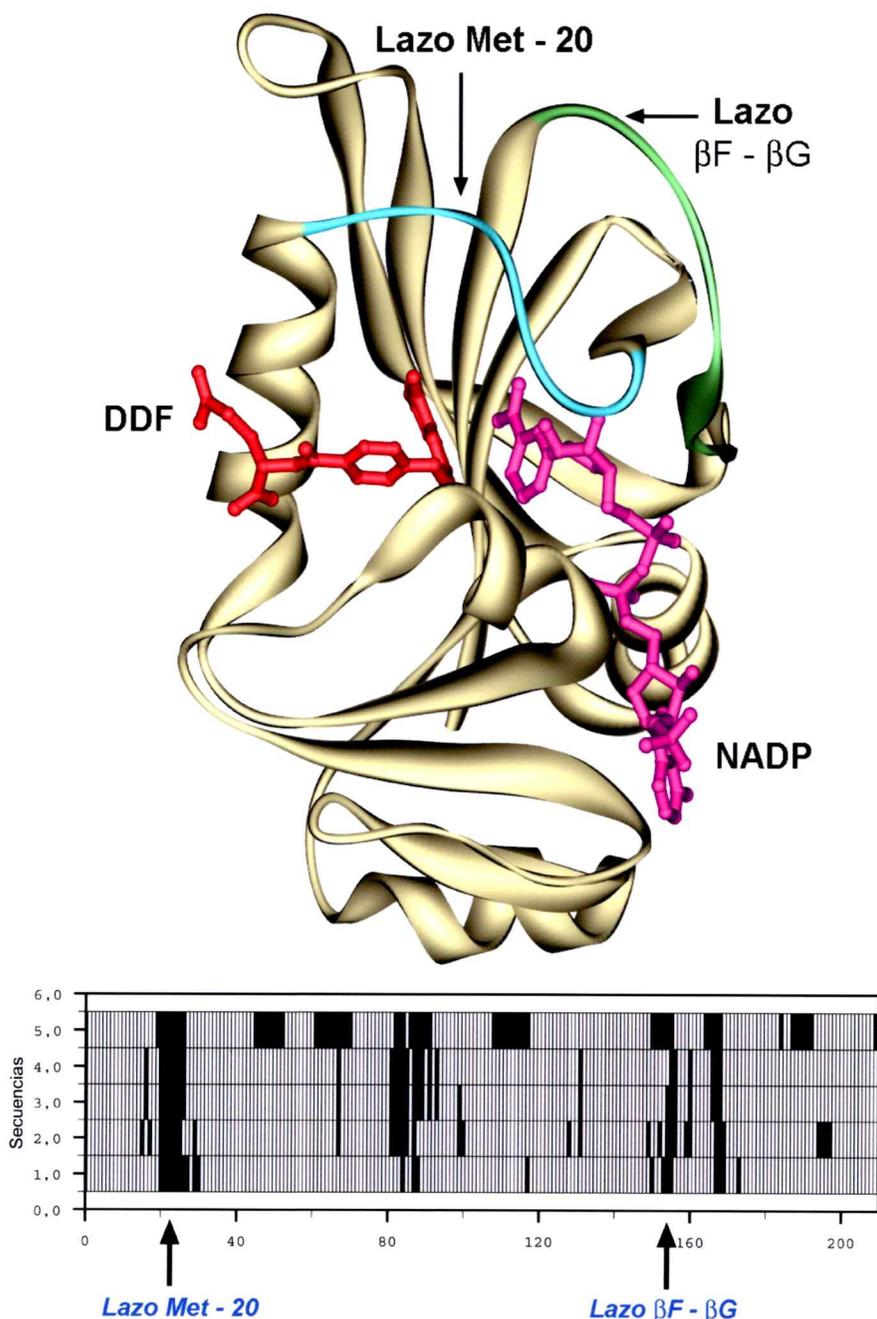


Figura 19. Regiones alejadas del SA que participan en LRCC en la familia DHFR. En A, esquema de cintas que muestra la enzima dihidrofolato reductasa de *Escherichia coli* (1rx2.pdb) con DDF y NADP en los sitios activos, representados como cables rojos y violeta respectivamente. En B, gráfico de densidad que muestra los LRCC para la unión del ligando al sitio de NADP con el sitio de DDF ocupado. En *itálica azul* se indican las zonas de cambios conformacionales en el gráfico de densidad, que corresponden a los lazos Met - 20 y $\beta F - \beta G$.

La abundancia de lazos en los LRCC, y la poca variabilidad en el contenido de estructura secundaria producida por la unión del ligando (resultados no se muestran), sugieren que los movimientos en la proteína inducidos por la unión del ligando son movimientos de cuerpos rígidos, donde elementos de estructura secundaria como hélices α y hojas β relativamente superficiales se mueven sin alterar el núcleo de la proteína, es decir la estabilidad global de la macromolécula (Hilser, Dowdy et al. 1998). La aspartato transcarbamilasa presenta este tipo de movimientos de cuerpos rígidos (Yon, Perahia et al. 1998). La presencia de movimientos de cuerpos rígidos en nuestra base de datos, no se aplica a proteínas no estructuradas o poco estructuradas (Dunker, Lawson et al. 2001; Dyson and Wright 2002; Gunasekaran, Tsai et al. 2003) pues pueden presentar otro tipo de movimientos que incluyan la variación en el contenido de estructura secundaria al unir el ligando. La pequeña población de este tipo de proteínas en la base de datos, debido a que son difíciles de cristalizar, no permitió dilucidar el problema.

Influencia de la flexibilidad sobre la función de FtsZ. La hidrólisis de GTP es esencial para la dinámica de polimerización de FtsZ (Nogales, Whittaker et al. 1999). Por esto cuatro estructuras cristalinas de FtsZ fueron elegidas para comparar las conformaciones GTP y GDP, debido a que tanto para MjFtsZ como para MtbFtsZ los grupos espaciales de los cristales en estado GTP y GDP coincidían. El hecho que ambas conformaciones tengan el mismo grupo espacial es esencial para evitar interferencias debidas a distintos contactos cristalográficos en la predicción de los cambios conformacionales. Sin embargo la inducción de la conformación GTP / GDP en MtbFtsZ se hizo difundiendo el nucleótido sobre la proteína cristalizada en presencia de citrato, lo que se consideró para el análisis de los resultados. De manera similar la conformación GTP / GDP en MjFtsZ se obtuvo difundiendo un análogo no hidrolizable de GTP sobre el cristal en conformación

GDP. Considerando estas restricciones, el análisis de los archivos PDB permitió determinar la flexibilidad implícita y funcional de FtsZ, para identificar aquellas regiones involucradas en los cambios conformacionales relacionados a la hidrólisis de GTP. Se analizó la estructura de MtbFtsZ unida a GTP (1RLU) y a GDP (1RQ7) (Leung, White et al. 2004) y la conformación de MjFtsZ cristalizada con GTP (1W58) y GDP (1FSZ). Los resultados mostraron que el intercambio de GDP por GTP induce pequeños movimientos (tabla 3) principalmente en regiones específicas en la superficie de las proteínas, que incluye regiones próximas al sitio de unión de GTP y otras regiones alejadas del sitio de unión del nucleótido, en el dominio carboxilo terminal. El análisis, mediante el factor térmico, de la flexibilidad intrínseca de FtsZ en distintos organismos mostró que son las mismas regiones de los dominios amino y carboxilo terminal las que presentaron flexibilidad, sugiriendo una propensión de estas regiones a sufrir rearrreglos con los cambios conformacionales. Coincidentemente, los elementos estructurales en el dominio amino y carboxilo terminal de dos miembros de la familia de las FtsZ, el mesófilo (MtbFtsZ) y el termófilo (MjFtsZ) mostraron una dinámica similar asociada al intercambio entre la conformación GTP y GDP. Se sabe que la hidrólisis de GTP es inducida por la interacción de la cara amino longitudinal de un monómero y la cara carboxilo longitudinal de otro monómero (Erickson 2000). La interacción longitudinal es transitoria mientras el nucleótido se encuentra como GTP lo que permite la despolimerización en estado GDP después de la hidrólisis. El intercambio del nucleótido GDP por GTP induce la polimerización de la proteína (Chen, Bjornson et al. 2005). Postulamos que la flexibilidad alrededor del sitio del nucleótido y en la región carboxilo terminal de la FtsZ permite los cambios conformacionales necesarios para que la conformación GTP cambie a la conformación

GDP después de la hidrólisis y debilita la estructura del polímero. Así, GTP estabiliza la conformación recta del polímero, mientras que GDP la desestabiliza. Por lo tanto una mutación que perturbe la flexibilidad alteraría el equilibrio entre estas conformaciones afectando la polimerización o despolimerización, dependiendo de que conformación se favorezca. Triptófano es el residuo con mayor volumen hidrodinámico entre todos los aminoácidos, por lo tanto una mutación puntual por triptófano en la proteína produciría una perturbación estérica y rearrreglos estructurales. Este efecto debiera ser más marcado en zonas rígidas, como se observa con la mutante F40W, donde se altera levemente el espectro de dicroísmo circular y la función. Si se introduce en zonas más flexibles no debiera observarse un efecto sobre la estructura. La función debiera afectarse si es un residuo que participa en la función, por ejemplo si se altera la flexibilidad y de este modo una conformación necesaria para la función. Así, la mutante puntual F135W, localizada en una zona flexible que participa en los LRCC, lazo S5 – H5 del dominio amino, presentó una disminución en la funcionalidad, la polimerización se redujo a un 32% y la actividad GTPásica a un 71% respecto a EcFtsZ wt, y mostró complementación *in vivo*, aunque la estructura y la estabilidad no se afectaron. La mutación I294W localizada en una región flexible del dominio carboxilo que participa en los LRCC, en la hebra β S9, afectó severamente la funcionalidad, pues favoreció el estado polimerizado, con una baja actividad GTPásica (13%) y una complementación *in vivo* residual. Los resultados apoyan la hipótesis de que una pequeña modificación en las zonas flexibles que afecten a los cambios conformacionales, afectan la funcionalidad de FtsZ. La actividad GTPásica de MjFtsZ se afectó cuando el residuo W319, localizado en la posición estructuralmente equivalente al aminoácido I294, se cambió por tirosina (Oliva, Huecas et al. 2003). Estos resultados y el hecho que la mutación I294W no afecta

la estabilidad estructural de FtsZ, como lo demostraron los espectros de fluorescencia y CD, sugieren un marcado efecto sobre la flexibilidad causado por la mutación en la posición I294. La estabilización de una conformación polimérica se debería a la formación de un puente por la interacción aromática entre el triptófano 294 y la fenilalanina 275, que estabilizaría la conformación GTP. La energía típica de una interacción aromática es de alrededor de 1 o 2 kcal/mol, mientras que la contribución a la estabilidad global de la proteína varía entre $-0,6$ y $+1,3$ kcal, dependiendo del ambiente del par aromático (Burley and Petsko 1985). Los aminoácidos aromáticos F275 y W294 presentan una accesibilidad intermedia al solvente (entre el 5 y el 40%), con un ambiente mayormente hidrofóbico, por lo tanto la interacción aromática debería ser más favorable, produciéndose una pérdida de flexibilidad en la región.

Otra posible explicación de los resultados obtenidos con la mutante I294W podría ser que el lazo T7, responsable de la inducción de la actividad GTPásica, estabilizara la interacción longitudinal entre los monómeros, aun en estado GDP al favorecer una conformación GTP. Sin embargo esta explicación no daría cuenta de un cambio concertado entre la hidrólisis del fosfato gama y el cambio conformacional a la conformación GDP. Una explicación razonable que daría cuenta de todos los resultados sería que la hélice H1 o H4 con su dipolo positivo ayude a estabilizar una conformación GTP. Al hidrolizarse el fosfato gama se pierde la interacción entre el fosfato gama con el fosfato beta dando libertad a la hélice, la que induciría un cambio conformacional donde estaría involucrado el lazo T3 del dominio amino de un monómero. Este cambio conformacional se transmitiría hacia el dominio carboxilo terminal del monómero adyacente a través de la superficie y el lazo T7 que llevaría a la conformación GDP. Al impedir este último cambio con la mutación I294 no se formaría la conformación GDP

aunque se hubiera producido la hidrólisis, dejando a la proteína más bien en una conformación GTP, como polímero. Esta hipótesis se sustenta en que el sitio activo de la mutante conservó su capacidad para hidrolizar GTP, demostrado al lavar los polímeros de FtsZ para eliminar el nucleótido libre y posteriormente denaturarlos con GdmCl, donde se aisló el nucleótido en una proporción mayor al 95% en estado GDP (Montecinos 2006). Estos resultados mostraron que la disminución en la actividad GTPásica se debió más bien a un lento intercambio de GDP por GTP, probablemente asociado a la polimerización y formación de manojos de filamentos (figura 17C) (Marrington, Small et al. 2004).

Proponemos un modelo donde diferentes conformaciones de FtsZ existen en el polímero y dependen del estado de fosforilación, donde GTP favorece las interacciones longitudinales y laterales en los pares de filamentos y GDP desestabiliza las interacciones longitudinales y laterales a través de un cambio conformacional en el dominio amino que se extiende de forma cooperativa, longitudinalmente al dominio carboxilo. La regulación fina de la interacción longitudinal y lateral se lleva a cabo mediante aminoácidos específicos (LRCC) en ambas interfases, dependiendo del estado de fosforilación del nucleótido (fig. 20). Este modelo requiere de zonas flexibles en el dominio amino y carboxilo terminal que debiliten las interacciones en el polímero de FtsZ, y no de un gran cambio conformacional de los monómeros para despolimerizar los protofilamentos. Nuestro modelo de regiones flexibles responsables de la funcionalidad de FtsZ puede explicar el efecto de las mutantes puntuales, encontradas en literatura, sobre la funcionalidad de EcFtsZ (fig. 1) (Lutkenhaus 1990; Lu, Stricker et al. 2001; Chen, Bjornson et al. 2005).

La caracterización estructural, además del análisis funcional de las mutantes de EcFtsZ provee evidencia experimental y teórica de que la rigidización del dominio carboxilo en conformación GTP, puede estabilizar los protofilamentos, como sucede con los polímeros de la mutante I294W debido a la formación de una interacción aromática.

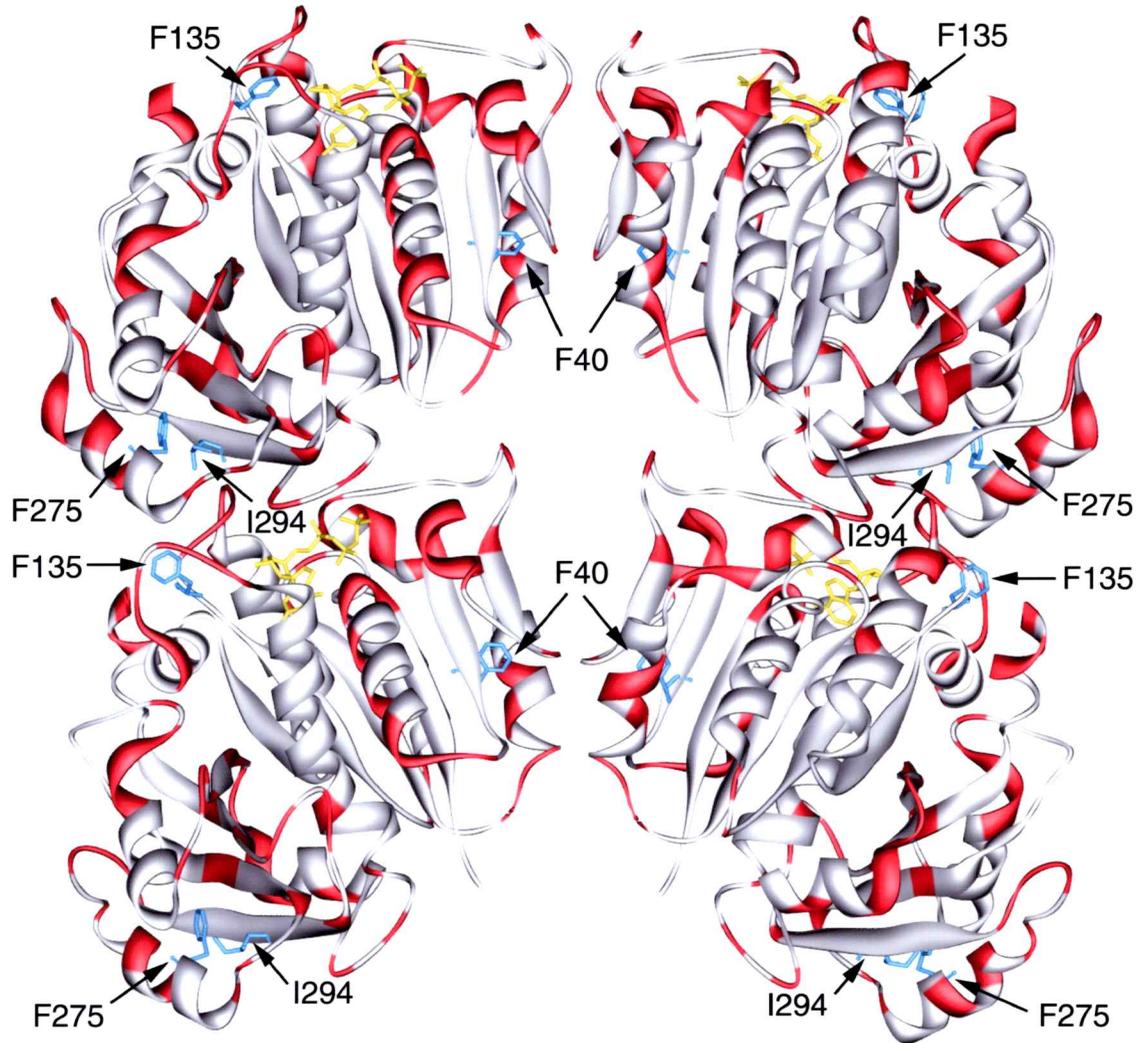


Figura 20. Modelo de interacción longitudinal y lateral en los doble filamentos de FtsZ. En esquema de cintas se muestra un tetrámero de EcFtsZ que representa verticalmente las interacciones longitudinales y horizontalmente las interacciones laterales en los polímeros de FtsZ, de acuerdo a Lowe y col. 1999 (Lowe and Amos 1999). El nucleótido se muestra con barras amarillas, las mutantes F40W, F135W e I294W se representan con sus cadenas laterales en azul. En rojo se muestran los LRCC.

CONCLUSIONES

1. En promedio un 3% de los residuos en las proteínas forman parte del sitio activo, y un 8% forman parte de los LRCC. La distribución de los LRCC es más importante alrededor del sitio activo pero puede alcanzar regiones más alejadas en la estructura.
2. La composición aminoacídica de los LRCC se caracteriza por la presencia de residuos como: G, S, D y N. Los elementos de estructura secundaria más frecuentes son vueltas β y lazos.
3. Los LRCC se caracterizan por su alta proporción de residuos expuestos al solvente, en cambio los residuos del sitio activo se caracterizan por una accesibilidad intermedia. El grado de exposición se correlaciona con la flexibilidad de los residuos, donde los LRCC muestran mayor flexibilidad expresada como factor térmico normalizado, a diferencia de otros residuos más rígidos del sitio activo.
4. Los LRCC para un mismo sitio activo son conservados en familias de proteínas e independientes de la naturaleza del ligando. Sitios activos distintos, dentro de una familia de proteínas, presentan LRCC distintos y característicos para cada sitio.

5. Análisis de LRCC y del factor térmico para FtsZ de *Mycobacterium tuberculosis* y *Methanococcus jannaschii*, mostraron regiones flexibles similares en la estructura que estarían participando en el intercambio entre la conformación GTP y GDP.

6. Las mutaciones puntuales que afectan los cambios conformacionales, perturban la función de la proteína sin un cambio significativo de su estructura 3D.

7. La mutante I294W favorece la conformación GTP en FtsZ, a través de una interacción aromática que rigidiza la estructura, favoreciendo las interacciones longitudinales y laterales en los polímeros y disminuyendo la actividad GTPásica.

8. Los cambios conformacionales inducidos por la hidrólisis de GTP en los polímeros de FtsZ, modificarían regiones puntuales de la estructura involucradas en las asociaciones longitudinales y laterales de FtsZ, produciendo la desestabilización de los polímeros.

BIBLIOGRAFIA

1. Agarwal, P. K., S. R. Billeter, et al. (2002). "Network of coupled promoting motions in enzyme catalysis." *Proc Natl Acad Sci U S A* **99**(5): 2794-9.
2. Aloy, P., E. Querol, et al. (2001). "Automated structure-based prediction of functional sites in proteins: applications to assessing the validity of inheriting protein function from homology in genome annotation and to protein docking." *Journal of Molecular Biology* **311**(2): 395.
3. Aloy, P. and R. B. Russell (2002). "Interrogating protein interaction networks through structural biology." *PNAS* **99**(9): 5896-5901.
4. Altschul, S. F., T. L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucl. Acids Res.* **25**(17): 3389-3402.
5. Andre, L. B. A., M. C. Nonato, et al. (2005). "A Molecular Mechanism for Lys49-Phospholipase A2 Activity Based on Ligand-induced Conformational Change." *Journal of Biological Chemistry* **280**(8): 7326.
6. Berman, H. M., J. Westbrook, et al. (2000). "The Protein Data Bank." *Nucleic Acids Research* **28**(1): 235.
7. Beuria, T. K., S. S. Krishnakumar, et al. (2003). "Glutamate-induced Assembly of Bacterial Cell Division Protein FtsZ." *J. Biol. Chem.* **278**(6): 3735-3741.
8. Bi, E. and J. Lutkenhaus (1990). "Analysis of ftsZ mutations that confer resistance to the cell division inhibitor SulA (SfiA)." *J Bacteriol* **172**(10): 5602-9.
9. Bramhill, D. and C. M. Thompson (1994). "GTP-dependent polymerization of Escherichia coli FtsZ protein to form tubules." *Proc Natl Acad Sci U S A* **91**(13): 5813-7.
10. Burley, S. K. and G. A. Petsko (1985). "Aromatic - Aromatic Interaction: A Mechanism of Protein Structure Stabilization." *Science* **229**(4708): 23 - 28.
11. Carugo, O. and S. Pongor (2002). "Protein fold similarity estimated by a probabilistic approach based on C[alpha]-C[alpha] distance comparison." *Journal of Molecular Biology* **315**(4): 887.
12. Chen, Y., K. Bjornson, et al. (2005). "A Rapid Fluorescence Assay for FtsZ Assembly Indicates Cooperative Assembly with a Dimer Nucleus." *Biophys. J.* **88**(1): 505-514.
13. Cordell, S. C., E. J. Robinson, et al. (2003). "Crystal structure of the SOS cell division inhibitor SulA and in complex with FtsZ." *Proceedings of the National Academy of Sciences of the United States of America* **100**(13): 7889.
14. Dai, K., A. Mukherjee, et al. (1994). "Mutations in ftsZ that confer resistance to SulA affect the interaction of FtsZ with GTP." *J Bacteriol* **176**(1): 130-6.
15. Díaz, J. F., A. Kralicek, et al. (2001). "Activation of cell division protein FtsZ. Control of switch loop T3 conformation by the nucleotide gamma-phosphate." *Journal of Biological Chemistry* **276**(20): 17307.
16. Domingues, F. S., J. Rahnenfuhrer, et al. (2004). "Automated clustering of ensembles of alternative models in protein structure databases." *Protein Eng Des Sel* **17**(6): 537-43.
17. Dunker, A. K., J. D. Lawson, et al. (2001). "Intrinsically disordered protein." *Journal of Molecular Graphics and Modelling* **19**(1): 26.

18. Dyson, H. J. and P. E. Wright (2002). "Coupling of folding and binding for unstructured proteins." *Current Opinion in Structural Biology* **12**(1): 54.
19. Erickson, H. P. (1995). "FtsZ, a prokaryotic homolog of tubulin?" *Cell* **80**(3): 367-70.
20. Erickson, H. P. (2000). "Dynamin and FtsZ. Missing links in mitochondrial and bacterial division." *J Cell Biol* **148**(6): 1103-5.
21. Erickson, H. P., D. W. Taylor, et al. (1996). "Bacterial cell division protein FtsZ assembles into protofilament sheets and minirings, structural homologs of tubulin polymers." *Proceedings of the National Academy of Sciences of the United States of America* **93**(1): 519.
22. Eswar, N., C. Ramakrishnan, et al. (2003). "Stranded in isolation: structural role of isolated extended strands in proteins." *Protein Eng.* **16**(5): 331-339.
23. Fan, H. and A. E. Mark (2003). "Relative stability of protein structures determined by X-ray crystallography or NMR spectroscopy: a molecular dynamics simulation study." *Proteins* **53**(1): 111-20.
24. Freire, E. (1999). "The propagation of binding interactions to remote sites in proteins: analysis of the binding of the monoclonal antibody D1.3 to lysozyme." *Proceedings of the National Academy of Sciences of the United States of America* **96**(18): 10118.
25. Gerstein, M. and N. Echols (2004). "Exploring the range of protein flexibility, from a structural proteomics perspective." *Current Opinion in Chemical Biology* **8**(1): 14.
26. Gerstein, M., N. Echols, et al. (2003). "MolMovDB: analysis and visualization of conformational change and structural flexibility." *Nucleic Acids Research* **31**(1): 478.
27. Goehring, N. W. and J. Beckwith (2005). "Diverse paths to midcell: assembly of the bacterial cell division machinery." *Curr Biol* **15**(13): R514-26.
28. Gunasekaran, K., B. Ma, et al. (2003). "Triggering Loops and Enzyme Function: Identification of Loops that Trigger and Modulate Movements." *Journal of Molecular Biology* **332**(1): 143.
29. Gunasekaran, K. and R. Nussinov (2004). "Modulating Functional Loop Movements: The Role of Highly Conserved Residues in the Correlated Loop Motions." *ChemBioChem* **5**(2): 224-230.
30. Gunasekaran, K., C.-J. Tsai, et al. (2003). "Extended disordered proteins: targeting function with less scaffold." *Trends in Biochemical Sciences* **28**(2): 81.
31. Gutteridge, A., G. J. Bartlett, et al. (2003). "Using A Neural Network and Spatial Clustering to Predict the Location of Active Sites in Enzymes." *Journal of Molecular Biology* **330**(4): 719.
32. Henrick, K. and J. M. Thornton (1998). "PQS: a protein quaternary structure file server." *Trends in Biochemical Sciences* **23**(9): 358.
33. Hilser, V. J., D. Dowdy, et al. (1998). "The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble." *Proceedings of the National Academy of Sciences of the United States of America* **95**(17): 9903.
34. Hubbard, S. J. and J. M. Thornton (1993). NACCESS. Department of Biochemistry and Molecular Biology, University College, London.
35. Jones, S. and J. M. Thornton (1996). "Principles of protein-protein interactions." *Proceedings of the National Academy of Sciences of the United States of America* **93**(1): 13.

36. Joshi, H. C., M. J. Palacios, et al. (1992). "Gamma-tubulin is a centrosomal protein required for cell cycle-dependent microtubule nucleation." *Nature* **356**(6364): 80-3.
37. Katz, B. A., J. M. Clark, et al. (1998). "Design of potent selective zinc-mediated serine protease inhibitors." *Nature* **391**(6667): 608.
38. Kern, D. and E. R. P. Zuiderweg (2003). "The role of dynamics in allosteric regulation." *Current Opinion in Structural Biology* **13**(6): 748.
39. Kleanthous, C. (2000). *Protein - Protein Recognition*, Oxford University Press.
40. Koradi, R., M. Billeter, et al. (1996). "MOLMOL: A program for display and analysis of macromolecular structures." *Journal of Molecular Graphics* **14**(1): 51.
41. Lakowics, J. (1999). *Principles of Fluorescence Spectroscopy*. New York, Kluwer Academic / Plenum Publishers.
42. Lee, B. and F. M. Richards (1971). "The interpretation of protein structures: Estimation of static accessibility." *Journal of Molecular Biology* **55**(3): 379.
43. Leung, A. K. W., L. E. White, et al. (2004). "Structure of Mycobacterium tuberculosis FtsZ Reveals Unexpected, G Protein-like Conformational Switches." *Journal of Molecular Biology* **342**(3): 953.
44. Lovell, S. C., J. M. Word, et al. (1999). "Asparagine and glutamine rotamers: B-factor cutoff and correction of amide flips yield distinct clustering." *PNAS* **96**(2): 400-405.
45. Lowe, J. (1998). "Crystal structure determination of FtsZ from *Methanococcus jannaschii*." *J Struct Biol* **124**(2-3): 235-43.
46. Lowe, J. and L. A. Amos (1998). "Crystal structure of the bacterial cell-division protein FtsZ." *Nature* **391**(6663): 203-6.
47. Lowe, J. and L. A. Amos (1999). "Tubulin - like protofilaments in Ca²⁺ - induced FtsZ sheets." *Embo J* **18**(9): 2364 - 2371.
48. Lu, C., M. Reedy, et al. (2000). "Straight and Curved Conformations of FtsZ Are Regulated by GTP Hydrolysis." *J. Bacteriol.* **182**(1): 164-170.
49. Lu, C., J. Stricker, et al. (2001). "Site-specific mutations of FtsZ - effects on GTPase and in vitro assembly." *BMC Microbiology* **1**(1): 7.
50. Luque, I. and E. Freire (2000). "Structural stability of binding sites: consequences for binding affinity and allosteric effects." *Proteins Suppl* **4**: 63-71.
51. Luque, I. and E. Freire (2002). "Structural parameterization of the binding enthalpy of small ligands." *Proteins: Structure, Function, and Genetics* **49**(2): 181-190.
52. Lutkenhaus, J. (1990). "Regulation of cell division in *E. coli*." *Trends in Genetics* **6**: 22.
53. Lutkenhaus, J. and S. G. Addinall (1997). "Bacterial cell division and the Z ring." *Annu Rev Biochem* **66**: 93-116.
54. Mandelkow, E. M., E. Mandelkow, et al. (1991). "Microtubule dynamics and microtubule caps: a time-resolved cryo-electron microscopy study." *J Cell Biol* **114**(5): 977-91.
55. Margolin, W. (2000). "Organelle division: Self-assembling GTPase caught in the middle." *Curr Biol* **10**(9): R328-30.
56. Marrington, R., E. Small, et al. (2004). "FtsZ Fiber Bundling Is Triggered by a Conformational Change in Bound GTP." *Journal of Biological Chemistry* **279**(47): 48821.

57. Mine, S., T. Ueda, et al. (2000). "Analysis of the internal motion of free and ligand-bound human lysozyme by use of ^{15}N NMR relaxation measurement: a comparison with those of hen lysozyme." *Protein Sci* **9**(9): 1669-1684.
58. Mingorance, J., S. Rueda, et al. (2001). "Escherichia coli FtsZ polymers contain mostly GTP and have a high nucleotide turnover." *Mol Microbiol* **41**(1): 83-91.
59. Mizoue, L. S. and W. J. Chazin (2002). "Engineering and design of ligand-induced conformational change in proteins." *Current Opinion in Structural Biology* **12**(4): 459.
60. Montecinos, F. (2006). Under graduate student. Santiago.
61. Mosyak, L., Y. Zhang, et al. (2000). "The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography." *Embo J* **19**(13): 3179-91.
62. Mukherjee, A., C. Saez, et al. (2001). "Assembly of an FtsZ mutant deficient in GTPase activity has implications for FtsZ assembly and the role of the Z ring in cell division." *J Bacteriol* **183**(24): 7190-7.
63. Nogales, E., K. H. Downing, et al. (1998). "Tubulin and FtsZ form a distinct family of GTPases." *Nat Struct Biol* **5**(6): 451-8.
64. Nogales, E., M. Whittaker, et al. (1999). "High-resolution model of the microtubule." *Cell* **96**(1): 79-88.
65. Nogales, E., S. G. Wolf, et al. (1998). "Structure of the alpha beta tubulin dimer by electron crystallography." *Nature* **391**(6663): 199-203.
66. Ohmura, T., H. Motoshima, et al. (2002). "Fluctuations in free or substrate-complexed lysozyme and a mutant of it detected on x-ray crystallography and comparison with those detected on NMR." *J Biochem (Tokyo)* **131**(5): 701-4.
67. Oliva, M. A., S. C. Cordell, et al. (2004). "Structural insights into FtsZ protofilament formation." *Nat Struct Mol Biol* **11**(12): 1243-50.
68. Oliva, M. A., S. Huecas, et al. (2003). "Assembly of archaeal cell division protein FtsZ and a GTPase-inactive mutant into double-stranded filaments." *Journal of Biological Chemistry* **278**(35): 33562.
69. Orengo, C. A., A. D. Michie, et al. (1997). "CATH - a hierarchic classification of protein domain structures." *Structure* **5**(8): 1093.
70. Osawa, M. and H. P. Erickson (2005). "Probing the domain structure of FtsZ by random truncation and insertion of GFP." *Microbiology* **151**(12): 4033-4043.
71. Parak, F. G. (2003). "Proteins in action: the physics of structural fluctuations and conformational changes." *Current Opinion in Structural Biology* **13**(5): 552.
72. Parthasarathy, S. and M. R. N. Murthy (1997). "Analysis of temperature factor distribution in high-resolution protein structures." *Protein Sci* **6**(12): 2561-2567.
73. Perona, J. J. and C. S. Craik (1995). "Structural basis of substrate specificity in the serine proteases." *Protein Sci* **4**(3): 337-360.
74. Pichoff, S. and J. Lutkenhaus (2002). "Unique and overlapping roles for ZipA and FtsA in septal ring assembly in Escherichia coli." *Embo J* **21**(4): 685-93.
75. Ponstingl, H., T. Kabir, et al. (2003). "Automatic inference of protein quaternary structure from crystals." *Journal of Applied Crystallography* **36**(5): 1116-1122.
76. Press, W. H. e. a. "Numerical Recipes in C++."
77. Radivojac, P., Z. Obradovic, et al. (2004). "Protein flexibility and intrinsic disorder." *Protein Sci* **13**(1): 71-80.

78. RayChaudhuri, D. and J. T. Park (1992). "Escherichia coli cell-division gene *ftsZ* encodes a novel GTP-binding protein." *Nature* **359**(6392): 251-4.
79. Redick, S. D., J. Stricker, et al. (2005). "Mutants of FtsZ Targeting the Protofilament Interface: Effects on Cell Division and GTPase Activity." *J. Bacteriol.* **187**(8): 2727-2736.
80. Rivas, G., A. Lopez, et al. (2000). "Magnesium-induced Linear Self-association of the FtsZ Bacterial Cell Division Protein Monomer. THE PRIMARY STEPS FOR FtsZ ASSEMBLY." *J. Biol. Chem.* **275**(16): 11740-11749.
81. Rod, T. H. and C. L. Brooks, 3rd (2003). "How dihydrofolate reductase facilitates protonation of dihydrofolate." *J Am Chem Soc* **125**(29): 8718-9.
82. Rod, T. H., J. L. Radkiewicz, et al. (2003). "Correlated motion and the effect of distal mutations in dihydrofolate reductase." *Proc Natl Acad Sci U S A* **100**(12): 6980-5.
83. Sali, A. and T. L. Blundell (1993). "Comparative Protein Modelling by Satisfaction of Spatial Restraints." *Journal of Molecular Biology* **234**(3): 779.
84. Sanchez, R. (2003). Assistant Professor. New York.
85. Scheffers, D. J., J. G. de Wit, et al. (2001). "Substitution of a conserved aspartate allows cation-induced polymerization of FtsZ." *FEBS Lett* **494**(1-2): 34-7.
86. Schlieper, D., M. A. Oliva, et al. (2005). "Structure of bacterial tubulin BtubA/B: evidence for horizontal gene transfer." *Proc Natl Acad Sci U S A* **102**(26): 9170-5.
87. Schuyler, S. C. and D. Pellman (2001). "Microtubule "Plus-End-Tracking Proteins": The End Is Just the Beginning." *Cell* **105**(4): 421.
88. Sippl, M. J. (1993). "Recognition of errors in three-dimensional structures of proteins." *Proteins* **17**(4): 355-62.
89. Smith, S. (1994). "The animal fatty acid synthase: one gene, one polypeptide, seven enzymes." *FASEB J.* **8**(15): 1248-1259.
90. Sontag, C. A., J. T. Staley, et al. (2005). "In vitro assembly and GTP hydrolysis by bacterial tubulins BtubA and BtubB." *J. Cell Biol.* **169**(2): 233-238.
91. Spiegel, M. (1988). *Estadística*, McGraw-Hill.
92. Sprang, S. R. (1997). "G PROTEIN MECHANISMS: Insights from Structural Analysis." *Annual Review of Biochemistry* **66**(1): 639-678.
93. Todd, A. E., C. A. Orengo, et al. (2001). "Evolution of function in protein superfamilies, from a structural perspective." *Journal of Molecular Biology* **307**(4): 1113.
94. Trusca, D., S. Scott, et al. (1998). "Bacterial SOS checkpoint protein Sula inhibits polymerization of purified FtsZ cell division protein." *J Bacteriol* **180**(15): 3946-53.
95. Vihinen, M. (1987). "Relationship of protein flexibility to thermostability." *Protein Eng* **1**(6): 477-480.
96. Walker, M. L., S. A. Burgess, et al. (2000). "Two-headed binding of a processive myosin to F-actin." *Nature* **405**(6788): 804.
97. Wang, H. W. and E. Nogales (2005). "Nucleotide-dependent bending flexibility of tubulin regulates microtubule assembly." *Nature* **435**(7044): 911-5.
98. Wang, X., J. Huang, et al. (1997). "Analysis of the interaction of FtsZ with itself, GTP, and FtsA." *J Bacteriol* **179**(17): 5551-9.
99. Weiss, D. S. (2004). "Bacterial cell division and the septal ring." *Mol Microbiol* **54**(3): 588-97.
100. Westbrook, J., Z. Feng, et al. (2002). "The Protein Data Bank: unifying the archive." *Nucleic Acids Research* **30**(1): 245.

101. Xiao, T., J. Takagi, et al. (2004). "Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics." *Nature* **432**(7013): 59-67.
102. Yon, J. M., D. Perahia, et al. (1998). "Conformational dynamics and enzyme activity." *Biochimie* **80**(1): 33.
103. Yuan, Z. Z., Ju; Wang, Zhi-Xin (2003). "Flexibility analysis of enzyme active sites by crystallographic temperature factors." *Protein Engineering* **16**(2): 109-114.
104. Závodszky, P., J. Kardos, et al. (1998). "Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins." *Proceedings of the National Academy of Sciences of the United States of America* **95**(13): 7406.