



Caracterización bioquímica y estructural de la chaperona DnaK de *Mycoplasma genitalium*

Maria Adell Moruno

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA

DEPARTAMENTO DE BIOQUÍMICA I BIOLOGIA MOLECULAR

PROGRAMA DE DOCTORADO DE BIOTECNOLOGÍA

**Caracterización bioquímica y estructural de la chaperona DnaK de
*Mycoplasma genitalium***

Memòria presentada per Maria Adell Moruno per optar al títol de doctor per la
universitat de Barcelona

Director: Dr. Ignasi Fita Rodríguez/ Dra. Bárbara M. Calisto

Doctoranda: Maria Adell Moruno

Tutor: Miquel Coll Capella

MARIA ADELL MORUNO 2015

A Diego, a Carmen, a Miquel i a Paquita,

A Miquel, a Mari Carmen i a Èlia

Agradecimientos

Perquè tot inici té un final ha arribat el moment de tancar aquesta etapa de la meua vida. Aquesta etapa que m'ha permès créixer tan professional com personalment.

Primer de tot voldria agrair al meu director de tesi el Dr. Ignasi Fita per permetre'm fer la tesi en el seu laboratori i per la llibertat a l'hora de portar el projecte. A la meua co-directora de tesi Dra. Bárbara M. Calisto perquè tot i haver-se unit a l'últim tram, m'ha aconsellat, m'ha ajudat i tant pacientment ha corregit la tesi a distància.

Al grup de cristal·lografia de l'ESRF (MXC group) en especial a Daniele de Sanctis per permetre que realitzés una estancia a l'ESRF. A Bárbara per ser la meua mentora, per supervisar els experiments al "camp de batalla" i per la super excursió amb raquetes.

A Martha Brennich per ensenyar-me y respondre quesevol dubte sobre SAXS. A la Montse S per aconsellar-nos amb el projecte. I com no agrair als que han fet que la meua vida al ESRF fos més fàcil i divertida (segur que em deixo algú!!): a la Martha, al Cristiano, al Benoit, al Michael, al Phillippo, a la Gaele, a la petita Rachel, al Xuelei, a l'Sri i a la Béné que no va dubtar en compartir la taula, la poiata, l'ordinador, les tasses, pel "tea-time/coffee-time" i per Lyon!

I com no a la meua "chica PAC" (ara "chica EMBL"), perquè Grenoble sense ella no hagués estat el mateix. Segueix sent com ets, perquè és el que et fa especial.

Al Marc Jamin per ajudar-nos a analitzar les dades de MALS.

Al grup del Jaume Piñol i de l'Enrique Querol (Alícia, Lluís G i Lluís G) per haver-me donat la possibilitat de treballar amb la DnaK de *Mycoplasma genitalium* i, en

especial, al Luis Gonzalez per preparar-me tots els constructes que li he demanat sense queixar-se!. I al Lluís Garcia per la raclette-fondue a Grenoble!

També crec que es mereixen una especial menció la Jenny Colom i la Isabel Latorre per la seva professionalitat durant aquests anys i per fer-nos més fàcil la vida a la sala de purificació.

I com no a "los chicos PAC" a la Sònia, al Robert, al Joan, la Xandra i l'Àlicia per fer que tot funcioni.

Y no me quería olvidar de Alfonso por solucionar, incluso a distancia, mis problemas con los ordenadores.

A la família dels CRIS!! Als Cri6 per fer-nos sentir com a casa quan pugem a buscar reactius, consells....; a la Cristina F per la seva saviesa i perquè vals molt! al Pablo "el nuevo" per la seva tranquil·litat i bon humor, a la Laieta perquè cada cop que la veig està més gran i més guapa, al Jordi per ser una gran company, a la Cristiña pels horts urbans!. Als que no hi son, al Damià i a l'Arnau (Rubiooo) ¡¡¡Somos rubias pero no tontas!!!

A los Cri7 ! a l'Arka por sus chistes en inglés, al Pablo "el viejo" por no tener pelos en la lengua, a la Cuppi por sus visitas al lab y por las merendolas con chocolate desecho. Y como no a la Anna Rubioooo per la teva alegria i per pensar en els demás!!

Als Cri1, per ser els nostres veïns inseparables, a l'Estercita per mai tenir un no quan li demanem plaques, cèl·lules, reactius....i per ser una tècnic exemplar! a l'Albert per la seva disposició ajudar als demás, al pack Cristina M-Montse per ser unes

"currantes", a Simone per ser la fiesta, ànims en la recta final! I no oblidar-me dels que no hi són el Salvatore, la Zuzanna, el Roeland, el Radek, a Robert, el Dani, la Naybe, el Carlo, la Mailys i d'altres que m'oblido pel camí...

A l'Ori per lindy-hop i les BBQ! i a les seves noies: la Nere per preocupar-se per nosaltres, a la Irene per la alegria que ha traído consigo, tu puedes con todo pequeña!!!! i pels "Oriolets" (els nens de pràctiques), al Marc, l'Anna i la Carlita!

A l'Esther "secre" per ser una més entre nosaltres!

I als meus Cri2!!! La meva familia al laboratori que sense ells no hagués estat el mateix, al Xavi per aguantar la meva música durant molts anys i pel gran llançament de canya de xocolata. A l'Àlicia pels moments de teràpia. A la Babi per ser un exemple de treball i constància. A la Queri per les seves empanadilles per anar al sincro, per cuidar-nos i per les seves "AGUJETAS". A la Mercè pels seus pastissos, pel viatge al sincro i pels agents d'aduanes! Y a mis "post-docs", ¿Queremos ser IP? Al Deivid per la seva serenitat, per mai dir que no a una festa o a unes birres i pel meu primer viatge al sincro. Al Luqui per ser como eres, por las horas resolviendo-refinando estructuras, por el *one shot one kill* y por los viajes al sincro, las cervezas y las cenas! Y per últim a la meva tween, a la Rouse, per cuidar-me des del dia 1 quan vaig arribar al lab, perquè som un pack, per l'EDTA, per les confidencies (si la sala de 20 graus parlés, jajajaj), pels retreats, perquè "somos rubias y nos gusta", i perquè he tingut el privilegi d'aprendre de "the best crystal maker in the world, by M.C".

I als de fora del lab:

A la Raichel per la zumba, pel mojito, pel red velvet i per moltes coses més!!!!

A les bio-gallines: la Laura S, la Maria M, la Núria Co, l'Eva, la Lorena (si ho sé, et dec un cafè!), l'Andrea, la Teresa, la Sara J i la Sara A (per les converses al cotxe després d'una nit de festa, per canviar-me les sabates, i per brindar per les relacions llargues!!!) i a les que tot i que no són "bio" són tan gallines com nosaltres: Laura M, Núria Ca. Pels amics invisibles, sopars, festes, "cotilleos" i fins i tot pels emoticos al mòvil!

A l'Anna Ra perquè les trobades són breus però intenses!

A los desiguales (Rocio, David) por ser tan originales. A Kasia por encontrar un momento en sus viajes relámpago a Barcelona para verme!

A Rub y a Yas por "preocuparse" por las proteínas, los cristales, la estructura y esas cosas a las que me dedico. Por siempre estar dispuestos a organizar una cenita, un Montmeló, un cámping y muchas cosas más....

A la güebitoh (Zaida) pels sopars, castanyades, caps d'any, pels passats i pels que ens queden per passar! Pel Pablito i pel petit Bruno.

A la Laia per tots aquests anys, que no són pocs!!! per la costura, pel viatge a sud de França i pels que ens queden per fer, començant pel de Venècia!! I no m'oblido del Sergi!

A Dara, a Pepe y a Dolores por considerarme una más.

Aquesta tesi li dedico en especial a la meva família perquè sense el seu recolzament no hagués estat possible i perquè, al cap i a la fi, és el més important. Als meus tiets i

cosins. Als meus avis: al Miquel, a la Paquita, i a la Carmen per pensar sempre en nosaltres. Als meus pares en Miquel i la Mari Carmen per ser un exemple de constància i superació, per mimar-me i per donar-me suport en totes les meves decisions. A l'Èlia per tenir temps de cuidar-nos a tots, per saber que puc comptar amb ella i per seguir tractar-me com la germana petita que sóc. Al Víctor pel suport, per llegir-se la tesi!, i per les seves bromes àcides. I als meus dos amors: la petita Ivet per ser la bruji-princesa de la casa i al meu clon, el Dídac per ser una personeta tan especial.

I finalment, i sobretot, al Jona que m'aporta la calma i la tranquil·litat que no tinc, per suportar aquests mesos d'escriptura que no han sigut fàcils. Perquè m'acceptes tal com sóc i perquè junts formem un gran equip. Fins l'infinit i més enllà. Gràcies per tot.

Índice

Resumen.....	1
Capítulo 1 Introducción.....	4
1.1 Mycoplasma	4
1.1.1. Introducción general a los Mycoplasmas	4
1.1.2 <i>Mycoplasma genitalium</i>	8
• Motilidad.....	14
• División celular.....	17
• Adhesión celular	18
• Patogenicidad	19
1.2 Chaperonas	20
1.2.1 Perspectiva histórica	20
1.2.2 Generalidades	22
1.2.3 Organización citosólica en los diferentes Dominios de la Biología	24
• Componentes de los sistemas de chaperonas en procariotas	26
1.2.4 Sistema de chaperonas en <i>Mycoplasma</i>	35
• Generalidades	35
• Sistema de chaperonas en <i>Mycoplasma genitalium</i>	39
1.2.5 Sistema DnaK-DnaJ-GrpE en <i>Mycoplasma genitalium</i>	42
Objetivos.....	48
Capítulo 2 Resultados y discusión.....	49
2.1 Estructura cristalina de la DnaK de <i>M. genitalium</i> (<i>MgDnaK</i>)	49
2.1.2 Clonaje, expresión y purificación de <i>MgDnaK</i>	49
2.1.3 Clonaje, expresión y purificación de <i>MgDnaKΔCt</i>	53
2.1.4 Cristalización de <i>MgDnaKΔCt</i>	58
2.1.5 Determinación de la Estructura de rayos-X de <i>MgDnaKΔCt</i>	60
2.1.6 Descripción de la estructura cristalina de <i>MgDnaKΔCt</i>	64
• Dominio de unión a nucleótido (NBD) de <i>MgDnaKΔCt</i>	66

• Dominio de unión a sustrato (SBD) de <i>MgDnaKΔCt</i>	66
• Enlazador	69
• Comparación de las estructuras <i>MgDnaKΔCt</i> :AMP-PNP vs <i>MgDnaKΔCt</i> :ATP	71
• Comparación de las estructuras <i>MgDnaKΔCt</i> :ATP vs <i>GkDnaK</i> :ADP	72
• Comparación de las estructuras <i>MgDnaKΔCt</i> :AMP-PNP vs <i>Sse1</i> :ATP de <i>S. cerevisiae</i>	74
• Modelo alostérico de chaperonas Hsp70	76
2.1.6 Mutantes de la <i>MgDnaKΔCt</i>	80
2.2 Descripción de la estructura cristalina de <i>MgNBD</i>	85
2.2.1 Introducción.....	85
2.2.2 Clonaje, expresión y purificación de <i>MgNBD</i>	86
2.2.3 Cristalización del NBD	89
2.2.4 Determinación de la Estructura de rayos-X de NBD.....	91
2.2.5 Descripción de la estructura cristalina del <i>MgNBD</i>	94
• Comparación de las estructuras <i>MgNBD</i> :AMP-PNP, <i>MgDnaKΔCt</i> :AMP-PNP	101
• Comparación de las estructuras <i>MgNBD</i> y <i>MgDnaKΔCt</i> en presencia de ATP	103
2.3 Estudio bioquímico y biofísico de las chaperonas de <i>M. genitalium</i>	104
2.3.1 Cristalización de GrpE de <i>M. genitalium</i> (<i>MgGrpE</i>).....	104
2.3.2 Estudio de la interacción entre <i>MgDnaKΔCt</i> , <i>MgDnaJ</i> y <i>MgGrpE</i>	110
• Estudio de la Interacción entre <i>MgDnaKΔCt</i> y <i>MgDnaJ</i> por SPR.....	111
• Estudio de la Interacción entre <i>MgDnaKΔCt</i> y <i>MgGrpE</i>	118
• Estudio de <i>MgDnaK</i> -NBD, <i>MgDnaKΔCt</i> y <i>MgGrpE</i> por dispersión de rayos-X a bajo ángulo.....	135
Capítulo 3 Conclusiones.....	142
Capítulo 4 Materiales y métodos.....	146
4.1 Materiales	146
4.1.1 Lista de materiales y equipos.....	146
4.2 Cepas bacterianas y vectores de expresión.....	148
4.2.1 Cepas bacterianas usadas para la amplificación del DNA y clonaje	148
4.2.2 Cepas bacterianas usadas para la expresión recombinante de proteínas.	148
4.2.3 Vectores de expresión	148

4.3 Procedimientos de biología molecular.....	149
4.3.1 Métodos microbiológicos	149
• Composición del medio de cultivo.....	149
• Antibióticos y suplementos	149
• Condiciones generales de cultivo de <i>E. coli</i>	150
• Transformación de células de <i>E. coli</i> por choque térmico.....	150
4.3.2 Tecnología del DNA recombinante	151
• Clonación de la variante de DnaK, DnaK Δ Ct.....	151
• Mutagénesis dirigida.....	152
4.4 Métodos de producción y análisis de proteínas	153
4.4.1 Preparación de extractos celulares totales	153
4.4.2 Electroforesis en geles de poliacrilamida en condiciones desnaturalizantes (SDS-PAGE).....	153
4.4.3 Tinción de geles de poliacrilamida con azul de Coomassie (CBB)	154
4.4.4 Técnicas cromatográficas de purificación de proteínas	154
4.4.5 Experimentos de proteólisis limitada	156
4.5 Métodos biofísicos para la caracterización proteica.....	157
4.5.1 Dispersión dinámica de luz (Dynamic Light Scattering , DLS).....	157
4.5.2 DSF	158
4.5.3 Dispersión de luz en múltiples ángulos-cromatografía de gel filtración (MALS-SEC)	160
4.6 Difracción de rayos-X en cristales de proteína	161
4.6.1 Cristalización de proteínas.....	161
4.6.2 Crioprotección de cristales de proteínas	164
4.6.3 Difracción de rayos-X.....	164
• Recogida, procesado y escalado de datos	164
• Resolución de la estructura tridimensional	165
• Construcción y refinamiento de los modelos moleculares	167
• Validación de la estructura	168
4.5.4 Resonancia superficial por plasmón (SPR).....	169
4.7 Dispersión de rayos-X a bajo ángulo (SAXS)	172
Bibliografía.....	175

Abreviaciones

ADP: adenosina 5'-difosfato

DNA: ácido desoxirribonucleico

Ala: alanina

AMP-PNP: β,γ -imidoadesina 5'-trifosfato de sal de litio hidratada

Arg: Arginina

ARN: ácido ribonucleico

Asn: Asparagina

Asp: Ácido aspártico

ATP: adenosina 5'-trifosfato

CBB: Azul de Coomassie

CIRCE: repeticiones invertidas que controla la expresión de chaperonas

ClpB: caseína proteinasa lítica B

CO₂: Dióxido de carbono

DSF: Fluorimetría diferencial de barrido

FPLC: cromatografía líquida rápida de proteínas

Glu: Ácido glutámico

His: Histidina

HPD = His-Pro-Asp: Histidina-Prolina-Ácido aspártico

HrcA: receptor regulador del calor en circe

Hsp: *heat shock proteins* o proteínas de choque térmico

HSPs: small *heat shock proteins* o proteínas de choque térmico pequeñas

KSCN: tiocianato de potasio

IMAC: Cromatografía de afinidad por ion metálico inmovilizado

Ile: Isoleucina

Leu: Leucina

Loop: lazo

LS: *Light Scattering* o dispersión de luz

MALS: *Multi-Angle-Light-Scattering* o Dispersión de luz de múltiples ángulos

MR: reemplazamiento molecular

MgCl₂: Cloruro de magnesio

MW: peso molecular

NBD: Dominio de unión a nucleótidos

NAC: *nascent-chain-associated complex* o complejo asociado a la cadena naciente

SBD: Dominio de unión a sustrato

DLS: *Dynamic light scattering* o Dispersión de luz dinámica

G/F: Glicinas y Fenilalaninas

G+C: Guaninas+Citocinas

Gln: Ácido glutámico

IPTG: Isopropil-β-D-tiogalactosido

OT: Organela Terminal

PCR: reacción en cadena de la polimerasa

PDB: banco de datos de proteínas

PFD: prefoldina

PMF: *Peptide Mass Fingerprinting*

Pro: Prolina

EM: Microscopía Electrónica

RAC: Complejo Asociado a ribosoma

Rg: radio de giro

RI: Índice refracción

NMR: *Nuclear Magnetic Resonance* resonancia magnetica nuclear

rmsd: *root mean square deviation* o desviación de la raíz cuadrada

RRAP: Regiones Ricas en residuos Acídicos y Prolinas

RRAG: Región Rica en residuos Aromáticos y Glicinas

SAXS: *Small Angle X-Ray Scattering* o dispersión de rayos-X de bajo ángulo

SEC: *Size Exclusion Chromatography* o cromatografía de exclusión molecular

SPR: *Surface Plasmon resonance* o resonancia de superficie de plasmon

TCP-1: complejo del polipéptido1

TF: *trigger factor*

Thr: Treonina

Tm: *Melting Temperature* o temperatura de fusión

TRiC: complejo anillo

Resumen

Mycoplasma genitalium es un patógeno humano considerado uno de los organismos autorreplicativos más pequeños que existen. Su genoma fue completamente secuenciado en 1995, y actualmente, se ha convertido en un intenso objeto de estudio porque ha sido descrito como modelo ideal de célula mínima. *M. genitalium* presenta una morfología celular caracterizada por una extensión de la membrana en uno de los polos de la célula, conocido como Organela Terminal (OT). La OT está involucrada en procesos celulares tales como adhesión, división celular y motilidad con implicaciones directas en patogenicidad y virulencia. A pesar de la importancia de la OT, hay muy pocas proteínas de esta organela que hayan sido caracterizadas estructuralmente, siendo, por tanto, reducido el actual conocimiento que se tiene de su estructura y función. Estudios de Microscopia Electrónica (del inglés *Electron Microscopy, EM*) muestran que la chaperona molecular DnaK se localiza principalmente en la OT. Además, estudios *in vivo* sugieren que existe una estrecha relación entre la proteína MG200, que contiene un dominio J y es la posible co-chaperona de DnaK, con las adhesinas mayoritarias de las células (P140 y P110), que son proteínas que intervienen en procesos de adhesión con superficies sólidas o con las células huésped y que son críticas para la motilidad. Por ello, el principal objetivo de esta Tesis doctoral fue caracterizar el sistema de chaperonas de *M. genitalium* e investigar su posible rol en la formación y funcionamiento de la OT. La chaperona molecular DnaK forma parte de uno de los sistemas procariotas de chaperonas (DnaK-DnaJ-GrpE) más estudiados pero su mecanismo alostérico aún no ha sido completamente esclarecido. Definida como

una proteína de choque térmico (Hsp70), la principal función de DnaK consiste en interactuar con polipéptidos que requieren procesos de plegamiento a través de ciclos de hidrólisis de ATP.

El primer paso para entender la posible función de DnaK, fue resolver la estructura cristalina de un constructo al que le faltan los últimos 74 residuos del C-terminal, por difracción de rayos-X a una resolución de unos 2 Å. Este constructo contiene un dominio de unión a nucleótido y un dominio de unión a sustrato (del inglés *Nucleotide Binding Domain* y *Substrate Binding Domain*, *NBD* y *SBD* respectivamente) unidos entre sí por un enlazador muy flexible. En la estructura de DnaK el NBD se encuentra unido a un cofactor (ADP o AMP-PNP) y el SBD a un péptido que mimetiza un sustrato sin plegar, esta estructura corresponde a una conformación cerrada. Para estudiar desde un punto de vista estructural el mecanismo alostérico de DnaK, se determinó la estructura tridimensional del NBD en su forma apo, en presencia de AMP-PNP, ATP o ADP. La comparación de estas estructuras indicó que los cambios estructurales consecuencia de la hidrólisis de nucleótido son sutiles. Sin embargo, se apreciaron diferencias importantes en la coordinación del ión magnesio, que es imprescindible en el mecanismo de hidrólisis del nucleótido.

El segundo paso fue estudiar las interacciones de DnaK con sus posibles co-chaperonas (MG200 y MG019) y el factor intercambiador de nucleótido GrpE, tanto por la técnica de Resonancia Superficial por Plasmón (del inglés *Surface Plasmon Resonance*, *SPR*) como por cromatografía de gel filtración. No se detectó ninguna interacción entre DnaK y sus posibles co-chaperonas, MG019 y MG200, debido a que las condiciones experimentales en las que se realizó el ensayo no fueron las idóneas probablemente

porque no se lograron reproducir las condiciones *in vivo*. Se detectó una fuerte interacción entre DnaK y GrpE. La estequiometría de este complejo fue determinada por dispersión de luz en múltiples ángulos (del inglés *MultiAngle Light Scattering, MALS*) y fue confirmada por Espectrometría de Masas Nativa (del inglés *Native Mass Spectrometry, Native MS*) y corresponde a una molécula de DnaK para dos de GrpE (1:2). La dificultades en obtener cristales de este complejo bien como de las múltiples versiones que se diseñaron tanto de DnaK como de GrpE, llevó al estudio de este complejo en solución por la técnica de difracción de rayos-X a bajo ángulo (del inglés *Small Angle X-ray Scattering, SAXS*) acoplada a una columna de gel filtración. Esta técnica permitió obtener modelos a baja resolución de los componentes del complejo *MgNBD* y *MgDnaKΔCt* y de *MgGrpE*. La estructura en solución de *MgNBD* es un dominio rígido y globular idéntico a su forma cristalina. *MgDnaKΔCt* presenta una conformación cerrada en solución que equivale a la conformación encontrada en el cristal, aunque existen variaciones en la orientación espacial entre los subdominios NBD y SBD. Este resultado refuerza la posible existencia de un paso intermedio extra en el modelo alostérico de DnaK propuesto hasta el momento. El modelo a baja resolución de *MgGrpE* fue validado por un modelo teórico tri-dimensional de *MgGrpE* puesto que las estructuras de otras GrpEs resueltas hasta el momento no se ajustaban al modelo en solución. Este modelo teórico de GrpE muestra una elevada flexibilidad de los dominios α-hélice de las dos subunidades que componen el dímero siendo el N-terminal de la proteína desestructurado pero con tendencia a formar estructuras helicoidales lo que parece confirmar la posibilidad de que el N-terminal desestructurado interactúe con el subdominio SBD de *MgDnaKΔCt* mientras se efectúa el intercambio de nucleótidos en el subdominio NBD *MgDnaKΔCt*.

1. Introducción

1.1 Mycoplasma

1.1.1. Introducción general a los Mycoplasmas

Los Mycoplasmas son los organismos autorreplicativos más pequeños que existen, midiendo entre 0.2 a 0.5 μm . Son procariotas de la clase de los *Mollicutes* que están relacionados filogenéticamente con bacterias Gram positivas y se cree que son el resultado de una evolución regresiva de parte de su genoma (Fraser, Gocayne et al. 1995). Los Mycoplasmas son además parásitos obligados de humanos, mamíferos, reptiles, peces, artrópodos y plantas y se caracterizan por su pequeño genoma (de 564 a más de 2000 kb), por el bajo contenido del DNA genómico en guaninas (G) y citosinas (C) y por la ausencia de pared celular. Poseen un código genético alternativo donde el codón UGA codifica para triptófano en vez de para el codón STOP, como ocurre en el código genético universal (Garnier, Foissac et al. 2001).

Las células bacterianas, en general, están rodeadas por una variedad de estructuras en la superficie celular que les permite sobrevivir en ambientes extremos. Los componentes de la envuelta celular y de la matriz extracelular son responsables de proporcionar a las células el soporte estructural y mediar la comunicación intercelular,

permitiendo la adhesión y el movimiento de las superficies celulares que protegen las células del ataque del sistema inmune o de antibióticos, y facilitan la obtención de nutrientes. Uno de los componentes más importantes de la pared celular bacteriana son los polisacáridos (Ginsberg, Walker 2008). Los Mollicutes, debido a la ausencia de pared celular, son sensibles a la lisis por choques osmóticos o por la presencia de detergentes o alcoholes. Por ese motivo, las células de Mycoplasma están recubiertas por una membrana compuesta por lipoproteínas, glicolípidos y lipoglucanos y en la que se encuentran proteínas ancladas como por ejemplo, proteínas responsables de la variación antigénica, lipoproteínas capaces de activar las citoquinas de las células del huésped, y colesterol (Razin, Argaman et al. 1963, Razin 1996). El reducido genoma de los Mycoplasmas implica que sus capacidades biosintéticas sean limitadas. Durante su evolución han perdido la capacidad de sintetizar aminoácidos que deben ser proporcionados bien por el huésped o bien por el medio de cultivo artificial donde crecen. Los cofactores (como las vitaminas necesarias para su supervivencia) y ácidos nucleicos (bases púricas y pirimidínicas) las obtienen *in vivo* por la acción de sus potentes nucleasas que degradan moléculas de DNA y de RNA de las células huésped. Los genes relacionados con la biosíntesis de ácidos grasos también están ausentes y aunque son capaces de sintetizar los fosfolípidos y glicolípidos de la membrana esto exige un elevado gasto energético para las células así que una gran parte de ellos los obtienen de fuentes exógenas. La regulación de la fluidez de la membrana celular es suplida por la incorporación de grandes cantidades de colesterol (Razin, Yogev et al. 1998).

Los Mycoplasmas se han convertido en el objeto de estudio de una parte de la comunidad científica porque han sido considerados como modelo natural de célula

mínima (Koonin 2000), a pesar de: i) presentar deficiencias metabólicas y, por lo tanto, tener complejos requerimientos nutricionales, ii) tener un bajo rendimiento de crecimiento y iii) no presentar marcadores seleccionables para su estudio mediante las herramientas genéticas más comunes. El modelo de célula mínima trata de distinguir los genes que son esenciales para la viabilidad celular de los que no lo son, en condiciones ideales (Koonin 2000, Glass, Assad-Garcia et al. 2006). Este se ha estudiado a partir de dos aproximaciones distintas: i) mediante la identificación de los genes no esenciales para el crecimiento de las células *in vitro* (Hutchison, Peterson et al. 1999, Glass, Assad-Garcia et al. 2006) y ii) mediante la síntesis química, de manera artificial, en el que se consiguió sintetizar, clonar, ensamblar y trasplantar 1.08Mpb del genoma de *M. mycoides* en células de *M. capricolum* de éstos genomas reducidos (Gibson, Benders et al. 2008, Gibson, Glass et al. 2010). En el primer caso, a través mutagénesis por transposición, se definieron entre 265-360 genes esenciales (Hutchison, Peterson et al. 1999), aunque análisis genómicos detallados demostraron que en realidad eran 387 los genes esenciales (Glass, Assad-Garcia et al. 2006). Estudios posteriores realizados en el laboratorio de los doctores Jaume Piñol y Enrique Querol (IBB-UAB (Barcelona, España) sugirieron la posibilidad de que esta lista fuera incompleta. Estos estudios, obtenidos utilizando la misma técnica, indicaron que algunos genes que habían sido considerados como esenciales no lo eran (Pich, Burgos et al. 2006, Pich, Burgos et al. 2006, Lluch-Senar, Vallmitjana et al. 2007, Pich, Burgos et al. 2008, Lluch-Senar, Querol et al. 2010). En el segundo caso, en el laboratorio del doctor Craig Venter, (Craig Venter Institute, La Jolla, California, USA) fueron capaces de construir un modelo de célula mínima, desarrollando una estrategia mediante la cual ensamblaron largas moléculas de DNA dando lugar al genoma sintético de *M.*

genitalium (Gibson, Benders et al. 2008). A los dos años consiguieron sintetizar, ensamblar, clonar y trasplantar 1.08 Mpb del genoma de *M. mycoides* en células de *M. capricolum* (Lartigue, Vashee et al. 2009, Gibson, Glass et al. 2010).

La biología de sistemas tiene como objetivo la representación de redes celulares a través de modelos matemáticos. Para poder representarlos se necesita que todas las dianas y sus reacciones estén identificadas, que las reacciones sean cuantificadas y que todas las relaciones que se establezcan puedan ser analizadas en el contexto fisiológico del organismo. Para ello se escogió a *M. pneumoniae* como modelo. Como sucede en otros Mollicutes, *M. pneumoniae* ha sufrido una reducción en su genoma que incluye 689 genes que codifican para proteínas, de las cuales, 231 no tienen función conocida, además puede ser cultivado *in vitro* de manera independiente y autónoma. La red metabólica de este microorganismo ha sido reconstruida y validada manualmente, y se ha estudiado su regulación complementando este estudio con el análisis del transcriptoma (Guell, van Noort et al. 2009) y con la descripción del proteoma (Kuhner, van Noort et al. 2009). Estos estudios han demostrado que a pesar de la simplicidad de *M. pneumoniae*, este muestra unas respuestas metabólicas y de adaptación similares a bacterias más complejas, dejando entrever posibles mecanismos de regulación que aún no se han de determinado (Yus, Maier et al. 2009).

1.1.2 *Mycoplasma genitalium*

El genoma de *M. genitalium* fue completamente secuenciado en 1995 (Fraser, Gocayne et al. 1995). Con un genoma de 580 kb y un contenido de G+C de ~32% este microorganismo contiene 482 genes que codifican para proteínas y de estos 387 parecen ser esenciales (Hutchison, Peterson et al. 1999, Glass, Assad-Garcia et al. 2006). *M. genitalium* presenta una morfología celular en la que esencialmente se distinguen cuatro partes: el botón y el cuello (que forman la Organela Terminal (OT), el cuerpo celular y la cola (Figura 1.1).

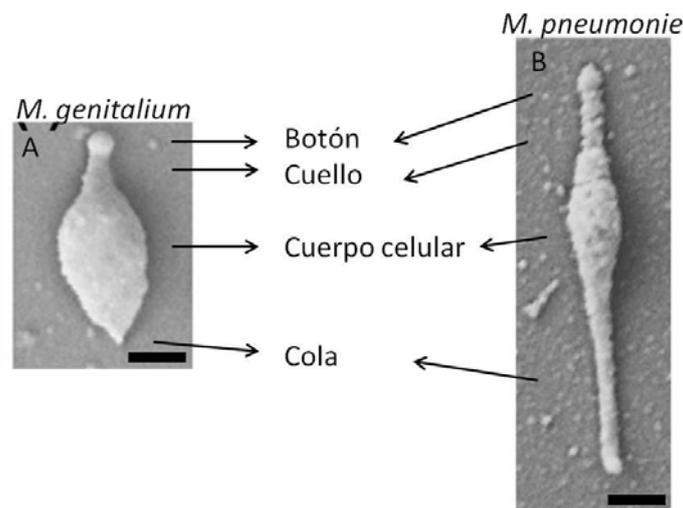


Figura 1.1 Microscopía electrónica de barrido de células de Mycoplasma crecidas en cubreobjetos de cristal. A) Imagen de *M. genitalium*. B) Imagen de *M. pneumoniae*. Escala 250nm. Imagen adaptada de (Hatchel and Balish 2008).

El llamado clúster de *M. pneumoniae* está formado por ocho especies agrupadas según la similitud en sus secuencias de ácido ribonucleico ribosómico (RNAr) 16s (Johansson

and Petersson 2002) y de la presencia de una organela terminal (Balish, Hahn et al. 2001, Krause and Balish 2001). Forman parte de este clúster *M. pneumoniae*, *M. genitalium*, *M. gallisepticum*, *M. imitans*, *M. amphoriforme*, *M. testudinis*, *M. pirum* y *M. alvi*, este último es el único que no presenta OT (Hatchel and Balish 2008).

La OT es una extensión de la membrana celular debida a la presencia de un complejo citoesqueleto que otorga a las células una forma asimétrica. La OT, además de ser un factor de virulencia, permite que las células se adhieran a la célula huésped y presenta funciones relacionadas con la división celular, el reconocimiento del huésped y la movilidad celular (Krause and Balish 2001).

Los estudios para determinar la estructura del citoesqueleto que compone la OT y de sus elementos asociados se realizaron hasta el momento mayoritariamente en *M. pneumoniae*. Estos consistieron en tratar las células con el detergente Triton X-100, lo que permitió recuperar el citoesqueleto intacto de las OT de las células, al cual se le denominó fracción insoluble en Triton X-100 (Regula, Boguth et al. 2001). En esta fracción se identificaron más de 100 proteínas, de las cuales las más abundantes son la DnaK (proteína de choque térmico), el factor de elongación *Tu*, la FtsH (proteína relacionada con la división celular) y la enzima tiorredoxina. De éstas proteínas hay que destacar que DnaK es una proteína estructuralmente homóloga a la actina (Flaherty, McKay et al. 1991) que fue asociada con la adhesina P1 de *M. pneumoniae* (P140 en *M. genitalium*) por "crosslinking" (Seto, Kenri et al. 2005).

Los estudios de EM en Mycoplasmas son particularmente difíciles de realizar porque la ausencia de pared celular hace que el cuerpo celular se deforme muy fácilmente y que la membrana se dañe durante la preparación de las muestras. Pero, estudios en células

de Mycoplasma realizados durante la última década mediante crio-EM han contribuido a la determinación de las ultra-estructuras que componen la OT (Henderson and Jensen 2006, Seybert, Herrmann et al. 2006). Estos estudios han revelado que la OT está formada por tres partes diferenciadas (Figura 1.2):

- El botón terminal (*terminal button*), localizado en la parte más distal del cuerpo celular, es más amplio que el núcleo electrodense (Figura 1.2, 1);
- El núcleo electrodense (*electrodense core*), de unos 300 nm, localizado longitudinalmente en el centro de la OT está compuesto por dos segmentos paralelos de diferentes largada y amplitud (Figura 1.2, 2);
- El complejo rueda (*wheel complex*) está localizado en la parte proximal del cuerpo celular (Figura 1.2, 3) y une el cuerpo celular a la OT y a la membrana celular.

Los estudios de EM también han revelado la presencia de proteínas en la superficie de la membrana localizadas mayoritariamente en la región de la OT que tienen forma de esferas de unos 5 nm de diámetro (denominadas *nap*), y que son responsables por la adhesión celular. Los *nap* son complejos heterodiméricos de las adhesinas MG191 y MG192 (Figura1.2, 4).

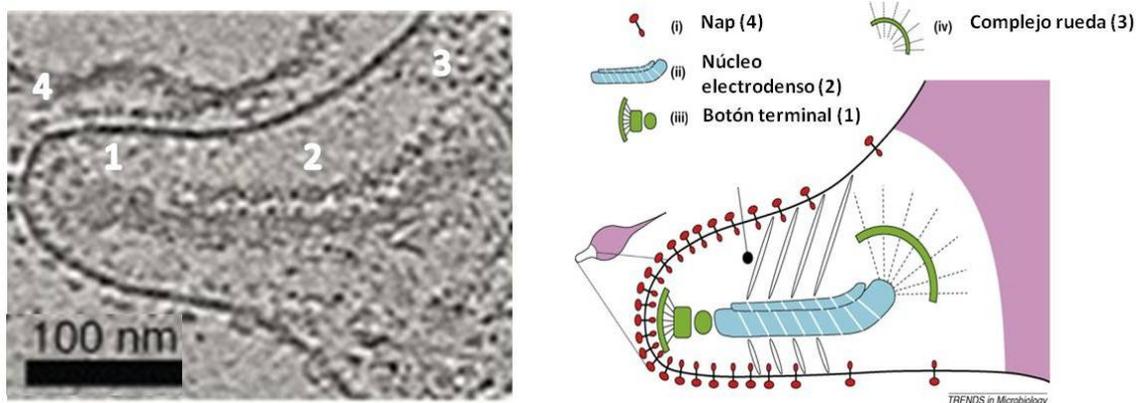


Figura 1.2 A) Micrografía de la organela terminal de *Mycoplasma pneumoniae* donde se indican las ultra-estructuras que componen la OT: 1- Botón terminal, 2- Núcleo electrodenso, 3- complejo rueda Imagen adaptada de (Henderson and Jensen 2006). B) Esquema de la OT donde se muestran las ultraestructuras que la forman, elaborado a partir de los datos de EM. Imagen adaptada de Miyata 2008.

Las proteínas relacionadas con citoadherencia que se identificaron en *M. pneumoniae* después de tratar las células con Triton X-100 se muestran en la Tabla 1.1.

Tabla 1.1 Proteínas de la organela terminal implicadas en motilidad y citoadherencia.

<i>M. genitalium</i>		<i>M.pneumonie</i>		Identidad de secuencia (%)
Gen	Proteína	Gen	Proteína	
MG191 (MgpA)	P140	MPN141	P1	44,9
MG192 (Mgp3)	P110	MPN142	P40/P90	48,9
	DNAJ-like protein (MG200)			
MG200		MPN119	TopJ	34,8
MG217	P65 homologo	MPN309(p65)	P65	10,2
MG218 (hmw2)	HMW2	MPN310(hmw2)	HMW2	57,0
MG491	MG491	MPN311	P41	53,2
MG219	MG219	MPN312	P24	17,0
MG312 (hmw1)	HMW1	MPN447(hmw1)	HMW1	32,5
MG317 (hmw3)	HMW3	MPN452(hmw3)	HMW3	33,4
MG318	P32	MPN453(p30)	P30	42,8
MG386	P200	MPN567	P200	29,2

La adhesina P1 de *M. pneumoniae* (P140 en *M. genitalium* es la P140, Tabla 1.1) tiene un peso molecular aproximado de 170 kDa y posee dos segmentos transmembrana. Esta proteína está asociada a las adhesinas P40 y P90 (P110 en *M. genitalium*, Tabla 1.1) teniendo el complejo una función relacionada con la motilidad y la adhesión de las células a superficies sólidas (Seto, Kenri et al. 2005, Burgos, Pich et al. 2006).

Las proteínas P65, HMW3 (proteína de alto peso molecular 3, del inglés *High Molecular Weight 3*, HMW3) y P30 de *M. pneumoniae* (los homólogos en *M. genitalium* son P65, HMW3 y P32, respectivamente) componen el botón terminal de la OT (Krause 1996, Seto and Miyata 2003, Miyata 2008). La proteína P65 de *M. genitalium* es intracelular (Burgos, Pich et al. 2007) a diferencia de su proteína homóloga en *M. pneumoniae* donde parte de ella está expuesta a la superficie (Proft, Hilbert et al. 1995, Jordan, Berry et al. 2001, Hasselbring, Sheppard et al. 2012). La función de la P65 de *M. genitalium* parece estar relacionada con la curvatura del botón terminal (Burgos, Pich et al. 2008) mientras que cuando se interrumpe el gen de la P65 de *M. pneumoniae* se observa una disminución de la movilidad celular y de la adherencia (Hasselbring, Sheppard et al. 2012). La proteína P30 de *M. pneumoniae* (P32 en *M. genitalium*) está relacionada con adherencia, movilidad y virulencia (Relich and Balish 2011). Finalmente, las HMW3s de ambos organismos se caracterizan por la presencia de Regiones Ricas en residuos Acídicos y Prolinas (RRAP) que también están presentes en las proteínas HMW1, P65 y P200 de *M. pneumoniae* así como en las proteínas homologas de *M. genitalium*.

Otra de las proteínas identificadas en la OT es la HMW1 que posee además de una RRAP, una Región Rica en residuos Aromáticos y Glicinas (RRAG). Estas regiones

también se han encontrado en las proteínas MG386 y MG200 de *M. genitalium* bien como en sus ortólogos en *M. pneumoniae* y *M. gallisepticum* (Balish, Hahn et al. 2001, Calisto, Broto et al. 2012). Esta proteína parece tener una función relacionada con la movilidad en *M. pneumoniae* (Burgos, Pich et al. 2007). En *M. genitalium* mutantes por transposición de MG386 o MG200 mostraron un fenotipo menos móvil y con un mayor porcentaje de células no móviles (Burgos, Pich et al. 2007).

La HMW2 es el principal componente del mayor elemento estructural de la OT, el núcleo electrodenso. Esta proteína se caracteriza por la presencia de repeticiones de siete aminoácidos, aminoácidos hidrofóbicos en las posiciones 1 y 4 e hidrofílicos en las restantes, que podrían formar las estructuras filamentosas del núcleo electrodenso (Pich, Burgos et al. 2008).

Las proteínas P41, P24, P200 y TopJ se localizan en el complejo rueda de la OT y tienen la función de conectar la OT con el cuerpo celular. El uso de proteínas fluorescentes fusionadas a las proteínas del complejo rueda permitió localizar P24 y P41 (Kenri, Seto et al. 2004) y así esclarecer sus roles en la motilidad (Hasselbring and Krause 2007). P41 es esencial para que P24 pueda realizar su función de posicionar correctamente la organela terminal naciente y en el movimiento celular. P24 es determinante para la mantener la frecuencia de formación de nuevas organelas terminales y P24 se ve además afectada por la ausencia de TopJ la delección de la cual presenta fenotipos no móviles y no adherentes (Cloward and Krause 2009). En *M. genitalium* las proteínas MG386 y MG200 están implicadas en la movilidad celular ya que mutantes en los que se ha delecionado ambas proteínas presentan fenotipos con una menor velocidad de

movimiento y el porcentaje de células móviles también es menor (Pich, Burgos et al. 2006).

- **Motilidad**

La OT juega un papel clave en la motilidad de *M. genitalium* lo que es importante para el proceso de infección ya que permitirá que las células se diseminen y proliferen. Los Mycoplasmas carecen de flagelos o pilis, elementos de motilidad comunes en otras bacterias, poseyendo a su vez un mecanismo único para moverse conocido como movimiento por deslizamiento (*gliding*) que les permite desplazarse sobre superficies sólidas (McBride 2001). Se han descrito dos mecanismos distintos de motilidad: el modelo del "cien pies" y el modelo de "oruga" que tienen como representantes *M. mobile* y *M. pneumoniae*, respectivamente (Fraser, Gocayne et al. 1995, Himmelreich, Hilbert et al. 1996, Chambaud, Heilig et al. 2001, Papazisi, Gorton et al. 2003, Jaffe, Miyata et al. 2004).

- El modelo de "cien pies" o de deslizamiento rápido está presente en los Mycoplasmas del clúster *mobile* que se mueve con una velocidad de entre 2 a 4.5 $\mu\text{m/s}$. El estudio de células de *M. mobile* por EM reveló una estructura en la cabeza celular denominada como estructura medusa que tiene forma de campana ovalada y está repleta de un enrejado hexagonal de 12 nm. Cada enrejado tiene decenas de tentáculos que están plegados de partículas 20 nm de diámetro que se extienden por todo el cuerpo celular (Figura 1.3 A). En la cabeza de las células de *M. mobile* se han identificado cuatro componentes: i) Gli123, que es el nexo de unión de todas las proteínas implicadas en el movimiento (Uenoyama and Miyata 2005), (ii) Gli349, que es la pierna del

mecanismo (Uenoyama, Kusumoto et al. 2004), iii) Gli521, que es el engranaje del mecanismo y iv) P42, que es una nucleótido trifosfatasa que es el motor presentando actividad ATPasa (Figura 1.4, B). El ciclo de movilidad está compuesto por seis estados (Figura 1.3 B (a-f)) conectados a 6 pasos (Figura 1.3 B (i-vi)). En el primer paso (a) la proteína Gli349 o pierna, se une estrechamente a la superficie y espera a la nueva molécula de ATP. La fuerza aplicada por la pierna desde la parte frontal desencadena un cambio conformacional en la parte superior de la pierna causando un golpe (i) que depende de ATP. Esta tensión aplicada en la pierna (b) da como resultado el movimiento (ii). A continuación la tensión aplicada por la pierna disminuye (c) y la pierna se liberará de la superficie (iii) devolviéndola a su conformación inicial (e). Esta conformación permitirá que se vuelva a unir a la superficie (f). Los pasos (iii) y (iv) ocurren cuando hay una acción sensible a la tensión.

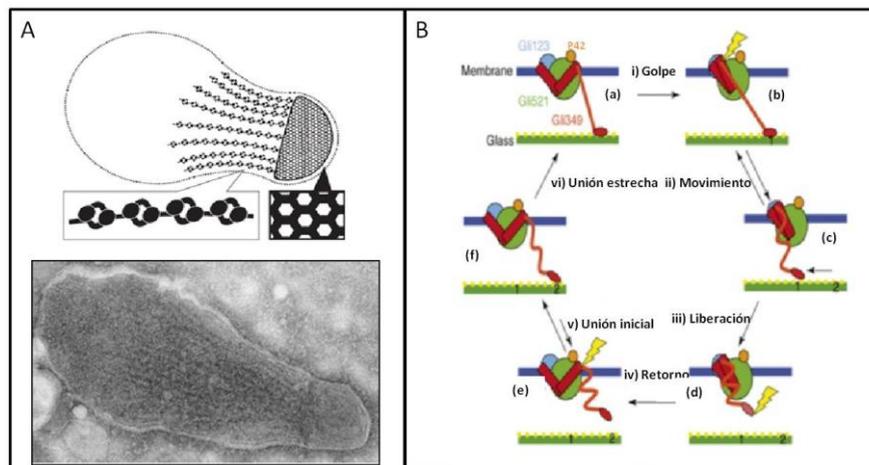


Figura 1.3 A) Esquema del citoesqueleto de *M. mobile*. Micrografía de microscopía electrónica de transmisión del citoesqueleto de *M. mobile* por tinción negativa. Figura adaptada de Nakane, 2007.(Nakane and Miyata 2007). B) Modelo del movimiento del "cien pies" propuesto por Miyata et al. (Miyata 2008).

El modelo de “oruga” o de deslizamiento lento está presente en *M. pneumoniae* que se mueve a una velocidad de entre 0.3 a 0.5 $\mu\text{m/s}$. En este modelo los Mycoplasmas usan su OT para engancharse y moverse invariablemente en una dirección sobre superficies sólidas. El modelo también propone que el botón terminal y el complejo rueda están unidos a la membrana celular y que es el núcleo electrodenso, que es altamente flexible, que cambia su longitud y curvatura causando la disociación, desplazamiento y asociación de las moléculas de la superficie de la membrana con la superficie sólida propulsando el cuerpo celular (Henderson and Jensen 2006, Seybert, Herrmann et al. 2006). Los cambios conformacionales que se prevén que realicen las proteínas filamentosas de la OT podrían ser las responsables por la propulsión de las células y por la división celular (Figura 1.4) (Wolgemuth, Igoshin et al. 2003).

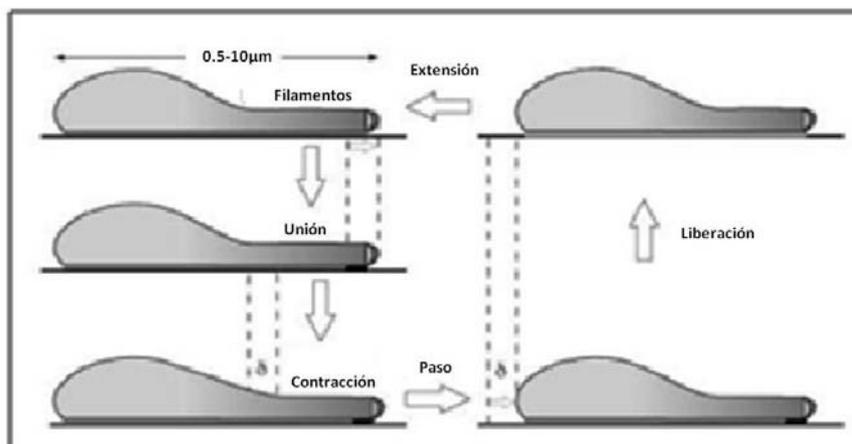


Figura 1.4 Esquema del modelo de oruga en *M. pneumoniae* propuesto por Wolgemuth et al., 2003. Figura adaptada de (Wolgemuth, Igoshin et al. 2003) .

En un momento en que aún no se ha identificado el motor molecular que promueve el movimiento celular en *M. genitalium* y basándonos en el hecho que otros organismos

utilizan chaperonas como motores para el desplazarse se creyó relevante estudiar el sistema de chaperonas de *M. genitalium*, en particular, DnaK, como posible candidato a ser el motor molecular del mecanismo de motilidad. Experimentos realizados por A. Broto (IBB-UAB, Barcelona, España) (en preparación) muestran que modificaciones en una posible co-chaperona de DnaK (MG200) resultan en una falta de expresión de adhesinas y consecuentemente, las células presentan un fenotipo no mótil, lo que implica una relación indirecta de la DnaK en la motilidad.

- **División celular**

Estudios *in vivo* de la división celular en *M. pneumoniae* mediante proteínas de la OT fusionadas a proteínas fluorescentes demostraron que la duplicación y separación de OTs ocurre inmediatamente antes de la citoquinesis, es decir, antes de que se dé la división física de las dos células (Seto, Layh-Schmitt et al. 2001). En 2006 Hasselbring y colaboradores observaron la presencia de más de dos organelas terminales antes que la célula se dividiera en dos, lo que indicó que la duplicación de la OT y la división celular estaban estrechamente relacionadas bajo ciertas condiciones de cultivo (Hasselbring, Jordan et al. 2006, Miyata 2008)(Figura 1.5).

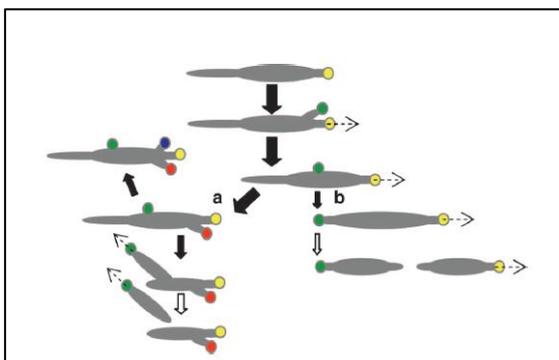


Figura 1.5 Esquema de la duplicación de la OT y ciclo de crecimiento de *Mycoplasma pneumoniae*. El círculo amarillo representa la OT inicial, el círculo verde, rojo y azul representan las OT nacientes. Las flechas discontinuas representan el movimiento de las OT, las blancas indican la citoquinesis mientras que las negras indican los pasos en el ciclo celular y la medida indica la frecuencia relativa. en muchos casos, hay múltiples duplicaciones de las OT antes de que la célula hija emerja (a), aunque en algunas células una única duplicación de la OT es seguida por la citoquinesis (Hasselbring, Jordan et al. 2006) de acuerdo con el modelo previamente descrito de la división celular de *M. pneumoniae*.

El mecanismo por fisión binaria parecía el más probable en la división de Mycoplasmas (Miyata and Seto 1999) pero se demostró que la proteína FtsZ, esencial para la división celular de muchas bacterias, no era esencial en *M. genitalium* ya que se podía realizar la citoquinesis utilizando la fuerza generada por las OTs localizadas en cada polo de la célula sin necesitar la proteína FtsZ, de manera que existe un mecanismo alternativo de división celular, en un microorganismo, independiente de FtsZ (Lluch-Senar, Querol et al. 2010). Estos datos indican el posible papel que tiene la OT de los Mycoplasmas en la división celular y en la segregación de las células hijas a partir de la célula madre. Además, está de acuerdo con la idea de que en los Mycoplasmas la duplicación del DNA se lleva a cabo antes de la división celular y casi al mismo tiempo que la duplicación de otras biomoléculas. Sin embargo, la segregación cromosómica en Mycoplasmas aún no se entiende completamente y por lo tanto es difícil especular sobre la función de la OT en la división celular.

- Adhesión celular

Los Mycoplasmas tienen que adherirse a las células del huésped para la colonización y la infección. La ausencia de pared celular y de los componentes a ella asociados hacen que el mecanismo de adhesión sea diferente al de otras bacterias. En este caso, son los componentes asociados a la membrana los que juegan un papel fundamental en la adhesión, y esencialmente son las adhesinas las encargadas de realizar esta función. Además se ha demostrado que la pérdida de adherencia de las células de Mycoplasma resulta en la ausencia de infectividad (Razin, Yogev et al. 1998).

Las adhesinas mejor estudiadas son las P1, P30 y P40/P90 de *M. pneumoniae* (homólogos de las proteínas P140, P32 y P110 en *M. genitalium*, respectivamente) las

cuales tienen una fuerte respuesta inmunológica. Éstas se unen a receptores de la membrana tales como sialo-oligosacáridos o glicolípidos sulfatados (Razin and Jacobs 1992) y además se concentran en la región de la OT estando relacionadas con su citoesqueleto y facilitando el movimiento celular (Rottem 2003).

- **Patogenicidad**

M. genitalium es un patógeno humano emergente de transmisión sexual (Waites, Katz et al. 2005) que se aisló por primera vez en 1981 en dos pacientes masculinos que mostraban síntomas de uretritis no gonocócica (Taylor-Robinson, Tully et al. 1981), y desde ese momento se consideró como un patógeno relacionado directamente con enfermedades de transmisión sexual. En hombres, la infección por *M. genitalium*, se ha relacionado con otras enfermedades como la prostatitis crónica o la epididimitis aguda, mientras que en mujeres se ha relacionado con uretritis, cervicitis, vaginitis bacteriana, enfermedad inflamatoria pélvica, y con algunos casos de infertilidad (Taylor-Robinson and Jensen 2011). Estas infecciones se tratan con un amplio espectro de antibióticos como la tetraciclina, los macrólidos o las quinolonas (Jensen 2004) ya que la ausencia de pared celular no permite el uso de antibióticos como la penicilina o los antibióticos betalactámicos (Suarez and Gudiol 2009).

El primer paso para la colonización de la célula huésped es la adhesión mediante las adhesinas que, como se ha descrito anteriormente, se localizan mayoritariamente en la zona de la membrana celular que envuelve la OT. La complejidad del proceso de adherencia en un proceso de infección hace necesaria la implicación de otras proteínas accesorias de la membrana localizadas en la OT que permiten a los Mycoplasmas

penetrar en el interior de las células (Dhandayuthapani, Rasmussen et al. 1998, Rottem and Yogev 2000, Ueno, Timenetsky et al. 2008).

1.2 Chaperonas

1.2.1 Perspectiva histórica

La conversión de la información genética en proteínas que lleva a cabo el programa genético es uno de los procesos centrales en biología. Los ribosomas tienen un rol esencial en éste proceso ya que son los encargados de traducir el ácido ribonucleico mensajero (RNAm) en polipéptidos. Sin embargo existe un proceso igualmente esencial más allá; el correcto plegamiento de la cadena polipeptídica lineal en una conformación tridimensional que permita a los péptidos recién sintetizados realizar su función apropiadamente. Las proteínas responsables por este proceso son las chaperonas moleculares que llevan a la formación de proteínas funcionales (Young, Agashe et al. 2004).

El descubrimiento de las chaperonas moleculares generó mucha controversia. Desde los años 60 que existen estudios dónde se había observado que el tratamiento mediante calor de algunos organismos inducía la expresión de diferentes familias de proteínas que se denominaron Hsp60, Hsp70, Hsp90 (Ritossa 1962, 1996). Su nombre es el acrónimo en inglés de “heat shock proteins” o proteínas de choque térmico

seguido por el valor del peso molecular de la proteína correspondiente. En los mismos estudios también se observó que estas proteínas se conservan entre especies, desde las bacterias hasta los seres humanos.

El término chaperona molecular apareció a finales de los 70 por Ron Laskey que lo utilizó para describir la función de la nucleoplasmina, una proteína nuclear necesaria para el correcto ensamblaje de los nucleosomas de huevos de *Xenopus*. Este trabajo sugería que la nucleoplasmina prevenía interacciones incorrectas entre histonas y DNA (Laskey, Honda et al. 1978).

La función de chaperona fue la última de las principales funciones de la célula en ser descrita y fue John Ellis quien generalizó el término en 1987. Tanto el dogma central de la biología como los experimentos de Christian Anfinsen en 1973 apoyaban la hipótesis que el plegamiento de las proteínas y su ensamblaje en complejos macromoleculares se producían de manera espontánea y sin asistencia de otras proteínas. Uno de los pasos que permitió que emergiera el término chaperona fue la interpretación por Hugh Pelham en 1986 de los efectos observados en células bacterianas sometidas a estrés por aplicación de calor. Pelham propuso que la función de las Hsp en células estresadas era la de deshacer agregados proteicos y prevenir su formación. El término chaperona se generalizó finalmente en 1987 cuando John Ellis, partiendo de sus observaciones del ensamblaje de la ribulosa 1-5 difosfato carboxilasa (más conocida como Rubisco), la enzima responsable de la asimilación del CO₂ en los cloroplastos, con su correspondiente proteína de unión (conocida como la subunidad grande de la proteína de unión a la Rubisco) descubierta en 1980. La función chaperona fue progresivamente demostrada en este sistema y estos estudios llevaron a la descripción de una clase de proteínas celulares cuya función es asegurar el correcto plegamiento

de ciertos polipéptidos y su ensamblaje en estructuras oligoméricas (Barraclough and Ellis 1980, Pelham 1984, Pelham 1985, Pelham 1986, Ellis 1987, Richards 1995).

1.2.2 Generalidades

Dentro de las células, la síntesis *de novo* de proteínas tiene el riesgo de formar agregados que resultan en especies tóxicas para las células (Hartl, Bracher et al. 2011). Para evitar estos peligros las células desarrollaron un complejo sistema de proteostasis que garantiza la integridad de las proteínas y que es necesario para que haya una buena funcionalidad del proteoma. El mantenimiento de este no sólo consiste en ayudar al plegamiento de las proteínas sino también en asistir al ensamblaje de complejos macromoleculares, al transporte, a la degradación, a la disociación de agregados y al plegamiento de proteínas desnaturalizadas por estrés celular. De esta manera, las proteínas realizan su función en el momento en que estén bien plegadas, en que sean funcionales y además estén localizadas dónde realizan su función. El cumplimiento de estas tres condiciones depende en gran parte del balance entre síntesis, plegamiento y degradación de proteínas (Ellis and Hemmingsen 1989, Young, Agashe et al. 2004, Hartl, Bracher et al. 2011, Taylor and Dillin 2011, Kim, Hipp et al. 2013, Saibil 2013).

Las chaperonas son de diferentes tipos y son clasificadas en diferentes grupos en base a la homología de secuencia. Muchas de las chaperonas son las que forman parte de las proteínas de choque térmico ya que su expresión es inducida bajo condiciones de

estrés tales como la temperatura. Desde el punto de vista evolutivo, las chaperonas tuvieron un papel fundamental en la evolución estructural de las proteínas por mantenerlas en un estado soluble y competente. La posible ocurrencia de mutaciones que impiden el buen plegamiento de proteínas recién sintetizadas es amortiguada por la función de la chaperona que, de esta manera, amplía la gama de proteínas sujetas a la selección darwiniana. En el caso de que aparezcan mutaciones favorables, es decir, que aumenten la eficiencia de plegamiento y la solubilidad, la proteína en cuestión es más independiente de la función de la chaperona. Es interesante el hecho que proteínas que son muy dependientes de chaperonas, sean menos abundantes y, con frecuencia, con funciones no esenciales a diferencia de las que no son muy dependientes de chaperonas (Tokuriki and Tawfik 2009, Taipale, Jarosz et al. 2010). Existen excepciones como los casos de la Rubisco, la actina o la tubulina que aunque sean proteínas muy abundantes en la naturaleza son dependientes de chaperonas para lograr su estado funcional. Parece ser que han quedado estancadas en la evolución, ya que mutaciones que reducen la dependencia de la chaperona son incompatibles para que realicen su función. (Hartl 1996, Bukau and Horwich 1998). Un estudio publicado recientemente pone de manifiesto que las chaperonas pueden tener una doble función: catalizar el plegamiento de proteínas y adaptar proteínas recién plegadas a su entorno disminuyendo las posibles variaciones genéticas entre fenotipos promoviendo así la capacidad de evolución (Cetinbas and Shakhnovich 2013).

1.2.3 Organización citosólica en los diferentes Dominios de la Biología

Existen dos mecanismos principales en el plegamiento de las proteínas recién sintetizadas. En un primer mecanismo hay una interacción con los péptidos recién sintetizados por el ribosoma y la chaperona lo que implica la existencia de alguna afinidad entre chaperona y ribosoma. Ejemplos de proteínas plegadas según este mecanismo son: el trigger factor (TF), el complejo asociado a la cadena nascente (del inglés *nascent-chain-associated complex*, NAC), y ciertas especializaciones de Hsp70s. En un segundo mecanismo los polipéptidos nacientes son reconocidos por un grupo de chaperonas, las cuales no presentan afinidad por el ribosoma, siendo este, el sistema clásico de la Hsp70, donde los péptidos son liberados en un estado competente para ser posteriormente plegados por un segundo grupo de chaperonas denominados macrocomplejos cilíndricos de chaperoninas o sistema Hsp90. Estos macrocomplejos proporcionan compartimentos físicamente definidos dentro de los cuales proteínas completas o dominios de éstas acaban de ser correctamente plegados. Recientes estudios bioinformáticos ponen de manifiesto la existencia de un interactoma de sustratos revelando una organización cooperativa de la red de chaperonas (Frydman, Nimmesgern et al. 1994, Kerner, Naylor et al. 2005, Albanese, Yam et al. 2006, Kim, Hipp et al. 2013).

La organización de las chaperonas en las células fue muy conservada a lo largo de la evolución y la encontramos tanto en Bacteria como en Arqueas, Eucariotas o Fungi pero a pesar de ello existen importantes diferencias entre dominios.

En el caso del dominio Bacteria, el TF se une a los péptidos recién sintetizados, aunque aproximadamente un 60 - 80% del total de proteínas se plegarán muy rápidamente sin necesidad de asistencia, y entre un 10 - 20% de los péptidos que poseen una cadena más larga, una vez liberados por el TF, necesitarán interactuar con DnaK, DnaJ (homólogos a Hsp70 y a Hsp40 de Eucariotas, respectivamente) y sus cofactores realizando uno o varios ciclos de unión y liberación de ATP. Finalmente, entre un 10 y un 15% de los péptidos acabarán de ser plegados por las chaperoninas GroEL y GroES (Figura 1.6 A).

En el dominio Arquea existe el NAC que se une a regiones hidrofóbicas de los segmentos que emergen de las cadenas recién sintetizadas por los ribosomas. Los Arquea carecen de Hsp70 pero poseen la prefoldina (PFD) que asiste al plegamiento de la cadena polipeptídica que está unida al NAC y la transfiere al termosoma que realiza la última etapa del plegamiento (Figura 1.6 B).

El dominio de los Eucariotas presenta sistemas de chaperonas más complejos que los otros dominios. Uno de estos sistemas es a través del NAC, que probablemente interactúa con las proteínas recién sintetizadas del mismo modo que el TF en Bacteria. Una vez han interactuado con el NAC una parte de estos péptidos interactuarán con el sistema de chaperonas Hsp70 y Hsp40, algunos de estos péptidos ya estarán plegados completamente y podrán realizar su función en la célula. Otros en cambio, serán transferidos a otros sistemas de chaperonas como la Hsp90, la PFD (prefoldina) y/o al sistema del complejo de polipeptido 1 (TCP-1)-complejo anillo (TRiC)/chaperonina que contiene TCP-1(CCT), para finalizar su plegamiento. (Figura 1.6 C).

El sistema de chaperonas en Fungi implica la unión al ribosoma del Complejo Asociado a Ribosoma (del inglés *Ribosome Associated Complex*, *RAC*) que es constituido por la proteína Ssz1 (homóloga a Hsp70 de Eucariotas) y la zuotin que actúa como una cochaperona Hsp40 y asiste al plegamiento de la cadena recién sintetizada juntamente con una isoforma de Hsp70 llamada Ssb (Figura 1.6 D).

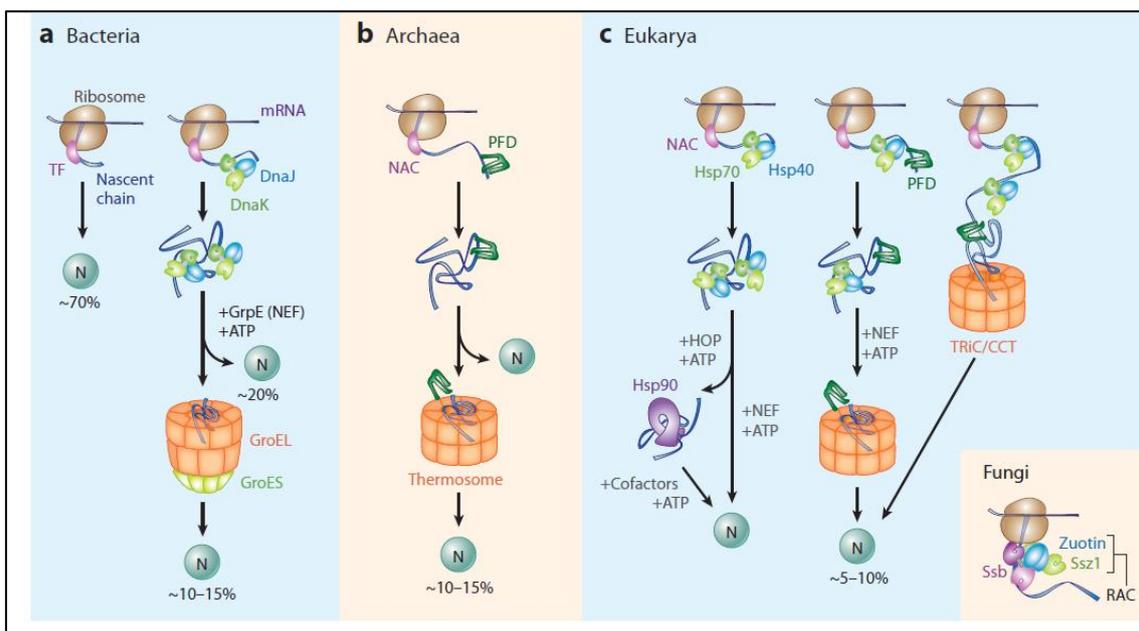


Figura 1.6 Esquema de los sistemas de chaperonas presentes en los diferentes Dominios. Imagen adaptada de (Hartl and Hayer-Hartl 2002, Kim, Hipp et al. 2013)

- Componentes de los sistemas de chaperonas en procariotas

Trigger Factor

El Trigger factor, de 48 kDa, es la única chaperona asociada al ribosoma conocida en Bacterias, a pesar de presentar una similitud funcional pero no estructural con chaperonas asociadas al ribosoma de Eucariotas (Wegrzyn and Deuerling 2005). El TF se expresa constitutivamente y de forma muy abundante en el citosol, interacciona

con la proteína L23, una de las múltiples proteínas que forman parte del ribosoma, con una estequiometría de 1:1 y enlaza directamente la síntesis proteica y el plegamiento de las proteínas asistido por las chaperonas. El TF se localiza directamente en el canal de salida de los péptidos recién sintetizados por el ribosoma, donde crecen las cadenas del polipéptido, abandonando el ribosoma a la vez que entran a formar parte del concurrido citosol y presenta una actividad peptidil-prolil cis/trans isomerasa (PPIasa) *in vitro*. El reconocimiento de péptidos por parte del TF es independiente de residuos de prolina y es mediado por secuencias de aminoácidos ricas en residuos básicos y aromáticos solapándose su función con la del principal sistema de chaperonas en Bacteria y Eucariota, respectivamente DnaK y DnaJ, y Hsp70 y Hsp40 que estabilizan las cadenas nacientes hasta que sean plegadas (Hoffmann, Bukau et al. 2010).

DnaK, DnaJ y factor intercambiador de nucleótidos (NEF)

De las chaperonas moleculares los aspectos estructurales y mecánicos que mejor se conocen son los del sistema Hsp70 que se compone de DnaK (Hsp70), de su cochaperona DnaJ (Hsp 40) y de su factor intercambiador de nucleótidos llamado GrpE.

DnaK es una proteína de aproximadamente 70 kDa que consiste en dos dominios muy conservados: i) el dominio N-terminal (~43 kDa) que une nucleótidos (Dominio de unión a nucleótidos, del inglés *Nucleotide Binding Domain*, NBD) y ii) el dominio C-terminal que une los sustratos (Dominio de unión a sustrato, del inglés *Substrate Binding Domain*, SBD). Estos dominios están conectados por un enlazador (del inglés *linker*) o unión muy flexible. Los sustratos parcialmente plegados o sin plegar son reconocidos por el SBD de DnaK que son transferidos a la proteína a través de la DnaJ. Entonces, mediante una

serie de ciclos de unión y liberación de ATP la DnaK cambia alternativamente de la conformación abierta a cerrada y vice-versa lo que conlleva al plegamiento del sustrato y liberación de un fosfato consecuencia de la hidrólisis del ATP. A continuación la GrpE se encarga de devolver la DnaK a un estado en que pueda unir ATP, o sea una conformación abierta (Figura 1.7)(Calloni, Chen et al. 2012).

DnaJ es una cochaperona de aproximadamente 40 kDa típicamente formada por cuatro dominios distintos: i) el dominio J, ii) una región rica en glicinas y fenilalaninas (región G/F), iii) una región con un dedo de zinc y iiiii) un dominio C-terminal. El rol más conocido de DnaJ es el de unir péptidos sin plegar y transferirlos a DnaK a la vez que acelera la hidrólisis de ATP. La existencia de diferentes tipos de DnaJs que no poseen todos los dominios característicos de estas proteínas ponen de manifiesto la posible variedad de funciones que tienen: i) al poseer los dominios de unión a polipéptidos los unen y transfieren de manera específica a la Hsp70, ii) pueden estabilizar los complejos formados por Hsp70 y el sustrato activando la hidrólisis del ATP y iii) miembros especializados de la familia de las Hsp40 son localizados en diferentes compartimentos celulares de manera que pueden haber uniones Hsp70-Hsp40 diferentes y únicas en estos compartimentos (Fan, Lee et al. 2003).

GrpE es un factor intercambiador de nucleótidos de unos 25 kDa. Su función es la de promover la disociación de la adenosina 5'-difosfato (ADP) unida al dominio de unión a nucleótido de DnaK permitiendo la entrada de la adenosín 5'-trifosfato (ATP). GrpE y DnaJ actúan conjuntamente para controlar el flujo de péptidos sin plegar dentro y fuera del dominio de unión a sustrato de DnaK regulando así la unión de nucleótidos a DnaK y su conformación (Harrison 2003).

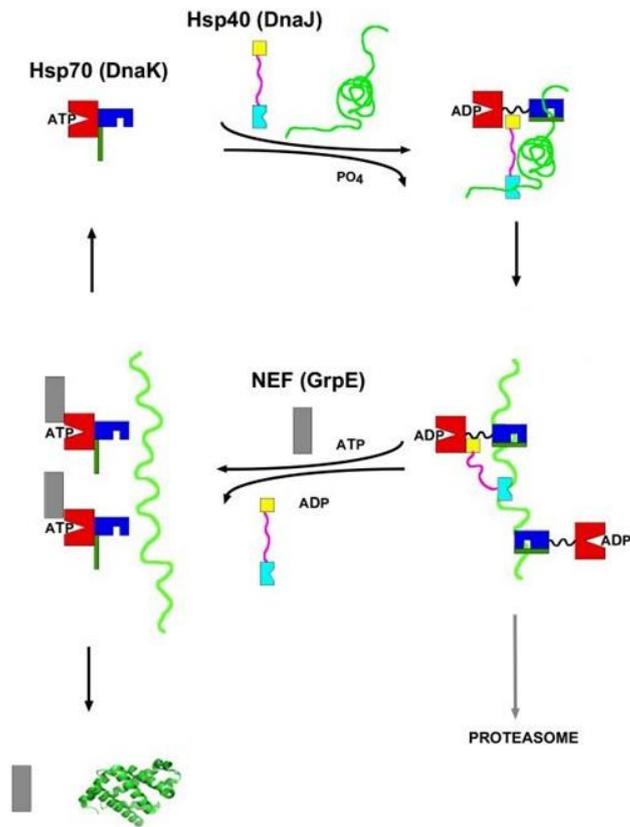


Figura 1.7 Ciclo de Hsp70-Hsp40 y GrpE. Figura adaptada de Zuiderweg (Zuiderweg, Bertelsen et al. 2013).

A pesar de los múltiples estudios del sistema Hsp70 aún existe controversia en relación a su mecanismo de acción por la elevada dinámica que posee. Actualmente existen tres hipótesis para tal mecanismo. Tanto la primera como la segunda hipótesis fueron resultado de la interpretación de los experimentos en que se evidenció que los sustratos unidos a la DnaK de *E.coli* estaban en una conformación más desplegada que en solución. La primera hipótesis de estos resultados se denominó como "modelo de selección conformacional". En este caso hay un equilibrio entre los sustratos que están más plegados y compactos y los que lo están menos. DnaK atraparía el sustrato en una conformación menos plegada por la elevada tasa de asociación de la conformación abierta y pasaría rápidamente a un estado de elevada afinidad por el sustrato.

A la segunda hipótesis se le llamó "modelo del desplegamiento inducido" en el que DnaK se uniría a la conformación plegada del sustrato e induciría su desplegamiento.

La tercera hipótesis ha aparecido en los últimos años y se ha demostrado, mediante técnicas biofísicas, el gran dinamismo y las continuas fluctuaciones en el ensamblaje de conformaciones que presentan las proteínas. Como muchos osciladores físicos, cuando una masa se añade al oscilador, éste cambiará intrínsecamente la frecuencia, la amplitud y la dirección de las oscilaciones. Cuando DnaK se une a proteínas puede cambiar la fluctuación y la amplitud, y los estados conformacionales accesibles a la proteína. Además, la DnaK puede formar parte del sistema oscilador por las fluctuaciones térmicas de las chaperonas y sus sustratos se acoplarían (Mayer 2013).

GroEL y GroES

Las proteínas GroEL y GroES forman parte de la familia de las chaperoninas. GroEL forma un gran complejo de dos anillos y su peso molecular es de unos 800-900 kDa. Su función consiste en plegar proteínas introduciéndolas en su interior pudiendo llegar a plegar proteínas de hasta 60 kDa. GroEL funciona conjuntamente con GroES, que forma la tapa del complejo de anillos. Los dominios apicales de GroEL presentan aminoácidos hidrofóbicos que se unen al sustrato en el centro del anillo. Seguidamente, el plegamiento depende del encapsulamiento global por parte de GroES. Una vez dentro del complejo el sustrato es plegado en unos 10 segundos que es el tiempo que tarda el ATP a ser hidrolizado por GroES (anillos cis). Cuando el péptido está plegado GroES se disocia de GroEL y el sustrato es liberado. Este proceso está regulado por la unión de ATP en el anillo opuesto (anillo trans). Cuando el sustrato no

está totalmente plegado vuelve a ser reintroducido en el anillo de GroEL que sigue plegándolo (Figura 1.8) (Hartl, Bracher et al. 2011).

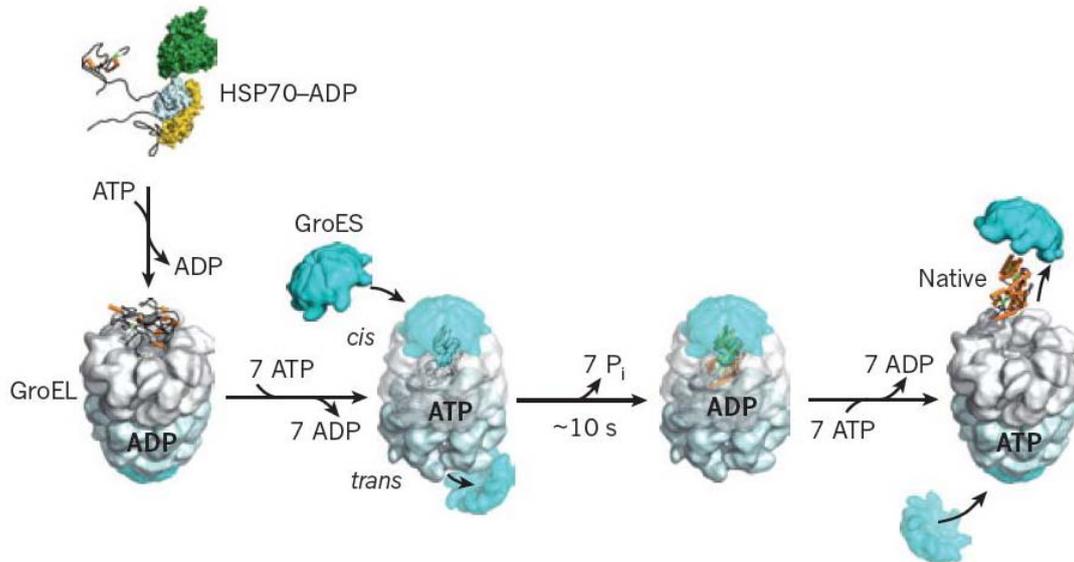


Figura 1.8 Mecanismo de plegamiento de GroEL-GroES. Figura extraída de (Hartl, Bracher et al. 2011)

ClpB

En *E. coli* se identificó la familia de las Hsp100 o caseína proteinasa lítica B (del inglés *caseinolytic peptidase B*, ClpB) que presenta un rol importante en el plegamiento, transporte y degradación de proteínas. Su peso molecular es de unos 100 kDa y la proteína forma unos cilindros hexaméricos (por ejemplo: ClpA, ClpX y HslU) que se unen coaxialmente a un complejo proteolítico también con forma de anillo (por ejemplo: ClpP y HslV). La actividad proteolítica está localizada en el interior del anillo y parece ser que los sustratos proteicos son desplegados dentro del anillo para

posteriormente ser degradados por los complejos proteolíticos (Reid, Fenton et al. 2001).

Proteína de altas temperaturas (HtpG)

La proteína de alta temperatura G (del inglés *High Temperature Protein G*, HtpG) es una chaperona homóloga a las Hsp90 de Eucariotas y tiene un peso molecular de unos 90 kDa. Estas proteínas sólo están presentes en Bacterias y Eucariotas y su función consiste en el plegamiento general de proteínas y en estabilizarlas porque puede unirse a péptidos intermediarios que no están plegados manteniendo un equilibrio entre sustratos plegados y sin plegar evitando la agregación irreversible de proteínas. Además parece tener un papel en la señalización celular (Spence, Cegielska et al. 1990).

Pequeñas proteínas de choque térmico (HSPs)

Las pequeñas proteínas de choque térmico (del inglés *Small Heat Shock Proteins*, HSPs) tienen un peso molecular comprendido entre 12 a 43 kDa y se caracterizan por la presencia de aproximadamente 80-100 aminoácidos muy conservados en el C-terminal, llamado "dominio α -cristalina", que está flanqueado por dos dominios menos conservados: el N-terminal y una extensión C-terminal. Estas proteínas pertenecen a una familia conservada de chaperonas moleculares independientes de ATP cuya su función consiste en unir a sustratos proteicos no nativos y facilitar su plegamiento en cooperación con chaperonas dependientes de ATP (Shi, Yan et al. 2014).

Proteasa Lon

La proteasa Lon es dependiente de ATP y contribuye a la regulación de funciones celulares importantes que incluye la resistencia a la radiación, la división celular, la filamentación, la producción de polisacáridos capsulares y la degradación proteolítica de proteínas reguladoras o mal plegadas (Fu, Smith et al. 1997).

La regulación de las Hsps

La regulación de la expresión de Hsps en procariotas se da de manera transcripcional, post-transcripcional y translacional. Estos mecanismos han sido muy estudiados en *E. coli*, como ejemplo de Gram negativo, y en *Bacillus subtilis* como ejemplo de Gram positivo. *E. coli* presenta un mecanismo de regulación positiva por el factor σ^{32} , a diferencia de *B. subtilis* que presenta tres clases sistemas de regulación: clase I, control negativo del factor inducible por calor, represor de la transcripción de HrcA; clase II, control positivo por el factor σ^B y clase III, regulación negativa por el regulador de la respuesta frente al estrés del tipo III (CtsR).

En el caso de *E. coli*, cuando las células crecen bajo condiciones de estrés, el nivel celular del factor σ^{32} incrementa de manera transitoria incrementando también la transcripción de los promotores de las proteínas del choque térmico. A esta primera etapa se le llama fase de inducción, que va seguida de la fase de adaptación. Esta segunda etapa consiste en una disminución de la síntesis de σ^{32} consecuencia del control negativo ejercido por DnaK, DnaJ y GrpE controlado a nivel de represión translacional, desestabilización y probablemente por inactivación del factor σ^{32} (Figura 1.9 A) (Yura and Nakahigashi 1999).

El mecanismo de regulación transcripcional negativa más extendido es el que involucra el receptor regulador del calor en CIRCE (HrcA), que interacciona con el DNA regulador, llamado elemento con repeticiones invertidas, que controla la expresión de chaperonas (del inglés *Conserved Controlling Inverted Repeat Chaperone Expression*, CIRCE). Cuando HrcA interacciona con el elemento CIRCE la transcripción queda reprimida. Normalmente, los elementos CIRCE se encuentran delante de los genes DnaK y GroEL donde se encuentran los operones que codifican para las chaperonas. El elemento cis en negativo, consiste en una repetición invertida de una secuencia conservada de 9 pb que contiene un espaciador también de 9 pb (TTAGCACTC-N9-GAGTGCTAA) y es la conservación de esta secuencia que determina la eficiencia de la represión. Recientemente, se ha demostrado que la región CIRCE controla la expresión génica por el incremento de la estabilidad de los transcritos (Figura 1.9 B)(Walsh, Bursac et al. 2004, Chang, Chen et al. 2008). En cambio, la regulación de la transcripción de HrcA se da por sí mismo, es decir, se autorregula, a pesar de también poder estar regulado por otros mecanismos. La represión de la actividad de HrcA está influenciada por su interacción con el sistema GroES/GroEL (Falah and Gupta 1997, Oshima, Kakizawa et al. 2004). Existen otros mecanismos de regulación negativa como la regulación de los operones clpC, clpE y clpP por CtsR el cual se une específicamente a la región consenso A/GGTCAAA NAN A/GGTCAAA. Este es el caso en *Listeria monocytogenes*, *Lactococcus lactis*, y *Streptococcus pneumoniae* (Derre, Rapoport et al. 1999, Nair, Derre et al. 2000). Otro represor de los genes *hsp* en bacteria es el HspR que se une al elemento HspR repetición invertida asociada, localizado detrás de la región promotora de *dnaK* en operones como los de *Streptomyces*.

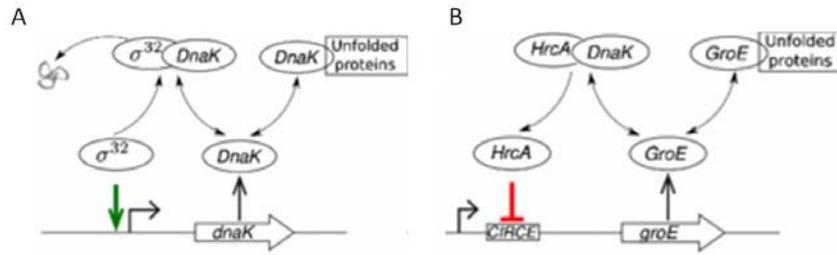


Figura 1.9 A) Esquema de la regulación de la transcripción A) positiva por el factor σ^{32} y B) negativa por HrcA Figura extraída de (Inoue, Mitarai et al. 2012).

1.2.4 Sistema de chaperonas en *Mycoplasma*

- Generalidades

El sistema de Hsp en *Mycoplasma* se descubrió en 1990 cuando se demostró la síntesis de Hsp en *Acholeplasma laidlawii* como respuesta a choque térmico. Después de incrementar la temperatura de 32 a 44 grados se pudo observar un aumento de la síntesis de polipéptidos de pesos moleculares de aproximadamente 20 a 70 kDa (Dascher, Poddar et al. 1990).

Los *Mycoplasmas*, en general, raramente están expuestos a cambios severos de temperatura excepto durante la colonización y procesos de infección. La presencia de Hsp en genomas tan reducidos como el de *Mycoplasma* demuestra el papel fundamental que tienen en todos los organismos, no sólo como respuesta frente a condiciones ambientales desfavorables sino también en condiciones normales (Narberhaus 2002).

El acceso a las secuenciaciones de los genomas de 35 especies de Mycoplasmas han permitido el estudio comparativo de los genes de estos Mycoplasmas y, concretamente, en el caso de las chaperonas, parece ser que el papel de DnaK, DnaJ, y GrpE podría ser multifuncional lo que es biológicamente razonable si nos basamos en la capacidad codificante limitada que presentan. Un ejemplo es el caso de *M. suis* en el cual la proteína análoga a DnaK está involucrada en la adhesión a eritrocitos. Otro ejemplo es el caso de la proteína TopJ de *M. pneumoniae* que posee un dominio J típico de las proteínas DnaJ, y que forma parte de la OT y tiene una función relacionada con la adhesión a células huésped, motilidad por deslizamiento y división celular (Vishnyakov, 2013). La MG200 de *M. genitalium* (proteína homóloga a la TopJ de *M. pneumoniae*), también se localiza en la OT y tiene una función relacionada con la motilidad pero no con la adhesión celular (Pich, Burgos et al. 2006).

Las sHSP cuya función, juntamente con otras chaperonas, es la de proteger las estructuras celulares, parecen, además, tener una función en el mantenimiento de la fluidez de las membranas y en su integridad cuando hay fluctuaciones térmicas en *A. Laidlawii* (Vishnyakov, 2013).

Las Hsp de Mycoplasma son muy inmunogénicas y pueden ser consideradas potenciales dianas para desarrollar nuevos test diagnósticos o nuevas vacunas (Wong and Houry 2004, Musatovova, Dhandayuthapani et al. 2006). Uno de los sistemas más inmunogénicos en Mycoplasma es el sistema GroEL/GroES pero, curiosamente, es uno de los sistemas que desapareció en algunas especies de Mycoplasma. Hay diversas hipótesis sobre la pérdida de este sistema, desde que la pérdida fue debida a que los Mycoplasmas a lo largo de la evolución invirtieron más en la degradación de las

proteínas que en prevenir su agregación (Wong and Houry 2004), a que la presencia del gen de GroEL en especies invasoras como *M.penetrans* podría ser el resultado de una transferencia horizontal de otra bacteria invasora como *Helicobacter pylori* (Williams and Fares 2010) o que debido a la reducción del genoma las proteínas que dependían de GroEL/GroES para su plegamiento también se hayan reducido en el curso de la evolución (Clark and Tillier 2010).

Como consecuencia el sistema Hsp70 (DnaK/DnaJ/GrpE) y el TF han pasado a tener una especial importancia en los Mycoplasmas a pesar de que especies como *M. haemominutum*, *M. wenyonii* y *M. suis* sobreviven perfectamente en ausencia de TF (Tabla 1.2) (Vishnyakov, 2013).

Mycoplasmas/HSP	CipB	CipA	CipX	HtpG	DnaK	DnaJ	GrpE	GroEL	GroES	FisH (HIB)	DegP (HtrA)	Protease La (Lon)	Hsp33 (HslO)	sHsp (hpsAB)	HrcA (repressor)	Trigger factor
<i>Mesoplasma florum</i> (AE017263.1)	-	-	-	+	+	3±	+	-	-	-	-	+	+	-	+	+
<i>Ureaplasma urealyticum</i> (CP001184.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Ureaplasma parvum</i> (AF222894.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma agalactiae</i> (CU179680.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma hominis</i> (FP236530.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma hyorhinis</i> (CP002170.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma hyopneumoniae</i> (AE017332.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma bovis</i> (CP002188.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma fermentans</i> (CP002458.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	2+
<i>Mycoplasma arthritidis</i> (CP001047.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma cynos</i> (HF559394.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma mycoides</i> (BX293980.2)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	-
<i>Mycoplasma conjunctivae</i> (FM864216.2)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma capricolum</i> (CP000123.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma leachii</i> (CP002108.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma mobile</i> (AE017308.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma putrefaciens</i> (CP003021.1)	+	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma crocodyli</i> (CP001991.1)	-	-	-	-	-	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma haemofelis</i> (FR773153.2)	-	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma pulmonis</i> (AL445566.1)	-	-	-	-	+	+	+	-	-	-	-	+	-	-	+	+
<i>Mycoplasma synoviae</i> (AE017245.1)	-	-	-	-	+	+	+	-	-	-	-	+	-	-	+	±
<i>Mycoplasma haemocanis</i> (CP003199.1)	-	-	-	-	+	+	1+, 1±	-	-	-	-	+	-	-	+	±
<i>Ca. Mycoplasma haemominutum</i> (HE613254.1)	-	-	-	-	+	+	+	-	-	-	-	+	-	-	+	±
<i>Mycoplasma suis</i> (FQ790233.1)	-	-	-	-	+	+	+	-	-	-	-	+	-	-	+	-
<i>Mycoplasma wenyonii</i> (CP003703.1)	-	-	-	-	+	+	+	-	-	-	-	+	-	-	+	-
<i>Ca. Mycoplasma haemolamae</i> (CP003731.1)	-	-	-	-	+	+	+	1+, 1±	-	-	-	+	-	-	+	-
<i>Mycoplasma genitalium</i> (L43967.2)	+	-	-	-	+	3+	+	+	+	+	-	+	-	-	+	+
<i>Mycoplasma pneumoniae</i> (U00089.2)	+	-	-	-	+	3+	+	+	+	+	-	+	-	-	+	+
<i>Mycoplasma penetrans</i> (BA000026.2)	+	-	-	-	+	+	+	+	+	+	-	+	-	-	+	+
<i>Mycoplasma gallisepticum</i> (AE015450.2)	+	-	-	-	+	8±	+	+	+	+	-	+	-	-	+	+
<i>Ca. Phytoplasma mali</i> (CU469464.1)	-	-	-	-	+	+	+	+	+	12+	-	+	-	-	+	+
<i>Ca. Phytoplasma asteris</i> OY-M (AP006628.2)	-	-	-	±	+	+	+	+	+	-	-	+	-	+	+	-
<i>Ca. Phytoplasma asteris</i> AY-WB (CP000061.1)	-	-	-	-	+	+	+	+	+	+	-	+	-	+	+	+
<i>Ca. Phytoplasma australiense</i> (AM422018.1)	-	-	-	-	+	+	+	+	+	+	-	+	-	+	+	+
<i>Acholeplasma laidlawii</i> (CP000896.1)	+	-	-	-	+	+	+	+	+	+	-	+	-	+	+	+

Tabla 1.2 Comparación de la presencia o ausencia de genes que codifican para las proteínas de la red de chaperonas en Mycoplasmas: chaperonas, cochaperonas, proteasas y proteínas reguladoras. Figura extraída de Vishnyakov, 2013.

En la Tabla 1.2 se muestran detalladamente las diferencias en la presencia o no de los diferentes genes que codifican para las Hsp en las diferentes especies de Mycoplasma. Cabe destacar que excluyendo los genes parálogos, existen diferentes especies de Mycoplasma en los que sólo hay 6 genes para Hsps, como por ejemplo en *M. crocodyli*, *M. haemominutum*, *M. wenyonii* y en *M. suis* pero además, en el caso de *M. crocodyli*, no existe el gen que codifica para el homólogo de una de las Hsp más importantes, la DnaK, aunque sí mantiene el de su cochaperona, la DnaJ, y su factor intercambiador de nucleótidos, GrpE. En cambio, en las especies *M. penetrans*, *M. gallisepticum*, *M. genitalium*, *M. pneumoniae*, *Candidatus phytoblasma asteris* y *Candidatus phytoblasma australiense* poseen 10 genes codificantes para Hsps (Vishnyakov, 2013). Por lo tanto, es importante tener en cuenta que los Mycoplasmas, en comparación con las Bacterias, a pesar de tener un genoma reducido y de poseer un menor conjunto de genes que codifican para chaperonas y proteasas, mantuvieron una representación amplia y variada de estas para poder sobrevivir en ambientes desfavorables. Se cree que la red múltiple de chaperonas responsable del estado de las proteínas es crucial para la viabilidad celular tanto en condiciones de estrés como en condiciones óptimas. Cuantos más componentes forman parte de esta red, más interacciones se producen y, por lo tanto, la capacidad amortiguadora de la red frente a cambios en el ambiente es mayor. Las Hsp son las proteínas principales de esta red múltiple de chaperonas y por lo tanto, las diferencias en el número de genes que codifican para las Hsp serían el resultado de la adaptación a condiciones desfavorables (Narberhaus 2002, Wong and Houry 2004, Csermely, Korcsmaros et al. 2008).

Otro factor destacable del análisis de los genomas de diferentes Mycoplasmas es la identificación de un único factor σ , el represor HrcA, y de las secuencias homólogas a

CIRCE, que indican el uso del mecanismo de regulación negativa mediado por HrcA como respuesta a un cambio de temperatura. Este podría ser el único mecanismo regulador presente en *Mycoplasmas* (Chang, Chen et al. 2008) y un ejemplo de eficiencia en la regulación. Sin embargo, los cambios en el nivel de transcripción de genes relacionados con el choque térmico donde sus promotores de regulación son aun desconocidos indican que quizás en *Mycoplasma* existen otros mecanismos de regulación que aun no se han estudiado. Un ejemplo es el hecho que algunos *Mycoplasmas* son capaces de inducir la síntesis de Hsps (*A. laidlawii* y *M. pneumoniae*) frente a los que son sintetizados constitutivamente (*M. genitalium* y *M. fermentas*) (Vonskii, Astvatsaturiants et al. 1993).

- Sistema de chaperonas en *Mycoplasma genitalium*

Sorprendentemente, *M. genitalium* revela una baja similitud con otras especies bacterianas con respecto a la regulación de la expresión génica. Su genoma presenta genes ortólogos para la mayor parte de reguladores bacterianos convencionales, como un factor σ alternativo y sistemas de transducción de dos componentes. A pesar de poseer ortólogos para los genes *hsp* *M. genitalium* incluye un gen que codifica para la proteína reguladora HrcA.

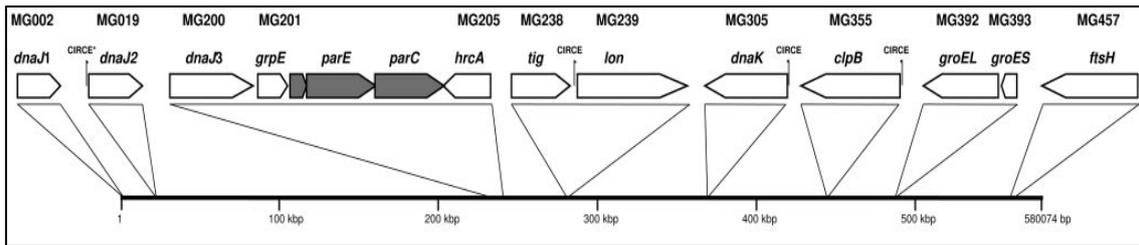


Figura 1.10. Esquema de la posición y organización de los genes relacionados con choque térmico en el genoma de *M. genitalium*. Los genes del choque térmico están representados mediante flechas blancas y su orientación representa la correspondiente hebra codificadora. Las flechas negras muestran genes que separan *hrcA* de *DnaJ3* y *GrpE*. Los genes son designados por el número de su locus y los nombres anotados en www.tigr.org. También se muestra su localización en el cromosoma. Figura extraída de Musatovova y colaboradores (Musatovova, Dhandayuthapani et al. 2006)

M. genitalium contiene los genes asociados a las proteínas de choque térmico dispersas en su genoma poseyendo tres genes parálogos que codifican para proteínas con un dominio J. El primer gen es el *dnaJ1* (locus MG002) que codifica para la DnaJ1, delante del cual se encuentra la subunidad β de la DNA polimerasa III (*dnaN*, locus MG001) y la subunidad β de la DNA girasa (*gyrB*, locus MG003) ambas asociadas con funciones de reparación y replicación del DNA. El segundo gen es el *dnaJ2* (locus MG019) que se encuentra flanqueado por el gen que codifica para la helicasa *Mot1* (locus MG018) que está involucrada en el procesamiento de RNA y el de la Prolina iminopeptidasa (*Pip*, locus MG020) que está involucrada en la degradación de proteínas, péptidos y glicopéptidos. El tercer gen es el *dnaJ3* (locus MG200), que es el único que tiene una relación aparente con el gen *hsp*, y que codifica para el factor intercambiador de nucleótidos GrpE (locus MG201). Cerca de esta región está también el gen que codifica para la proteína reguladora de las proteínas de choque térmico HrcA (locus MG205) sin embargo, la dirección de transcripción es la opuesta a la de GrpE. Otros genes relacionados son: *lon* (locus MG239), *dnaK* (locus MG305), y *clpB* (locus MG355) que codifican respectivamente para la proteasa dependiente de ATP Lon, DnaK y para la subunidad ATPasa de la proteasa dependiente de ATP ClpB. En las

regiones promotoras de estos genes, se encuentran los genes del elemento CIRCE, indicando que estos están regulados por HrcA. Además, los genes que codifican para la chaperona GroEL (locus MG393) y para la cochaperona GroES (locus MG392) están localizados de manera adyacente a las anteriores tal y como se observa en la gran mayoría de especies bacterianas.

A pesar de los genes convencionales relacionados con el choque térmico, *M. genitalium* posee otros genes relacionados con el sistema Hsp70, como por ejemplo: *tig* (locus MG238) que codifica para el TF que cataliza el plegamiento de proteínas asociadas al ribosoma (Bang, Pecht et al. 2000, Vogtherr, Jacobs et al. 2002). Este gen se localiza delante del locus MG239 que codifica para la proteasa Lon. El último locus relacionado con las Hsp es el MG457 que codifica para la proteína FtsH que es una proteasa asociada a Hsps.

Musatovova y colaboradores demostraron en 2006 que la inducción de las proteínas del choque térmico en respuesta a un estrés por calor era selectiva e irreversible. Entre los genes inducidos aparecieron *lon*, *dnaK*, *clpB*, *dnaJ3*, y *grpE*, concretamente, los que presentaban una respuesta mayor fueron los que codificaban para las proteínas DnaK, Lon y ClpB. En cambio, los genes *dnaJ1*, *dnaJ2*, *groES*, *groEL*, *tig*, *ftsH*, *hrcA*, *rpoD*, y *tuf* mantuvieron el mismo nivel de expresión o menor (Musatovova, Dhandayuthapani et al. 2006). Ninguna de las secuencias reguladoras que se conocen fue encontrada en el promotor del gen *dnaJ3* a pesar del incremento después del choque térmico.

M. genitalium posee varias copias muy parecidas a la secuencia CIRCE (TTAGCATC-N9-GAGTGCTAA) delante de los promotores de *dnaK*, *clpB*, y *lon*. También se encontró otra secuencia muy parecida a la CIRCE delante del promotor de *dnaJ* (TTAaCACTC-N9-

aAGTgATT, en minúsculas se muestran los nucleótidos no conservados), su secuencia está menos conservada que la anterior y además no hay un cambio en respuesta a un choque térmico. Estudios de unión de HrcA con secuencias CIRCE demostraron que *dnaK*, *clpB*, y *lon* tenían una regulación transcripcional mediada por *HrcA*, mientras que *dnaJ2* implicaba otros factores independientes del sistema HrcA-CIRCE como podría ser la posición relativa de la secuencia CIRCE con el inicio de la transcripción

En el caso de los genes *tig* y *lon* su transcripción es combinada, aunque *lon* posee su propio promotor. Por último, en el operon *groEL/groES* también existe una secuencia CIRCE pero no hay cambios en su expresión frente a un choque térmico (Musatovova, Dhandayuthapani et al. 2006).

1.2.5 Sistema DnaK-DnaJ-GrpE en *Mycoplasma genitalium*

Numerosas funciones se han descrito para la chaperona DnaK y para las cochaperonas moleculares DnaJ y GrpE, las cuales intervienen en procesos celulares como la división celular, el plegamiento, la secreción proteica, la síntesis de RNA y DNA, la regulación del choque térmico y, además, en la motilidad celular.

La participación de chaperonas moleculares en mecanismos de motilidad fue descrita por Shi y colaboradores en 1992 demostrando que la síntesis del flagelo bacteriano requería la participación de DnaK, DnaJ y GrpE (Shi, Zhou et al. 1992). En 1998 se demostró que *Myxococcus xanthus*, un organismo no flagelar, requería una proteína homóloga a DnaK, no regulada por choque térmico, para moverse (Yang, Geng et al.

1998). Este fenómeno no se observó solamente en procariotas sino también en eucariotas cuando, en 1995, Karl Johnson identificó una isoforma de Hsp70 en el flagelo de *Chlamydomonas* que se localizaba mayoritariamente en el lugar del ensamblaje de los microtúbulos. Los autores propusieron un modelo en el cual componentes como la tubulina y otras proteínas que forman parte del flagelo (axonemales) se asociaban con una isoforma de Hsp70, específica del flagelo, durante su síntesis en el citoplasma. Estos complejos formados entre las proteínas flagelares y la chaperonas eran translocados hasta el extremo del compartimento flagelar donde se disociaban para poder montar los extremos distales de los axonemas (Johnson and Rosenbaum 1992). Estas posibles nuevas funciones de las Hsp70 son coherentes con las características que se conocían para estas proteínas que también fueron estudiadas en otros organismos (Georgopoulos and Welch 1993). Dado que las Hsp70, la tubulina, los microtúbulos y la estructura flagelar se han conservado a lo largo de la evolución, desde los eucariotas inferiores hasta los humanos, la función de las proteínas de choque térmico podría ser clave en el montaje del citoesqueleto (Georgopoulos and Welch 1993).

La localización de DnaK, MG200 (proteína con un dominio J) y GrpE en la OT de *M. genitalium* ha sido el punto de partida para iniciar la investigación sobre el posible papel de las chaperonas de *M. genitalium* en la motilidad celular. Más concretamente, DnaK y MG200 fueron detectadas en el complejo rueda, y GrpE en el Núcleo electrodenso de la OT (A. Mariscal y L. Gonzalez manuscrito en preparación).

Como se describió en apartados anteriores la actividad de DnaK está controlada por DnaJ y por GrpE. La familia de las proteínas J se clasifica en tres subfamilias según su

secuencia primaria: i) proteínas J de tipo I, contienen todos los dominios estructurales definidos para DnaJ, un dominio formado por cuatro hélices α que corresponde al dominio J, una región G/F, un dominio de dedo de zinc y un dominio Carboxi-terminal; ii) proteínas J de tipo II, contienen un dominio J, una región(G/F) y un dominio C-terminal y iii) proteínas J de tipo III, sólo contienen un dominio J y un dominio C-terminal (Figura 1.11 A)(Cheetham and Caplan 1998, Walsh, Bursac et al. 2004).

En *M. genitalium* existen tres proteínas que contienen dominios J: MG019, MG002 y MG200. MG019 correspondería a una proteína J de tipo I siendo, a nivel de secuencia, la proteína homóloga a la DnaJ de *E.coli*, MG002 correspondería al tipo III y MG200 al tipo II. MG200 se clasifica así desde hace relativamente poco porque la región G/F ha sido descrita muy recientemente (A. Broto et al., manuscrito en preparación) (Figura 1.11 A). El análisis de la estructura secundaria de proteínas J y el conocimiento de la estructura tridimensional de varios dominios J determinados por cristalografía de rayos-X y/o RMN indican que los dominios J están formados por cuatro hélices α (I a IV). La hélice II es rica en residuos básicos y el bucle entre las hélices II y III contiene el motivo formado por los residuos His-Pro-Asp (HPD) que son muy conservados y esenciales para la interacción con DnaK (Greene, Maskos et al. 1998). Además, éstas hélices (II y III) contienen residuos hidrofóbicos que las estabilizan Figura 1.11 B y C (Pellecchia, Szyperski et al. 1996, Walsh, Bursac et al. 2004). Mientras que las distinciones estructurales entre las DnaJ de tipo I y II están claramente definidas sus funciones no lo están. Parece que la DnaJ de tipo I es la homóloga a *E. coli* y su función sería la de actuar como cochaperona de DnaK, unir péptidos sin plegar y evitar la agregación de estos en las células, al igual que sucede con las DnaJ de tipo II (Gierash 2014). En cambio, las DnaJ de tipo III pertenecen funcionalmente a otra categoría y

estimulan la actividad de Hsp70 en otros compartimentos celulares. Estudios en levaduras localizan DnaJ principalmente en el citosol, pero también en mitocondrias, núcleo o unidas al retículo endoplasmático. Estas DnaJ tienen secuencias específicas para el reconocimiento de determinados sustratos que son entregados a DnaK (Walsh, Bursac et al. 2004). La presencia de múltiples genes que codifican para proteínas DnaJ no es por lo tanto inusual en *M. genitalium*. En el citosol de levaduras, por ejemplo, hay tres isoformas de Hsp70 que coexisten con al menos 14 DnaJ distintas (Walsh, Bursac et al. 2004). En los Mycoplasmas sólo hay dos casos, aparte de *M. genitalium*, en el que existe más de un gen que codifica para proteínas J, en *M. gallisepticum* y en *M. pneumoniae* (Musatovova, Dhandayuthapani et al. 2006).

A diferencia de DnaJ, GrpE tiene una función muy bien definida. Cuando DnaK está unida a ADP y el SBD se encuentra ocupado por un péptido (estado de alta afinidad por el sustrato), GrpE puede unirse a DnaK causando la disociación del ADP y promoviendo la unión de DnaK a ATP (estado de baja afinidad por el sustrato) y la liberación del péptido. La estructura del dímero de GrpE de *Thermus thermophilus* adopta una forma de ángel y cada protómero está formado por tres dominios: el N-terminal, que está formado por una larga hélice α , el dominio central, que está formado por 2 hélices α , y el dominio C-terminal, que está formado por una hoja β . En el dímero el Nterminal de los dos dominios forma un coiled-coil (Figura 1.11 D) (Harrison, Hayer-Hartl et al. 1997, Harrison 2003, Nakamura, Takumi et al. 2010)

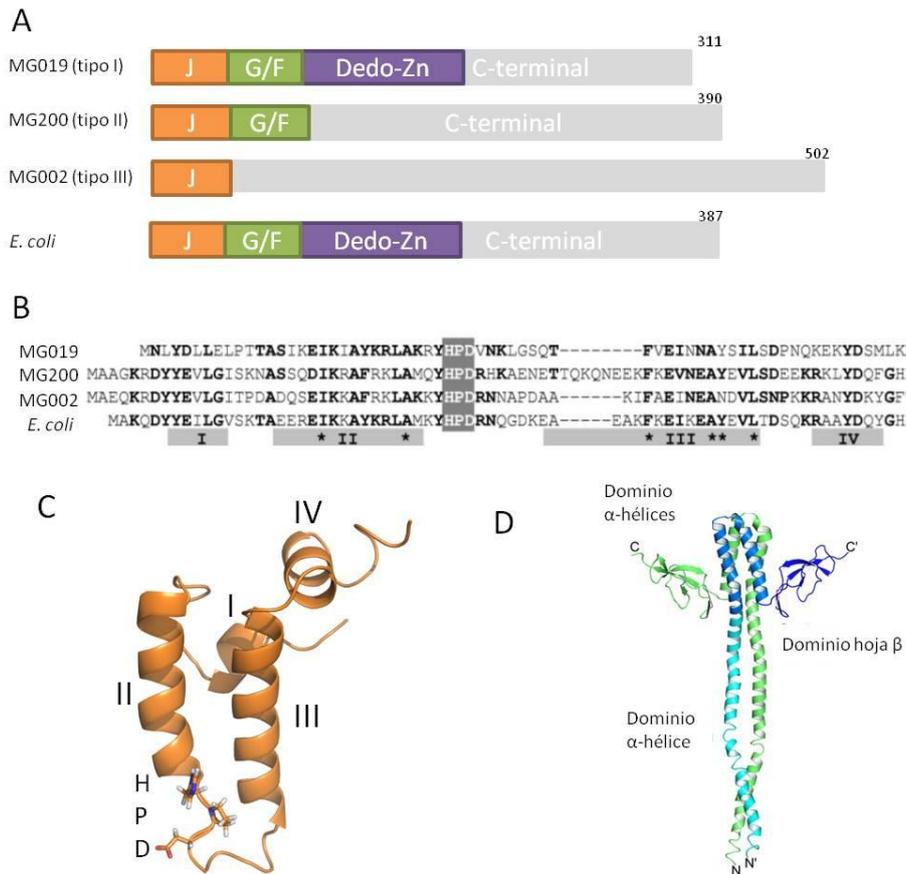


Figura 1.11 A) Clasificación de MG019, MG002 y MG200 según la presencia y organización de determinados dominios en la secuencia proteica. Adaptación de la imagen de Musatovova y colaboradores (Musatovova, Dhandayuthapani et al. 2006) B) Aliniamiento de los dominios J de las proteínas que contienen este dominio en *M. genitalium* (MG019, MG002, MG200) y el dominio J de la proteína DnaJ de *E.coli*. En la parte inferior de la imagen, los rectángulos grises indican las 4 hélices α que forman el dominio J. Imagen extraída de (Pellecchia, Szyperski et al. 1996). C) Estructura del dominio J de la DnaJ de *E.coli* (código PDB 1XBL). D) Estructura del dimero de GrpE de *T. thermophilus* (código PDB 3A6M) Adaptación de la imagen de Nakamura et al. (Nakamura, Takumi et al. 2010).

Los procesos de motilidad y adhesión en *M. genitalium* han sido relacionados con la virulencia de este microorganismo, que es un patógeno humano, pero todavía se conoce relativamente poco sobre la maquinaria implicada en este proceso, concretamente se desconoce cual es su motor. Como *M. genitalium* es considerado un modelo de célula mínima y que estudios en otros organismos revelaron la posible participación de chaperonas en sistemas de motilidad, se llevó a cabo un estudio

detallado del sistema de chaperonas de *M. genitalium* con el objetivo de determinar sus peculiaridades y su posible implicación en el proceso de motilidad celular.

Objetivos

Para determinar la posible implicación del sistema de chaperonas en la motilidad de *Mycoplasma genitalium*, y teniendo como principal objetivo la determinación estructural y funcional de la chaperona DnaK, se establecieron los siguientes objetivos concretos:

- 1.- Determinación de la estructura tri-dimensional de la chaperona DnaK de *M. genitalium* mediante cristalografía de rayos-X en su forma libre y con sustrato.
- 2.- Caracterización del mecanismo de hidrólisis de ATP por el subdominio de unión a nucleótidos de la chaperona DnaK de *M. genitalium*.
- 3.- Caracterización del complejo de DnaK con el factor intercambiador de nucleótidos GrpE de *M. genitalium*.
- 4.- Identificación y caracterización biofísica de las posibles interacciones entre DnaK y las posibles cochaperonas presentes en *M. genitalium*, MG019 y MG200, y GrpE.
- 5.- Determinación de la estructura en solución de la chaperona DnaK, del dominio NBD de DnaK, de MgGrpE y del complejo DnaK-GrpE de *M. genitalium* por SAXS.

2. Resultados y discusión

2.1 Estructura cristalina de la DnaK de *M. genitalium* (MgDnaK)

2.1.2 Clonaje, expresión y purificación de MgDnaK

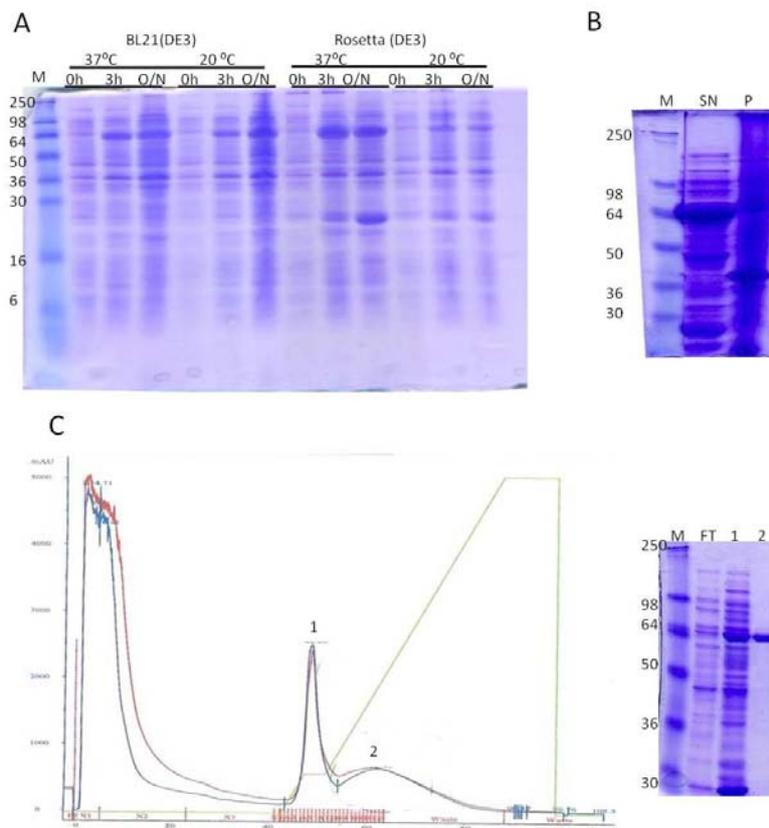
El vector que se utilizó para expresar la MgDnaK entera (código Uniprot: P47547) fue el pET21d (Novagen, Madison, WI, USA) que añade una cola de seis histidinas en el C-terminal de la proteína.

Para establecer el protocolo de expresión se realizaron varias pruebas en las que se utilizaron dos cepas bacterianas distintas, *E. coli* BL21 (DE3) y *E. coli* Rosetta2 (DE3), y se testaron dos temperaturas (20°C y 37°C) y varios tiempos de expresión una vez añadido el inductor isopropil- β -D-tiogalactósido (IPTG) a una concentración de 1 mM.

El peso molecular teórico de MgDnaK corresponde a 65 kDa y se calculó con el servidor ProtParam (Wilkins, Gasteiger et al. 1999, Bairoch, Apweiler et al. 2005). Los extractos crudos resultantes de la disrupción de las células se analizaron por gel de poliacrilamida SDS-PAGE y la banda mayoritaria correspondía a la expresión de MgDnaK en *E. coli* Rosetta2 a 37°C durante 3h. Además, la mayor parte de la proteína producida era soluble en esas condiciones (Figura 2.1 A). Una vez establecido el

protocolo de expresión se procedió a la expresión y purificación a gran escala. La purificación consistió en un primer paso de purificación por cromatografía de afinidad por metales inmovilizados. Se utilizó una columna HisTrap 5mL (GE Healthcare Life Sciences, Uppsala, Sweden) equilibrada con el tampón 0.5 M NaCl, 0.02 M Tris-HCl pH 8 y utilizando un gradiente lineal de imidazol a una concentración final de 0.5M. La proteína eluyó en dos picos distintos como se observa en el análisis de un gel de desnaturizante SDS-PAGE al 12% (v/v) acrilamida donde se cargaron las fracciones que se recogieron (Figura 2.1 C). El pico 1 contenía más contaminantes que el pico 2 y por ello se seleccionaron las fracciones correspondientes al pico 2 para posterior purificación. Estas fueron concentradas mediante un dispositivo de ultrafiltración (Amicon (Darmstadt, Germany)) con un punto de corte de 10 kDa y la solución así obtenida fue inyectada en una columna de gel filtración (Superdex 200 10/300, GE Healthcare Life Sciences, Uppsala, Sweden) previamente equilibrada con el tampón 0.5 M NaCl, 0.02 M Tris-HCl pH 8. Se analizaron las fracciones resultantes de la purificación en un gel de poliacrilamida desnaturizante al 12% (v/v) acrilamida donde se observó que la muestra aún presenta posibles contaminantes (Figura 2.1 D). El volumen de elución del pico 2 en la columna de gel filtración (Figura 2.1 D) fue de 12.9 mL que corresponde a un peso molecular aproximado de 65 kDa según la calibración de la columna con proteínas estándar con pesos moleculares conocidos, indicando que la proteína en solución está en un estado monomérico. Como aún se detectó la presencia de posibles contaminantes en la solución de *MgDnaK* después de la columna de gel filtración (Figura 2.1 D) se intentó mejorar la purificación de la proteína añadiendo un último paso de purificación a través de una columna de intercambio iónico (MonoQ, GE Healthcare Life Sciences, Uppsala, Sweden) (Figura 2.1 E). A pesar que con esta

técnica tampoco se consiguió obtener una muestra totalmente pura se realizaron ensayos de cristalización de *MgDnaK* a una concentración de 10 mg/mL utilizando solamente las fracciones más puras de proteína. De los ensayos de cristalización de *MgDnaK* hechos en ausencia o presencia de 1mM ATP y 1mM MgCl₂ no se obtuvieron cristales. El contaminante mayoritario, marcado con una flecha en la parte inferior del gel SDS-PAGE 12% (v/v) acrilamida de la Figura 2.1 D, fue analizado por la técnica de identificación de proteínas mediante su huella peptídica (del inglés *Peptide Mass Fingerprinting*, PMF) que indicó que la banda contaminante correspondía a una forma truncada de la *MgDnaK* de *M. genitalium*.



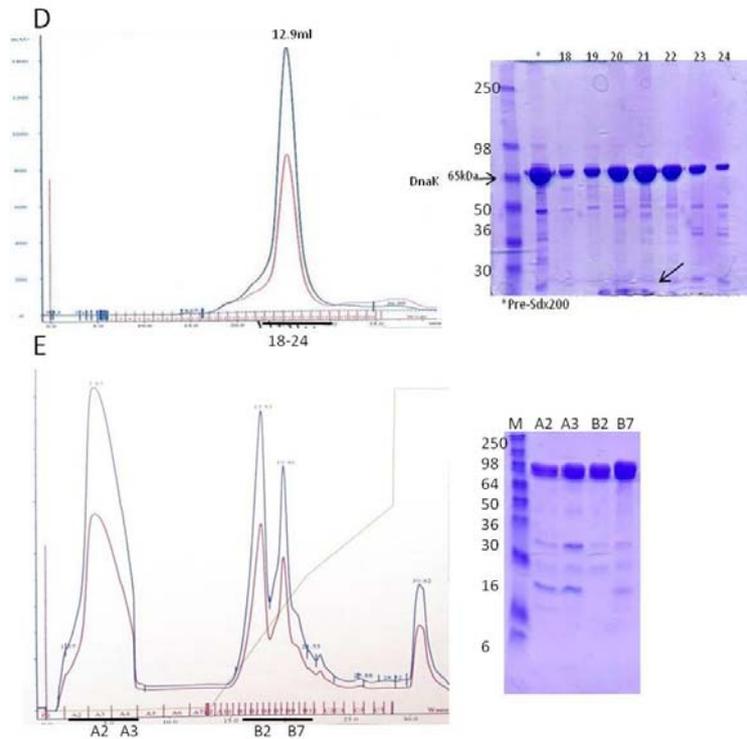


Figura 2.1 Cromatogramas y geles SDS-PAGE 12% (v/v) de acrilamida de la expresión y purificación de *MgDnaK*. A) SDS-PAGE 12% (v/v) de acrilamida de la expresión de *MgDnaK* a diferentes temperaturas utilizando dos cepas bacterianas distintas. B) SDS-PAGE 12% (v/v) de acrilamida del test de solubilidad de la condición de expresión de *MgDnaK* en *E.coli* Rosetta2 (DE3) durante 3h a 37°C. P corresponde a la fracción insoluble de las células mientras que SN corresponde al sobrenadante después de la lisis celular. C-E) Perfiles cromatográficos con su correspondiente análisis por gel de acrilamida SDS-PAGE de: C) columna de afinidad HisTrap 5mL en la que se cargó el sobrenadante de la lisis celular, D) columna de gel filtración Superdex200 10 300 (GE Healthcare, Upsala, Sweden) en la que se cargó la muestra correspondiente al pico 2 de la columna de afinidad anterior, E) columna de intercambio iónico Mono Q (GE Healthcare, Upsala, Sweden) en la que se inyectó la muestra proveniente de la columna de gel filtración anterior.

La dificultad en la obtención de proteína suficientemente pura y consecuentemente de cristales, juntamente con el conocimiento que la proteína se degrada llevó a que se decidiera intentar la cristalización de la proteína por dominios.

2.1.3 Clonaje, expresión y purificación de *MgDnaK* Δ Ct

Mediante ensayos de proteólisis limitada, en que la degradación de la proteína se da, principalmente, en segmentos expuestos y flexibles, es posible identificar los subdominios más estables y globulares de una proteína. Este tipo de estudios fueron realizados en la muestra de *MgDnaK* obtenida anteriormente y fue tratada con las proteasas: tripsina, quimotripsina y papaína. Cuando se llevó a cabo a 37°C el experimento de proteólisis limitada con papaína, tripsina y quimotripsina el resultado era una digestión prácticamente total de la proteína mientras que a 20°C la digestión con quimotripsina y tripsina era parcial permitiendo por lo tanto, identificar los subdominios más estables y globulares de la proteína. La proteólisis de *MgDnaK* con quimotripsina dio a lugar a un fragmento de aproximadamente 50 kDa (Figura 2.2). Este fragmento, indicado con una flecha en la Figura 2.2, a pesar de no haber sido identificado mediante secuenciación de Edman sugirió que correspondería al dominio de unión a nucleótido de *MgDnaK* porque se conocía que el NBD había sido cristalizado con anterioridad, que era un dominio globular y que el peso obtenido en el gel podría corresponder de manera aproximada al esperado. Se consideró entonces como un buen candidato para ensayos de cristalización (Bertelsen, Chang et al. 2009).

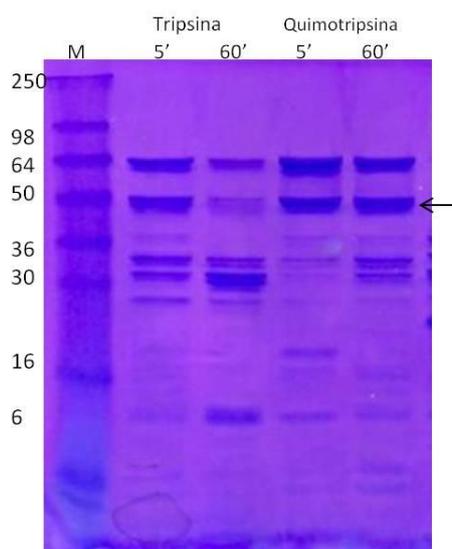


Figura 2.2 Perfil electroforético de un gel desnaturalizante SDS-PAGE 12% (v/v) acrilamida de los productos de digestión de *MgDnaK* con tripsina y quimotripsina a 20°C. El primer carril corresponde al marcador de peso molecular. El segundo y el tercero corresponden al perfil obtenido después de la proteólisis con tripsina a 5 y a 60 minutos. Los dos últimos corresponden al perfil de la proteólisis con quimotripsina a 5 y 60 minutos. Los números de la parte superior del gel indican el tiempo de la reacción de digestión. La flecha indica el fragmento que podría corresponder al dominio *MgNBD*.

Se realizó, además, una búsqueda en el PDB de todas las estructuras de DnaK que se habían resuelto mediante difracción de rayos-X que sirvió, juntamente con los resultados de proteólisis limitada, de base para diseñar nuevos constructos. La DnaK más larga que se ha cristalizado hasta el momento es la de *Geobacillus kaustophilus* (*GkDnaK*) (código PDB: 2V7Y) con 509 residuos (Chang, Sun et al. 2008) y que tiene una identidad del 62% con la *MgDnaK* según lo calculado con el programa Lalign (Gasteiger, Gattiker et al. 2003). Se utilizó el programa ClustalW (Thompson, Higgins et al. 1994) para realizar un alineamiento de las secuencias de las dos proteínas (Figura 2.3). Se decidió preparar el clon en el que se eliminarían los últimos 74 aminoácidos de la *MgDnaK* entera, basándonos en la estructura resuelta de la *GkDnaK*. A este clon se le denominó *MgDnaKΔCt*.

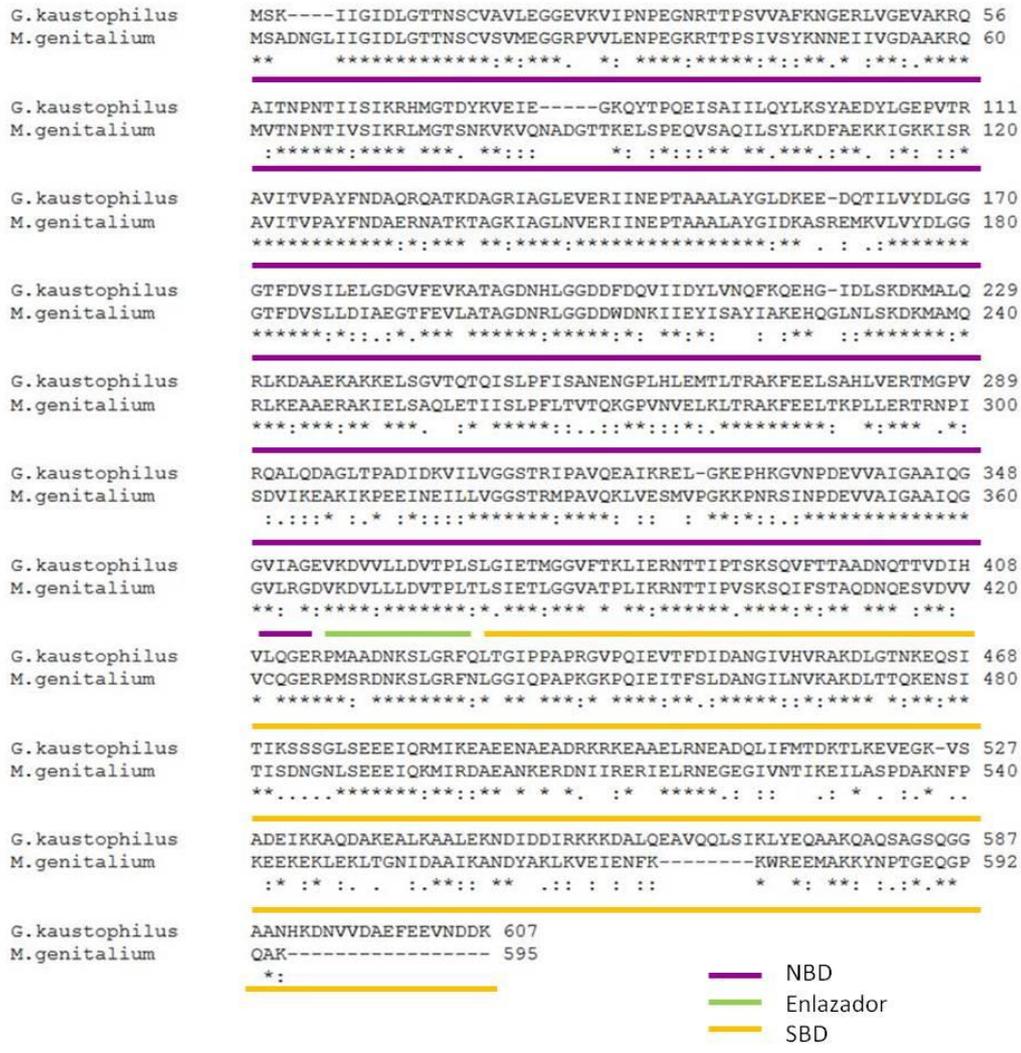


Figura 2.3 Alineamiento de la secuencia entera de DnaK *G. kaustophilus* y *M. genitalium*. Las líneas de colores indican los dominios de la DnaK: NBD (morado), SBD (verde) y el enlazador (naranja). El * (asterisco) indica las posiciones que contienen un único residuo completamente conservado. Los: (dos puntos) indican la conservación entre residuos con propiedades similares (valor>0.5) calculadas con la matriz Gonnet PAM 250. El . (punto) indica la conservación entre residuos con unas propiedades poco conservadas (valor<0.5) calculadas con la matriz Gonnet PAM 250.

Los cebadores que se usaron para el clonaje fueron:

Tabla 2.1 Cebadores utilizados para realizar el nuevo constructo de *MgDnaK*

Secuencia	
Nombre	
DnaKΔCt F	5' ATATACATATGAGTGCAGACAATGGTTT3'
DnaKΔCt R	3'TACGTAATGAAGGGGAACCTCGAGCAC 5'

El fragmento de *MgDnaK* amplificado con los cebadores descritos en la Tabla 2.1 se clonó entre los sitios de restricción *NdeI* y *XhoI* del vector de expresión pET21d (Novagen, Madison, WI, USA). El vector resultante fue transformado en células *E. coli* BL21 (DE3). Los ensayos de expresión realizados indicaron que la mejor condición de expresión es inducir el cultivo celular cuando la $OD_{600} \sim 0.6-0.8$ con 1mM IPTG a una temperatura de 20°C y durante toda la noche (Figura 2.4 A).

La purificación de *MgDnaK* Δ Ct se realizó mediante dos pasos como en el caso de la *MgDnaK*. En un primer paso el sobrenadante proveniente de la lisis celular se cargó en una columna HisTrap™ HP de 5 mL (GE Healthcare Life Sciences, Uppsala, Sweden), equilibrada en 0.5 M NaCl, 0.02 M Tris-HCl pH 8, y la elución se realizó mediante la implementación de dos escalones en el gradiente de imidazol; el primero a un porcentaje del 8%, donde la mayor parte de las proteínas contaminantes fueron eluidas y el segundo a un porcentaje del 18%, donde eluyó la *MgDnaK* Δ Ct según el análisis de las fracciones por SDS-PAGE (Figura 2.4 B). En un segundo paso las fracciones más enriquecidas en *MgDnaK* Δ Ct resultantes de la elución de la columna de afinidad se cargaron en una columna de filtración en gel (Superdex 200 16/60, GE Healthcare Life Sciences, Uppsala, Sweden) equilibrada con 0.5 M NaCl, 0.02 M Tris-HCl pH 8 (Figura 2.4 C y D). El volumen de elución de *MgDnaK* Δ Ct en la Superdex 200 16/60 es de 80 mL que corresponde, según la calibración de la columna con proteínas estándar de pesos moleculares conocidos, a un monómero en solución.

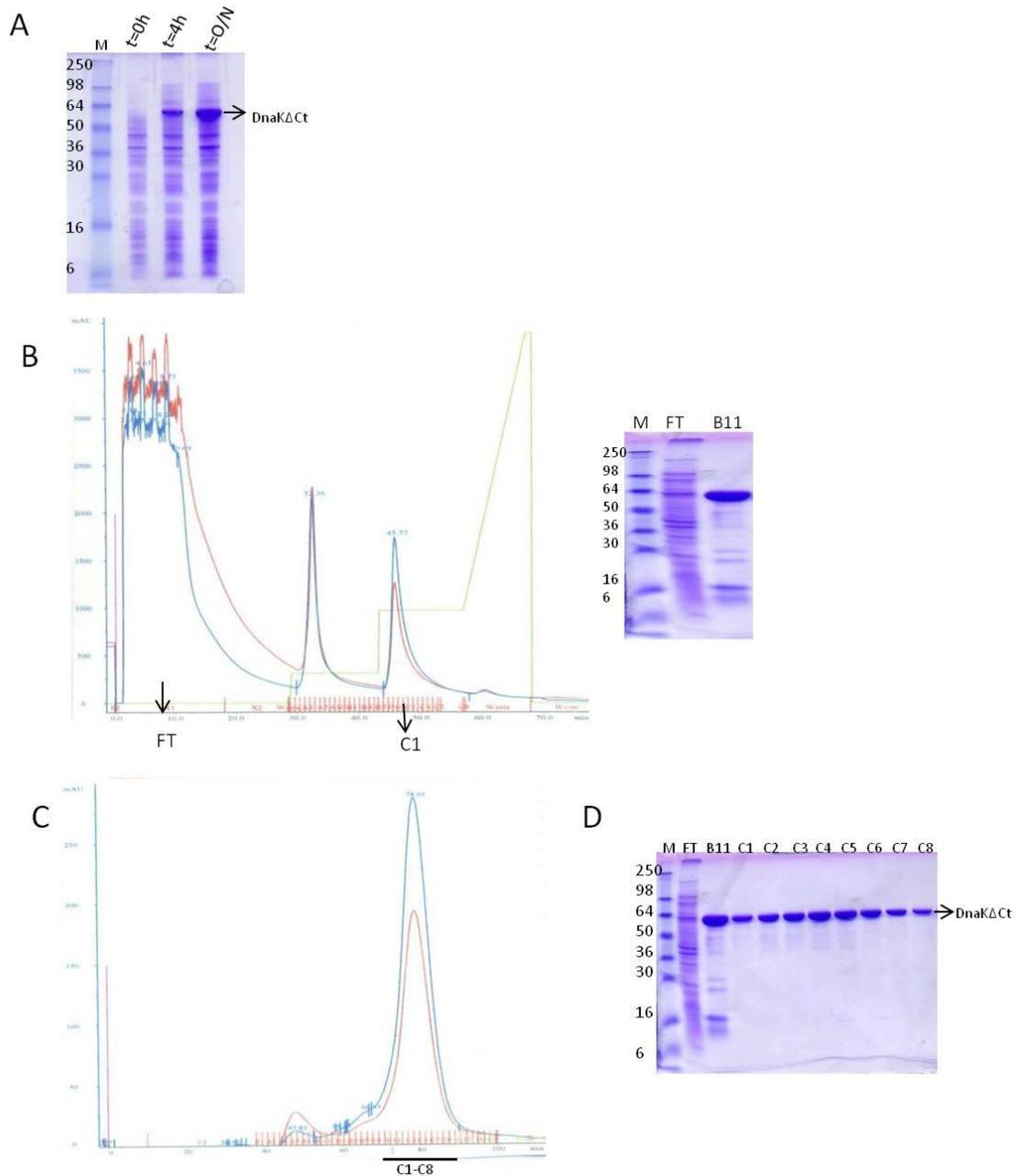


Figura 2.4 Cromatogramas y geles SDS-PAGE 12% (v/v) de acrilamida de la expresión y purificación de *MgDnaK* A) Perfil del gel de poliacrilamida de las pruebas de expresión de *MgDnaKΔCt*. Se muestra la expresión a 3 tiempos distintos: basal, a las 4h y durante toda la noche. B) Perfil de la cromatografía de afinidad. C) Perfil de la cromatografía de Ni. D) SDS-PAGE de la muestra eluida de la Superdex 200 16/60. En ambas cromatografías el eje de las X corresponde a mAU mientras que el de las Y al volumen de elución.

Las fracciones más enriquecidas en *MgDnaKΔCt* resultantes de la elución en Superdex 200 16/60 se concentraron con un dispositivo de ultrafiltración (Amicon Darmstadt,

Germany) hasta 10 mg/mL y a continuación la proteína fue incubada con 1 mM ATP y 1 mM MgCl₂ durante una hora. Seguidamente se iniciaron los cribados de cristalización a escala nanoescala.

2.1.4 Cristalización de *MgDnaKΔCt*

Los ensayos de cristalización por la técnica de difusión de vapor de *MgDnaKΔCt* se realizaron tanto a 20°C como a 4°C. Los primeros cristales que se obtuvieron tenían forma de placas (Figura 2.5 A) y aparecieron en presencia de ATP en 0.1M citrato de sodio pH 5, 5% (v/v) butanol y 28.75% (w/v) PEG MME5000 (condición del cribado de cristalización PAC23 de la Plataforma Automatizada de Cristalización del PCB-IRB-CSIC). Se optimizaron las condiciones de cristalización variando el pH y las concentraciones de sal y de precipitante alrededor de las condiciones donde habían aparecido los primeros cristales para obtener más cristales y de mayor tamaño. Estas optimizaciones ya no se realizaron en nanogota sino en microgota en placas de 24 pocillos y se obtuvieron unos cristales en forma de prisma (Figura 2.5 B y C). Una vez obtenidos cristales de *MgDnaKΔCt* con ATP, el cual se esperaba que fuera hidrolizado en el cristal, se hicieron ensayos de co-cristalización con 1 mM AMP-PNP (análogo de ATP no hidrolizable), para obtener la estructura de *MgDnaK* en dos estados diferentes. Estas co-cristalizaciones dieron como resultado unos pequeños cristales en forma de prisma en la condición obtenida anteriormente, que también se optimizaron en placa de 24 pocillos (Figura 2.5 D).

Para proteger a los cristales de los daños producidos por la radiación cuando estos son expuestos al haz de rayos X, se utilizó una solución crioprotectora, compuesta por los mismos componentes que la solución de cristalización pero añadiendo un 20% (v/v) de glicerol, en la que los cristales se sumergieron y posteriormente fueron congelados rápidamente en nitrógeno líquido.

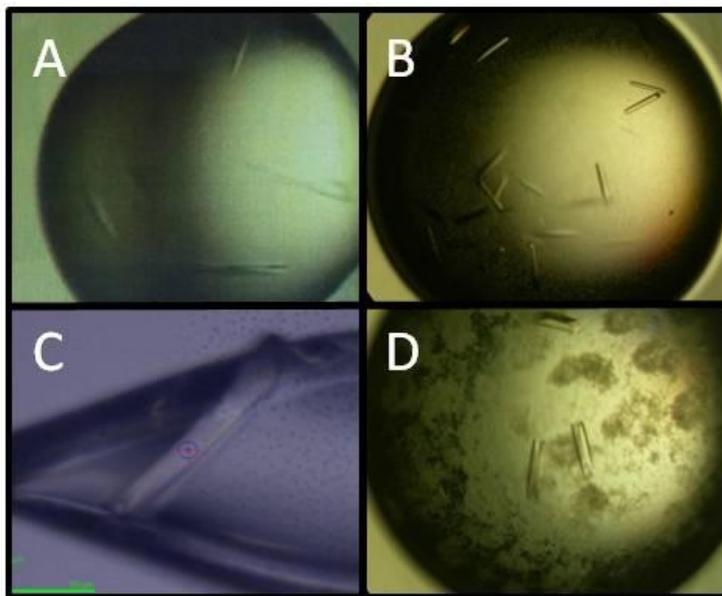


Figura 2.5 Cristales de *MgDnaKΔCt*. A-B) Cristales de *MgDnaKΔCt* en presencia de ATP. C) Cristal de *MgDnaKΔCt* en presencia de ATP montado en la línea de luz ID23eh1 del ESRF en un loop de 0.2 μm . D) Cristales de *MgDnaKΔCt* en presencia AMP-PNP.

2.1.5 Determinación de la Estructura de rayos-X de *MgDnaKΔCt*

Los cristales de *MgDnaKΔCt* tanto en presencia de AMP-PNP como en presencia de ATP se difractaron en la línea de luz ID23eh1 del Laboratorio Europeo de Radiación Síncrotrón (del inglés *European Synchrotron Radiation Facility*, ESRF, Grenoble, Francia), hasta 1.8 Å y 2.4 Å de resolución respectivamente. Los datos recogidos fueron indexados como $P4_32_12$ con parámetros de celdilla $a=b=133$ Å y $c=64$ Å. El análisis de los coeficientes de Mathews indicó que en ambos casos *MgDnaKΔCt* contiene una molécula por unidad asimétrica y que los cristales tienen con un contenido de solvente medio del 56%. Los datos se procesaron con el programa XDS (Kabsch 2010) y se escalaron con SCALA (Evans 2006).

El problema de las fases del cristal de *MgDnaKΔCt* en presencia de AMP-PNP se solventó por Reemplazo Molecular (del inglés *Molecular Replacement*, MR) (Rossmann 1990) que es un método para determinar las fases que usa la información de estructuras conocidas que están relacionadas o son homólogas a los componentes del cristal. Por ese motivo, no se requieren experimentos o datos adicionales, y, además, simplifica la construcción de modelos. El método de MR suele ser el método de elección para la determinación de estructuras por difracción de rayos-X cuando un modelo de búsqueda adecuado está disponible.

El MR se realizó con el programa MolRep del paquete CCP4 (Vagin and Teplyakov 2010). Primero se utilizó como modelo de búsqueda la versión de *GkDnaK* (código PDB: 2V7Y) a la que le falta el C-terminal (Chang, Sun et al. 2008) pero con los parámetros utilizados no se encontró ninguna solución. Se lanzó otra vez el programa pero utilizando como modelos de búsqueda los dominios NBD y SBD de la *GkDnaK*. Los valores estadísticos $tfz = 8.2$ obtenidos entonces fueron prometedores indicando que el programa había calculado una solución. Con el programa Coot (Emsley and Cowtan 2004) y Refmac5 se hicieron varios ciclos de construcción y refinamiento del modelo (Murshudov, Vagin et al. 1997). El modelo final consiste en 509 aminoácidos (desde la Asn4 hasta la His524, incluyendo las histidinas de la cola C-terminal que provienen del vector de expresión y que son utilizadas para la purificación de la proteína), no se pudieron trazar por falta de densidad electrónica 6 aminoácidos que se encuentran entre la Ser483 y el Gln490 muy cerca del enlazador. El modelo contiene además 522 moléculas de agua, una molécula de AMP-PNP y un ion magnesio. La estructura se consideró finalizada con un R_{factor} de 18.1% y un R_{free} de 22.9%. Todos los datos referentes al procesamiento y refinamiento de la estructura se encuentran en la Tabla 2.2.

En el caso de *MgDnaKΔCt* en presencia de ATP el problema de las fases se solventó también mediante MR usando como modelo de búsqueda la *MgDnaKΔCt* de *M. genitalium* en presencia de AMP-PNP y se utilizó el programa Phaser (Adams, Afonine et al. 2010). Después de varios ciclos de construcción y refinamiento respectivamente por Coot y Refmac se obtuvo la estructura final de *MgDnaKΔt-ATP* con un R_{factor} de 22.6% y un R_{free} de 26.2% (Tabla 2.2). El modelo de *MgDnaKΔCt-ATP* incluye del aminoácido Asn4 hasta His524, una molécula de ADP y un fosfato inorgánico. Cuando

se modeló el ADP se confirmó que el ATP usado para la co-cristalización fue hidrolizado dando a lugar a una molécula de ADP y un fosfato inorgánico que también fue modelado. Después del análisis de los datos y de la inspección de los mapas se comprobó que estos últimos no eran suficientemente claros en algunas regiones indicando que el cristal difractaba de forma ligeramente anisotrópica. La difracción anisotrópica se da cuando existe una dependencia direccional en la calidad de la difracción. Al ser un problema muy sutil no fue necesario corregirlo pero posiblemente debido ello y a la resolución (2.5-3Å) no fue posible modelar moléculas de agua ni otros posibles componentes de cristal como sería el ion de magnesio que si se modeló en la estructura de *MgDnaKΔCt*:AMP-PNP. Los datos estadísticos de la recogida de datos y del refinamiento de ambas estructuras se muestran en la Tabla 2.2.

El servidor MolProbity (Chen, Arendall et al. 2010) fue utilizado para evaluar los modelos finales de *MgDnaKΔCt*:AMP-PNP y *MgDnaKΔCt*:ATP. En la estructura de *MgDnaKΔCt*:AMP-PNP el 98.6% de los residuos se encuentran en zonas permitidas del diagrama de Ramachandran mientras que en la estructura de *MgDnaKΔCt*:ATP hay 95.2% residuos en estas zonas. Los valores del diagrama de Ramachandran de las zonas moderadamente permitidas y de las no permitidas se detallan en ambos casos en la Tabla 2.2

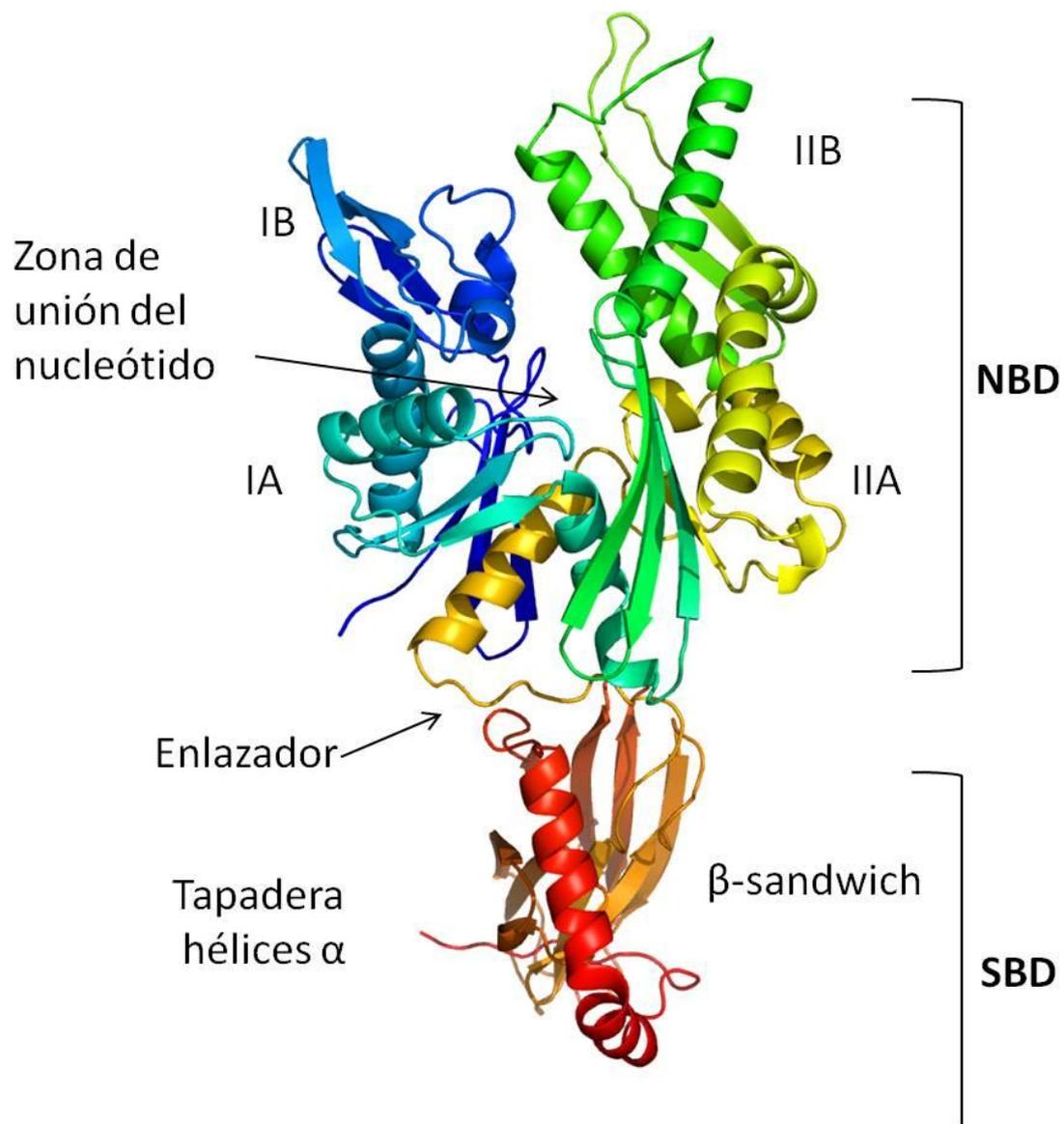
Tabla 2.2 Datos estadísticos de la recogida de datos de y del refinamiento de las estructuras *MgDnaKΔCt:AMP-PNP* y *MgDnaKΔCt:ATP*. Entre paréntesis se muestran los resultados de la capa de mayor resolución.

Conjunto de datos	<i>MgDnaKΔCt:AMP-PNP</i>	<i>MgDnaKΔCt:ATP</i>
Fuente rayos-X	ESRF ID23eh1	ESRF ID23eh1
Longitud de onda(Å)	0.9763	0.9763
Grupo espacial	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Dimensiones celdilla:		
a,b,c (Å)	a=b= 133.71 c= 64.07	a=b= 133.55 c= 64.52
α,β,γ(°)	α=β=γ=90	α=β=γ=90
Resolución (Å)	47.31-2.00(2.12-2.00)	47.22-2.49 (2.63-2.49)
R _{merge} (%)	0.26 (0.57)	0.164 (1.196)
I/σ(I)	5.6 (4)	0.172(1.251)
Complejitud (%)	99.6 (97.3)	99.7 (98)
Nº reflexiones únicas	41211(5787)	20864 (2917)
Multiplicidad	13.1	11.5 (11.6)
Wilson B-factor (Å ²)	21	32
R _{pim}	0.07(0.16)	0.05 (0.36)
Refinamiento		
R _{factor} /R _{free} (%)	18.1/22.9	22.6/26.2
Rmsd		
• Longitud enlace (Å)	0.012	0.008
• Ángulo (°)	1.523	1.284
B-valores(Å ²)		
• Proteína	18.32	40.5
• Péptido	32.3	20.7
• Aguas	18.0	-----
Ramachandran plot		
• Más favorables(%)	98.64	95.2%
• Permitidas(%)	100	98.8%
• No permitidas(%)	0	1.15%

2.1.6 Descripción de la estructura cristalina de *MgDnaK* Δ Ct

La estructura de *MgDnaK* Δ Ct se organiza en dos dominios claramente diferenciados: el dominio de unión a nucleótido (NBD) y el dominio de unión a sustrato (SBD) unidos por una región constituida por aminoácidos hidrofóbicos, denominada como enlazador.

Tanto el NBD como el SBD están divididos en subdominios (Figura 2.6).



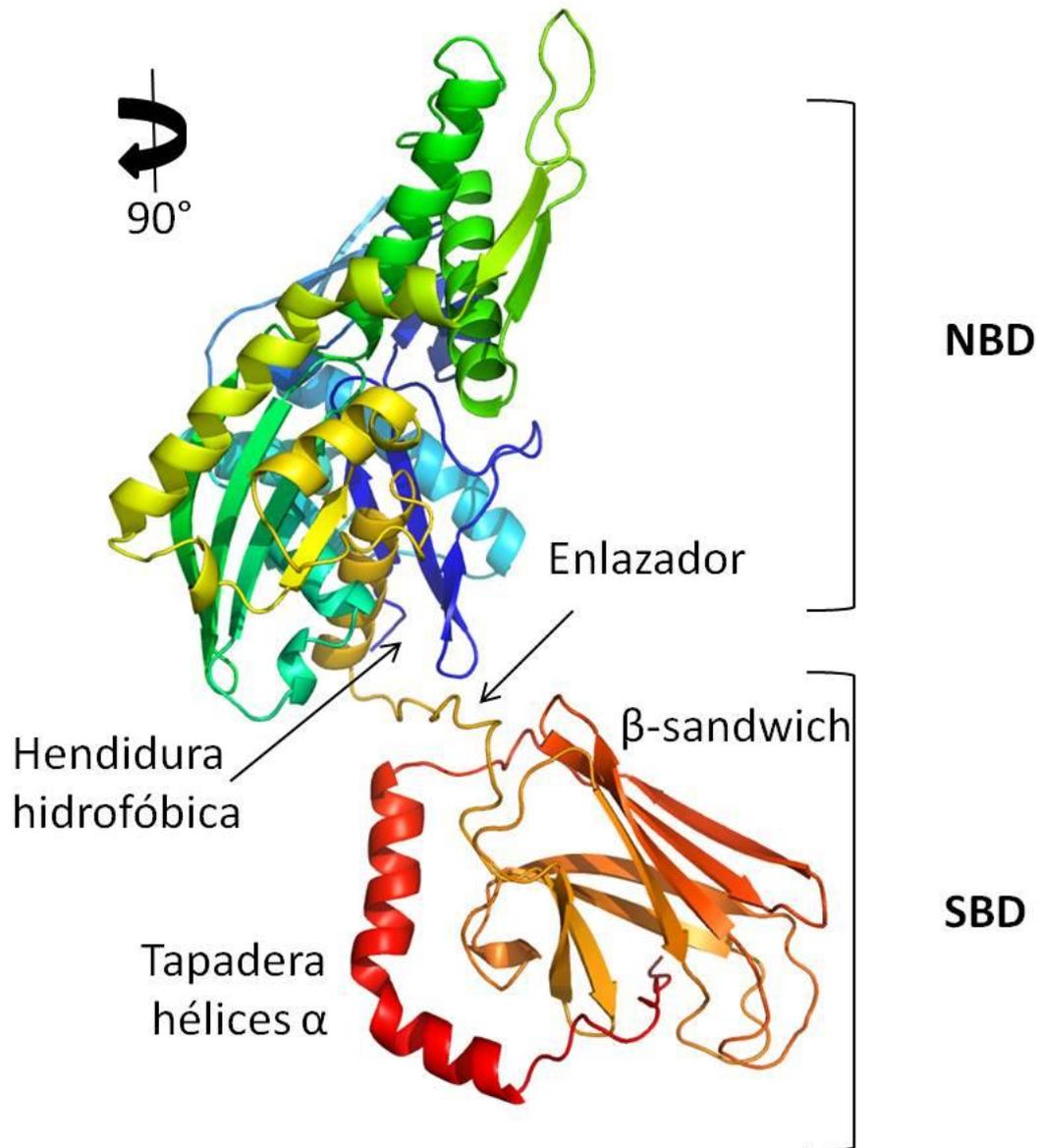


Figura 2.6 Estructura general de *MgDnaK Δ Ct*. La imagen de la izquierda corresponde a una vista frontal de la proteína y la de la derecha corresponde a una rotación de 90° de la primera imagen.

- Dominio de unión a nucleótido (NBD) de *MgDnaKΔCt*

El NDB, compuesto por 365 aminoácidos, está formado por dos lóbulos, I y II, que a su vez se dividen en dos subdominios: IA y IB, y IIA y IIB. El nucleótido se une en la forma de V que forman los subdominios IA-IB y IIA-IIB, los cuales están involucrados en la coordinación del nucleótido (Figura 2.6). El NBD presenta un plegamiento típico de las proteínas de la familia de la actina (Dominguez 2004) y el análisis estructural detallado del dominio se describe en la sección 2.2.

- Dominio de unión a sustrato (SBD) de *MgDnaKΔCt*

El SDB está formado por tres subdominios que han sido descritos en otros trabajos (Zhu, Zhao et al. 1996, Bertelsen, Chang et al. 2009) : un subdominio β -sándwich que contiene un bolsillo hidrofóbico en el que interactuaran los péptidos aún sin plegar, un subdominio tapadera compuesto por hélices α y un subdominio C-terminal desestructurado. El constructo de *MgDnaKΔCt* de *M. genitalium* usado para los ensayos de cristalización está formado por 144 aminoácidos que corresponden al subdominio β -sándwich y a las primeras dos hélices (hélice A y hélice B) de la tapadera que está formada por hélices α (Figura 2.7 A y B). El subdominio β -sándwich está formado por un par de láminas β antiparalelas y enfrentadas, cada una de las cuales contiene cuatro hebras β antiparalelas. La lámina β de la parte inferior, formada por las hebras $\beta 3$ - $\beta 6$ - $\beta 7$ - $\beta 8$ realiza una torsión de la lámina β con un giro de la horquilla que conecta las hebras. La lámina β superior, formada por las hebras $\beta 2$ - $\beta 1$ - $\beta 4$ - $\beta 5$, es irregular y se relaciona con los bucles que sobresalen de la estructura y que ayudan a encajar el sustrato en su lugar de unión (Figura 2.7 A). Este subdominio está conectado al subdominio tapadera mediante un pequeño enlazador. A su vez, el subdominio

tapadera, que normalmente consiste en 5 hélices α que forman una cavidad con un núcleo hidrofóbico estable (Figura 2.7 B y 2.8), no se observa entero en el caso del constructo usado para cristalizar la *MgDnaK Δ Ct*, que contiene solamente la hélice A y la hélice B. Estas dos hélices tienen la función de tapar el subdominio β -sándwich. La hélice B abarca todo el subdominio β -sándwich y estabiliza el sustrato en el bolsillo de unión, como se observa en la estructura de DnaK de *E. coli* (*EcDnaK*) (código PDB 1DKZ) (Zhu, Zhao et al. 1996). La *MgDnaK Δ Ct* a pesar de sólo tener una parte de la hélice B también actúa como tapadera (Figura 2.8 A) pero la interacción con el bucle formado entre las hebras $\beta 2$ y $\beta 3$ del β -sándwich no se produce por puentes salinos como en el caso de la estructura de *EcDnaK* (código PDB 1DKZ) (Zhu, Zhao et al. 1996), sino que la interacción y estabilización de esta hélice es consecuencia de la interacción de la cola de histidinas, que fue añadida por el clonaje, en el C-terminal de la *MgDnaK Δ Ct* con el SBD. En este caso la cola de histidinas está mimetizando un posible sustrato que interacciona con el bolsillo hidrofóbico del SBD (Figura 2.7 Figura 2.8 A B), que está formado por las cadenas laterales de los residuos Ile382, Phe407, Val417, Val419, Ile453, Ile455 y está ocupado por la Leu522, aminoácido añadido como consecuencia de la clonación de la *MgDnaK Δ Ct* en pET21d.

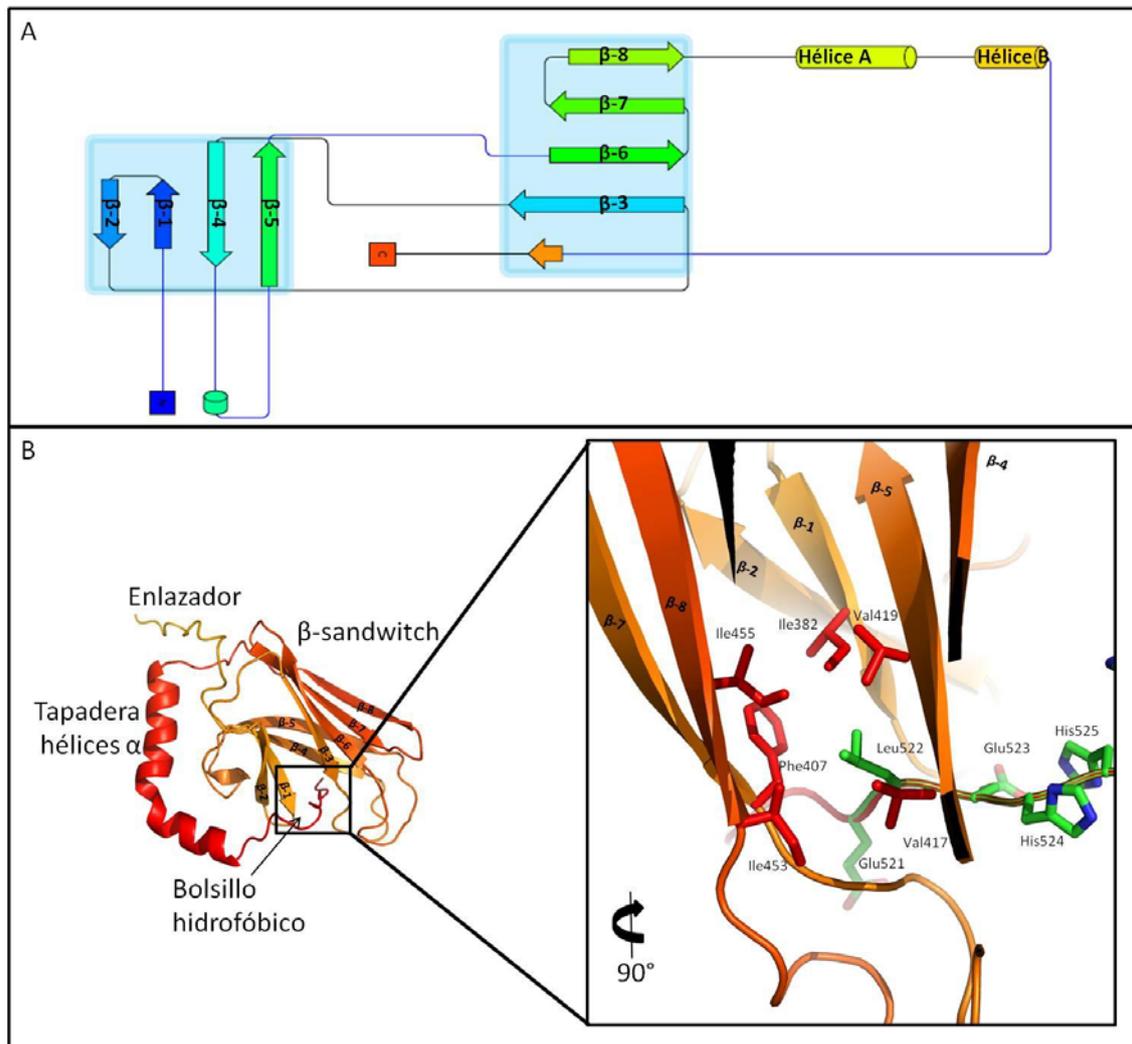


Figura 2.7 Representación del dominio SBD A) Topología del SBD generado por el servidor Pro-origami (Stivala, Wybrow et al. 2011) donde se detallan la disposición de las láminas β del β -sándwich y de las hélices α de la tapadera. B) Representación del SBD donde se muestra en detalle el bolsillo hidrofóbico de unión a sustrato. En particular, se detalla la interacción entre los residuos que forman el bolsillo hidrofóbico (rojo), y la Leu522 que mimetiza un posible sustrato (verde).

Hasta el momento se han descrito dos formas alternativas de cómo un péptido desplegado se puede unir al SBD, en función de la orientación con la que el péptido interactúa con el SBD: i) hacia delante o ii) hacia atrás (Zhu, Zhao et al. 1996, Zahn, Berthold et al. 2013). La comparación entre las estructuras que presentan un modo de unión hacia delante (códigos PDB: 1DKZ, 4EZW, 4EZX, 3DPP, 3DPO, 4HY9, 4HYB, 4EZN, 3QNJ, 4EZO, 4EZP, 4E81 y 4F00) y hacia atrás (códigos PDB: 4EZQ, 4EZR, 4EZT, 4E22,

4EZY y 4F01)) permitieron determinar que la *MgDnaK* Δ Ct une el péptido hacia delante (Figura 2.8 B y C) y no hacia atrás como se muestra en la Figura 2.8 D y E.

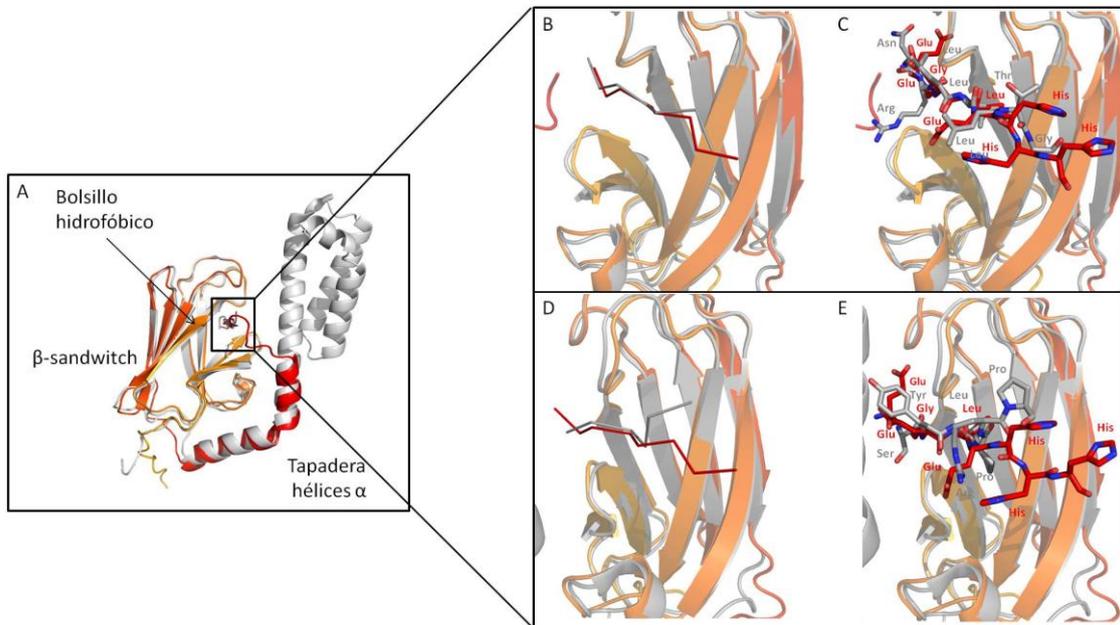


Figura 2.8 A) Representación del SBD donde se aprecia la tapadera de hélices α (rojo), y el β -sándwich (naranja). Superposición de la estructura del SBD de *E. coli* vs SBD *M.genitalium* (código PDB 1DKZ). B-C) Representación de la unión hacia delante del péptido LEHHHHHHH al bolsillo de unión a sustrato, en cintas (B) y en bastones (C). Superpuesto a esta estructura está representado en gris el SBD de *E. coli* (código PDB 1DKZ) con el péptido NRLLLTG ocupando el bolsillo de unión a sustrato. D-E) Representación de la unión hacia atrás del péptido SYLPRP al bolsillo de unión a sustrato del SBD de *E.coli*, en cintas (B) y en bastones (C) en gris (código PDB 4EZN). Superpuesto a esta estructura está representado en gris el SBD de *M.genitalium* con el péptido LEHHHHHHH interaccionando en el bolsillo de unión a sustrato.

- Enlazador

La parte principal del enlazador está formado por ocho residuos, normalmente hidrofóbicos, que tienen la siguiente secuencia consenso D/E-V/I/L-L-L-D-V-*P (Figura 2.9) (Vogel, Mayer et al. 2006). Esta secuencia está muy conservada entre especies y mutaciones en sus residuos, principalmente en los hidrofóbicos, derivan en una pérdida de la función chaperona (Laufen, Mayer et al. 1999, Han and Christen 2001).

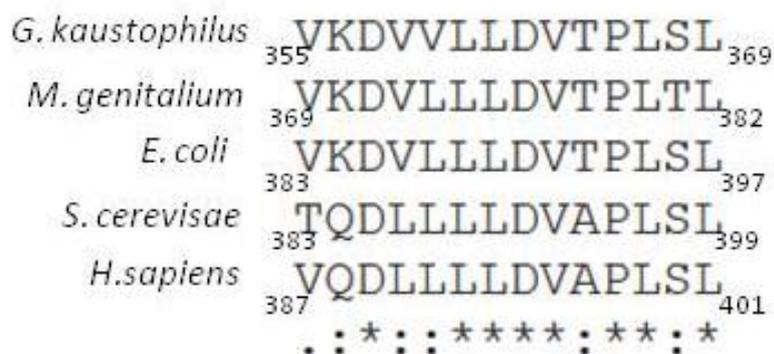


Figura 2.9 Alineamiento múltiple de secuencias que corresponden al enlazador de DnaKs de diferentes especies.

La comparación de la estructura de la *MgDnaKΔCt* con la *GkDnaK* (código PDB: 2V7Y) tanto en presencia de ATP como de AMP-PNP, permitió observar que el modelo de la *MgDnaKΔCt* presenta una conformación más contraída en comparación con *GkDnaK* (Figura 2.10 A). Igual que sucede con la comparación estructural de *MgDnaKΔCt* en presencia de ATP con la *MgDnaKΔCt* en presencia de AMP-PNP.

El enlazador presenta diferentes conformaciones según el nucleótido y/o sustrato que estén unidos a la proteína: se encuentra expuesto al solvente cuando adopta una conformación extendida en presencia de ADP y sustrato (conformación cerrada)(Figura 2.10 B), o se encuentra menos expuesto al solvente cuando adopta una conformación contraído en presencia de ATP y ausencia de sustrato (Buchberger, Valencia et al. 1994, Swain, Dinler et al. 2007, Bertelsen, Chang et al. 2009, Zhuravleva and Gierasch 2011, Zhuravleva, Clerico et al. 2012) acoplándose en este caso a la hendidura hidrofóbica entre los subdominios IA-IIA (conformación abierta)(Bhattacharya, 2009) (Figura 2.10 C).

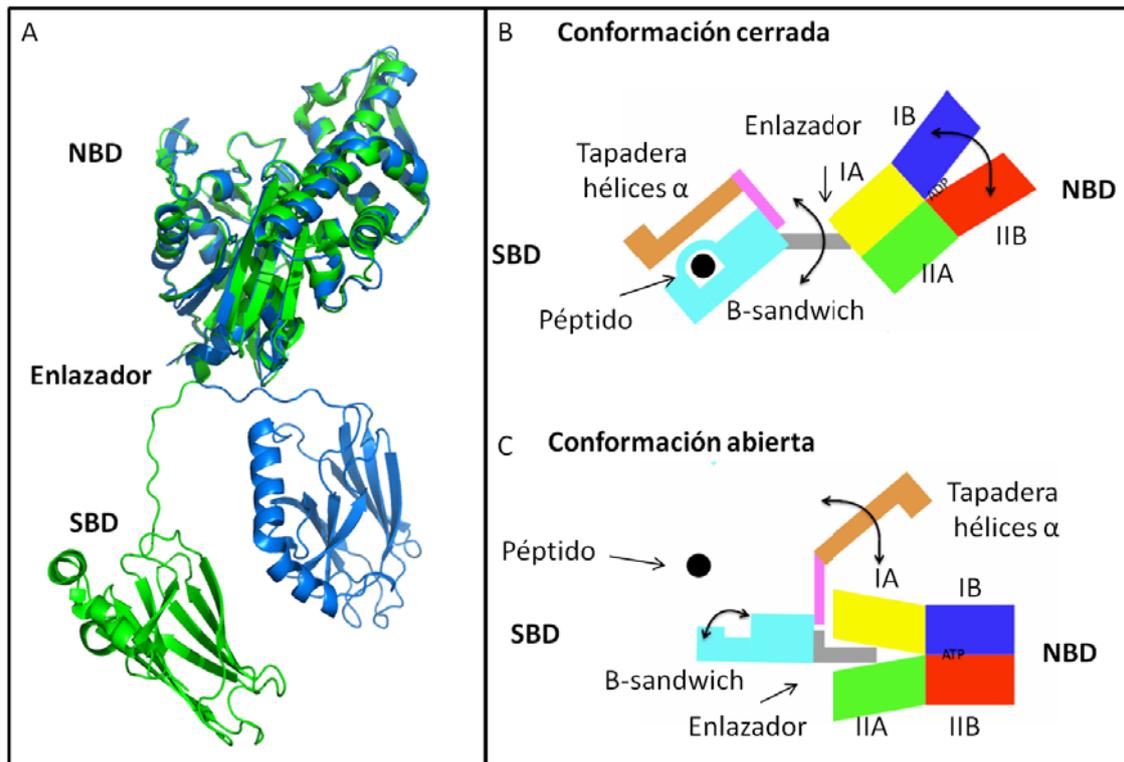


Figura 2.10 A) Comparación estructural de *MgDnaKΔCt*-ATP (azul) vs *GkdnaK* (gris). Se observa la diferencia en la disposición del enlazador de ambas estructuras. C-D) Representación esquemática de la conformación abierta y cerrada que presenta la estructura de DnaK. Los lóbulos del NBD están representados por los colores azul, rojo, amarillo y verde. El enlazador se representa en gris. La tapadera de hélices α se representa en rosa y en marrón, esta última está ausente en la estructura de *MgDnaKΔCt*. El péptido se representa por un círculo negro. C) Conformación cerrada. D) Conformación abierta. Figura adaptada de Zuiderweg y colaboradores (Zuiderweg, Bertelsen et al. 2013).

- Comparación de las estructuras *MgDnaKΔCt*:AMP-PNP vs

MgDnaKΔCt:ATP

La superposición estructural realizada con Coot (Emsley and Cowtan 2004) entre la *MgDnaKΔCt* en presencia de AMP-PNP y de ATP hidrolizado resulta en una media cuadrática (del inglés *root mean square deviation, r.m.s.d*) de 0.32 Å, siendo mínimas las variaciones entre las estructuras. Las variaciones se deben a cambios de conformación de las cadenas laterales de algunos aminoácidos expuestos no existiendo cambios significativos ni en el centro activo del NBD ni en el del SBD. La disposición de los nucleótidos es la misma en ambos casos, pero en el caso en el que se

añadió ATP se observó que la molécula fue hidrolizada dando lugar a una molécula de ADP y un fosfato inorgánico (Figura 2.11 A). La resolución de la estructura de *MgDnaKΔCt* en presencia de ATP no permitió modelar la coordinación del ion magnesio que es necesario para la hidrólisis de ATP y que aparece muy claramente en la estructura de la *MgDnaKΔCt* en presencia de AMP-PNP, que será discutida más detalladamente en la sección 2.2.3 (Figura 2.11 B).

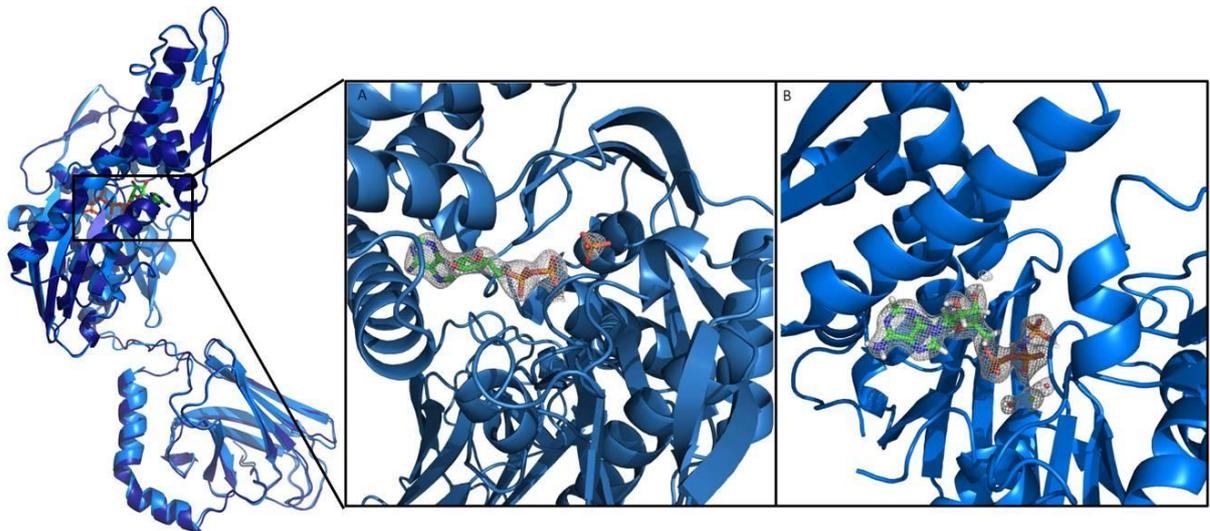


Figura 2.11 Superposición de *MgDnaKΔCt* : ATP y *MgDnaKΔCt* : AMP-PNP. Representación del centro activo del NBD de *MgDnaKΔCt* en presencia de ATP hidrolizado (A Azul marino) y de AMP-PNP (B, azul). En cada imagen se muestra la densidad electrónica correspondiente a cada nucleótido, en el primer caso ATP hidrolizado y en el segundo AMP-PNP.

- Comparación de las estructuras *MgDnaKΔCt*:ATP vs *GkDnaK*:ADP

En las estructuras de la *MgDnaKΔCt*:ATP y en la *GkDnaK*:ADP hay solamente una molécula de proteína por unidad asimétrica. La principal diferencia entre estas dos estructuras se encuentra en la disposición del enlazador que está más extendido en la

estructura de *G. kaustophilus* que en la de *M. genitalium*, resultando esta última en una rotación del SBD en relación al NDB (Figura 2.12 A). Para comparar los dos dominios de DnaK se compararon primero los NBD y luego los SBD, por separado. El r.m.s.d. obtenido de la superposición de los NBDs es de 0.84 Å con una identidad de secuencia del 61%, mientras que es de 1.01 Å con una identidad de secuencia del 67% cuando se superponen los SBDs. La mayor diferencia se encuentra en el bucle expuesto formado entre los aminoácidos Val82 a Glu92 de la estructura de *M. genitalium* que es mayor que el presente en *G. kaustophilus* donde está formado por los aminoácidos Ile79 a Gln82.

Tanto en la estructura de *G. kaustophilus* como en la de *M. genitalium* el SBD está unido a un péptido desplegado. En el primer caso el SBD está ocupado por el enlazador de una molécula simétrica (Figura 2.12 B) mientras que en el segundo está ocupado por la cola de histidinas (Figura 2.8). Esto explicaría que haya una disposición extendida del enlazador en la *GkDnaK:ADP* y más contraída en la *MgDnaKΔCt:ATP*. Otras diferencias entre las dos estructuras se encuentran en la disposición de las cadenas laterales de los aminoácidos altamente conservados cercanos al lugar de unión del sustrato (Asp418, Ile453, Leu440, Thr409, Gln414, Gln405 y Arg437 en la *MgDnaKΔCt:ATP* y Asp406, Ile441, Gln427, Thr397, Gln402, Gln 405, Arg427 en la *GkDnaK:ADP*, y son debidas a la ocupación del SBD por parte del péptido que mimetiza el sustrato.

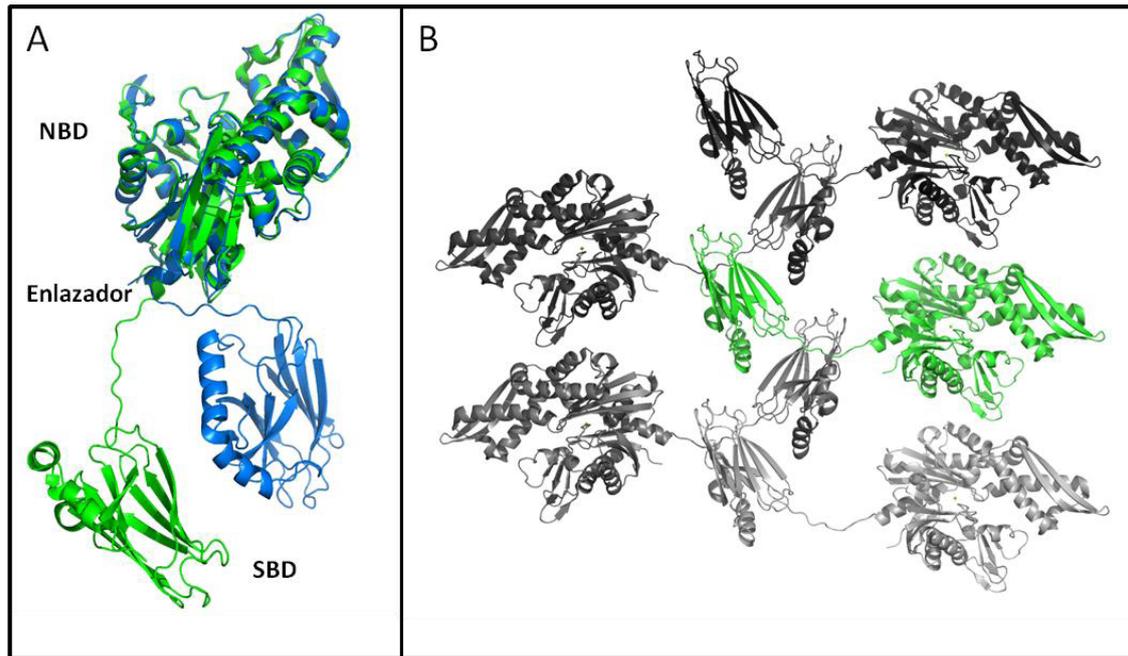


Figura 2.12 A) Comparación estructural de *MgDnaKΔCt-ATP* (azul) vs *GkDnaK* (verde). Se observa la diferencia en la disposición del enlazador de ambas estructuras. B) Estructura de *GkDnaK* donde el SBD está ocupado por el enlazador de una molécula vecina. Imagen adaptada de (Chang, Sun et al. 2008).

- Comparación de las estructuras *MgDnaKΔCt:AMP-PNP* vs *Sse1:ATP* de *S. cerevisiae*

Una vez comparadas las estructuras entre *MgDnaKΔCt:ATP* y *GkDnaKΔCt:ATP* se buscó una estructura con la que comparar *MgDnaKΔCt:AMP-PNP*. Hasta el momento, no se han resuelto estructuras que contengan los dos dominios, NBD y SBD, de DnaK en presencia de AMP-PNP. Por ese motivo, se escogió para comparación la estructura de la proteína *Sse1* de *S. cerevisiae* que fue cristalizada en presencia de ATP (Liu and Hendrickson 2007) (Figura 2.13 A)

Sse1 es una Hsp110 de *Saccharomyces cerevisiae* que estructuralmente posee los dominios presentes en las Hsp70, sugiriendo su posible papel como chaperona Hsp70. Sin embargo, la capacidad de plegar péptidos en Hsp70 recae en cambios conformacionales entre NBD y SBD según la unión de ATP/ADP, lo que es inexistente

en Sse1. En Sse1 el SBD siempre permanece cerrado independientemente de la presencia de ATP o ADP sugiriendo que la posible unión de un péptido en el SBD no tendría una función chaperona (Andreasson, Fiaux et al. 2008, Andreasson, Fiaux et al. 2008) sino que se cree que podría tener una función relacionada con la de un factor intercambiador de nucleótidos (NEF) (Dragovic, Broadley et al. 2006) y/o tener un papel en la propagación de priones en levadura (Moran, Kinsella et al. 2013). El alineamiento de secuencias realizado con Lalign (Huang, et al. 1991) muestra una identidad del 16% con *MgDnaK*. A pesar de ello esta estructura se utiliza como modelo de DnaK en presencia de ATP y en una conformación abierta (Liu and Hendrickson 2007).

La estructura de Sse1 es un dímero que resulta de la interacción entre la tapadera de hélices α de cada monómero y se considera que la interfaz NBD-SBD de Sse1 puede ser utilizada como modelo del estado *MgDnaK:ATP* (Figura 2.13 B). La estructura muestra el NBD y el SBD juntos, de acuerdo con el modelo esperado de *MgDnaK:ATP*, donde el SBD se une al dominio IB del NBD. El enlazador, en este caso, se encuentra formando una hebra adicional de las dos hojas β del subdominio IIA del NBD. Uno de los puntos débiles de considerar representativa la estructura de Sse1 es que la secuencia del enlazador está muy poco conservada (la secuencia del enlazador de Sse1 corresponde a DIHPY) mientras que está muy conservada en otras especies (Figura 2.9). De manera que la interacción entre el enlazador y el NBD de Sse1 podría ser ligeramente diferente al esperado (Liu and Hendrickson 2007).

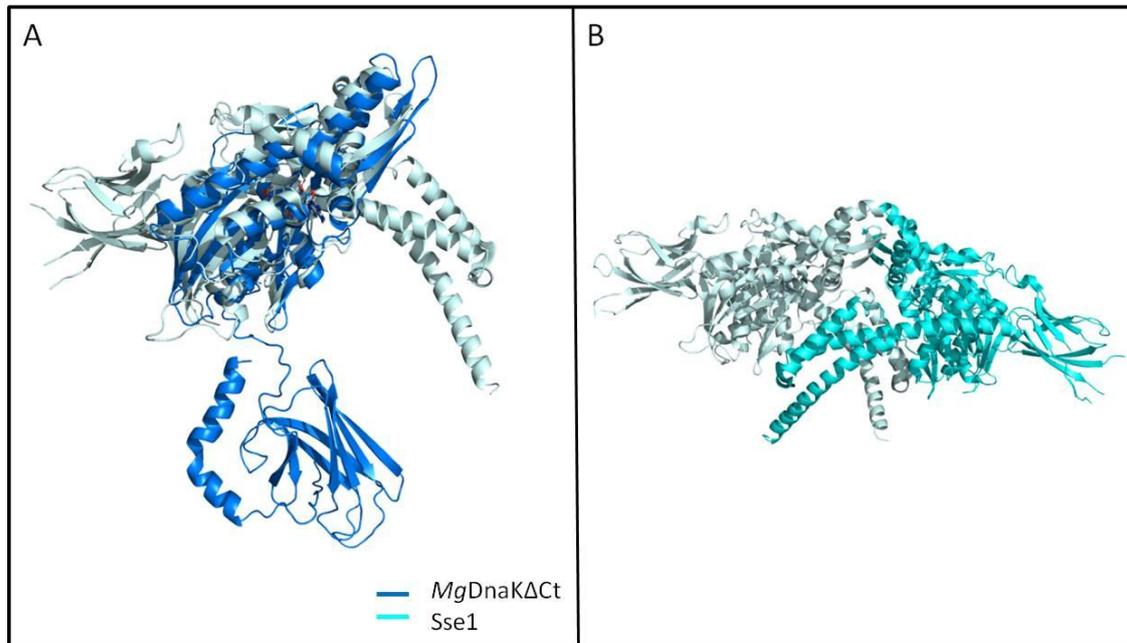


Figura 2.13 A) Comparación estructural de *MgDnaKΔCt*-ATP (azul) vs *Sse1* (turquesa). Se observa la diferencia en la disposición del enlazador de ambas estructuras. B) Dímero de *Sse1* formado por la tapadera de las hélices α de cada monómero.

- Modelo alostérico de chaperonas Hsp70

El modelo alostérico de las chaperonas Hsp70 propuesto por Bhattacharya y colaboradores en 2009 (Bhattacharya, Kurochkin et al. 2009) propone que cuando en el NBD de DnaK está unido a ATP y el SBD está libre los dos dominios están cerca porque el enlazador se dispone bajo la hendidura hidrofóbica del NBD encargándose de transmitir la señal alostérica del NBD al SBD (Figura 2.10). El NBD es un dominio relativamente rígido excepto cuando hidroliza el ATP, momento en el que hay una relajación entre las regiones IA - IB y IIA - IIB (Figura 2.6). Durante la hidrólisis de ATP ocurre una rotación del subdominio IIA-IIB en sentido horario que promueve el cierre de la hendidura entre los subdominios del lóbulo I y por consiguiente se produce la expulsión del enlazador y un distanciamiento entre NBD y SBD. Este modelo también sugiere que la energía de la hidrólisis de ATP se utiliza para compensar la pérdida de

energía de la unión entre NBD-SBD. Mientras tanto, el subdominio IIB se predispone a unirse a una molécula de GrpE, que es el factor intercambiador de nucleótidos cuya función es la de intercambiar la molécula de ADP por una de ATP, estabilizando a su vez el estado relajado del enlazador y la conformación cerrada del SBD de DnaK en presencia de ADP. La unión del NBD a otra molécula de ATP provoca la reaproximación entre NBD y SBD llevando a la expulsión del sustrato. Por último, los autores de este modelo también sugieren que la proximidad entre NBD y SBD compensa energéticamente la pérdida de energía consecuencia de la unión del sustrato (Bhattacharya, Kurochkin et al. 2009).

El mecanismo según el cual GrpE controla el estado conformacional de DnaK dependiendo del intercambio de ADP por ATP fue recientemente publicado (Melero, Moro et al. 2015). En este modelo la chaperona reconocería y se uniría a péptidos sin plegar (conformación abierta con ATP) (Figura 2.14, i). Seguidamente, se produciría la hidrólisis de ATP que induce la conformación cerrada de DnaK, atrapando el sustrato en el SBD (Figura 2.14, ii). Para la liberación del sustrato es necesario el intercambiar ADP por ATP, lo que induce la conformación abierta de DnaK. Esta transición es energéticamente desfavorable y la unión del ATP supera esta barrera energética gracias a dos estados intermedios de DnaK. Este modelo también propone que GrpE facilitaría esta transición modulando el equilibrio de las distintas conformaciones de DnaK. En concreto, GrpE estabilizaría una conformación intermedia de DnaK en la que se daría el intercambio de nucleótidos, ADP por ATP, mientras el sustrato se mantendría unido al SBD hasta la finalización de la reacción (Figura 2.14, iii). El siguiente paso correspondería a la liberación del sustrato (Figura 2.14, iv). Entre estos dos últimos pasos GrpE adoptaría un estado intermedio en que GrpE estabilizaría DnaK

gracias a la unión de los 33 primeros residuos de GrpE al SBD que actúan mimetizando el sustrato de DnaK (Figura 2.14, v) (Melero, Moro et al. 2015) .

La estructura de *MgDnaKΔCt:ATP* representaría la conformación cerrada del modelo de Bhattacharya y también correspondería al paso ii de la Figura 2.14 del modelo de Melero y colaboradores. En esta estructura el enlazador presenta una conformación extendida, el NBD interacciona con una molécula de ADP y el SBD está ocupado por un péptido. Si se compara con la estructura de *GkDnak*, que también representaría la conformación cerrada, se observa que el enlazador de *MgDnaKΔCt:ATP* presenta una conformación más contraída que la que presenta *GkDnak*. Esto se debe a que el enlazador de *GkdnaK* interacciona con el bolsillo hidrofóbico del SBD de una molécula vecina, y por lo tanto, el enlazador está más extendido. En el caso de *MgDnaKΔCt:AMP-PNP* presenta una conformación que no encajaría totalmente con el modelo alostérico propuesto por Bhattacharya y colaboradores o Melero y colaboradores una vez que el enlazador de DnaK en presencia de ATP debería interaccionar con una región del NBD, en concreto, bajo el cruce de hélices α del NBD (conformación abierta). Además, según su modelo, el SBD debería estar libre en presencia de ATP y en el caso de la *MgDnaKΔCt:AMP-PNP* está ocupado por un péptido que mimetiza al sustrato (ver sección 2.1.4). Por lo tanto, la estructura de *MgDnaKΔCt:AMP-PNP* podría representar un paso intermedio del ciclo de chaperonas *MgDnaK-MgDnaJ-MgGrpE* (Figura 2.14), correspondiente al momento en el que *MgDnaK* ya ha reconocido el péptido sin plegar pero la hidrólisis de ATP aún no ha tenido lugar (en este caso no ocurre pues está unido a AMP-PNP). Este paso intermedio se encontraría entre los pasos (i) y (ii) de la Figura 2.14 y correspondería al paso (i') de esta figura. Si fuera así, este intermedio indicaría que el cambio

conformacional del enlazador no sólo sería debido a la presencia de ATP sino también a la presencia/ausencia del sustrato. El hecho que ambas estructuras, *MgDnaKΔCt:AMP-PNP* y *MgDnaKΔCt:ATP* presenten la misma conformación podría ser resultado del empaquetamiento de las proteínas en el cristal (Figura 2.11).

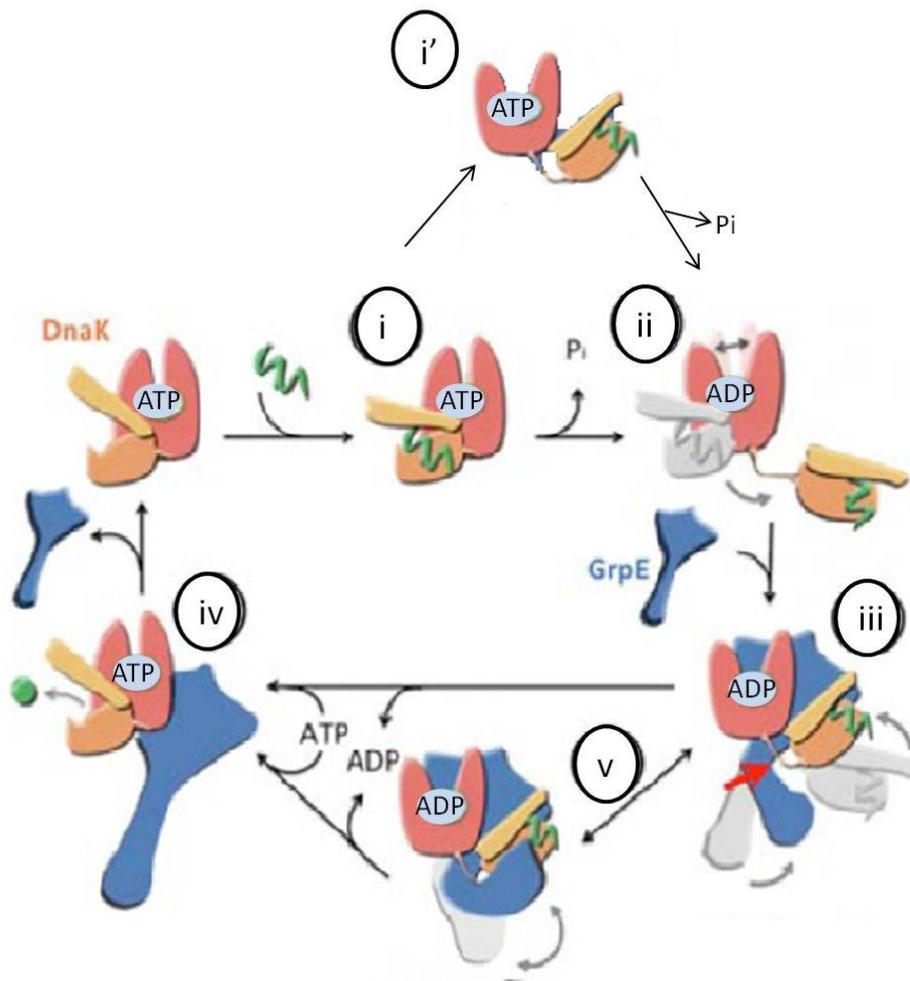


Figura 2.14 Modelo alostérico propuesto por Melero y colaboradores. El paso (i') podría corresponder al nuevo paso intermedio observado en la estructura de *MgDnaK:AMP:PNP*. Figura adaptada de (Melero, Moro et al. 2015).

2.1.6 Mutantes de la *MgDnaKΔCt*

El estudio estructural del SBD de la *MgDnaKΔCt* de *M. genitalium* reveló la presencia de la cola de histidinas del C-terminal en el sitio de unión de sustrato y la importancia de la presencia de aminoácidos hidrofóbicos en el bolsillo de unión a sustrato (Figura 2.7) para la cristalización de *MgDnaKΔCt*. Para averiguar los cambios estructurales producidos cuando el SBD se encuentra libre de sustrato se diseñaron tres mutantes de *MgDnaKΔCt* en los cuales se mutó la Leu522, aminoácido clave para el reconocimiento del péptido por parte del SBD y que no forma parte de la secuencia de la proteína sino que fue añadida como consecuencia de la clonación, a glicina (el aminoácido más pequeño), a ácido aspártico (aminoácido cargado negativamente) o a lisina (aminoácido cargado positivamente). Estas mutaciones se escogieron considerando las características físico-químicas de estos aminoácidos para poder determinar el rol de la leucina en el reconocimiento del sustrato por parte del SBD. Las mutaciones puntuales se realizaron a partir de los cebadores descritos en la Tabla 2.3

Tabla 2.3 Cebadores utilizados para preparar los tres mutantes

Nombre	Secuencia
<i>MgDnaKΔCt-L522D F</i>	5'-TGAAGGGGAA GAC GAGCACCACCAC-3'
<i>MgDnaKΔCt-L522D R</i>	3'-TGAAGGGGAA GAC GAGCACCACCAC-5'
<i>MgDnaKΔCt-L522G F</i>	5'-TGAAGGGGAA GGC GAGCACCACCAC-3'
<i>MgDnaKΔCt-L522G R</i>	3'-TGAAGGGGAA GGC GAGCACCACCAC-5'
<i>MgDnaKΔCt-L522K F</i>	5'-TGAAGGGGAA AAG GAGCACCACCAC-3'
<i>MgDnaKΔCt-L522KR</i>	3'-TGAAGGGGAA AAG GAGCACCACCAC-5'

Una vez obtenidos los mutantes puntuales estos se transformaron, expresaron y purificaron siguiendo el protocolo definido para producir *MgDnaKΔCt* (ver sección 4.4). El volumen de elución de los diferentes mutantes inyectados en una columna de filtración en gel Superdex 200 16/60 (GE Healthcare, Uppsala, Sweden) fue de 63 mL (Figura 2.15), mientras que la elución de la *MgDnaKΔCt* purificada a través de la misma columna es de 80mL (Figura 2.4 C). Estos resultados indican que el estado de oligomerización de los mutantes en comparación con *MgDnaKΔCt* es diferente y además el pico de elución de los mutantes presenta un hombro, ausente en la purificación de *MgDnaKΔCt*, seguramente consecuencia de la heterogeneidad de la muestra.

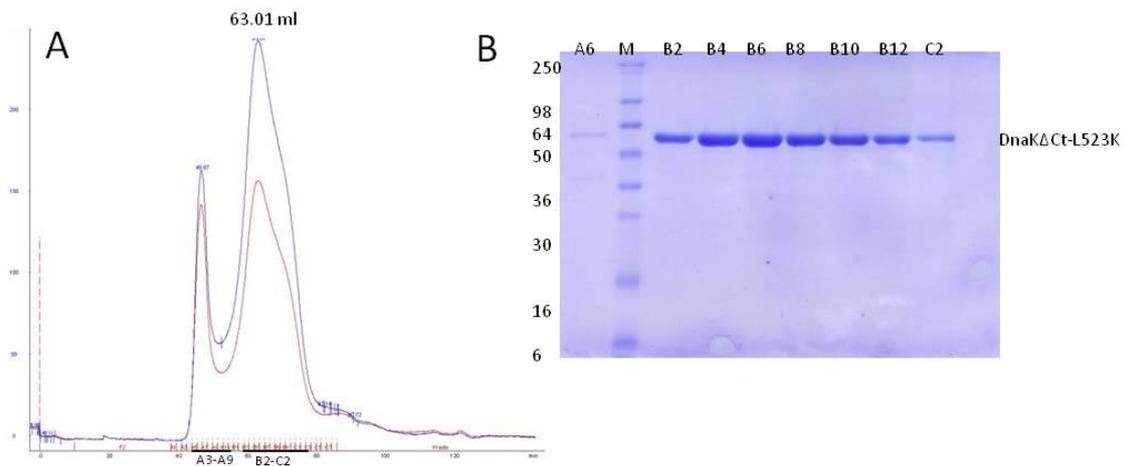


Figura 2.15 A) Cromatografía de filtración por gel filtración en una Superdex 200 16/60 del mutante *MgDnaKΔCt-L522K*. B) Gel de acrilamida al 12% SDS-PAGE de la muestra eluída de la Superdex 200 16/60. La muestra eluída corresponde a las fracciones de los picos observados en el cromatograma.

Se realizaron cribados de cristalización con los diferentes mutantes a una concentración de 10 mg/mL en presencia de ATP o AMP-PNP. En estos, más de un 80% de las gotas estaban precipitadas por lo cual se repitieron los ensayos a concentraciones menores de proteína y también se testaron dos temperaturas, 20°C y 4°C, y en ningún caso se obtuvieron cristales. La ausencia de cristales fue relacionada con el hecho que las muestras no eran suficientemente estables y homogéneas, condiciones muy importantes para la cristalización, lo que llevó a que se estudiara su estabilidad mediante la técnica de cromatografía de exclusión molecular y dispersión de luz en múltiples ángulos (del inglés *multiangle light scattering*).

Sólo se estudiaron por esta técnica los mutantes *MgDnaKΔCt-L522D* y *MgDnaKΔCt-L522K*, porque el mutante *MgDnaKΔCt-L522G* era más inestable y agregaba más, además no se obtuvo con un grado de pureza suficiente a pesar de que su perfil cromatográfico es idéntico al de los otros mutantes. Por lo tanto, los mutantes de *MgDnaKΔCt* fueron inyectados individualmente en la columna de exclusión molecular

Superdex 200 10/300 previamente equilibrada en tampón 0.5 M NaCl, 0.02 M Tris pH8, y analizados en línea por un instrumento de MALS cuyos resultados permiten determinar la masa molecular absoluta bien como la polidispersidad de las moléculas en solución. En los resultados de MALS referentes al análisis de los mutantes *MgDnaKΔCt-L522D* y *MgDnaKΔCt-L522K* se observa que ambas muestras son heterogéneas. El pico obtenido en ambos casos en la cromatografía de gel filtración después de su purificación a través de una columna de afinidad, con el paso del tiempo, se transformó en diferentes especies dando a lugar a varios picos en la columna de gel filtración. Cuando estas muestras se analizaron por MALS se detectaron tres picos indicativos de la falta de homogeneidad de las muestras. Estos resultados no permitieron calcular con precisión la masa molecular de las diferentes especies encontradas en solución pero evidenciaron que las muestras eran polidispersas. Esta heterogeneidad y consecuente inestabilidad puede ser debida a que en el caso de estos mutantes, *MgDnaKΔCt-L522D* y *MgDnaKΔCt-L522K*, el SBD no reconoce al péptido y el bolsillo de unión al sustrato está libre. Esto es debido a que el carácter altamente hidrofóbico del bolsillo impide que el Asp522 y la Lys522 de los mutantes que contiene cadenas laterales cargadas positivamente interaccionen con el SBD. Esto hace que el SBD no esté ocupado por ningún péptido y por lo tanto, que sea muy inestable. Parece ser que sólo cuando se encuentra ocupado es estable. Otro motivo que puede provocar la inestabilidad de la muestra es que tal y como sucede en la *GkDnaK*, el enlazador de una molécula vecina pasa a mimetizar el sustrato resultando en una mezcla de especies con diferentes estados de oligomerización, lo que no favorece la cristalización de estos mutantes ya que habría una mezcla de especies distintas.

Para intentar obtener la estructura de la *MgDnaKΔCt* sin que la cola de histidinas añadidas por clonaje interactuara con el bolsillo del SBD se diseñó otro mutante. En este nuevo constructo la secuencia de la *MgDnaKΔCt* fue introducida en el vector pET28a (Novagen, Madison, WI, USA) que tiene la cola de histidinas en el N-terminal seguida de una diana de corte de Trombina antes de la secuencia de la *MgDnaKΔCt*. Este constructo permitiría la purificación de *MgDnaKΔCt* por cromatografía de afinidad a través de la cola de histidinas que sería eliminada a posteriori por la digestión con Trombina en *batch*, permitiendo la obtención de la *MgDnaKΔCt* en solución sin que la cola de histidinas interactuara con el SBD. Después de un primer paso de purificación por cromatografía de afinidad la solución de proteína obtenida fue digerida con Trombina durante toda la noche y se volvió a cargar la muestra por la columna de afinidad donde la proteína quedó retenida, indicando la falta de accesibilidad de la diana de corte por parte de la proteasa. Se siguieron otros pasos de optimización de la reacción de digestión con Trombina en columna que tampoco funcionaron por lo cual se interrumpieron los intentos de obtener *MgDnaKΔCt* sin cola de histidinas.

2.2 Descripción de la estructura cristalina de *MgNBD*

2.2.1 Introducción

Los estudios estructurales llevados a cabo con el dominio de unión a nucleótido de la *MgDnaK* se iniciaron por dos motivos principalmente: estudiar este dominio en complejo con GrpE (ver sección 2.3.2) y estudiar el mecanismo de hidrólisis del ATP, ya que en ausencia de nucleótidos no se habían obtenido cristales de la *MgDnaKΔCt* posiblemente debido a su inestabilidad y heterogeneidad. La inestabilidad y heterogeneidad pueden ser debidas por una parte a la ausencia de nucleótido en el NBD que provoca que el SBD esté más expuesto y lo que conlleva a unos mayores cambios conformacionales entre NBD y SBD dificultando así su cristalización. Por otra parte, la hélice B del SBD, que en nuestro constructo no está completa, y cuya función se ha demostrado que permite la estabilización del SBD en presencia de un péptido podría proporcionar cierta estabilidad en la DnaK. Es muy conocido el hecho que DnaK es una proteína que tiene tendencia a agregar tanto en presencia de nucleótidos como en ausencia de estos y es principalmente, la ausencia de péptido en el SBD lo que provoca una agregación mucho mayor. (Zuiderweg, Bertelsen et al. 2013).

2.2.2 Clonaje, expresión y purificación de MgNBD

El NBD que a partir de ahora se la denominará como *MgNBD*, fue clonada en el vector de expresión pET21d (Novagen, Madison, WI, USA) que añade una cola de seis histidinas en el C-terminal de la proteína. El clon se transformó en células de *E. coli* BL21 (DE3) que fueron posteriormente plaqueadas en medio de cultivo LB/agar suplementado con ampicilina. Para establecer el protocolo de expresión de la proteína de interés se realizaron varias pruebas en medio LB variando la temperatura y el tiempo de expresión una vez añadido el inductor IPTG. Como se observa en el gel desnaturizante de poliacrilamida de la Figura 2.14 donde se cargaron extractos celulares de la expresión a dos tiempos distintos ($t=3h$ y $t=0/N$) posterior a la inducción y a dos temperaturas (a $20^{\circ}C$ y a $37^{\circ}C$). En el caso de la expresión a $20^{\circ}C$ solo se observó expresión de la proteína después de incubar el cultivo toda la noche en presencia de 1 mM IPTG, a diferencia de lo observado cuando las células se indujeron a $37^{\circ}C$ donde se observó expresión de la proteína de interés a partir de las 3 horas de inducción. El peso molecular teórico del *MgNBD*, calculado con ProtParam (Bairoch, Apweiler et al. 2005) es de 40.6 kDa, equivalente al peso molecular de la banda proteína que fue expresada (Figura 2.16), indicando que se trata de la proteína de interés.

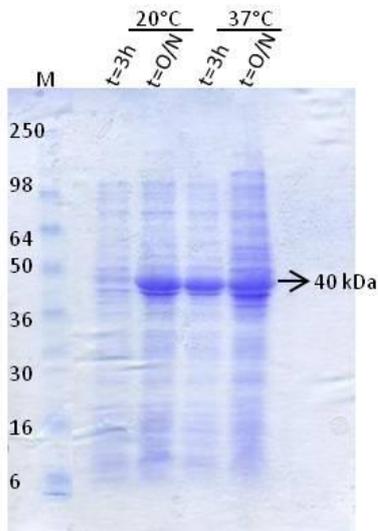


Figura 2.16 Geles desnaturalizantes SDS-PAGE 12% acrilamida. Resultados de las pruebas de expresión del *MgNBD* a 20°C y a 37°C. Se observa una banda de unos ~40kDa que corresponde a la expresión de NBD a diferentes temperaturas y tiempos de expresión.

Se decidió trabajar con las condiciones de expresión a 20°C porque parece que la cantidad de proteína expresada es ligeramente superior que cuando es expresada a 37°C. Se procedió a la expresión y purificación de la proteína a gran escala.

El sobrenadante obtenido de la lisis celular de 1L de cultivo *E. coli* se inyectó en una columna de afinidad HisTrap™ HP de 1 mL (GE Healthcare, Uppsala, Sweden), equilibrada en 0.5 M NaCl, 0.02 M Tri- HCl pH 8, y la elución se realizó mediante la aplicación de un gradiente lineal de imidazol (la concentración final de imidazol fue de 0.5 M). La eficiencia del primer paso de purificación se evaluó mediante análisis de las fracciones recogidas en un gel de poliacrilamida desnaturalizante y se escogieron las fracciones más puras para proceder al siguiente paso de purificación a través de una columna de gel filtración Superdex 200 16/60 (GE Healthcare, Uppsala, Sweden) equilibrada en el mismo tampón que la HisTrap™ HP (GE Healthcare, Uppsala, Sweden). El volumen de elución de *MgNBD* en la columna gel filtración corresponde a 85.2 mL lo que indica un estado monomérico (~40 kDa) de la proteína en solución,

según la calibración de la columna con una recta patrón de los volúmenes de elución de proteínas estándares. El grado de pureza de la muestra también se evaluó por análisis de las fracciones recogidas en un gel de poliacrilamida desnaturalizante (Figura 2.17).

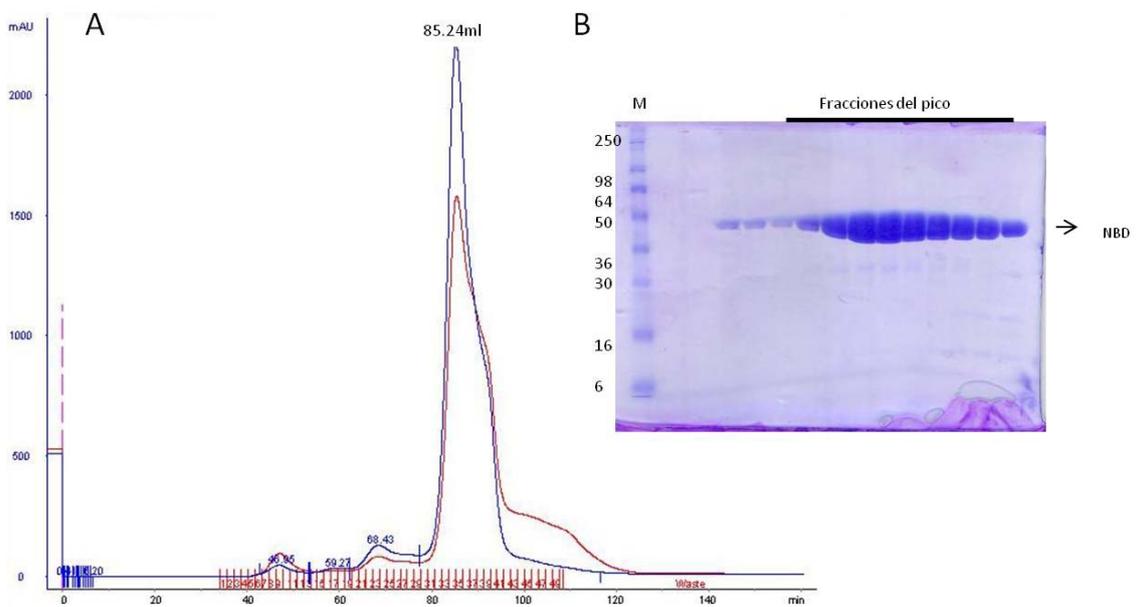


Figura 2.17 A) Perfil de la cromatografía de gel filtración del *MgNBD*. B) Análisis de las fracciones correspondientes a la elución de la muestra de *MgNBD* en una columna de gel filtración mediante un gel de poliacrilamida SDS-PAGE de 15%.

Las fracciones correspondientes a la elución de *MgNBD* se concentraron hasta 11 mg/mL utilizando un dispositivo de ultrafiltración con un punto de corte de 10 kDa (Amicon Ultra MWCO 10k Da, Darmstadt, Germany).

2.2.3 Cristalización del *Mg*NBD

Los ensayos iniciales de cristalización del *Mg*NBD se realizaron por la técnica de difusión de vapor a 20°C y a 4°C en placas Greiner Bio-One (Monroe, North Carolina, USA) de 96 condiciones en presencia de ATP. La proteína se incubó alrededor de una hora con 1mM ATP (Sigma Aldrich) y 1 mM $MgCl_2$ y se realizaron los cribados de cristalización ofertados por la Plataforma Automatizada de Cristalización (PCB-IRB-CSIC). La proteína cristalizó a los tres días en la condición 0.1 M bicina pH 9, 0.02 M tiocianato de potasio (KSCN) y 2.6 M sulfato de amonio a 4°C (Figura 2.18 A). Estos cristales se optimizaron variando el pH, la sal y el precipitante alrededor de las condiciones donde aparecieron los primeros cristales, para obtener más cristales y de mayor tamaño, en una placa de 24 pocillos. Los cristales que mejor difractaron se obtuvieron en la condición 0.1 M HEPES pH 7, 0.02 M KSCN y 2.4 M sulfato de amonio (Figura 2.18 B). Partiendo de esta condición de cristalización se realizaron ensayos de cristalización en ausencia o presencia de otros nucleótidos. Como anteriormente, primero se incubó durante una hora la proteína con 1 mM AMP-PNP (Sigma Aldrich) y 1 mM $MgCl_2$ a 4°C, y los primeros cristales en esta condición se obtuvieron a 20°C (Figura 2.18 C). Después de algunos pasos de optimización se obtuvieron los mejores cristales en la condición 0.1 M Tris-HCl pH 8, 0.02 M KSCN y 2.2 M sulfato de amonio (Figura 2.18 D).

Los ensayos de cristalización con ADP se podrían haber realizado sin la necesidad de añadir 1mM $MgCl_2$. Normalmente se añade el magnesio porque es necesario para que se dé la hidrólisis del ATP pero en este caso sería innecesario. Sin embargo se añadió

para intentar minimizar las diferencias entre las cristalizaciones del *MgNBD* con ATP o AMP-PNP. Partiendo de la condición de cristalización obtenida para *MgNBD*:ATP y *MgNBD*:AMP-PNP en placas Limbro de 24 pocillos (Hampton research, Aliso viejo, California, USA) se formuló un nuevo cribado de cristalización para *MgNBD* en presencia de ADP (Figura 2.18 E). Los primeros cristales obtenidos fueron optimizados variando la proporción proteína: solución de cristalización de 1:1 a 2:1 y los cristales aparecieron en 0.1 M Tris-HCl pH 7.5, 0.02 M KSCN y 2.5 M sulfato de amonio (Figura 2.18 F).

El último estado de *MgNBD* que faltaba por cristalizar era al que no se le añadían nucleótidos. Como en el caso anterior, partiendo de una placa de optimización de la condición 0.1 M bicina pH 9, 0.02 M tiocianato de potasio (KSCN) y 2.6 M sulfato de amonio se iniciaron los ensayos de cristalización. *MgNBD* sin nucleótido cristalizó en dos formas cristalinas distintas, en la condición 0.1 M Tris-HCl pH 7.5, 0.02 M KSCN y 2.4M sulfato de amonio: la primera aparecía a los 3 días y los cristales formaban un aglomerado de placas (Figura 2.18 G) mientras que la segunda aparecía a los 10 días y los cristales tenían forma de prismas (Figura 2.18 H).

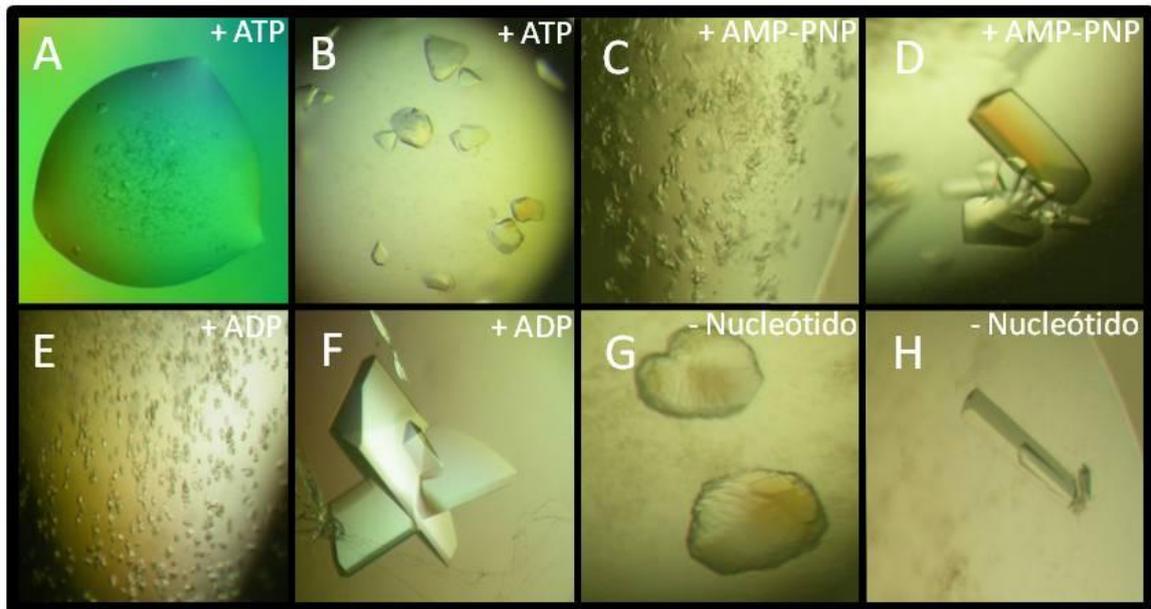


Figura 2.18 Imágenes de los diferentes cristales de *MgNBD* obtenidos según el nucleótido añadido: A-B) ATP, C-D) AMP-PNP, E-F) ADP y G-H) sin nucleótido.

2.2.4 Determinación de la Estructura de rayos-X de *MgNBD*

Los cristales de *MgNBD*:ATP se difractaron en la línea ID23eh2 del ESRF (Grenoble, Francia) mientras que los otros complejos se difractaron en el sincrotrón ALBA (Cerdanyola del Vallès, Barcelona, España) en la línea BL13-XALOC. Los diferentes cristales se indexaron con el programa XDS (Kabsch 2010) en el grupo espacial H3 con los siguientes parámetros de celda: $a=b=157.9 \text{ \AA}$, $c=57 \text{ \AA}$ y $\alpha=\beta=90^\circ$ y $\gamma=120^\circ$. Los cristales contienen una molécula por unidad asimétrica y un contenido de solvente de 65.5 % calculado según el coeficiente de Matthews (Matthews 1968).

Para resolver el problema de las fases del cristal de *MgNBD*:ATP se utilizó el método de MR. Al disponer de la estructura refinada de *MgDnaKΔCt* en presencia de AMP-PNP a una resolución de 1.8 \AA se utilizó el dominio de unión a nucleótido de esta estructura

como modelo de búsqueda. La solución única que se encontró mediante el programa Phaser (Adams, Afonine et al. 2010) permitió la construcción del modelo que se realizó intercalando ciclos de refinamiento con Refmac (Murshudov, Vagin et al. 1997) y de construcción manual con Coot (Emsley and Cowtan 2004) resultando en una estructura con valores de R_{factor} y R_{free} de 17.0% y 18.9%, respectivamente. Para resolver las estructuras de *MgNBD:AMP-PNP*, *MgNBD:ADP* y *MgNBD:APO* se utilizó como modelo de búsqueda la estructura de *MgNBD:ATP*, previamente resuelta, y el mismo procedimiento que se utilizó para resolver su estructura. Los datos estadísticos de la recogida de datos y del refinamiento de estas estructuras se muestran en la Tabla 2.4.

Tabla 2.4 Datos estadísticos de la recogida de datos y del refinamiento de las varias estructuras de MgNBD. Entre paréntesis se muestran los resultados para la capa de mayor resolución.

Conjunto de datos	MgNBD: ATP	MgNBD: AMPPNP	MgNBD:ADP	MgNBD:apo
Fuente rayos-X	ESRF ID23eh2	ALBA Xaloc	ALBA Xaloc	ALBA Xaloc
Longitud de onda (Å)	0.8729	0.9795	0.9795	0.9795
Grupo espacial	H3	H3	H3	H3
Dimensiones celdilla:				
a,b,c (Å)	a=b=157.91	a=b= 157.07	a=b= 157.03	a=b= 157.92
	c= 57.504	c= 56.7	c= 56.55	c= 64.70
$\alpha,\beta,\gamma(^{\circ})$	$\alpha=\beta=90 \gamma=120$	$\alpha=\beta=90 \gamma=120$	$\alpha=\beta=90 \gamma=120$	$\alpha=\beta=90 \gamma=120$
Resolución (Å)	45.59-1.65 (1.74-1.65)	45.34-1.30 (1.37-1.30)	45.33-1.28 (1.35-1.28)	47.00-1.78 (1.88-1.78)
R_{merge} (%)	0.09 (0.52)	0.05(0.58)	0.04(0.53)	0.03 (0.51)
I/σ(I)	8.5 (2.0)	12.0 (2.1)	12.0 (1.9)	16.0 (1.7)
Completitud (%)	99.7 (98.9)	99.5(97.0)	99.6 (97.8)	99.4 (98.0)
Nº reflexiones únicas	64132(9296)	127585(18174)	133004(19073)	57432(8302)
Multiplicidad	4.0 (3.8)	3.5 (3.4)	3.3 (3.1)	2.9 (2.8)
Wilson B-factor (Å²)	13.7	14.0	13.5	25.8
R_{pim}	0.05 (0.60)	0.03(0.36)	0.03 (0.35)	0.02(0.36)
Refinamiento				
R_{factor}/R_{free} (%)	17.00/18.8	16.27/17.80	16.82/18.01	16.32/19.32
Rmsd				
• Longitud enlace (Å)	0.027	0.036	0.036	0.024
• Ángulo (°)	2.753	3.209	3.193	2.417
B-valores(Å²)				
• Proteína	11.2	8.6	7.1	19.2
• Péptido	23.4	16.5	17.7	-----
• Agua	2.7	5.0	2.0	27.0
Ramachandran plot				
• Más favorables(%)	98.7	98.7	98.7	98.7
• Permitidas(%)	99.7	99.7	99.7	99.7
• No permitidas(%)	0.27	0.27	0.27	0.27

2.2.5 Descripción de la estructura cristalina del *Mg*NBD

La estructura cristalina de *Mg*NBD:ADP incluye los aminoácidos desde Ala3 a His375. El *Mg*NBD, del mismo modo que lo observado en la estructura de *Mg*DnaKΔCt, se compone de dos lóbulos, I y II, que se dividen en dos subdominios: IA y IB por el lóbulo I, y IIA y IIB por el lóbulo II. Los nucleótidos se unen al vértice de la V formada por los subdominios IA-IB y IIA-IIB, estando todos los subdominios involucrados en la coordinación del nucleótido (Figura 2.19 A y B).

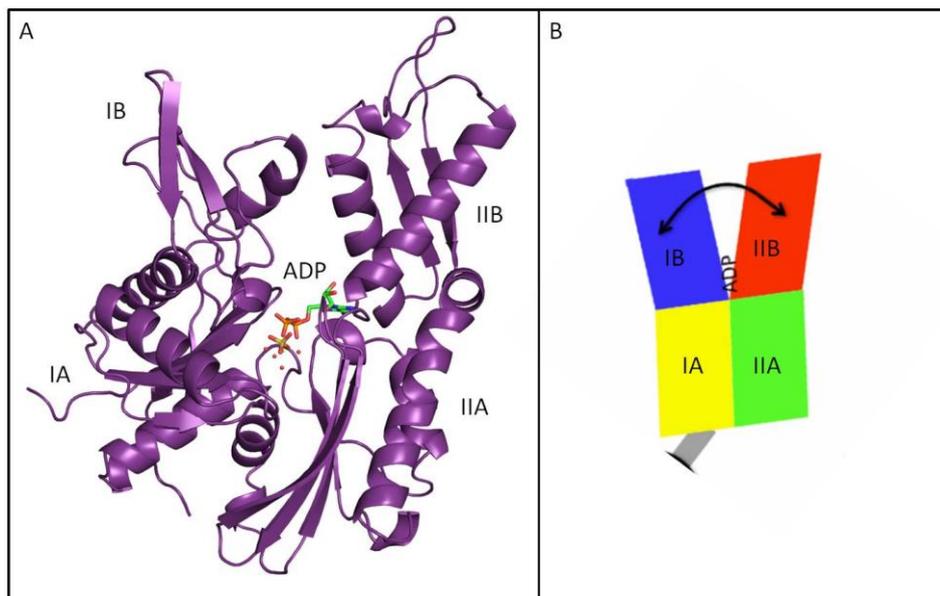


Figura 2.19 A) Estructura global de *Mg*NBD en presencia de ADP. Se indican los subdominios IA-IB y IIA-IIB y se representa la molécula de ADP unida al bolsillo de unión a nucleótido. B) Esquema global del NBD donde se muestra la V donde se unen los nucleótidos. Esquema adaptado de Zuiderweg y colaboradores (Zuiderweg, Bertelsen et al. 2013).

El nucleótido se dispone transversalmente a lo largo del *Mg*NBD, la adenosina y la desoxirribosa se encajan en la interfaz de los subdominios IIA y IIB. Los fosfatos α y β

del nucleótido interaccionan principalmente con los residuos del subdominio IIA mientras que el fosfato γ interacciona principalmente con el subdominio IA (Figura 2.19). Los residuos implicados en la hidrólisis del nucleótido fueron muy bien descritos por O'Brien y colaboradores que demostraron que la Lys71 de la Hsp70 de *Bos taurus* (código PDB 1KAY) (correspondiente a la Lys70 en el caso de la *EcDnaK* (código PDB 1DKZ) y Lys72 en el caso de la *MgDnaK*) y que pertenece al subdominio IIB, es esencial para este proceso. En estos estudios también se observaron que el grupo ϵ -amino de la lisina está involucrado en la activación de una molécula de agua que lleva a cabo el ataque nucleofílico al fosfato γ (Wilbanks, DeLuca-Flaherty et al. 1994, O'Brien, Flaherty et al. 1996), y demostraron que la hidrólisis no sólo depende de un único residuo (Lys70) sino que es el resultado de la configuración de los residuos que lo forman e incluso, de los iones que allí se coordinan. En el caso del *MgNBD* el nucleótido está coordinado solamente por un ion magnesio del que se hablará a continuación, mientras que en el caso de la Hsp70 de *B. taurus* el nucleótido aparece coordinado por tres iones potasio y un magnesio.

Comúnmente, la coordinación de un magnesio con un nucleótido es octaédrica como se observa en la estructura de la *MgNBD:ADP* donde la coordinación se produce entre el fosfato β del ADP, una molécula de sulfato (probablemente proveniente del sulfato de amonio presente en la condición de cristalización) y cuatro moléculas de agua, siendo la distancia promedio entre el magnesio y los átomos que lo coordinan de ~ 2.1 Å. La comparación entre la estructura de *MgNBD:ADP* y de *MgNBD:ATP* indica que después de la hidrólisis de ATP, el fosfato γ queda retenido en la estructura de la misma manera que el sulfato en el caso de la *MgNBD:ADP* (Figura 2.20).

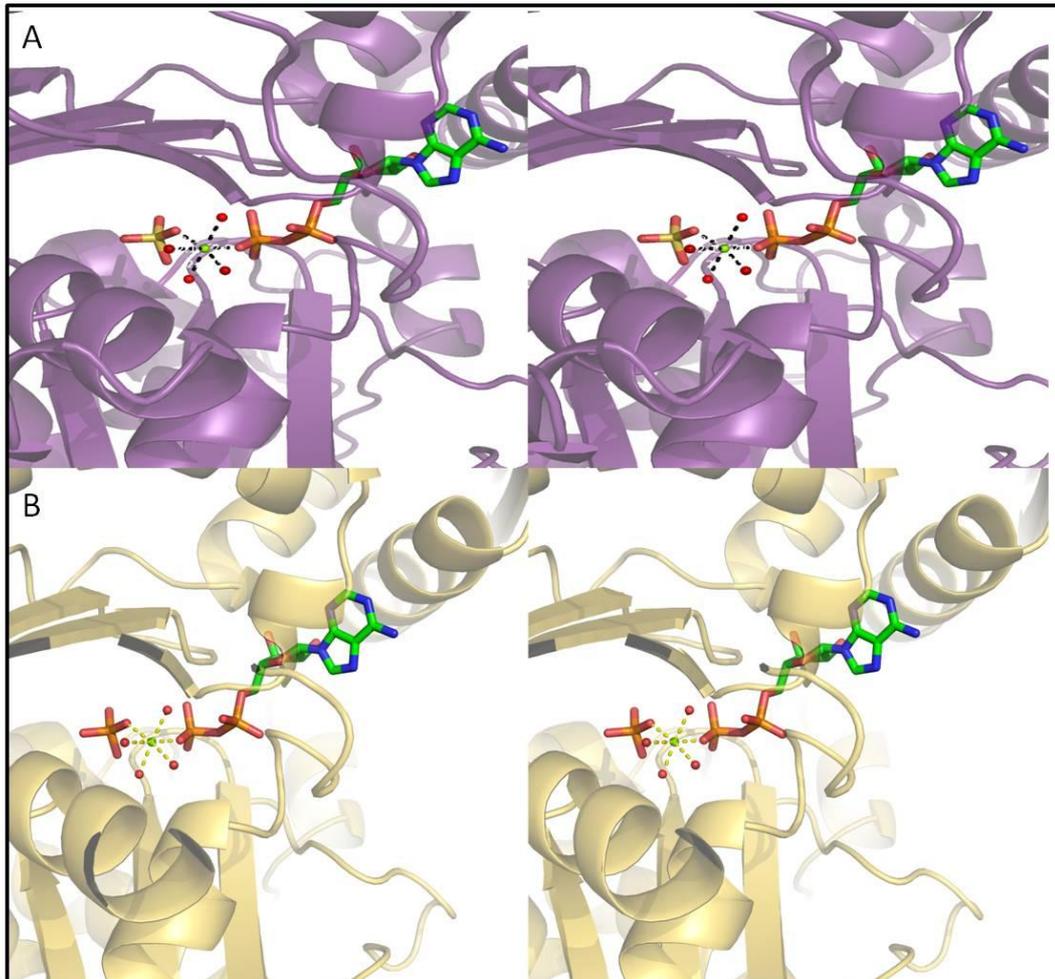


Figura 2.20 A) Representación estereoscópica de *MgNBD* en presencia de ADP. Se muestra el detalle de la coordinación octaédrica del magnesio a cuatro moléculas de agua, a un ion sulfato y al fosfato β del ADP. B) Representación estereoscópica de *MgNBD* co-cristalizado con ATP.

La comparación entre las estructuras globales de *MgNBD*:APO, *MgNBD*:ATP y *MgNBD*:AMP-PNP muestran que el plegamiento de las tres estructuras es idéntico, apreciándose a penas un ligero desplazamiento de los subdominios IB y IIB en la estructura de *MgNBD*:APO (Figura 2.21 A) sugiriendo que este desplazamiento es producido por la falta del nucleótido.

Sin embargo, cuando se comparó la estructura *MgDnaK Δ Ct*:AMP-PNP vs *MgNBD*:AMP-PNP se observó el mismo desplazamiento de los subdominios IB y IIB que se había

observado anteriormente, indicando por lo tanto, que la rotación de los subdominios IB y IIB es posiblemente consecuencia del empaquetamiento de la proteína en el cristal y es independiente del nucleótido que esté unido al *MgNBD*. Por este motivo se comparó *MgDnaKΔCt*:AMP-PNP vs *MgNBD*:APO confirmando que los dominios IB y IIB no muestran desplazamientos significativo entre ellos (Figura 2.21 B). También se comparó la estructura *MgDnaKΔCt*:AMP-PNP con *MgNBD*:ATP y *MgNBD*:AMP-PNP lo que volvió a mostrar, por una parte, que *MgNBD*:ATP y *MgNBD*:AMP-PNP son idénticos y que es la *MgDnaKΔCt*:AMP-PNP la que presenta un movimiento de los subdominios IB y IIB consecuencia del empaquetamiento que presenta la proteína en cristal (Figura 2.21 C).

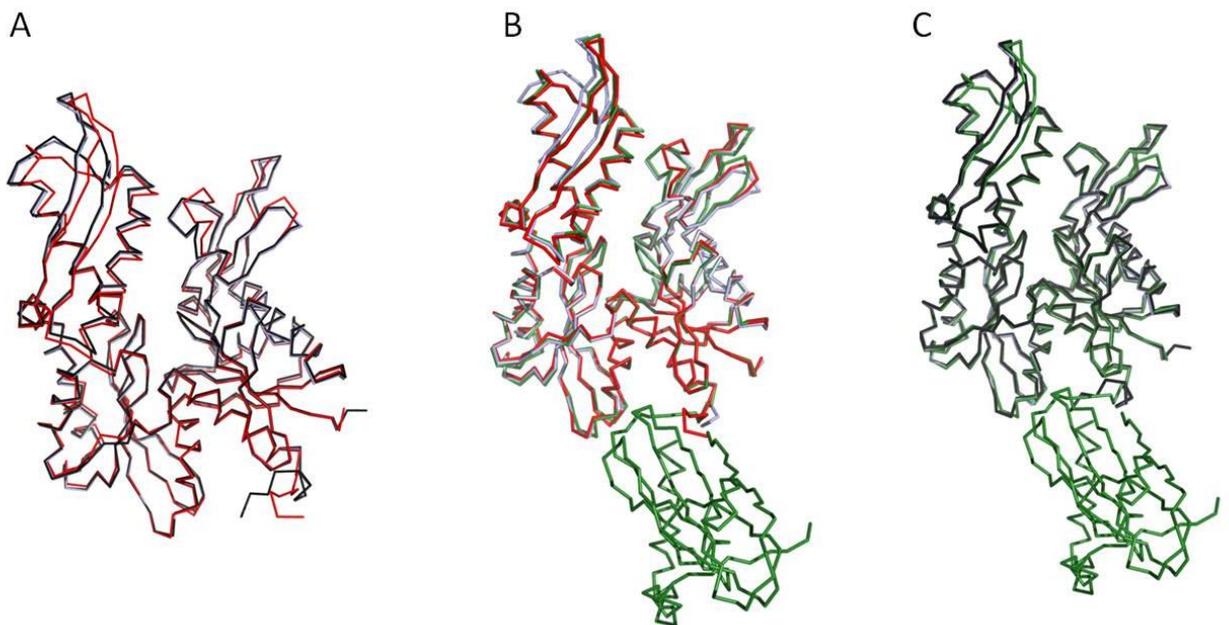


Figura 2.21 Superposición de diferentes estructuras de *MgDnaKΔCt* y/o *MgNBD* utilizando representación en cintas. A) *MgNBD*: ATP (negro), *MgNBD*: AMP-PNP (gris), *MgNBD*:APO (rojo). B) *MgNBD*:APO (rojo), *MgNBD*:AMP-PNP (gris), *MgDnaKΔCt*:AMP-PNP (verde). C) *MgNBD*: ATP (negro), *MgNBD*: AMP-PNP (gris), *MgDnaKΔCt*:AMP-PNP (verde).

Otra manera de estudiar las diferencias entre las estructuras descritas anteriormente fue calculando los r.m.s.d entre las diferentes estructuras que proporciona el valor aproximado de la distancia entre los átomos equivalentes de las estructuras superpuestas (Tabla 2.5).

Tabla 2.5 Cálculo de las diferencias observadas entre las diferentes estructuras de DnaKΔCt y NBD

r.m.s.d. (Å)	<i>Mg</i> NBD:ATP	<i>Mg</i> NBD:AMP-PNP	<i>Mg</i> NBD:APO	<i>Mg</i> NBD:ADP	<i>Mg</i> DnaKΔCt:AMP-PNP
<i>Mg</i> NBD:ATP	-	0.19	1.044	0.19	0.81
<i>Mg</i> NBD:AMPPNP		-	1.05	0.11	0.79
<i>Mg</i> NBD:APO			-	1.05	0.88
<i>Mg</i> DnaKΔCt:ADP				-	0.81

Se observó que la mayor diferencia se da entre la estructura de *Mg*NBD:APO y la estructura de este mismo dominio en presencia de diferentes tipos de nucleótidos. Cuando se compararon más detalladamente las estructuras de *Mg*NBD:ADP vs *Mg*NBD:AMP-PNP se observaron pequeños cambios en aminoácidos que interactúan con el nucleótido, como es el caso de la Asn347 o el Asp349, donde en la estructura *Mg*NBD:AMP-PNP muestran dobles conformaciones. Otra diferencia es que en el caso de *Mg*NBD:ADP existen pequeñas variaciones en la red de moléculas de solvente que se observan alrededor del nucleótido, donde apareció una zona de densidad extra que se pudo modelar como una molécula de glicerol. La coordinación del magnesio en *Mg*NBD:ADP pasa de octaédrica a una inusual coordinación tetragonal piramidal, en el caso de *Mg*NBD:AMPPNP. Curiosamente, en esta última estructura el ión magnesio no

coordina con los fosfatos β o γ del nucleótido (Figura 2.22 A) como sería de esperar, pasando a ser coordinado por 4 moléculas de agua y el fosfato β del AMP-PNP (Figura 2.22 B).

Normalmente el ión magnesio en presencia de ATP o ADP se coordina con los dos fosfatos terminales del nucleótido, tanto con el fosfato $\alpha\beta$ como con el $\beta\gamma$, según si el ATP haya sido o no hidrolizado promoviendo la transferencia del grupo fosforilo cuando se da el ataque nucleofílico (Figura 2.22, A) (Harding 1999). El ión magnesio se encuentra en la misma posición, en *MgNBD:ADP* y en *MgNBD:AMP-PNP* pero, en el caso de *MgDnaK Δ Ct:AMP-PNP*, su ocupación es menor. Se hipotetiza entonces, que la coordinación inusual del ión magnesio que se observa en la estructura *MgNBD:AMP-PNP* se debe al hecho de que el magnesio no puede coordinarse con el nitrógeno que sustituye al oxígeno en el enlace entre los fosfatos β y γ , del AMP-PNP. A pesar, incluso, de que el magnesio ocupa la posición que le correspondería para realizar una correcta coordinación octaédrica esta resulta en una coordinación tetragonal piramidal (Figura 2.22 B).

En la estructura del *MgNBD:APO* no se aprecian cambios estructurales muy significativos y se han podido modelar dos moléculas de sulfato en la posición donde normalmente se posicionan los nucleótidos, posiblemente provenientes de la condición de cristalización. Estos sulfatos presentan un ligero desplazamiento en relación a la configuración observada para los fosfatos de los nucleótidos presentes en las otras estructuras (Figura 2.22 C). La hendidura central en forma de V del NBD muestra un movimiento de los dos lóbulos del subdominio IIA-IIB, alrededor de los aminoácidos Gln240-Lys225 y Leu264-Leu278. Estas diferencias parecen indicar que el

intercambio entre ADP y ATP, es lo que provoca un cambio estructural importante en el NBD posiblemente provocado por la interacción con GrpE y no por la hidrólisis del nucleótido, lo que también fue sugerido por Melero y colaboradores (Melero, Moro et al. 2015)

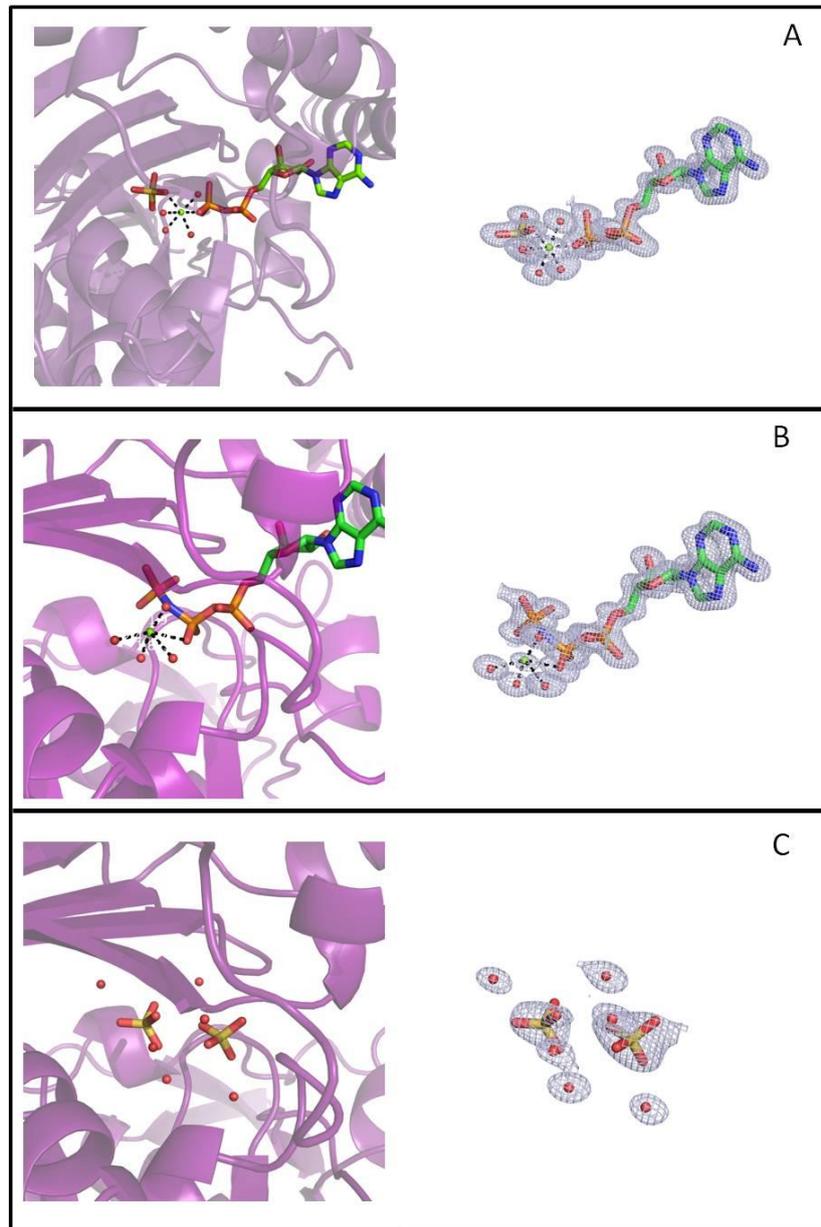


Figura 2.22 Detalle del bolsillo de unión a nucleótido en las diferentes estructuras del NBD donde se muestra la coordinación del magnesio y la densidad electrónica para los nucleótidos. A) *MgNBD:ADP*. B) *MgNBD:AMP-PNP*. C) *MgNBD:APO*.

- Comparación de las estructuras *Mg*NBD:AMP-PNP, *Mg*DnaKΔCt:AMP-PNP

Cuando se comparan las estructuras *Mg*DnaKΔCt:AMP-PNP y *Mg*NBD:AMP-PNP se observa que los cambios entre ambas proteínas son muy sutiles y que los cambios más significativos se localizan en la posición de los dos bucles flexibles presentes en el subdominio IIB (Figura 2.23 A y B).

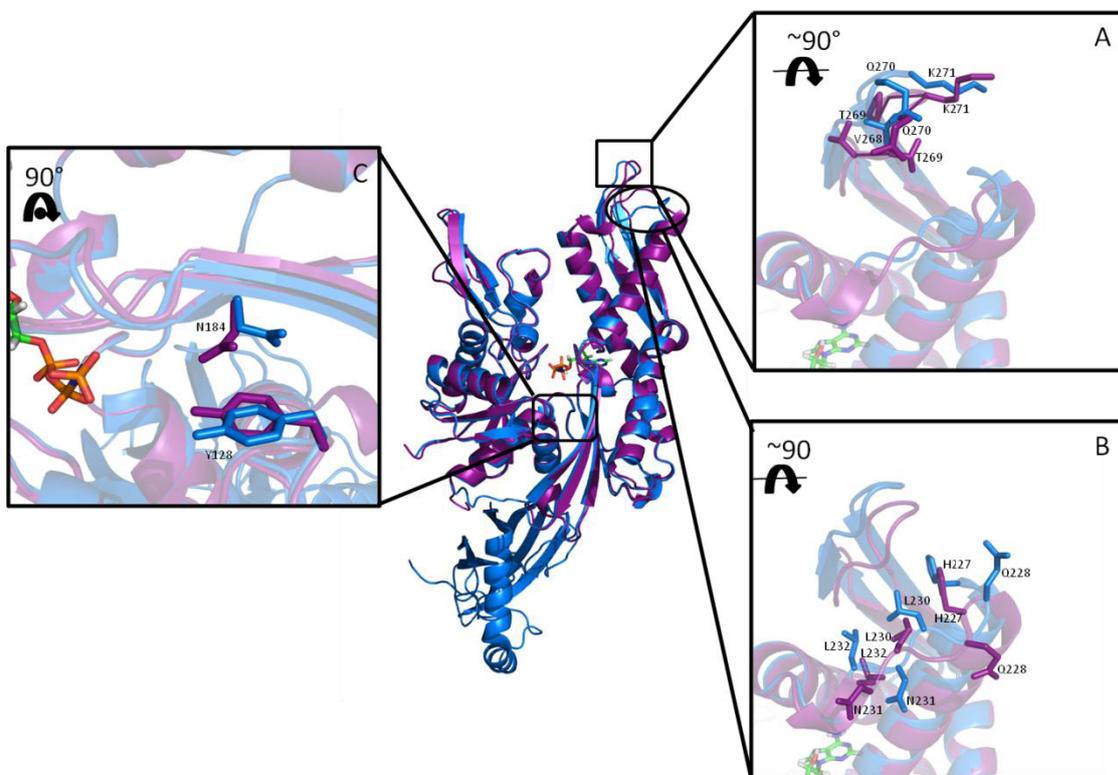


Figura 2.23 Comparación de las estructuras globales de *Mg*DnaKΔCt:AMP-PNP (azul) y *Mg*DnaK-NBD:AMP-PNP (morado). A-B) Detalle del movimiento de los bucles flexibles del dominio IIB. A) Detalle de los aminoácidos expuestos Val268, Tyr269, Gln270, Lys271. B) Detalle de los aminoácidos His227, Gln228, Leu230, Asn231, Leu232 C) Detalle de los aminoácidos expuestos Tyr128 y Asp184 que presentan diferentes orientaciones en las dos estructuras.

Pero los cambios más relevantes en las estructuras de *Mg*NBD:AMP-PNP y de *Mg*DnaKΔCt:AMP-PNP se encuentran en la coordinación del ión magnesio con el nucleótido. Como se ha comentado anteriormente, comúnmente, cuando un ión

magnesio se coordina con un nucleótido su coordinación habitual es octaédrica. Se puede observar un ejemplo de esta coordinación en la figura 2.21 A (*MgNBD:ADP*). Cuando se modeló el ión magnesio en la estructura de la *MgDnaKΔCt:AMP-PNP* se vio que su coordinación era tetragonal y que coordina con una molécula de agua, con el fosfato β del nucleótido y con los aminoácidos Asp12 y Asn17 (Figura 2.24 A). En cambio, el ión magnesio de la estructura de *NBD:AMP-PNP* presenta una geometría de coordinación diferente, en concreto tetragonal piramidal, apareciendo coordinado a cuatro moléculas de agua y al fosfato β del nucleótido (Figura 2.22 B y 2.24 B). La coordinación tetragonal piramidal de *MgNBD:AMP-PNP* (Figura 2.23 B), en realidad, es prácticamente una coordinación octaédrica pero le falta coordinarse con el nitrógeno presente entre los fosfatos β y γ del AMP-PNP lo que genera cierta inestabilidad y desorden en el lugar de unión al nucleótido. En el caso de *MgDnaKΔCt:AMP-PNP* la geometría de coordinación del magnesio es posiblemente consecuencia de la inestabilidad y desorden en el lugar de unión al nucleótido provocado por el AMP-PNP, juntamente con los cambios observados en los aminoácidos expuestos Tyr128 y Asp184 (Figura 2.24 C) debidos a contactos cristalinos que provocaron el desplazamiento del ión magnesio. Este desplazamiento resultó en una nueva coordinación observada para el magnesio en estas estructuras, la coordinación tetragonal que es muy posiblemente un tipo de coordinación artefactual porque el AMP-PNP no mimetiza perfectamente el ATP.

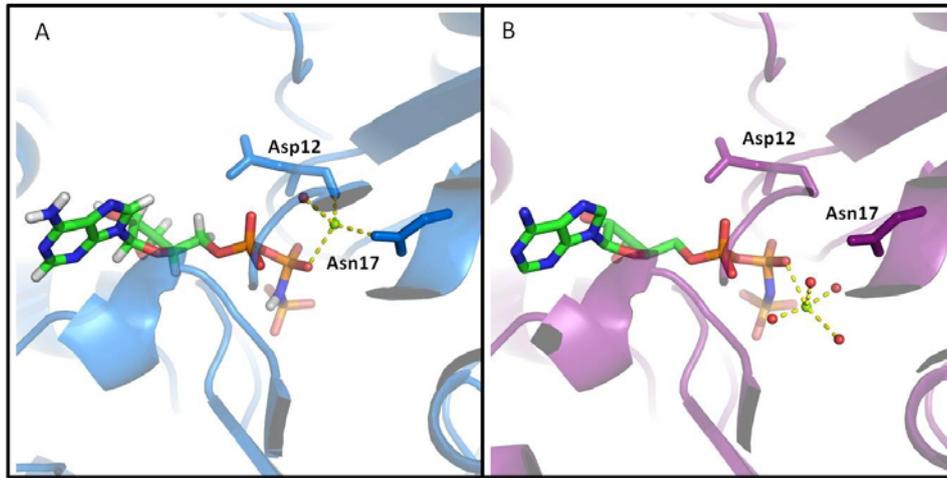


Figura 2.24 Detalle del lugar de unión a nucleótido donde se muestran las diferencias en la coordinación del magnesio observadas en las estructuras de A) *MgDnaKΔCt*:AMP-PNP y B) *MgNBD*:AMP-PNP.

- Comparación de las estructuras *MgNBD* y *MgDnaKΔCt* en presencia de ATP

Las diferencias observadas entre *MgDnaK-NBD*:ATP la y de *MgDnaKΔCt*:ATP parecen debidas al empaquetamiento de las moléculas en el cristal. Las principales diferencias ocurren en la disposición de los bucles expuestos: Tyr45 a Ile49 y Asp84 a Lys90 y Glu226 a Leu231. La resolución a la cual se resolvió la estructura de *DnaKΔCt*, 2.5 Å de resolución, impidió la observación y modelación del magnesio, aunque la disposición del nucleótido es idéntica en ambas estructuras. A parte de estos movimientos, debidos al empaquetamiento de las moléculas en el cristal, las estructuras son prácticamente iguales. En el sitio de unión a nucleótido se observan movimientos poco significativos en las cadenas laterales de los residuos Arg73, Tyr128, Asp177 y Asp349 y se pudo modelar una molécula de glicerol en la proximidad del fosfato y del ATP que había sido hidrolizado.

2.3 Estudio bioquímico y biofísico de las chaperonas de *M. genitalium*

Una vez determinada la estructura de *MgDnaK* Δ Ct y de su dominio de unión a nucleótido en presencia o ausencia de diferentes tipos de nucleótidos se empezó el estudio de la co-chaperona *MgDnaJ* y del factor intercambiador de nucleótidos *MgGrpE* sin los cuales *MgDnaK* no podría realizar su función.

2.3.1 Cristalización de GrpE de *M. genitalium* (*MgGrpE*)

La *MgGrpE* fue clonada en el vector de expresión pET21d (Novagen, Madison, WI, USA) y el clon se transformó en células de *E. coli* BL21 (DE3). Las mejores condiciones de expresión establecidas para la *MgGrpE* fueron induciendo el cultivo con 1 mM IPTG durante toda la noche a 20°C. La proteína se purificó mediante un primer paso de cromatografía de afinidad utilizando una columna HisTrap™ HP (GE Healthcare, Uppsala, Sweden) después del cual las fracciones enriquecidas en la proteína de interés se cargaron en una columna de gel filtración Superdex 200 16/60 (GE Healthcare, Uppsala, Sweden) equilibrada en el tampón 0.5 M NaCl, 0.02 M Tris-HCl pH 8 (Figura 2.25 A). El grado de pureza de la muestra se analizó por gel de poliacrilamida desnaturizante (Figura 2.25 B). El volumen de elución de *MgGrpE* en la Superdex 200 16/60 es de 74.9 mL que corresponde, según la calibración de la columna con

proteínas estándar de pesos moleculares conocidos, a un dímero en solución. Con esta muestra de proteína se realizaron múltiples cribados de cristalización ofertados por la Plataforma Automatizada de cristalografía (PCB-IRB-CSIC) tanto a 20°C como a 4°C y a diferentes concentraciones, llegando a alcanzar los 40mg/mL. Prácticamente resultaron el 80% de las gotas transparentes. A pesar de que el perfil cromatográfico mostraba una proteína estable con un elevado grado de pureza (Figura 2.25 A y B) los ensayos de cristalización no dieron resultados positivos, posiblemente debido, a la elevada solubilidad de la proteína posiblemente proveniente de su alto contenido en lisinas (14.7%).

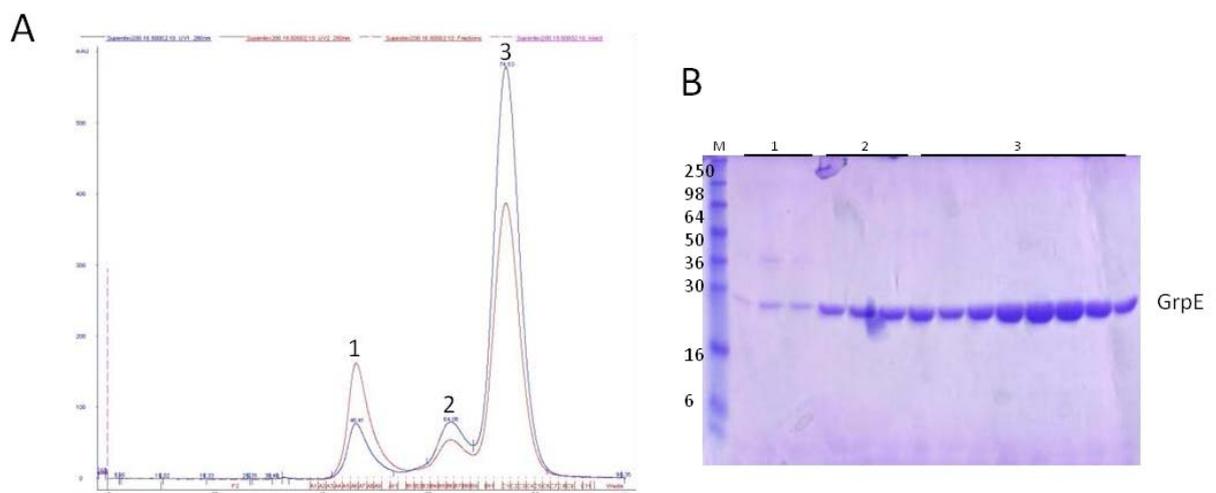


Figura 2.25 A) Perfil cromatográfico de *MgGrpE* en una columna de gel filtración Superdex 200 16/60. B) Análisis de las fracciones eluídas de la columna de gel filtración por gel desnaturalizante SDS-PAGE de acrilamida al 15%.

Las dificultades en la obtención de cristales después de los múltiples cribados de cristalización hizo que se buscaran alternativas para cristalizar *MgGrpE*. Así, se realizó un alineamiento de secuencias entre la *MgGrpE* y GrpEs de otros organismos que ya

fueron cristalizadas (Figura 2.26). Este alineamiento nos sirvió de base para preparar nuevos constructos que favorecieran la cristalización de *MgGrpE*. La secuencia aminoacídica de la *MgGrpE* fue alineada con las secuencias de las tres GrpE cristalizadas hasta el momento: la de *E. coli* (código PDB 1DKG) (*EcGrpE*)(Harrison, Hayer-Hartl et al. 1997), la de *G. kaustophilus* (código PDB 4ANI) (*GkGrpE*) (Wu, Naveen et al. 2012) y la de *Thermus thermophilus* (código PDB 3A6M) (Nakamura, Takumi et al. 2010) (*TtGrpE*) (Figura 2.24). El alineamiento fue realizado utilizando el servidor de alineamiento de secuencias múltiples T-coffee (Di Tommaso, Moretti et al. 2011) y llevó a que se identificaran las regiones más conservadas de esta familia de proteínas. En el alineamiento se diferencian claramente dos regiones que se denominaron como: región N-terminal, donde la homología entre GrpEs es mucho menor debido a la presencia de inserciones o deleciones y donde las zonas homologas presentan menor identidad; y la región C-terminal, donde se observa una mayor identidad de secuencia (Figura 2.26).

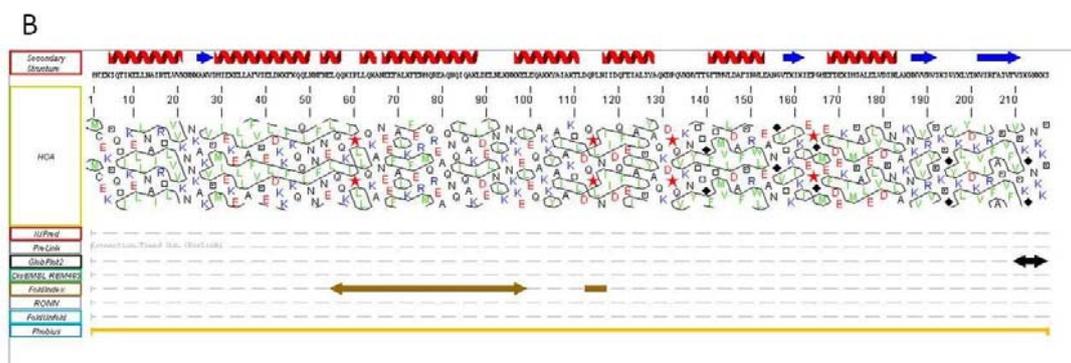


Figura 2.27 A) Predicción de la estructura secundaria realizada por el servidor PsiPred de la secuencia de GrpE *M. genitalium*. PsiPred indica los elementos de estructura secundaria: los cilindros rosas indican las hélices α y las flechas amarillas hebras β . B) Predicción realizada por el servidor MeDor que predice regiones desestructuradas. Las hélices α se representan en forma de hélices rojas y las láminas β en forma de flechas azules.

Partiendo de esta información y de la comparación de las estructuras tridimensionales de las varias GrpEs resueltas hasta el momento se diseñaron dos nuevos constructos con el objetivo de favorecer la cristalización de la *MgGrpE*:

- Al primero se deletaron los primeros 26 aminoácidos del N-terminal porque, tal como se puede apreciar en el alineamiento realizado por T-coffee (Figura 2.26) y los resultados obtenidos por MeDor y PsiPred (Figura 2.27 A y B) estos residuos están poco conservados entre las diferentes GrpEs, a pesar que esta región presenta en todos los casos, una estructura secundaria en forma de hélice; y porque no fue posible modelar esta región en las estructuras cristalinas de estas GrpEs, sea porque esta fue previamente deletada o porque está desordenada en el cristal. Esta nueva versión de la proteína, que se denominó *MgGrpE-Ct_26-217*, fue clonada en el vector de expresión pET21d (Novagen, Madison, WI, USA) y expresada en *E. coli* BL21 (DE3). La proteína fue entonces purificada del mismo modo que la proteína salvaje. Todos los intentos de cristalización fueron frustrados.

- Al segundo constructo, a la *MgGrpE*, se le delecionaron 58 residuos del N-terminal, en particular, se eliminó una larga hélice α responsable de la dimerización de la proteína (Nakamura, Takumi et al. 2010). Este constructo fue denominado *MgGrpE-Ct_58-217* y se expresó y purificó de la misma manera que el anterior. Pero este constructo resultó en una proteína muy inestable la cual, no nos permitió realizar más estudios, seguramente, debido a que la hélice responsable de la dimerización proporciona estabilidad a la proteína. Hasta el momento la GrpE en solución siempre ha sido dimérica y la delección de la hélice provocaría que la GrpE fuera monomérica, la cual parece ser muy inestable.

2.3.2 Estudio de la interacción entre *MgDnaK* Δ Ct, *MgDnaJ* y *MgGrpE*

El primer cribado para la detección de interacciones binarias proteína-proteína se realizó mediante la técnica de Resonancia Superficial por Plasmón (del inglés, Surface Plasmon Resonance, SPR) una técnica muy sensible que detecta cambios en el índice de refracción del orden de los 100nm y que permite detectar interacciones en tiempo real sin que haya que marcar la proteína y además, también de permite estudiar la cinética de una interacción (Jonsson, Fagerstam et al. 1991).

Considerando el ciclo realizado por las chaperonas (ver sección 2.1.6) no parece trivial que la cola de histidinas esté unida al dominio SBD mimetizando un péptido sin plegar. Por ese motivo, el primer cribado se realizó inmovilizando *MgDnaK* Δ Ct y dos de sus

mutantes, *MgDnaKΔCt-L523K* y *MgDnaKΔCt-L523D*, en el sensor CM5 (GE Healthcare, Uppsala, Sweden). Los dos mutantes fueron diseñados expresamente para evitar la presencia de la cola de histidinas en el sitio de unión del SBD (ver sección 2.1.6).

- Estudio de la Interacción entre *MgDnaKΔCt* y *MgDnaJ* por SPR

Una de las razones que llevó a que se testaran las interacciones entre las proteínas de *M. genitalium* que contienen un dominio J y DnaK fue para investigar sus posibles implicaciones en el mecanismo de motilidad de *M. genitalium*, del mismo modo que ocurre en otros organismos como *Myxococcus* o *Chlamydomonas* (Johnson and Rosenbaum 1992, Yang, Geng et al. 1998). Los experimentos de localización de *MgDnaK* (A. M. Mariscal y L. Gonzalez-Gonzalez, en preparación) muestran que *MgDnaK* se localiza principalmente en la OT, sugiriendo que *MgDnaK* puede estar implicada motilidad. Por otro lado los experimentos realizados *in vivo* por A. Broto (resultados en preparación) que consistieron en preparar tres modificaciones del N-terminal de la proteína MG200; la primera consistió en la supresión del dominio J de MG200, la segunda consistió en la sustitución del dominio J de MG200 por el dominio J de MG019 y el tercero consistió en la sustitución del dominio J y el dominio rico en G/F de MG200 por el dominio J y el dominio rico en G/F de MG019. El estudio del fenotipo de estos mutantes reveló la importancia del dominio J de MG200 en la expresión de las principales adhesinas celulares. Concretamente, se observó que en células donde se había delecionado este dominio no había expresión de adhesinas y, consecuentemente, las células mutantes presentaban un fenotipo de células no móviles y no adherentes. También se observó que cuando se sustituía la región G/F de

MG200, por la región equivalente de MG019, las células mutantes eran viables pero las adhesinas tampoco eran expresadas. Conjuntamente estos datos indican que podría haber una relación indirecta entre el proceso de motilidad celular y el sistema de chaperonas de DnaK-DnaJ-GrpE. Con los experimentos de SPR se pretendió verificar cuál de las proteínas de *M. genitalium*, que contiene un dominio J, funciona como co-chaperona de DnaK.

Como se describe en la sección 1.2.5 *M. genitalium* contiene tres proteínas que poseen un dominio J: MG200, MG019 y MG002 (Figura 1.11). MG019 se compone de cuatro regiones claramente diferenciadas: el dominio J, la región G/F, un dedo de zinc y por último, un dominio C-terminal. Estos dominios son los que normalmente forman las co-chaperonas activas de *MgDnaK*, que son homólogas a la co-chaperona DnaJ de *E.coli*. MG200, en cambio, de estas cuatro regiones sólo contiene el dominio J, la región G/F y el dominio C-terminal. Por último, MG002 sólo posee el dominio J y un dominio C-terminal. En este estudio no se testó la interacción de *MgDnaK* con el dominio J de MG002 porque aún no se había determinado *in vivo* su implicación en el funcionamiento de la OT. Además MG002 para poder ser expresada de manera recombinante en *E. coli* se deben mutar los 6 codones UGA, que codifican para triptófanos en *M. genitalium* a UGG, ya que en *E. coli* el codón UGA codifica para un codón STOP. Además, MG002, que aún no ha sido localizada en *M. genitalium*, está clasificada como una DnaJ de tipo III (ver sección 1.2.5) y su función varía según el compartimento celular en el que se encuentra localizada (Walsh, Bursac et al. 2004).

Se prepararon entonces cuatro constructos distintos de diferentes regiones de MG019 y de MG200: i) el dominio J de MG019 (aminoácidos Met1 a Glu79, MG019-J), ii) el

dominio J y la región rica en glicinas y fenilalaninas (G/F) de MG019 (aminoácidos Met1 a Asn120, MG019-J-G/F), iii) el dominio J de MG200 (aminoácidos 1 a 74, MG200-J) y iv) el dominio J y la región rica en G/F de MG200 (aminoácidos 1 a 120, MG200-J-G/F). La región rica en G/F (~residuos 75 a 110) de la DnaJ de *E.coli* es una región muy desordenada y dinámica cuya función está relacionada con permitir la flexibilidad del dominio J. Este dinamismo permite la reorientación del dominio J facilitando que la región de reconocimiento del sustrato del dominio J pueda interactuar con el sustrato. Esta supuesta flexibilidad podría permitir que mientras el dominio J de la DnaJ cede el sustrato al SBD de la DnaK pueda a la vez haber una interacción con el NBD de DnaK y con otra región de DnaJ. Favoreciendo la transferencia del sustrato de DnaJ a DnaK (Huang, Flanagan et al. 1999). Recientemente, se ha evidenciado que la región rica en G/F interactúa con el sitio de unión a sustrato del SBD de DnaK en ausencia de sustrato, y este podría ser un segundo sitio de unión entre DnaJ y DnaK, posiblemente facilitando una posible interacción secuencial, es decir, una detrás de la otra. (Ahmad, Bhattacharya et al. 2011). Además, la región rica en G/F también ha sido descrita como una región capaz de diferenciar conformaciones del sustrato que puede adaptar para que se formen complejos entre el dominio J de DnaJ y el sustrato (Perales-Calvo, Muga et al. 2010).

Los experimentos de interacción entre *MgDnaK* y las diferentes versiones de las diferentes *MgDnaJs* se realizaron en presencia de ATP o AMP-PNP porque *MgDnaKΔCt* siempre tiene unido un nucleótido (ATP o ADP) durante el ciclo de plegamiento de los sustratos (ver sección 2.1.6) y porque además es muy inestable en ausencia de estos.

Para realizar los estudios de interacción proteína-proteína por SPR se inmovilizó en un biosensor CM5 (GE, Healthcare Life sciences, Uppsala, Sweden) *MgDnaKΔCt* y dos de sus mutantes, *MgDnaKΔCt-L522D* y *MgDnaKΔCt-L522K*. A uno de los canales del sensor no se le unió ninguna proteína porque se utilizó como blanco de los experimentos. Para detectar las interacciones se hicieron fluir por los canales del sensor diferentes concentraciones de MG019 (dominio J y J-G/F) y de MG200 (dominio J y J-G/F). Estos experimentos se realizaron en un tampón 0.15 M NaCl, 0.02 M Tris pH8, 1 mM MgCl₂, 0.005% (v/v) surfactante P20 y 0.5mM de AMP-PNP o ATP según el experimento. Las mediciones se realizaron a temperatura ambiente.

Los resultados obtenidos por SPR para detectar una posible interacción entre *MgDnaK* y sus mutantes y alguno de los dominios J de las distintas *MgDnaJ* no revelaron ninguna interacción (Figura 2.28 A). La ausencia de interacción entre los mutantes de *MgDnaKΔCt* estudiados y las diferentes versiones de las *MgDnaJs* puede deberse a que el SBD de *MgDnaK* esté ocupado por el enlazador de alguna molécula vecina. Ya se vio anteriormente que estos mutantes en solución no son muy estables y acaban formando oligómeros en solución (ver sección 2.1.4) lo que impediría una posible interacción con DnaJ. Por lo tanto, considerando que el SBD de todas las versiones de *MgDnaK* testadas muy posiblemente está ocupado, las condiciones experimentales ensayadas no fueron las idóneas para detectar la interacción.

El ciclo de las chaperonas es muy complejo (ver sección 1.2.3.) y cumplir *in vitro* los requisitos necesarios para detectar la formación de complejo DnaK-DnaJ no es trivial. Primero, hay que tener en cuenta que en el constructo de DnaKΔCt con el que se trabajó (y del cual se ha resuelto la estructura) la cola de histidinas en el C-terminal,

utilizada para purificar la proteína, está unida al sitio de unión a sustrato mimetizando un sustrato y bloqueando el sitio de unión. Si nos basamos en la literatura disponible sobre el ciclo de las chaperonas (Buchberger, Valencia et al. 1994, Laufen, Mayer et al. 1999, Swain, Dinler et al. 2007, Mayer 2013, Zuiderweg, Bertelsen et al. 2013, Melero, Moro et al. 2015) se espera que DnaJ unida a un sustrato lo acerque a DnaK la cual está unida a ATP. Cuando se hidroliza el ATP, la DnaK pasa a un estado conformacional de alta afinidad para el sustrato. En las condiciones experimentales en que se testaron las interacciones DnaK-DnaJ, las diferentes versiones de DnaK en presencia de AMP-PNP estaban posiblemente en una conformación con baja afinidad para el sustrato lo que explicaría por qué no se detectó ninguna interacción. Otra posibilidad sería que a elevadas concentraciones de DnaJ, esta podría ejercer una función de sustrato, uniéndose al SBD y estimulando la hidrólisis de ATP, proceso artefactual y no fisiológico (Laufen, Mayer et al. 1999). Además, DnaJ también puede unirse al SBD de DnaK Δ Ct por la región rica en G/F. Este fenómeno fue observado por NMR, evidenciando la posibilidad que la interacción de DnaJ a DnaK sea secuencial (Li, Qian et al. 2003). Los ensayos de SPR realizados en este estudio y descritos anteriormente, no permitieron la detección de ninguna interacción, indicando que estos estudios no deberían hacerse con una DnaK en que su dominio SBD se encuentre ocupado. Con todo, a pesar de los múltiples intentos, hasta el momento no fue posible obtener *MgDnaK Δ Ct* con el SBD libre, razón por la cual, se dejaron de realizar estudios de interacción de DnaK con los dominios J de las diferentes *MgDnaJs*.

Han sido múltiples los estudios para intentar detectar la interacción entre DnaK-DnaJ a lo largo de los años, muchos de ellos han sido frustrados. Suh y colaboradores y Gross y colaboradores, respectivamente en 1998 y 1999, demostraron que constructos de

dominios J unidos a sus respectivas regiones ricas en G/F de DnaJ unían a DnaK del mismo modo que DnaJ salvaje o entera lo hace, mientras que no se detectaba unión de DnaK con constructos de DnaJ que sólo contienen el dominio J. Estos estudios sugirieron que DnaJ interacciona en dos lugares distintos de DnaK, en el NBD y en el SBD. Los dominios J están formados por cuatro hélices (I a IV). El bucle formado entre las hélices II y III contiene el motivo His-Pro-Asp (HPD) , el cual, está muy conservado. Este motivo ha sido descrito como parte esencial en la interacción con DnaK. (Pellecchia, Szyperski et al. 1996, Greene, Maskos et al. 1998, Walsh, Bursac et al. 2004) (ver sección 1.2.5 y Figura 1.11). Parece ser que para la interacción con DnaK, la hélice II del dominio J que se caracteriza por su alto contenido en lisinas y argininas, que son muy sensibles a mutaciones y que también está muy conservada, juega un papel importante (Wall, Zylicz et al. 1994, Maler, Blankenship et al. 2000). Se ha propuesto que la interacción que se da entre DnaK-DnaJ es electroestática y la función del motivo HPD es estabilizarla (Wittung-Stafshede, Guidry et al. 2003, Ahmad, Bhattacharya et al. 2011) mediante la interacción de DnaJ con la hendidura existente entre los dos lóbulos del NBD de DnaK (Suh, Burkholder et al. 1998).

Otro factor importante a tener en cuenta cuando se estudia la interacción entre DnaK y DnaJ es en presencia de qué nucleótido hay que realizar el ensayo. La interacción que se pretende detectar depende de la presencia de ATP (Wawrzynow, Banecki et al. 1995) y según la literatura no ocurre en presencia de ADP (Wittung-Stafshede, Guidry et al. 2003). La presencia de ATP y/o ADP ha generado controversia después del análisis de cinco estructuras de complejos DnaK-DnaJ (códigos PDB 2QWO, 2QWP, 2QWQ, 2QWR y 1XBL)(Jiang, Taylor et al. 2003, Jiang, Maes et al. 2007, Ahmad, Bhattacharya et al. 2011, Zuiderweg, Bertelsen et al. 2013).

Cuatro de estas estructuras fueron resueltas por cristalografía de rayos-X, varios constructos del NBD de DnaK de *Bos taurus* fueron co-cristalizadas con el dominio J de la proteína auxilina de *Bos taurus* también. En estas estructuras el NBD estaba ocupado por ADP (código PDB : 2QWO y 2QWP) o por AMP-PNP (código PDB 2QWQ y 2QWR) y la formación de los complejos se consiguió uniendo covalentemente el NBD y el dominio J un puente químico formado entre una asparagina y una cisteína. La disposición del NBD con respecto al dominio J en la estructura tridimensional del complejo parece ser artefactual siendo el resultado de la unión covalente por el puente disulfuro y no siendo la conformación que presentaría este complejo en solución (Jiang, Taylor et al. 2003, Jiang, Maes et al. 2007).

La estructura tridimensional del complejo *EcDnaK-EcDnaJ* en presencia de ADP y de un péptido (código PDB 1XBL) resuelta por RMN, las dos proteínas no fueron artificialmente unidas de manera covalente como en el caso anterior. En este complejo se observó que el dominio J se unía a través de su hélice II cargada positivamente a una área cargada negativamente del NBD de DnaK, en particular en la región ²⁰⁶EIDEVDGEKTFEVLAT²²¹. (Ahmad, Bhattacharya et al. 2011, Zuiderweg, Bertelsen et al. 2013). En *M. genitalium* cuando realizamos un aliniamento de secuencia encontramos sólo que la región conservada corresponde a ¹⁹⁴TFEVLAT²⁰¹ siendo esta una región conservada mucho menor que en la *EcDnaK* de *E. coli* y posiblemente, puede ser la causa de que la interacción entre *MgDnaK*-DnaJ sea menos específica y más dinámica.

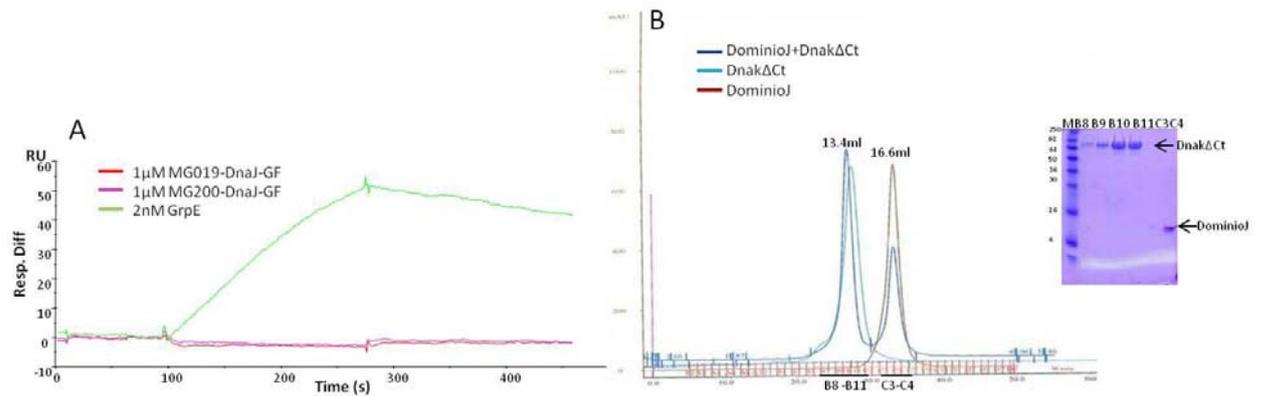


Figura 2.28 A) Estudios de la interacción entre *MgDnaKΔCt*, *MgDnaKΔCt-L523K* y *MgDnaKΔCt-L523D*, que fueron inmovilizadas en un sensor, y diferentes constructos de fragmentos conteniendo el dominio J de diferentes proteínas de *M. genitalium*. Se muestra la respuesta en RU cuando se hace fluir por la celdilla MG019-J-G/F o MG200-J-G/F.. B) Perfil cromatográfico de *MgDnaKΔCt*, MG200-J y de *MgDnaKΔCt*- MG200-J en una columna de gel filtración Superdex 200 10/300. Análisis de las fracciones eluidas de la columna de gel filtración por gel desnaturalizante SDS-PAGE de acrilamida al 18%.

- Estudio de la interacción entre *MgDnaKΔCt* y *MgGrpE*

Estudio de la interacción entre MgDnaKΔCt y MgGrpE por SPR

La siguiente interacción que se estudió fue entre *MgDnaKΔCt* y su factor intercambiador de nucleótidos, *MgGrpE*. Como se explica en la sección 1.2.5, cuando DnaK se encuentra en un estado de ADP y con el SBD ocupado por un sustrato, GrpE induce un cambio conformacional en DnaK que promueve el intercambio de nucleótido (de ADP a ATP) y la liberación del sustrato. Para realizar este experimento se utilizó un sensor CM5 (GE Healthcare Life Science, Uppsala, Sweden) donde se habían inmovilizado *MgDnaKΔCt*, *MgDnaKΔCt-L522D* y *MgDnaKΔCt-L522K* (ver secciones 2.3.2 y 4.5.4). Se hizo fluir por los distintos canales del sensor diferentes concentraciones de *MgGrpE* usando el tampón 0.15 M NaCl, 0.02 M Tris-HCl pH8, 1 mM MgCl₂, 0.005% (v/v) surfactante P20 y 0.5 mM ATP o AMP-PNP para realizar el ensayo. El experimento que se realizó a temperatura ambiente.

Los resultados obtenidos en el ensayo de SPR permitieron detectar una interacción muy fuerte entre *MgDnaKΔCt*, *MgDnaKΔCt-L523K* y *MgDnaKΔCt-L523D* y *MgGrpE* en presencia de ATP (Figuran 2.29 A). Los resultados obtenidos no permitieron calcular las constantes de asociación y disociación del complejo porque las curvas de asociación entre las diferentes variantes de DnaK y GrpE son bifásicas ocurriendo en un primer momento una asociación muy rápida que luego se ralentiza. Cuando se producen esta clase de respuestas no es posible ajustar el tipo de interacción a uno de los modelos estándar (Rich and Myszka 2007). Lo que también imposibilitó el cálculo de la constante de afinidad del complejo fue la lenta disociación entre las proteínas (Figura 2.29 A). Por ello tuvieron que regenerarse las superficies de cada uno de los canales del sensor después de cada inyección con 0.05% (w/v) SDS, lo que contribuyó a una desnaturalización parcial de las moléculas inmovilizadas y pérdida de material en cada uno de los canales. Las curvas de interacción obtenidas para el mutante *MgDnaKΔCt-L523D* y para *MgDnaKΔCt*, después de la inyección de *MgGrpE* en el sensor son idénticas. Por el contrario, la respuesta cuando *MgGrpE* se inyecta sobre el mutante *MgDnaKΔCt-L523K* es ligeramente menor, lo que puede deberse a una menor cantidad de proteína inmovilizada en el sensor (ver materiales y métodos sección 4.7.1). Así, el hecho que la interacción entre *MgDnaKΔCt* y sus mutantes sea idéntica indica que los mutantes, al igual que la *MgDnaKΔCt* (ver sección 2.1.6), tienen el SBD ocupado.

A continuación se repitió el mismo experimento descrito anteriormente pero esta vez en presencia de AMP-PNP e inyectando GrpE a 5 y 10 nM. Se detectó una interacción todavía más fuerte en presencia de AMP-PNP (Figura 2.29 B) que en presencia de ATP (Figura 2.29 A), lo que también impidió el cálculo de la cinética de la reacción. En estas mediciones se alargó el tiempo de asociación/disociación de *MgGrpE* para que se

percibiera mejor la curva de respuesta del complejo (Figura 2.29 B). Aunque no se pudieron calcular las constantes cinéticas de la reacción estos ensayos permitieron diseñar nuevas estrategias para los ensayos de cristalización del complejo ya que en presencia de AMP-PNP la interacción es más fuerte.

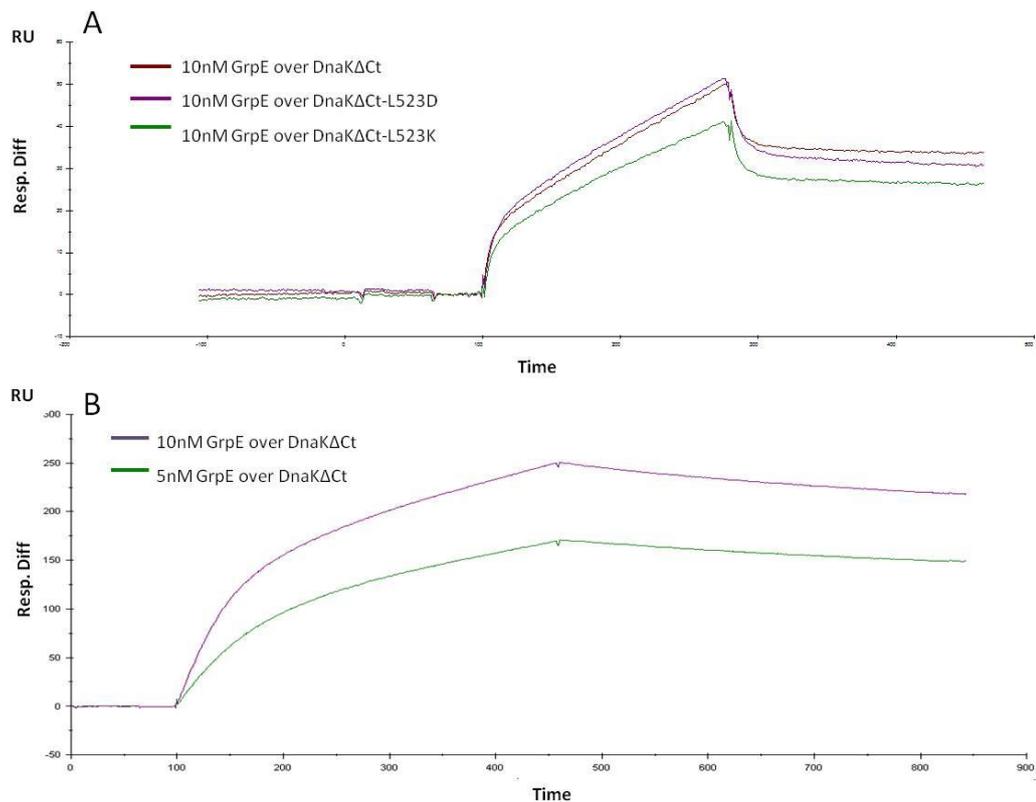


Figura 2.29 Estudios de la interacción entre *MgDnaKΔCt*, *MgDnaKΔCt-L523K* y *MgDnaKΔCt-L523D*, que fueron inmovilizadas en un sensor, y *MgGrpE*. Se muestra la respuesta en RU cuando se hace fluir por los canales del sensor *MgGrpE*. A) Curvas de respuesta obtenidas al inyectar 10 nM *MgGrpE* en presencia de ATP. B) Curvas de respuesta obtenidas después de inyectar 5 nM y 10 nM de *MgGrpE* en presencia de AMP-PNP en el canal donde se inmovilizó *MgDnaKΔCt*.

Purificación y cristalización del complejo MgDnaKΔCt-GrpE y de sus variantes

- Complejo MgDnaKΔCt-GrpE

Una vez detectada la interacción entre *MgDnaKΔCt* y *MgGrpE* por SPR esta se validó y estudió por cromatografía de gel filtración. El complejo se preparó purificando cada proteína separadamente, siguiendo los protocolos de expresión y purificación que ya se habían establecido (ver secciones de material y métodos 4.4), mezclando los dos componentes y dejando incubar a 4°C durante una hora. Como *MgDnaKΔCt* es más estable en presencia de nucleótidos los experimentos se realizaron en presencia de estos. La mezcla de las dos proteínas fue entonces inyectada a una columna de gel filtración Superdex 200 10 /300 (GE Healthcare, Uppsala, Sweden) y el complejo entre *MgDnaKΔCt-GrpE* eluyó en un solo pico con volumen de elución de 12.34 mL. Se analizaron las fracciones resultantes en un gel desnaturizante SDS-PAGE 12% acrilamida que evidenció la formación del complejo (Figura 2.30 A).

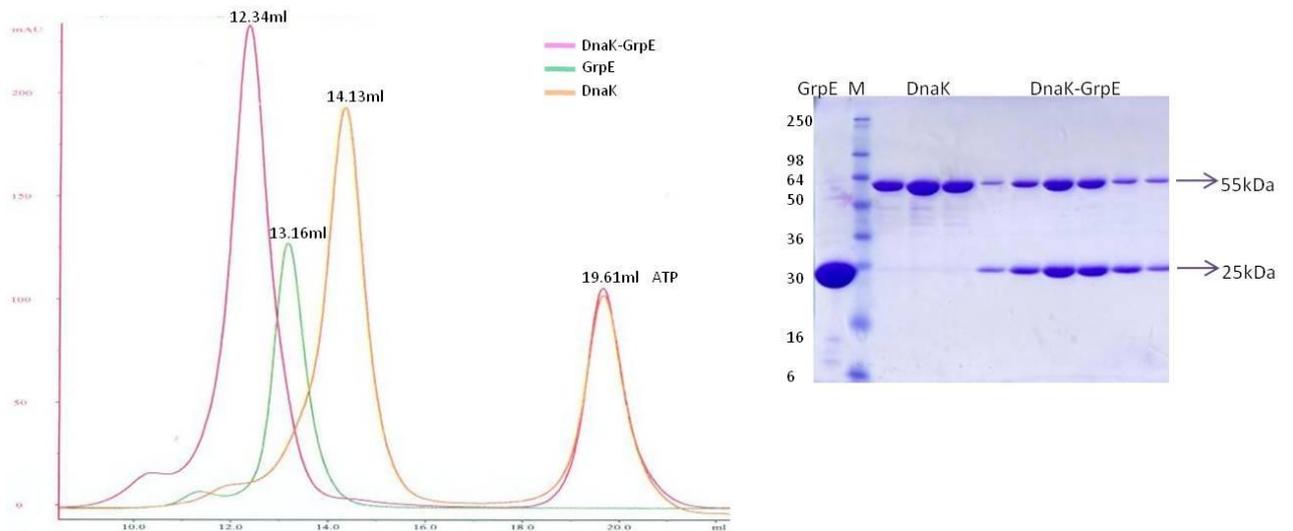


Figura 2.30 Perfiles cromatográficos obtenidos para *MgDnaKΔCt*, *MgGrpE* y *MgDnaKΔCt-GrpE* inyectadas en una columna de gel filtración Superdex 200 10/300 y análisis de las fracciones recogidas durante la elución de los picos en un gel desnaturalizante SDS-PAGE al 15% acrilamida

La fracciones correspondientes al complejo se concentraron hasta 10 mg/mL en un tampón compuesto por 0.5 M NaCl y 0.02 M Tris-HCl pH 8 utilizando un dispositivo de ultrafiltración con un punto de corte de 10 kDa (Amicon Ultra MWCO 10 kDa Darmstadt, Germany). Los ensayos iniciales de cristalización se realizaron a 20°C y a 4°C en placas Greiner de 96 condiciones en presencia de ATP y también en ausencia de nucleótido. No se obtuvieron cristales en ninguno de los casos. Se decidió entonces, cambiar el protocolo de purificación del complejo ya que en los diferentes intentos de purificar el complejo se verificó que el volumen de elución del pico no era reproducible y en el perfil cromatográfico siempre se observaba un exceso de alguno de los componentes del complejo. A pesar de la adición de pasos de purificación por gel filtración, con el objetivo de eliminar el exceso de alguna de las proteínas esto no fue posible. Así que, para evitar el exceso de uno de los componentes cuando se prepara la mezcla de las dos proteínas y lograr obtener una muestra más homogénea se optó por co-purificar las proteínas. Esta co-purificación consistió en expresar las proteínas

de manera independiente pero la lisis celular se realizaba de manera conjunta. En un primer paso de purificación del complejo por cromatografía de afinidad utilizando una columna HisTrap™ 5 mL (GE Healthcare, Uppsala, Sweden) el complejo fue eluído mediante un gradiente lineal de imidazol. Las fracciones correspondientes al complejo se cargaron en una columna de gel filtración Superdex 200 16/60 (GE Healthcare, Uppsala, Sweden) y las fracciones correspondientes al pico se analizaron por gel de poliacrilamida desnaturante (Figura 2.31). Varios ensayos de purificación se realizaron en tampón 0.5 M NaCl, 0.02 M Tris-HCl pH 8, en el que también se realizaron múltiples ensayos de cristalización a diferentes concentraciones del complejo y a diferentes temperaturas, 4°C y 20°C, sin que se obtuvieran cristales. La dificultad en obtener cristales llevó a que se realizaran nuevos ensayos de cristalización pero esta vez, el tampón en el que se purificó por gel filtración el complejo fue 0.05 M citrato de sodio pH 5.5, el mismo tampón en que se encontraba el complejo *GkDnaK-GkGrpE* (código PDB: 4ANI) cuando cristalizó (Wu, Naveen et al. 2012).

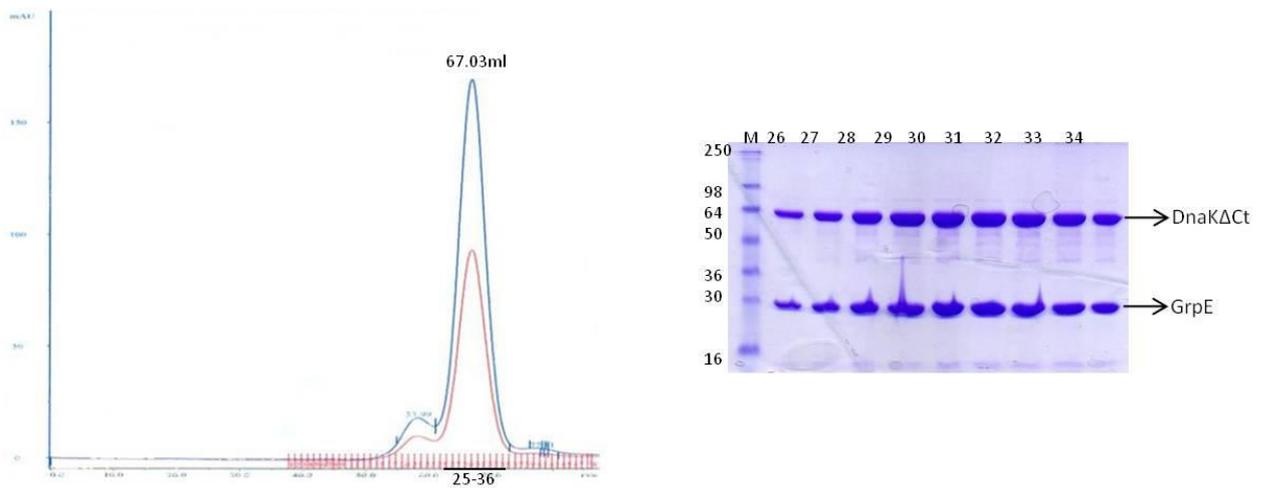


Figura 2.31 Perfil cromatográfico de la purificación del complejo *GkDnaKΔCt-GrpE* a través de una columna de gel filtración Superdex 200 10/300 en tampón 0.5 M NaCl, 0.02 M Tris-HCl pH 8. Las fracciones correspondientes al pico se analizaron en un gel desnaturalizante SDS-PAGE 15% acrilamida.

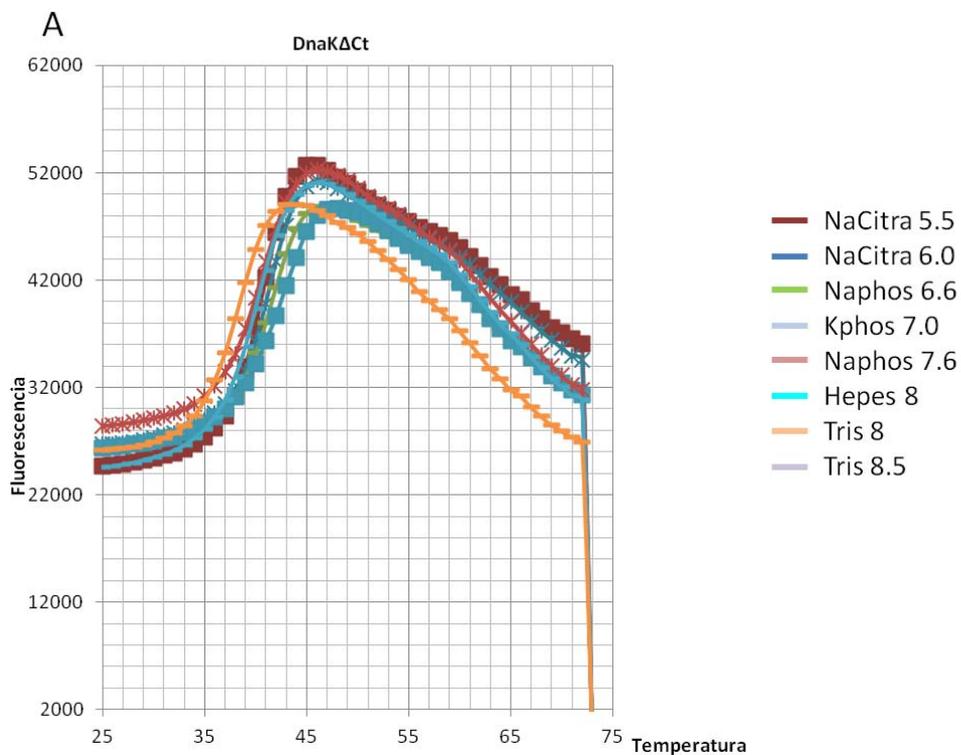
Este complejo está formado por un constructo de DnaK en el que se le ha deletado la parte C-terminal (código PDB 2V7Y de hecho, es la estructura que se usó de base para diseñar el constructo de *MgDnaKΔCt* que se cristalizó) y un constructo que expresa la *MgGrpE* entera, y fue preparado utilizando las proteínas purificadas por separado y que posteriormente fueron mezcladas y re-purificadas conjuntamente.

La estabilidad y homogeneidad del complejo *MgDnaKΔCt-MgGrpE* obtenido por co-purificación fue analizado por dispersión de luz dinámica (de inglés *Dynamic Light Scattering, DLS*) en 0.05 M citrato de sodio pH 5.5 después que la muestra hubiera sido congelada a -20°C . El índice de polidispersidad de la muestra fue del 32.1% calculado dividiendo la polidispersidad de la muestra por el radio hidrodinámico (R_H). Como valores mayores al 20% indican que las proteínas no son monodispersas el resultado del DLS mostró que el complejo no era una muestra homogénea.

Las dificultades en obtener cristales del complejo *MgDnaKΔCt-MgGrpE* hicieron que se decidiera hacer un cribado de diferentes tampones para ver como afectaban a la

estabilidad del complejo. Idealmente, este ensayo se debería haber realizado con el complejo formado con las dos proteínas, pero los resultados hubieran sido muy difíciles de interpretar. Por esa razón, se decidió hacer el cribado de tampones con *MgDnaKΔCt* y *MgGrpE* por separado y encontrar un tampón común, en el que ambas proteínas presentaran una mayor estabilidad. Para realizar este ensayo se utilizó la técnica de la fluorimetría diferencial de barrido (del inglés *Differential Scanning Fluorimetry, DSF*). En los ensayos se midió la variación de la fluorescencia en función de la temperatura que se fue incrementando desde los 25°C hasta los 75°C. El estudio se realizó añadiendo 50 μg de proteína a una concentración final de 0.2mg/mL y 1x del fluoróforo Sypro Orange® (Life Technologies, Grand Island, New York, USA) en cada tampón que se quería testar. A medida que aumenta la temperatura la proteína se desnatura dejando expuestas áreas hidrofóbicas donde se une el colorante que entonces emite una señal de fluorescencia. Los ensayos de DSF permitieron determinar que la temperatura de desnaturación de *MgDnaKΔCt* es de 40°C (Figura 2.32 A) mientras que la de *MgGrpE* es de 45°C (Figura 2.32 B), temperaturas que indican que ambas proteínas son inestables en las condiciones experimentales utilizadas porque presentan temperaturas bajas de desnaturación. En ambos casos el pH en el que son más estables está alrededor de 8 aunque en tampones con diferente naturaleza química. En el caso de *MgDnaKΔCt* el mejor tampón fue considerado el tampón fosfato de potasio a un pH7-7.6 y en el caso de *MgGrpE* el mejor tampón se consideró el Tris-HCl a pH8. Parece poco razonable que proteínas que funcionan en condiciones de choque térmico, es decir, que tienden a expresarse cuando las células son expuestas a condiciones "estresantes" o menos favorables no tengan una temperatura de fusión más elevada. Se esperaba que los resultados de DSF

indicasen que las dos proteínas fueran más estables, exhibiendo temperaturas de desnaturalización más elevadas. Dos posibles explicaciones son que *M. genitalium*, como patógeno humano, no acostumbra estar expuesto a elevadas temperaturas y que las condiciones experimentales en las que se manipularon *MgDnaKΔCt* y *MgGrpE* no son idóneas para que las proteínas sean más estables. Finalmente se decidió seguir trabajando con el complejo utilizando el tampón Tris-HCl a pH 8.



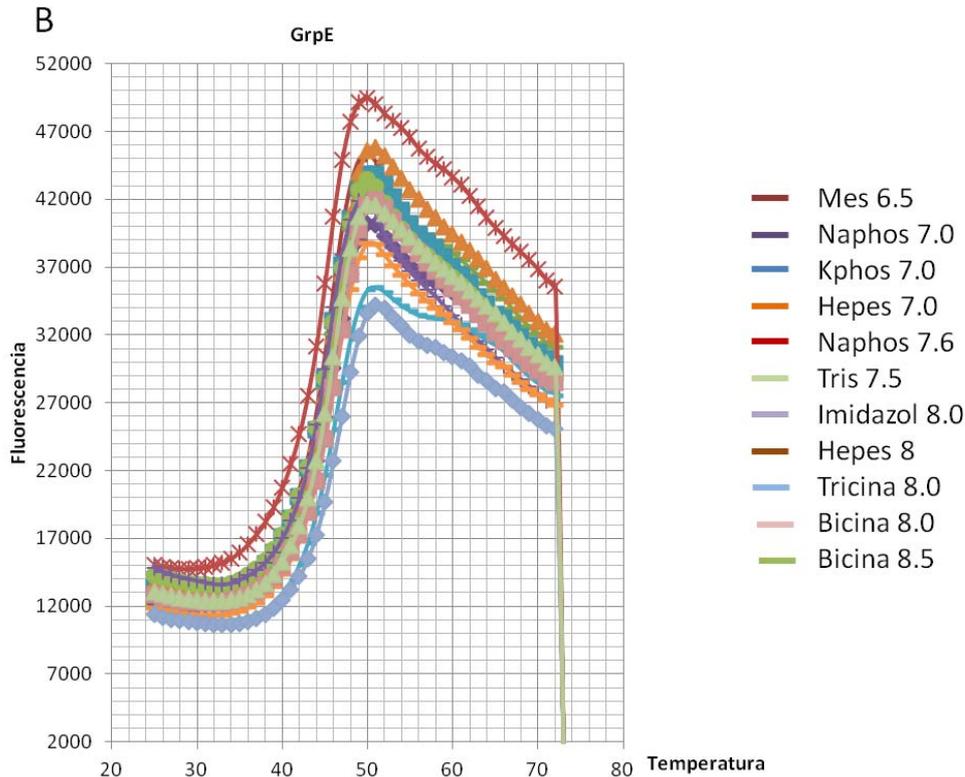


Figura 2.32 Curvas de fluorescencia de (A) DnaKΔCt y (B) GrpE en función de la temperatura y en presencia de diferentes tampones resultantes del ensayo de DSF. El rango de pHs en que ambas proteínas son más estables es de 5.5 a 8.5.

- Estudio del Complejo *MgDnaKΔCt-MgGrpE* por MALS

El intento de estimar el peso molecular aparente, y consecuentemente la estequiometría del complejo *MgDnaKΔCt-MgGrpE* mediante columna de gel filtración y utilizando una recta patrón con proteínas de peso molecular conocidos no dio un resultado muy preciso debido a la falta de reproducibilidad de las purificaciones del complejo. Por eso, se utilizó la técnica de MALS que permite calcular pesos moleculares con más precisión. Estos experimentos consistieron en cargar la muestra del complejo *MgDnaKΔCt-MgGrpE* en una columna de gel filtración y a medida que la muestra era eluída directamente pasaba por el detector de MALS. La estequiometría del complejo *MgDnaKΔCt-MgGrpE* en presencia de AMP-PNP correspondería a una

molécula de *MgDnaKΔCt* para dos de *MgGrpE*, siendo su peso molecular absoluto de aproximadamente 120 kDa ligeramente superior al peso molecular teórico que se esperaría de este complejo y que corresponde a 106 kDa (Figura 2.33). Este resultado sugiere que el complejo *MgDnaKΔCt-MgGrpE* presenta la misma estequiometría que el complejo *EcDnaK-EcGrpE* (código PDB: 1DKG). Los estudios de MALS permitieron indicar que este pico es monodisperso y que las muestras en solución son homogéneas. Tanto el peso molecular como la polidispersidad de la muestra fueron calculados utilizando el programa ASTRA6®(Wyatt-Technologies).

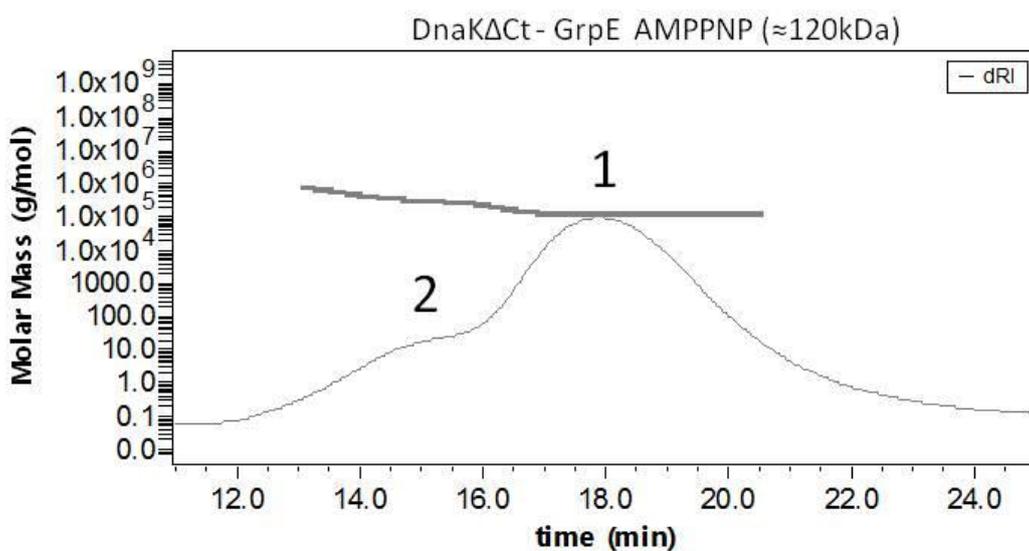


Figura 2.33 Gráfica del peso molecular y del índice de refracción del complejo *MgDnaKΔCt-MgGrpE* en presencia de AMP-PNP en función del tiempo, determinada mediante el ensayo de MALS. Se puede ver un pico principal (pico 1) que corresponde a un peso molecular de ~120 kDa. Para el pico 2, que es minoritario y corresponde muy posiblemente a formas agregadas del complejo, no se pudo calcular el peso molecular una vez que este no es constante a lo largo del pico.

A pesar de que la muestra fue considerada óptima para posteriores estudios de SAXS y que la estequiometría de complejo correspondía a 1:2, en la técnica de MALS se observó una leve agregación de la muestra. En la Figura 2.33 se puede apreciar por los

valores calculados para el peso molecular del agregado que corresponde al pico número 2, no es constante a lo largo del pico, decreciendo a lo largo de la elución. Estos resultados llevaron a que se confirmara la estequiometría del complejo mediante espectrometría de masas (del inglés *Mass Spectrometry, MS*).

- MS del complejo *MgDnaKΔCt-MgGrpE*

Se realizó primero una espectrometría de masas por ionización por electroesprai (del inglés *Electrospray Ionisation-Mass Spectrometry, ESI-MS*) (Ho, Lam et al. 2003) para identificar la presencia de los componentes del complejo *MgDnaKΔCt-MgGrpE* y si su peso era el esperado. Se detectó que ambas proteínas estaban intactas y que sus pesos moleculares correspondían a 57.6 kDa en el caso de *MgDnaKΔCt* y 25.9 kDa en el caso de *MgGrpE* ligeramente menores que los pesos moleculares teóricos calculados correspondientes a 57.71 kDa y 26.09 kDa, respectivamente. Una vez verificado que las muestras estaban intactas se procedió a realizar una espectrometría de masas nativa técnica que proporciona información sobre la composición, dinámica y propiedades estructurales de complejos proteicos (Lorenzen and van Duijn 2010). Los resultados obtenidos, tal y como se muestra en la Figura 2.34, confirman que la estequiometría del complejo en solución es 1 *MgDnaKΔCt*:2*MgGrpE*, lo que valida los resultados de MALS. También se detectaron vestigios de otros oligómeros del complejo 2:4 y 3:6.

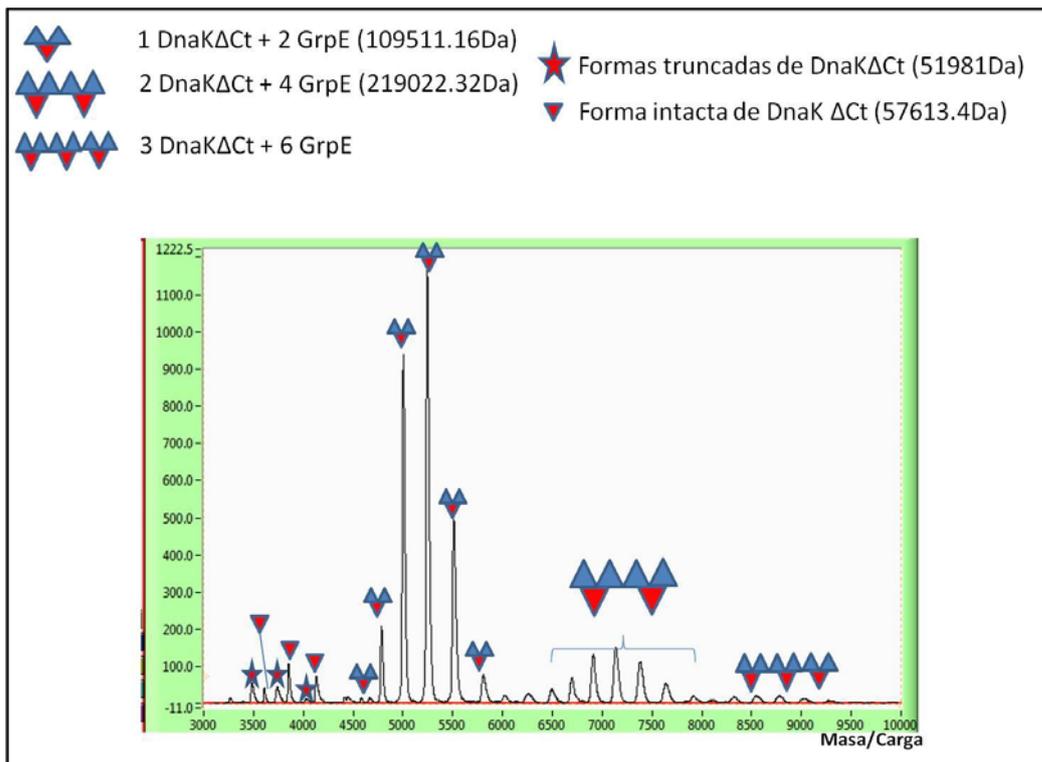


Figura 2.34 Gráfica obtenida de la espectrometría de masa nativa. Sobre cada pico se ha dibujado el oligómero correspondiente del complejo *MgDnaKΔCt-MgGrpE*. Y también de las formas truncadas e intactas de *MgDnaKΔCt*.

- Ensayos de Proteólisis Limitada del Complejo *MgDnaKΔCt-MgGrpE*

Las dificultades en la cristalización del complejo hicieron que se planteara la posibilidad de que hubieran ciertas regiones flexibles que pudieran estar desfavoreciendo la cristalización del complejo *MgDnaKΔCt-MgGrpE*. Por ello, se realizó una proteólisis limitada con diferentes proteasas para poder detectar la forma más globular del complejo. En un primer experimento se utilizaron las proteasas tripsina, quimotripsina y papaína incubando 1 μ L de complejo a 10 mg/mL con 0.2 μ L de tampón de digestión concentrado 10x de cada proteasa y 1 μ L de proteasa a diferentes concentraciones (entre 0.01mg/mL y 1mg/mL). Los ensayos se llevaron a cabo a dos temperaturas, 20°C y 37°C, y se bloquearon las reacciones a diferentes tiempos, de 1' a 60'. Las reacciones fueron detenidas añadiendo tampón de carga de geles de

poliacrilamida desnaturalizante. Después del primer cribado realizado con las tres proteasas se vio que los mejores resultados se obtuvieron con tripsina y quimotripsina así que se optimizaron las condiciones de digestión con estas dos proteasas a 20°C incubando las reacciones durante 60, 90 o 120 minutos. Los productos de digestión fueron analizados en geles de poliacrilamida desnaturalizantes (Figura 2.35).

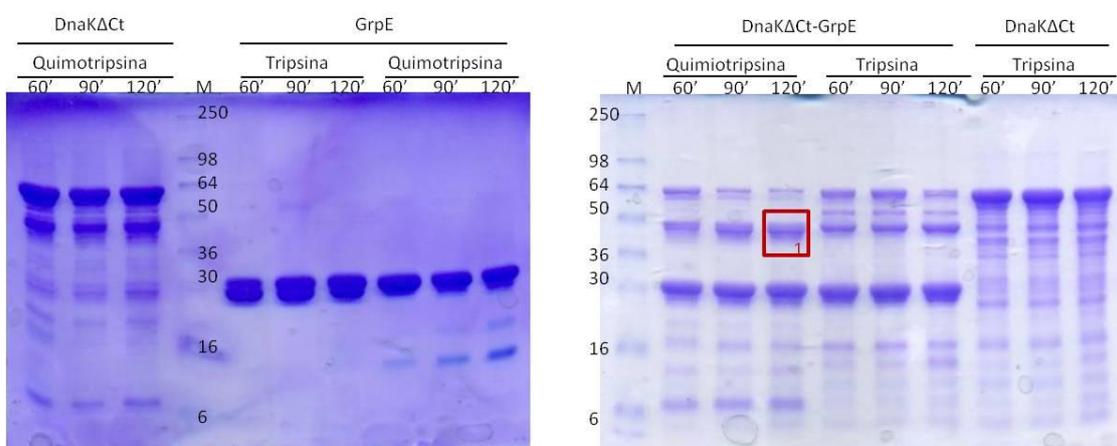


Figura 2.35 Geles desnaturalizantes SDS-PAGE de acrilamida al 12% donde se cargaron los productos de la proteólisis limitada de *MgDnaKΔCt*, *MgGrpE* y *MgDnaKΔCt-MgGrpE* con las proteasas tripsina y quimotripsina a diferentes tiempos (60, 90 y 120 minutos). Rectángulo rojo 1 muestra un fragmento más globular supuestamente de DnaK.

Las banda más prometedora, resultante de la digestión del complejo *MgDnaKΔCt-MgGrpE* fue la que se señala en la Figura 2.33 como 1 y que fue obtenida en la digestión con quimotripsina durante 120 minutos. La banda 1 podría corresponder al dominio *MgNBD*, estudiado anteriormente, indicando que quizás el complejo formado por el *MgNBD-MgGrpE* podría ser más globular y un candidato para futuros ensayos de cristalización. Justamente esta estrategia fue utilizada por Zhu y colaboradores en 1996 para cristalizar el complejo de *E.coli* formado por el dominio *EcNBD* de DnaK y *EcGrpE* que a su vez había sido digerido por una proteasa que deletionaba parte del N-

terminal de *EcGrpE* (código PDB: 1DKG) (Zhu, Zhao et al. 1996). La banda de la digestión proteolítica no fue secuenciada por secuenciación EDMAN, lo cual, nos hubiera proporcionado los 5 primeros aminoácidos del fragmento digerido y se hubiera podido deducir con exactitud el lugar de corte en *MgDnaKΔCt*. Antes de producir nuevos constructos se intentó obtener el complejo digerido y homogéneo partiendo del complejo co-purificado a 10mg/mL. Para ello, se realizó una digestión proteolítica a gran escala con quimotripsina (0.01 mg/mL) a 20°C, 2h después se inactivó la proteasa utilizando Pefabloc (Sigma Aldrich, Saint Louis, USA) a una concentración final de 1 mM. El producto de digestión fue inyectado en una columna de gel filtración Superdex 200 10/300 (GE Healthcare, Uppsala, Sweden) con el objetivo de separar los fragmentos generados de la digestión parcial del complejo, es decir, las zonas más flexibles y menos globulares que hayan sido digeridas. Contrariamente a lo esperado, no fue posible aislar una muestra homogénea de complejo lo que llevó a que se prepararan nuevos constructos de GrpE.

- Complejo *MgNBD*: *MgGrpE*-Ct 26-217

Partiendo de los resultados de la digestión proteolítica del complejo *MgDnaKΔCt*: *MgGrpE* y de la cristalización que realizaron del complejo de *E. coli* Zhu y colaboradores (Zhu, Zhao et al. 1996) se intentó cristalizar el complejo de *MgNBD* con una forma truncada de *MgGrpE*, en particular, una en la que se delecionaron los 26 primeros aminoácidos (constructo *MgGrpE*-Ct₂₆₋₂₁₇). Tanto *MgNBD* (en presencia de ATP, AMP-PNP o en ausencia de nucleótidos) como *MgGrpE*-Ct₂₆₋₂₁₇ se purificaron según los protocolos previamente establecidos (ver secciones 4.4.4 y 2.3.1), se mezclaron con un ratio 1:2 (*MgNBD*: *MgGrpE*-Ct₂₆₋₂₁₇) y se incubó durante

1 h a 4°C. La formación del complejo fue entonces analizada por cromatografía de gel filtración en una columna Superdex 200 10/300 (GE Healthcare, Uppsala, Sweden) y las fracciones recogidas fueron analizadas por un gel desnaturizante SDS-PAGE 15% acrilamida. La *MgNBD* eluye a 15.83 mL, y *MgGrpE-Ct_26-217* a 13.6 mL mientras que el complejo que se formó eluyó a un volumen de 13.1 mL. El hecho que *MgGrpE-Ct_26-217* y el complejo *MgNBD: MgGrpE-Ct_26-217* sean eluidos en aproximadamente mismo volumen de elución en la columna de gel filtración impidió la purificación del complejo por la dificultad en la separación de este de un posible exceso de *MgGrpE-Ct_26-217*, lo que resulta en una muestra heterogénea no óptima para realizar ensayos de cristalización. Para lograr resolver la posible contaminación del complejo *MgNBD: MgGrpE-Ct_26-217* del exceso de *MgGrpE-Ct_26-217* se diseñó un nuevo constructo de GrpE al que no se añadió la cola de histidinas en el C-terminal, lo que permitiría la co-purificación de *MgNBD* y *MgGrpE-Ct_26-217* sin arrastrar un exceso de *MgGrpE-Ct_26-217*. Como *MgNBD* sería el único componente con cola de histidinas este se uniría a la columna de afinidad y sólo la *MgGrpE-Ct_26-217* que estuviera formando el complejo también quedaría retenida en la columna.

- Complejo *MgNBD:MgGrpE-Ct_26-217* STOP

Se preparó entonces el nuevo constructo de *MgGrpE* añadiendo un codón STOP entre el último aminoácido de la proteína y la cola de histidinas del constructo pET21d *MgGrpE-Ct_26-217*, el cual se designó como *MgGrpE-Ct_26-217_STOP*. Se co-purificaron las proteínas *MgNBD* y *MgGrpE-Ct_26-217_STOP* del mismo modo que se había hecho con *MgDnaKΔCt-MgGrpE*. El perfil cromatográfico de esta muestra inyectada en una columna de gel filtración Superdex 200 16/60 (GE Healthcare,

Uppsala, Sweden) mostró la existencia de dos picos, uno correspondiente al complejo y el otro a un exceso de *MgNBD*. El volumen de elución del complejo en una columna Superdex 200 16/60 es de 70.28 mL lo que equivale aproximadamente a un complejo de ~100 kDa, indicativo de que la estequiometría de esta versión del complejo también es de 1:2 (*MgNBD*: *MgGrpE-Ct_26-217_STOP*). Con esta muestra se procedió a realizar nuevos cribados de cristalización a 20°C y a 4°C a dos concentraciones de proteína 10 y 20 mg/mL, con y sin añadir nucleótido. Se obtuvieron unos pequeños cristales pero resultaron ser un falso positivo ya que lo que cristalizó fue la *MgNBD*, seguramente debido a un posible exceso de *MgNBD* en la solución del complejo. A partir de ese momento se utilizó el *MgNBD* como control negativo en todos los ensayos de cristalización. A pesar de los múltiples ensayos de cristalización realizados en los que además de los factores variados anteriormente también se varió por ejemplo la concentración de cloruro de sodio en el tampón (0.5 M NaCl a 0.1 M NaCl) con el objetivo que el complejo fuera menos soluble no se obtuvieron cristales e ningún caso.

- Complejo *MgDnaKΔCt*: *MgGrpE-Ct_26-217_STOP*

Paralelamente a las otras pruebas de cristalización del complejo, se preparó el complejo que sería equivalente al complejo *GkDnaK-GkGrpE* cristalizado (código PDB 4ANI)(Wu, Naveen et al. 2012). Este complejo está formado por *GkDnaK* a la se le han delecionado los últimos 98 aminoácidos del C-terminal y *GkGrpE* entera. A pesar de que produjeron el complejo con *GkGrpE* entera en la estructura cristalina no lograron trazar los 58 primeros aminoácidos. Con base a estas observaciones y paralelamente al complejo de *M. genitalium*, se produjo un constructo de *MgGrpE* al que se le deleccionaron los primeros 58 aminoácidos, designado como *MgGrpE-Ct_58-217*. Este

nuevo constructo resultó ser muy inestable lo que impidió que se procediera a su purificación y a estudios posteriores. Por este motivo, se utilizó el constructo *MgGrpE-Ct_26-217_STOP* para proseguir. Se co-purificó entonces *MgGrpE-Ct_26-217_STOP* con *MgDnaKΔCt*, siguiendo el mismo protocolo descrito en la sección 4.4. Con el complejo de las dos proteínas se realizaron múltiples ensayos de cristalización en los que en ningún caso se obtuvieron cristales. La dificultad en obtener cristales de las variantes del complejo *MgDnaKΔCt: MgGrpE* llevó a que, paralelamente, se estudiara el complejo por dispersión de rayos-X de bajo ángulo para determinar si se trata de un complejo globular y si existen regiones desordenadas y/o flexibles. Además, estos estudios permitirían confirmar la estequiometría del complejo estequiometría y obtener un modelo del complejo a baja resolución.

- Estudio de *MgDnaK-NBD*, *MgDnaKΔCt* y *MgGrpE* por dispersión de rayos-X a bajo ángulo

Todos los intentos de obtener cristales para poder determinar la estructura del complejo *MgDnaKΔCt-MgGrpE* fallaron como se describe en apartados anteriores. Para obtener más información sobre la estabilidad del complejo en solución sobre los posibles cambios conformacionales que ocurren en los componentes del complejo, *DnaK* y *GrpE*, y con el objetivo de estudiar el proceso de señalización alostérica en *MgDnaK* se llevaron a cabo estudios de SAXS.

Los datos de SAXS se recogieron en la línea de luz BM29 del ESRF (Grenoble, Francia). Para la adquisición de datos se acopló una columna de gel filtración Superdex 200 increase 3.2/300 (GE Healthcare Life Sciences, Uppsala, Sweden) equilibrada con el tampón 0.1 M NaCl, 0.02 M Tris-HCl pH8, en línea con el instrumento de SAXS para

medir la muestra en el momento en que es eluída de la columna. Este método permite recoger datos de la proteína de interés totalmente separada de agregados que se forman durante el período de almacenamiento y procesar las imágenes que corresponden a una sola especie en solución, que es naturalmente monodispersa. Se inyectaron en la columna 10 μL de *MgNBD*, *MgDnaK Δ Ct* y *MgGrpE* preparadas como descrito en la sección 4.7 en presencia de AMP-PNP y a concentraciones de entre 10 a 15 mg/mL e a 20°C. Como blanco se utilizaron las primeras imágenes tomadas de la elución de la columna de gel filtración, correspondiente al tampón de la columna, y su promedio fue sustraído a la curva de dispersión de rayos-X de la muestra. Las curvas de cada una de las muestras así corregidas permitieron entonces calcular los valores de la intensidad de dispersión de ángulo 0 ($I(0)$) y el radio de giro (R_g) de cada una de las muestras mediante la aproximación de Guinier.

La curva experimental obtenida para *MgNBD:AMP-PNP*, que permitió determinar que el radio de giro del subdominio es de 2.4 nm (Figura 2.36, A), fue ajustada a la curva teórica del modelo de rayos-X (descrito en la sección 2.7) mediante el programa CRY SOL (Petoukhov, Franke et al. 2012) (Figura 2.36 B) (Schmidt, König et al. 1995). La curva teórica se ajusta a la curva experimental con un χ^2 de 0.798 indicando que la forma monomérica del dominio en solución en presencia de AMP-PNP es idéntico a la forma cristalina de la proteína y que por lo tanto este dominio globular es muy rígido. Esta posibilidad también es validada por la inspección de la superposición del modelo *ab initio* de la forma de la proteína obtenido con DAMMIF (Petoukhov, Franke et al. 2012) con la estructura cristalina de *MgNBD:AMP-PNP* calculada con el programa SUPCOMB (Kozin, et al. 2001)(Figura 2.36, C).

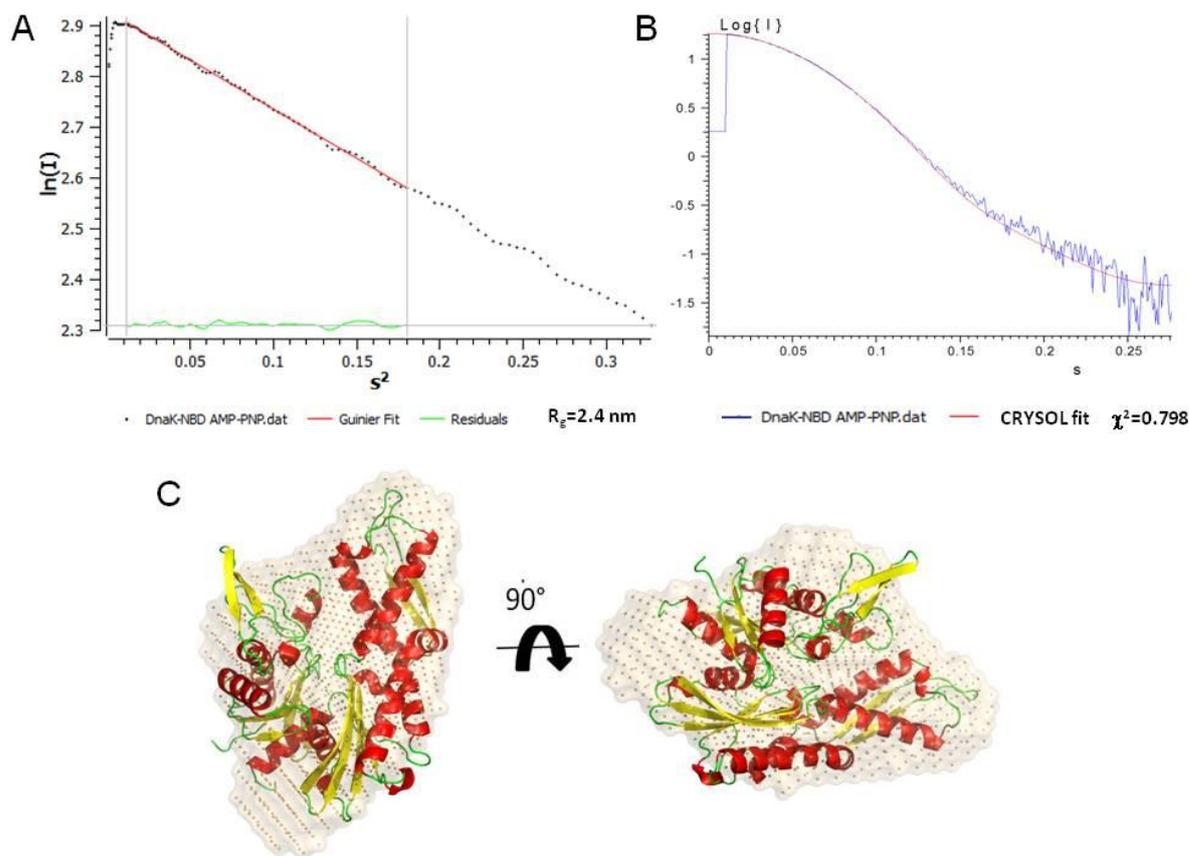


Figura 2.36 Estudio estructural mediante SAXS de *MgNBD:AMP-PNP*. A) Diagrama de Guinier donde se representa la intensidad de dispersión en escala logarítmica (círculos negros) en función de s . B) Ajuste mediante el programa CRYSOLOG de la curva teórica obtenida para la estructura cristalina de *MgNBD:AMP-PNP* (línea roja) a la curva de SAXS experimental (línea azul). C) Modelo de cuentas a baja resolución de *DnaK-NBD:AMP-PNP* obtenido con el programa DAMMIF (superficie beige)(Petoukhov, Franke et al. 2012) superpuesto mediante el programa SUPCOMB (Kozin, et al. 2001) a la estructura cristalina de *MgNBD:AMP-PNP*.

El radio de giro del constructo *MgDnaKΔCt* calculado mediante la aproximación de Guinier es de 3.2 nm. Del mismo modo que se procedió anteriormente, cuando se estudió el *MgNDB*, se utilizó el programa CRYSOLOG para ajustar la curva teórica de la estructura cristalina de *MgDnaKΔCt:AMP-PNP* con respecto a la curva de dispersión experimental obtenida para la misma muestra (Figura 2.37 A y B). El ajuste tiene un χ^2 de 2.5 indicando que parte de la proteína en solución podría adoptar una conformación idéntica a la encontrada en la forma cristalina pero que podrían existir

otras formas conformacionales de la proteína en solución. Para determinar cuáles son las conformaciones posiblemente adoptadas por la proteína se procedió del siguiente modo. Primero se ajustó con el programa CRY SOL la curva teórica calculada para la estructura de *EcDnaK*, proteína homóloga a *MgDnaK* y que presenta una conformación cerrada (código PDB 2KHO), a la curva experimental de *MgDnaK* Δ Ct:AMP-PNP y se obtuvo un χ^2 de 9.3 lo que sugiere que la *MgDnaK* Δ Ct podría sufrir el mismo tipo de cambio conformacional observado en la estructura de *EcDnaK*. Como el ajuste obtenido no es óptimo se ajustó de la misma manera la curva experimental de *MgDnaK* Δ Ct:AMP-PNP a la curva teórica calculada para la estructura de *GkDnaK* (código PDB 2V7Y), que también es homóloga a la *MgDnaK*. En este caso el χ^2 es un ligeramente más elevado indicando que es menos probable que en *MgDnaK* la disposición de los dominios NBD y SBD sea idéntica a la observada en la estructura de *GkDnaK*. Una vez comparados los datos de SAXS de *MgDnaK* Δ Ct con proteínas homólogas de otros procariontes lo que se hizo fue utilizar los subdominios NBD y SBD de la estructura de *MgDnaK* Δ Ct eliminando el enlazador y disponerlos de forma idéntica a la disposición de estos dominios en la estructura de *EcDnaK* (código PDB 2KHO). El ajuste, ejecutado con el programa CRY SOL (Svergun 1999), de la curva de dispersión de luz teórica calculada para este modelo a la curva experimental de *MgDnaK* Δ Ct:AMP-PNP presenta un χ^2 de 1.4 (Figura 2.37 C y D) indicando que *MgDnaK* Δ Ct:AMP-PNP presenta en solución una conformación cerrada reforzando la existencia del paso intermedio extra propuesto a raíz de los estudios por difracción de rayos-X de la *MgDnaK* descritos en la sección 2.1.6 como modificación a los modelos alostérico de Bhattacharya y colaboradores y Melero y colaboradores (ver sección 2.1.6).

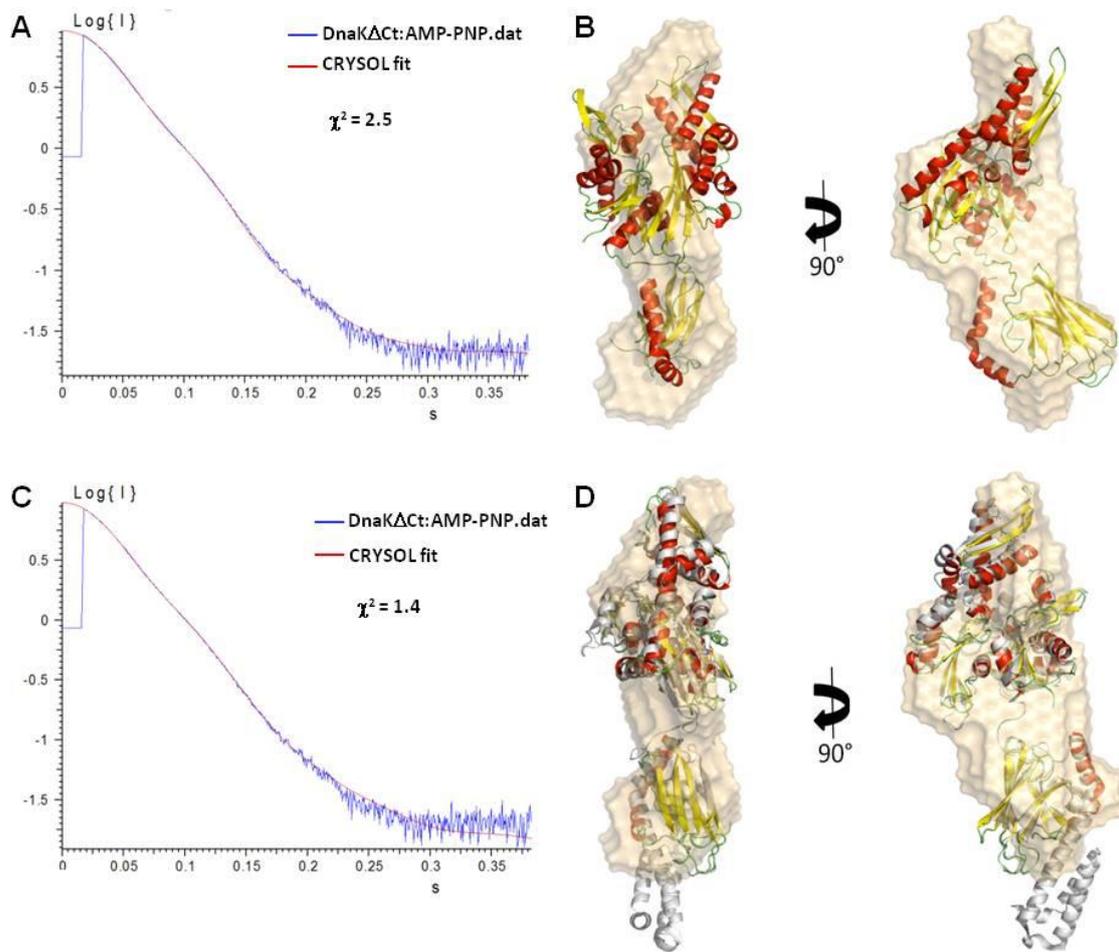


Figura 2.37 Estudio estructural mediante SAXS de *MgDnaKΔCt:AMP-PNP*. A) Ajuste mediante el programa CRY SOL de la curva teórica obtenida para la estructura cristalina de *MgDnaKΔCt:AMP-PNP* (línea roja) y su curva de SAXS experimental (línea azul). B) Modelo de cuentas a baja resolución de *MgDnaK-NBD:AMP-PNP* obtenido con el programa DAMMIF (superficie beige) (Petoukhov, Franke et al. 2012) superpuesto mediante el programa SUPCOMB (Kozin, et al. 2001) a la estructura cristalina de *MgDnaKΔCt:AMP-PNP* (modelo de cintas con los elementos de estructura secundaria en verde (lazos), amarillo (hebras β) y rojo (hélices α)). C) Ajuste mediante el programa CRY SOL (Svergun 1999) de la curva teórica obtenida para el NBD y SBD de la estructura de *MgDnaKΔCt:AMP-PNP* dispuestos de forma idéntica a la observada para los dominios de *EcDnaK* (código PDB 2KHO, línea roja) a la curva de SAXS experimental obtenida para *MgDnaKΔCt:AMP-PNP* (línea azul). D) Modelo de cuentas a baja resolución de *MgDnaKΔCt:AMP-PNP* obtenido con el programa DAMMIF (superficie beige) superpuesto mediante el programa SUPCOMB a la estructura cristalina de *MgDnaKΔCt:AMP-PNP* con los dominios NBD y SBD dispuestos de forma idéntica a la observada para los dominios de *EcDnaK* (modelo de cintas con los elementos de estructura secundaria en verde (lazos), amarillo (hebras β) y rojo (hélices α)). Se representa la estructura de *EcDnaK* (modelo de cintas en gris) también superpuesta a estos modelos.

Del mismo modo que se procedió para obtener datos estructurales de *MgDnaKΔCt* también se midieron datos de SAXS de *MgGrpE* que presenta un R_g de 4.8 nm. En una primera aproximación, y debido a que no fue posible hasta el momento determinar la

estructura cristalina de *MgGrpE*, se intentó ajustar mediante el programa CRY SOL la curva experimental de dispersión de luz de *MgGrpE* a curvas teóricas calculadas a partir de las estructuras cristalinas las GrpEs de *E. coli*, *G. kaustophilus* y *T. thermophilus* (códigos PDB 1DKG, 4ANI y 3A6M, respectivamente) que presentan una identidad de secuencia de 15-20% con *MgGrpE*. Finalmente, lo que se observó fue que ninguno de estos modelos se ajusta a los datos experimentales pero, de la superposición de estas estructuras (Figura 2.38 A), se puede intuir que estas proteínas presentan una gran flexibilidad y variabilidad a nivel de la disposición del dominio α -hélice (Figura 1.11 D). En estos momentos, para validar el modelo *ab initio* de baja resolución obtenido con DAMMIF (Petoukhov, Franke et al. 2012) se preparó un modelo teórico tri-dimensional de *MgGrpE* con el programa IntFold3 (McGuffin, Atkins et al. 2015) (Figura 2.38 A). Este modelo, además de presentar las características estructurales de las estructuras cristalinas conocidas de GrpEs de otros organismos, se ajusta bien al modelo *ab initio* determinado por DAMMIF (Petoukhov, Franke et al. 2012), como se puede observar en la superposición calculada mediante SUPCOMB (Kozin, et al. 2001) de este modelo con el modelo teórico de la proteína (Figura 2.38 B). Estos resultados apuntan a que los dominios α -hélice de las dos subunidades que componen el dímero de *MgGrpE* son muy flexibles lo que, junto con el hecho que el N-terminal de la proteína es desestructurado pero con tendencia a formar estructuras helicoidales como se ha determinado por diferentes programas que predicen la estructura secundaria de las proteínas (Figura 2.27 A y B), permite especular que esta región de GrpE tiene las propiedades necesarias para poder plegarse cuando se da la interacción con DnaK de modo a ocupar su SBD mientras se efectúa el intercambio de nucleótido, de acuerdo con lo sugerido por Melero y colaboradores.

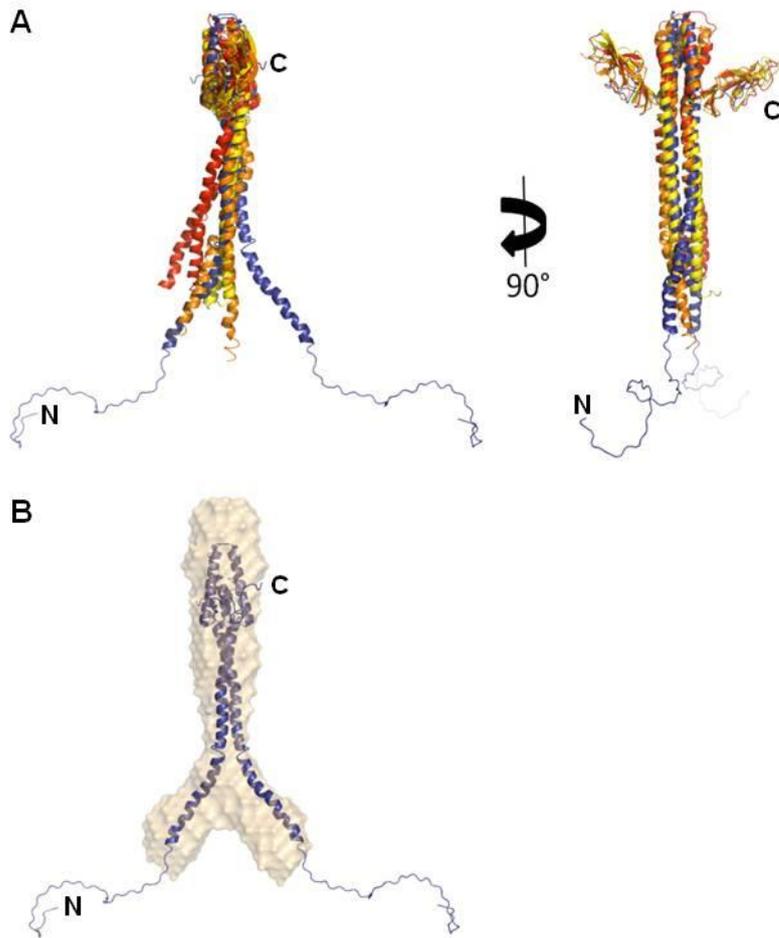


Figura 2.38 Estudio estructural mediante SAXS de *MgGrpE*. A) Superposición de las estructuras cristalinas de las GrpEs de *E. coli* (amarillo), *G. kaustophilus* (naranja), *T. thermophilus* (rojo) códigos PDB 1DKG, 4ANI y 3A6M, respectivamente) y el modelo teórico obtenido con IntFold3(McGuffin, Atkins et al. 2015) de *MgGrpE* (azul). B) Modelo de cuentas a baja resolución de *MgGrpE* obtenido con el programa DAMMIF (superficie beige)(Petoukhov, Franke et al. 2012) superpuesto mediante el programa SUPCOMB (Kozin, et al. 2001) al modelo teórico de *MgGrpE* (modelo de cintas en azul).

Los datos de SAXS adquiridos de las diferentes versiones del complejo *MgDnaK-MgGrpE* no pudieron ser procesados debido a la dificultad en preparar muestras suficientemente homogéneas y polidispersas, lo que también demuestra la dificultad en obtener cristales del complejo que permitan proseguir con los estudios estructurales.

3.- Conclusiones

La posible implicación del sistema de chaperonas DnaK-DnaJ-GrpE en el mecanismo de motilidad de *M. genitalium* observada *in vivo* por el grupo de colaboradores de este proyecto fue la motivación inicial para estudiar estas proteínas. Esta Tesis doctoral ha contribuido a entender más detalladamente el sistema de chaperonas del pequeño pero complejo microorganismo *M. genitalium*. Las principales conclusiones se pueden resumir en los siguientes puntos:

1.- La *MgDnaK* se sobreexpresó recombinantemente en células bacterianas y se purificó. Esta muestra se reveló inestable y susceptible a degradación, proceso que no pudo ser controlado. La imposibilidad de obtener cristales de dicha muestra llevó a que se diseñaran nuevos constructos de la proteína.

2.- Un nuevo constructo de *MgDnaK* (*MgDnaKΔCt*), al que se delecionaron los últimos 74 aminoácidos del dominio C-terminal, fue expresado, purificado y cristalizado en presencia de ATP y AMP-PNP. Se determinó la estructura tridimensional de estos dos tipos de cristales mediante reemplazo molecular y a una resolución de 1.8 y 2.5 Å, respectivamente.

3.- La estructura de *MgDnaKΔCt* está formada por dos dominios: el dominio de unión a nucleótido (NBD) y el dominio de unión a sustrato (SBD). Estos se conectan entre sí mediante un enlazador muy flexible que está involucrado en la regulación alostérica de la proteína. La disposición adoptada por el enlazador en las estructuras *MgDnaK*

corresponderían un estado extendido, confirmando una conformación cerrada del SBD estructura.

4.- En la estructura de *MgDnaKΔCt:AMP-PNP*, un ion magnesio se coordina con una molécula de agua, con el fosfato β del nucleótido y con los aminoácidos Asp12 y Asn17 adoptando una geometría tetragonal. Este estado de coordinación es muy inusual pudiendo ser consecuencia de la ausencia de un enlace entre el fosfato y del nucleótido y el magnesio.

5.- La disposición estructural extendida del enlazador en el caso de *MgDnaKΔCt:AMP-PNP* podría sugerir que es consecuencia no sólo del nucleótido presente en el NBD sino también de la presencia/ausencia de péptido en el SBD.

6.- El SBD de las estructuras *MgDnaKΔCt:ATP* y *MgDnaKΔCt:AMP-PNP* se encuentra ocupado por el péptido C-terminal del constructo que mimetiza un sustrato. La Leu522 interacciona con el bolsillo hidrofóbico que está constituido por los residuos Ile382, Phe407, Val417, Val419, Ile453 y Ile455. Y el modo de unión del sustrato al bolsillo de unión es "hacia delante".

7.- Se diseñaron tres nuevos mutantes para intentar que el sitio de unión del sustrato en sbd no se encontrara ocupado. El análisis de los tres mutantes mediante MALS reveló la alta heterogeneidad y polidispersidad y la tendencia a formar agregados.

8.- Se determinó la estructura tridimensional del dominio *MgNBD* en presencia de AMP-PNP, ADP, ATP y sin nucleótido mediante reemplazo molecular a una resolución de 1.30, 1.28, 1.40 y 1.80 Å, respectivamente. El ion magnesio presente en la estructura de *MgNBD:ADP* está coordinado de forma octaédrica, mientras que el

magnesio presente en la estructura de *MgNBD:AMP-PNP* se coordina con una geometría tetragonal piramidal.

9.- Los cambios estructurales producidos por la presencia de diferentes nucleótidos en *MgDnaKΔCt* son muy sutiles, siendo el movimiento del subdominio IB-IIB en la estructura apo de *MgNBD* el más pronunciado.

10.- Se expresó y purificó *MgGrpE*, *MgGrpE-Ct_26-217*, *MgGrpE-Ct_26-217_STOP* y *MgGrpE-Ct_58-217*. Estudios biofísicos demostraron que *MgGrpE* y el constructo *MgGrpE-Ct_26-217* eran los más estables y homogéneos, comportándose como dímeros en solución.

11.- No se detectaron interacciones entre *MgDnaKΔCt* con los dominios J o J+región G/F de MG200 ni con los dominios J o J+región G/F de MG019 por SPR. Con todo, no se puede descartar completamente que *DnaKΔCt* no interaccione con alguno de estos dominios, en condiciones muy específicas.

12.- Se identificó mediante SPR una fuerte interacción entre *MgDnaKΔCt* y *MgGrpE* en presencia de ATP o AMP-PNP. El complejo *MgDnaKΔCt: MgGrpE* i algunas de sus variantes fueron purificados por cromatografía de gel filtración: *MgNBD: MgGrpE-Ct_26-217*, *MgNBD: MgGrpE-Ct_26-217_STOP* y *MgDnaKΔCt: MgGrpE-Ct_26-217_STOP*. La técnica de MALS permitió determinar que la estequiometría del complejo es de 1 molécula de *MgDnaKΔCt* y dos 2 de *MgGrpE* correspondiente a un peso molecular aproximadamente de ~120 kDa. Este resultado fue también validado por la espectrometría de masas nativa.

13.- Se determinó la estructura en solución de *MgNBD:AMP-PNP* y *MgDnaKΔCt:AMP-PNP* y *MgGrpE* por dispersión de rayos X a bajo ángulo (SAXS). El modelo en solución de *MgNBD* muestra un dominio globular y rígido igual a la forma cristalina. El modelo en solución de *MgGrpE*, que se ajusta a un modelo teórico tri-dimensional de *MgGrpE*, indica que su dominio N-terminal es flexible. Esta propiedad junto con la tendencia de este dominio a formar estructuras helicoidales sugiere que esta región de la proteína podría plegarse cuando se da la interacción con DnaK a modo de ocupar su SBD mientras se efectúa el intercambio de nucleótido, de acuerdo con el modelo alostérico de DnaK propuesto por Melero y colaboradores.

4. Materiales y métodos

4.1 Materiales

4.1.1 Lista de materiales y equipos

- Dispensador Cartesian Honeybee (Hamilton, Reno, NV, USA)
- Columnas para purificación de proteínas de GE Healthcare Life Sciences (Uppsala, Sweden)
- Dispensador Crystal Phoenix (ARInstruments, Sunnyvale, CA, USA)
- Enzimas de restricción de Fermentas (Vilnius, Lithuania)
- Termociclador Eppendorf Mastercycler® (Eppendorf, Hauppauge, NY, USA)
- Gene Genius Bio Imaging System (Syngene, Cambridge, UK)
- Kit para purificación de bandas de DNA: GFX™ PCR DNA and gel band Purification Kit (GE Healthcare, Buckinghamshire, UK)
- Termociclador tiempo real iCycler5 (BioRad, Hercules, CA, USA)
- Polimerasa de DNA Pfu (Fermentas, Vilnius, Lithuania)
- Polimerasa de DNA KOD (Novagen)
- Espectrofotometro Nano-drop® (Thermo Scientific, Wilmington, DE, USA)
- Los oligonucleótidos utilizados fueron sintetizados por Sigma Aldrich (St Louis, MO, USA)

- Marcador de peso molecular de DNA Perfect DNA™ 1 kbp Ladder (Novagen, Madison, WI, USA)
- Kit de Miniprep QIAprep® Spin (QIAGEN, Hilden, Germany)
- Marcador de peso molecular de proteínas SeeBlue® (Invitrogen, Carlsbad, CA, USA)
- Teñidor de DNA SYBR® Safe (Invitrogen, Carlsbad, CA, USA)
- Teñidor SYPRO® Orange (Invitrogen, Carlsbad, CA, USA)
- Instrumento de DLS Zetasizer Nano (Malvern, Worcestershire, UK)
- Sonificador ultrasonico Vibra-Cell VCX 750W (Sonics, Newtown, USA)
- Instrumento de resonancia de plasmón superficial BIAcore® 3000 (GE Healthcare)
- Biosensor de resonancia de plasmón superficial CM5 (GE Healthcare)
- Instrumentos de cromatografía líquida rápida de proteínas (FPLC): ÄKTApurifier 10, ÄKTApurifier 100, ÄKTA Explorer, ÄKTA Basic (GE Healthcare Life Sciences, Uppsala, Sweden)
- Instrumento de caracterización SEC-MALS Dawn® Heleos II (Wyatt, Santa Barbara, CA, USA)
- Proteasas Tripsina, Quimotripsina y Papaína (Sigma-Aldrich, St Louis, MO,)
- Inhibidor Pefablock (Sigma-Aldrich, St Louis, MO,)
- Cribados de cristalización (Plataforma Automatizada Cristalografía, Barcelona, España)
- Complete EDTA-free (Roche, Basilea, Switzerland)

4.2 Cepas bacterianas y vectores de expresión

4.2.1 Cepas bacterianas usadas para la amplificación del DNA y clonaje

E. coli DH5 α . Genotipo: F⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ 80dlacZ Δ M15 Δ (lacZYA-argF)U169, hsdR17(r_K⁻ m_K⁺), λ -

E. coli XL1 Blue. Genotipo: endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F' [::Tn10 proAB⁺ lacI^q Δ (lacZ)M15] hsdR17(r_K⁻ m_K⁺)

4.2.2 Cepas bacterianas usadas para la expresión recombinante de proteínas.

E. coli BL21(DE3). Genotipo: F⁻ ompT hsdS_B(r_B⁻ m_B⁻) gal dcm (DE3)

E. coli Rosetta™ (DE3). Genotipo: F⁻ ompT hsdS_B (r_B⁻ m_B⁻) gal dcm (DE3) pRARE² (Cam^R)

4.2.3 Vectores de expresión

pET21d (Novagen, Madison, WI, USA) añade una cola de 6 histidinas en el C-terminal de la proteína de interés.

pET28a (Novagen, Madison, WI, USA) añade una cola de 6 histidinas en el N-terminal de la proteína de interés.

Ambos vectores se utilizaron para expresar proteínas de manera recombinante en *E.coli*

4.3 Procedimientos de biología molecular

4.3.1 Métodos microbiológicos

- Composición del medio de cultivo

Composición del medio de cultivo Luria-Bertani (LB)/agar (1L):

10 g triptona, 5 g extracto de levadura, 10 g NaCl y 15 g agar. El medio se autoclava, se deja enfriar hasta unos 60°C y se añaden los antibióticos apropiados y se distribuye el medio en placas de petri, añadiendo aproximadamente unos 30 mL.

Composición del medio de cultivo LB (1 L):

10 g triptona, 5 g extracto de levadura y 10 g NaCl (autoclavado).

- Antibióticos y suplementos
- Ampicilina 1000x: 100 mg/mL de ampicilina se preparó en agua y se filtró a través de un filtro de 0.22 µm.

- Isopropil- β -D-tiogalactosido (IPTG) 1000x: 1M de IPTG se preparó en agua y se filtró a través de un filtro de 0.22 μ m. Se utilizó para inducir la expresión de los genes que están bajo el promotor *lacZ*.
- Condiciones generales de cultivo de *E. coli*

Las células de *E. coli* se cultivaron en placas de LB con agar en presencia del antibiótico apropiado durante toda la noche a 37°C. Se seleccionó posteriormente una colonia que creció aislada y se inoculó en 4 mL de medio LB durante toda la noche a 37°C con agitación (250 rpm) a modo de obtener un cultivo supersaturado. Este precultivo se inoculó en 0.5 L de medio LB (dilución 1/200) y se crecieron las células a 37°C con agitación (250 rpm) hasta que el cultivo alcanzara una OD₆₀₀ ~0.6-0.8. Las células se dejaron atemperar hasta 20°C, se indujeron entonces con una concentración final de IPTG de 0.5 mM y se incubaron con agitación durante 16 h. Los cultivos fueron suplementados con ampicilina.

- Transformación de células de *E. coli* por choque térmico

Las células competentes de *E. coli* de la cepa conveniente se descongelan en hielo y se añaden 50 μ L de células a 1 μ L de muestra en un eppendorf estéril y previamente enfriado. La mezcla se incuba en hielo durante 20 minutos después de los cuales es sumergida en un baño maría a 42°C durante 45 segundos seguido de una incubación de 2 minutos en hielo. A las células transformadas se añaden entonces 0.5 mL de LB y se incuba durante una hora a 37°C. El producto de la transformación es plaqueado a continuación en placas de LB/agar suplementadas con el antibiótico apropiado que se incuban a 37°C durante toda la noche (Sambrook, 2001).

4.3.2 Tecnología del DNA recombinante

- Clonación de la variante de DnaK, DnaK Δ Ct

Se amplificó el cDNA de DnaK mediante la PCR usando los oligonucleótidos: 5'-ATATACATATGAGTGCAGACAATGGTTT-3' (dirección hacia delante) y 3'-TACGTAATGAAGGGGAACTCGAGCAC-5' (dirección reversa) diseñados con respecto al fragmento amplificado y al vector en el que sería insertado: pET21d (Novagen Madison, WI, USA). La reacción de PCR también se utilizó para incorporar mutaciones puntuales en el gen de DnaK por mutagénesis dirigida.

Las reacciones de PCR estándar se llevaron a cabo en un volumen final de 50 μ L conteniendo: 0.2 mM dNTPs, 0.4 μ M de cada encebador, 10 ng DNA plantilla y 2.5 U/ μ L polimerasa KOD Hifi (Novagen Madison, WI, USA) en tampón 1 x 0.02 Tris-HCl pH 8.8, 0.01 M KCl, 0.06 M (NH₄)₂SO₄, 1% (v/v) Triton X-100, 0.01% (w/v) BSA. Se aplicaron entonces los siguientes ciclos de temperatura: un ciclo de desnaturalización de 30 segundos a 98°C, 30 ciclos de los siguientes pasos: paso de desnaturalización de 30 segundos a 98°C, paso de hibridación de 15 segundos a 56°C y paso de elongación de 4 minutos a 72°C. Este último paso fue modificado según la longitud del fragmento a ser amplificado. En general, los productos amplificados por PCR y el vector de expresión pET21d se digirieron con las enzimas de restricción *Nde*I y *Xho*I (Fermentas, Vilnius, Lituania) según el protocolo proporcionado por Fermentas. Los productos de digestión se mezclaron y se ligaron también según el protocolo proporcionado por la compañía Fermentas (Vilnius, Lituania) y transformaron en células de *E. coli* DH5 α para amplificar el DNA plasmídico que fue extraído mediante el kit QIAprep® Spin MiniPrep Kit (

QIAGEN, Hilden, Germany). Los productos de PCR fueron posteriormente secuenciados para confirmar la secuencia.

- Mutagénesis dirigida

La técnica de mutagénesis dirigida fue utilizada para introducir mutaciones puntuales en la secuencia nativa del DNA de *DnaKΔCt* y para ello se diseñaron oligonucleótidos específicos (ver Tabla 2.3 En la sección de resultados).

Se llevaron a cabo una serie de reacciones de PCR estándar según el protocolo descrito en el apartado 4.3.2.1 de la sección de Materiales y métodos. Los ciclos de temperatura aplicados fueron los siguientes: un ciclo de 30 segundos a 95°C y 16 ciclos de 30 segundos a 95°C, 1 minuto a 55°C y 14 minutos a 68°C. Los productos de PCR se enfriaron en hielo hasta aproximadamente 37°C, temperatura en la que se incubaron entonces los productos de PCR 1h en presencia de *DpnI* (Fermentas, Vilnius, Lituania). La función de *DpnI* es reconocer y digerir el DNA metilado que proviene del DNA parental y no los productos provenientes de la PCR. Finalmente, 10μL de los productos de estas reacciones se transformaron en *E.coli* XL1 Blue y se verificaron las secuencias de DNA por una reacción de secuenciación.

4.4 Métodos de producción y análisis de proteínas

4.4.1 Preparación de extractos celulares totales

En general los cultivos celulares se centrifugaron a 4500 xg durante 20 minutos a 4°C y los precipitados se resuspendieron en tampón 0.5 M NaCl, 0.02 M Tris-HCl pH 8 que contenía una mezcla de inhibidores de proteasas ((Complete free-EDTA, Roche, Basilea, Switzerland) en una proporción de 15 mL de tampón de lisis por 1L cultivo). A *posteriori* se lisaron la células con un sonicador Vibra-Cell VCX 750W (Sonics, Newtown, USA) y los sobrenadantes, que contienen las proteínas de interés, se obtuvieron centrifugando los extractos celulares 20 minutos a 48000 xg y 4°C. Los sobrenadantes se filtraron mediante filtros de 0.22 μm antes de ser inyectados en alguna columna de cromatografía líquida.

4.4.2 Electroforesis en geles de poliacrilamida en condiciones desnaturalizantes (SDS-PAGE)

La electroforesis en gel de poliacrilamida se realizó según los protocolos estándares (Shapiro, Vinuela et al. 1967, Weber and Osborn 1969). Las muestras se desnaturalizaron por calor (95°C) durante 10 minutos en presencia de 5% (w/v) dodecilsulfato sódico (SDS) y 0.025% (w/v) azul de bromofenol y luego se cargaron en

el gel. El gel se corrió a un voltaje constante de 150mA. El marcador de peso molecular que se usó fue el SeeBlue® (Invitrogen, Carlsbad, CA, USA) que utiliza como estándares de proteína pre-teñidos: miosina (250 kDa), BSA (98 kDa), glutamato deshidrogenasa (64 kDa), alcohol deshidrogenasa (50 kDa), anhidrasa carbónica (36 kDa), mioglobina (30 kDa), lisozima (16 kDa), aprotinina (6 kDa) y cadena B de la insulina (4 kDa).

4.4.3 Tinción de geles de poliacrilamida con azul de Coomassie (CBB)

La tinción de geles de poliacrilamida con el teñidor CBB es un método cuantitativo y específico que detecta de 0.3 a 1 µg de proteína por banda. Para la tinción se sumerge el gel al menos 15 minutos en una solución de 0.1% (w/v) CBB R-250, 7% (v/v) ácido acético y 15% (v/v) etanol. Posteriormente, el gel se destiñe con una solución de 7% (v/v) ácido acético, 15% (v/v) etanol para poder observar las bandas.

4.4.4 Técnicas cromatográficas de purificación de proteínas

Para llevar a cabo estudios de biología estructural se deben aislar las proteínas de interés. El elevado grado de pureza necesario al análisis de las proteínas por técnicas de biología estructural se consigue generalmente, a través de pasos de purificación por cromatografía líquida. En cada uno de estos pasos una parte de la proteína se

pierde de manera que el protocolo de purificación de cada proteína consiste en conseguir el máximo grado de pureza empleando el menor número de pasos posibles. Estos son seleccionados dependiendo del tamaño de la proteína, de su carga eléctrica y solubilidad. Los métodos cromatográficos se realizaron en equipos automatizados de FPLC.

Cromatografía de afinidad por ion metálico inmovilizado (IMAC)

En la cromatografía por IMAC el lisado celular (ver sección 4.4.1) es inyectado en una columna HisTrap™ HP (GE Healthcare Life Sciences, Uppsala, Sweden) previamente equilibrada en tampón de lisis (0.5 M NaCl, 0.02 M Tris-HCl pH8). En este caso la proteína diana que contiene una cola C-terminal de 6 histidinas se une específicamente a los iones de Ni²⁺ de la columna. La elución de la proteína se lleva a cabo mediante un gradiente lineal de 0 a 0.5 M imidazol.

Cromatografía de intercambio iónico

La cromatografía líquida de intercambio iónico permite separar proteínas según su carga eléctrica. La columna Mono Q (GE Healthcare Life Sciences, Uppsala, Sweden) está preparada para el intercambio aniónico conteniendo una fase estacionaria con carga positiva que atrae a las moléculas cargadas negativamente. La elución de las proteínas retenidas por la matriz de la columna se consigue así aplicando un gradiente de sal, por ejemplo de 0 a 1 M NaCl, resultando en la neutralización o en el cambio de los grupos funcionales de la proteína.

Cromatografía de gel filtración (SEC)

La SEC es un método para separar macromoléculas basado en su radio hidrodinámico en solución. El volumen en el cual la macromolécula es eluída es directamente proporcional al tiempo en que la macromolécula es retenida en la columna e inversamente proporcional a su peso molecular (acrónimo del inglés *Molecular Weight*, MW). La SEC se utilizó siempre como último paso de purificación permitiendo la evaluación de la homogeneidad y del estado de oligomerización de la muestra. Se utilizaron columnas de gel filtración Superdex (GE Healthcare Life Sciences, Uppsala, Sweden) y el rango de separación se escogió según el tamaño de la proteína a purificar.

4.4.5 Experimentos de proteólisis limitada

Los experimentos de proteólisis limitada se utilizaron para la identificar fragmentos de la proteína de interés que se comportan como dominios globulares más estables que la proteína entera y, por lo tanto, tienen más posibilidades de cristalizar. Estos experimentos se basan en el hecho que las proteínas más globulares son más resistentes a la proteólisis y que un simple análisis de los productos de digestión en gel de poliacrilamida desnaturalizante permite una fácil identificación de estos fragmentos (Fontana, Fassina et al. 1986). Para llevar a cabo estos experimentos se mezclaron 10 µg de la proteína en presencia de las siguientes diluciones de las proteasas tripsina (10 mg/mL, quimiotripsina (1 mg/mL) y papaina (10 mg/mL): 1/10, 1/100, 1/500 y/o 1/1000, y se incubaron las reacciones a 20 y/o 37°C. La proteólisis con tripsina se realizó en el tampón 0.045 M Tris-HCl pH8, con quimiotripsina se utilizó el tampón 1 M

CaCl₂, 0.08 M Tris-HCl pH7.8 y con papaina se utilizó 1 M K₂HPO₄, 0.025 M EDTA, 0.03 mM β-mercaptoetanol. Las reacciones se bloquearon con tampón de carga de geles de poliacrilamida desnaturalizante.

4.5 Métodos biofísicos para la caracterización proteica

La determinación estructural de proteínas requiere entender el comportamiento de éstas en solución; su estabilidad, propensión a sufrir cambios conformacionales o a agregar, en función de condiciones experimentales, tales como: la composición del tampón utilizado, la temperatura y el tiempo. Estos son los principales motivos por los cuales se utilizaron en diferentes técnicas biofísicas que permitieron caracterizar las proteínas en solución.

4.5.1 Dispersión dinámica de luz (Dynamic Light Scattering , DLS)

La técnica de DLS permite evaluar la monodispersidad de soluciones proteicas puras y estimar el radio hidrodinámico (Rh) de partículas en solución de una manera rápida y sensible (Berne, Pecora, 2000). Los ensayos de DLS se realizaron a temperatura

ambiente con 60 μ L del complejo formado por DnaK Δ Ct-GrpE a \sim 10 mg/mL en tampón 0.5 M NaCl, 0.02 M Tris-HCl pH8 o 0.05 M citrato de sodio pH 5.5 utilizando el instrumento de DLS Zetasizer Nano (Malvern Instruments, Worcesterchire, UK).

4.5.2 DSF

Mediante DSF se puede comprobar la termoestabilidad de las proteínas en solución y si la adición de compuestos de bajo peso molecular con diferentes naturalezas químicas les confieren mayor estabilidad (Niesen, Berglund et al. 2007). Hay varios parámetros que pueden afectar a la estabilidad de una proteína como la temperatura, la composición del tampón, su pH, la presencia de sales, detergentes y/o ligandos que se pueden unir a la proteína de forma específica o inespecífica, llevando a una variación en su energía de Gibbs. Esta diferencia en la energía de Gibbs puede ser aprovechada para detectar cambios de estabilidad en la solución de proteína debido a cambios en su temperatura de fusión (T_m , del inglés melting temperature). En la técnica de DSF se utiliza una sonda fluorescente (Sypro Orange[®], Life Technologies, Grand Island, New York) que se une a las zonas hidrofóbicas de la proteína. Al aplicar un gradiente de temperatura a la muestra, esta pierde estabilidad y se desnaturaliza, proceso que es seguido a través de la monitorización de fluorescencia emitida en función de la temperatura. El experimento resulta en una curva sigmoidea cuyo punto de inflexión corresponde a la T_m (Niesen, Berglund et al. 2007).

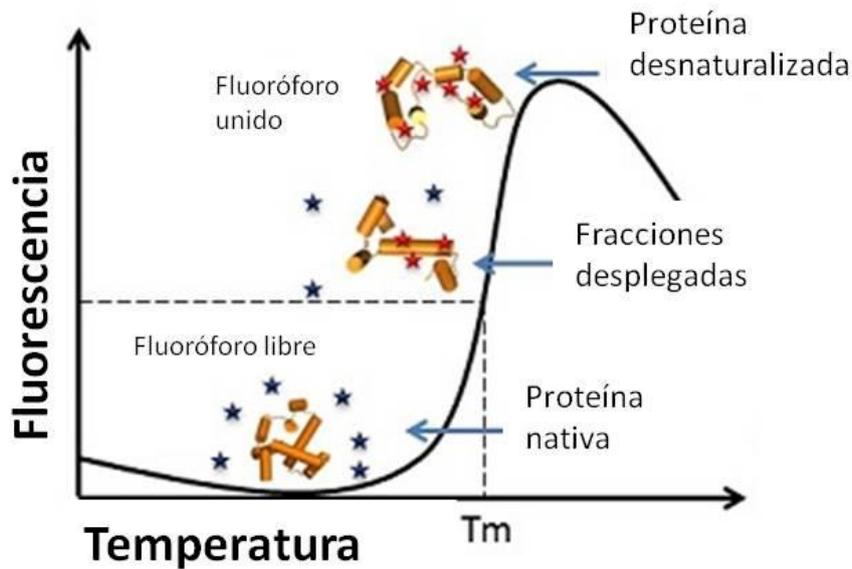


Figura 4.1 Curva tipo en los ensayos de DSF. En la primera parte de la curva se observa que el fluoróforo está libre y la proteína está plegada. Una vez aumenta la temperatura la proteína se va desnaturalizando, permitiendo la unión del fluoróforo. Cuando la proteína está desnaturalizada totalmente, el fluoróforo se ha unido dando la máxima señal. Imagen adaptada de <http://www.jenabioscience.com/cms/en/1/catalog/2570>

Los ensayos de DSF se realizaron en placas de 96 pozos conteniendo cada uno de ellos la mezcla: 5 μ L proteína a 10mg/ml (DnaK Δ Ct o GrpE) en 0.5 M NaCl, 0.02M Tris-HCl pH 8, 7.5 μ L SYPRO® orange a una concentración final 1x (Invitrogen, Carlsbad, CA, USA), preparado en el mismo tampón que la proteína a analizar, y 12.5 μ L del compuesto a ser probado 2x concentrado. El cribado de aditivos consistieron en tampones con diferentes composiciones químicas y abarcando un rango de pHs de 3.5 a 10 en presencia de 0.1 o 0.5 M NaCl. Este fue diseñado basándonos en el trabajo de Reinhard *et al.* (Nettlehip, Brown *et al.* 2008, Reinhard, Mayerhofer *et al.* 2013). Las medidas de DSF se realizaron en un termociclador de tiempo real iCycler5 (BioRad, Hercules, CA, USA) y la temperatura del barrido se incrementó de 1°C por minuto desde los 25 hasta los 85°C.

4.5.3 Dispersión de luz en múltiples ángulos-cromatografía de gel filtración (MALS-SEC)

Generalmente se ha utilizado la cromatografía de gel filtración para caracterizar de manera aproximada el peso molecular de proteínas. Este peso molecular se puede determinar cuando se usan como referencia proteínas de peso molecular conocido pero hay limitaciones en la técnica como el hecho que la elución de la proteína no sólo depende de su masa sino también de la forma y de la tendencia de la proteína a interactuar con la matriz de la columna. La técnica de la dispersión de luz (Light Scattering (LS)) también se ha utilizado durante muchos años como método para determinar el peso molecular, la forma y las oligomerizaciones de proteínas en solución (Zimm 1948)). Cuando la luz pasa a través de un medio como por ejemplo, un solvente, la interacción con las moléculas produce una dispersión del haz de luz. Esta dispersión puede ser detectada por un detector y su intensidad puede ser medida en diferentes ángulos del haz de luz. La ventaja de usar MALS acoplado a un sistema de cromatografía de exclusión molecular es que permite determinar el peso molecular absoluto, el tamaño (*radio de giro* R_g) y la conformación de las proteínas porque se puede calcular la cantidad de luz dispersada de cada ángulo, solventando los problemas asociados a detectores de bajo ángulo. La cantidad de luz dispersada se relaciona con la masa según la ecuación :

$$K^*c/R(\Phi,c)=1/M_wP(\Phi)+2A_2c$$

Donde $R(q,c)$ es el exceso del ratio Raleigh de la solución en función del ángulo dispersado q y la concentración c . Es directamente proporcional a la intensidad de la luz dispersada por el solvente. c es la concentración del soluto. M_w es la media del peso molecular. A_2 es el segundo coeficiente de expansión virial de la presión osmótica. K es la constante $4\pi^2(dn/dc)^2 n_0^2/Na^{104}$. N_a es el número de Avogadro y la $P(q)$ describe la dependencia angular de la luz dispersada y se puede relaciona con rms (Zimm 1948).

Para realizar este experimento se cargaron entre 10-50 μ l de la proteína a una Superdex 200 10 300 o una Superdex 200 5 150 Increase (GE, Healthcare,) previamente equilibrada con el tampón 0.5M NaCl, 0.02M Tris pH8 o 0.1M NaCl, 0.02M tris pH8. Posteriormente, se analizaron los datos obtenidos con el programa ASTRA^{®6} (Wyatt tecnologías, Sta. Barbara, California) y con la ayuda del Dr. Marc Jamin (Unit for Virus Host Cell Interactions, Université Grenoble Alpes-EMBL-CNRS, Grenoble, France).

4.6 Difracción de rayos-X en cristales de proteína

4.6.1 Cristalización de proteínas

La obtención de cristales de proteínas cuya difracción por rayos-X dé una resolución adecuada (3-3.5Å) que permita resolver la estructura es el paso limitante de la cristalografía de proteínas. Para obtener cristales se realizaron ensayos cuyo objetivo

es provocar una sobresaturación de la solución proteica para que se dé una precipitación ordenada en forma de cristal. En la cristalización básicamente se distinguen dos pasos: la nucleación y la fase de crecimiento del cristal. La fase de nucleación es favorecida por niveles de sobresaturación elevados, mientras que la fase de crecimiento del cristal es favorecida por niveles de saturación menores, ya que el objetivo es que el crecimiento del cristal sea lento y ordenado (Jancarik, Kim 1991)(Figura 2.2) .

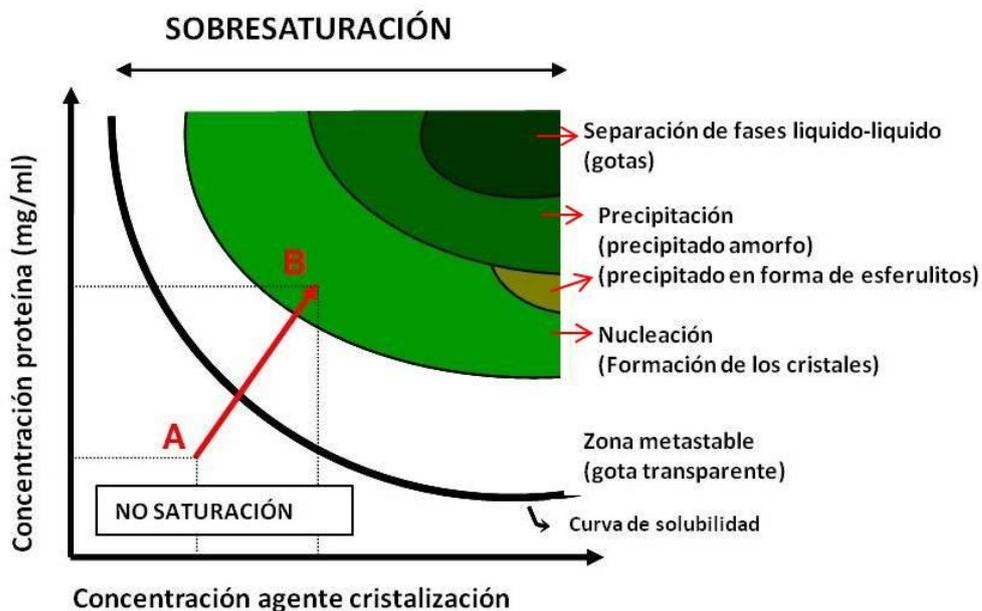


Figura 4.2 Diagrama de fases de los diferentes estados de saturación de una solución proteica, en función de la proteína y del agente precipitante. La flecha roja indica el cambio que sufre la solución proteica, a causa de la variación de las condiciones de la gota, hacia un estados en el que se pueden formar los núcleos cristalinos. Adaptación de la figura de <http://www-structmed.cimr.cam.ac.uk>

Los ensayos de cristalización se realizaron mediante la técnica de difusión de vapor que consiste en establecer un equilibrio de vapor en una gota que contiene la proteína y la solución de cristalización que está en el reservorio. La gota de proteína con la solución de cristalización se cierra herméticamente sobre el reservorio para que se pueda dar el equilibrio. Con las soluciones de cristalización se pueden testar diferentes

pH, precipitantes y aditivos. Normalmente, la concentración del agente precipitante es mayor que el de la gota de proteína y, por lo tanto, posee una presión de vapor inferior que el de la solución proteica, de manera, que habrá una evaporación de la solución proteica hacia el reservorio resultando en la sobresaturación de la gota de proteína. Si las condiciones son favorables, esta sobresaturación dará a lugar a la formación de cristales. Existen dos maneras de llevar a cabo el método de difusión de vapor: el de gota sentada y el de gota colgada. En el método de la gota sentada la mezcla proteína: reservorio se deposita en un pequeño pocillo más elevado que el reservorio y se sella la placa con cinta adhesiva especial no porosa, que impide la evaporación hacia el exterior. Esta técnica se utilizó para los cribados iniciales de cristalización que fueron realizados en placas de 96 condiciones en los cuales 100nL de proteína + 100nL de la solución de cristalización del reservorio fueron depositadas por el robot dispensador Cartesian Honeybee (Hamilton, Reno, NV, USA) o bien por el Crystal Phoenix (ARInstruments, Sunnyvale, CA, USA). Cuando se obtuvieron resultados prometedores de los cribados iniciales se optimizaron las condiciones de cristalización a gran escala y manualmente, utilizando placas Limbro de 24 pocillos (Hampton research) y usando el método de gota colgante. Con esta técnica la solución proteica con la solución de cristalización se deposita en un cubreobjetos que sellará el reservorio. Las cristalizaciones se realizaron tanto a 20 como a 4°C y las optimizaciones manuales se diseñaron variando las concentraciones de los componentes de la solución de cristalización, el pH, el precipitante, los aditivos o variando la proporción proteína: solución de cristalización en la gota (Chayen 1998, Dessau and Modis 2011).

4.6.2 Crioprotección de cristales de proteínas

Los cristales proteicos cuando son expuestos a un haz de rayos-X pueden sufrir daño por esta radiación y provocar que los datos de la difracción pierdan resolución y calidad. Por ese motivo, es necesario que los cristales se congelen muy rápidamente (*flash-cooling*) y se debe evitar la formación de hielo cristalino en la muestra, ya que distorsionaría el orden interno del cristal e interferiría en el patrón de difracción obtenido. Por ese motivo se añaden crioprotectores que provocan un estado vítreo del solvente. En nuestro caso, los cristales se sumergieron en una solución compuesta por la solución de cristalización a la que se le añadió un 30% de glicerol. Seguidamente, se pescaron con un *loop* de dimensiones apropiadas y se congeló muy rápidamente en nitrógeno líquido.

4.6.3 Difracción de rayos-X

- **Recogida, procesado y escalado de datos**

Los cristales de DnaK Δ Ct se recogieron en la línea ID23eh1 del ESRF (Grenoble, Francia) mientras que los cristales de DnaK-NBD en presencia de ATP se difractaron en la línea ID23eh2 del ESRF y los otros cristales de las dos variantes de la proteína en presencia o ausencia de otros nucleótidos se difractaron en la línea BL13-XALOC del ALBA (Cerdanyola del Vallès, Barcelona, España). La estrategia para la recogida de datos se

realizó de manera automática con los programas disponibles en cada una de las líneas donde se difractaron los cristales.

La calidad y adecuación de los datos fueron evaluadas considerando indicadores estadísticos como R_{merge} , R_{sym} , cociente $I/\sigma(I)$ y completitud de los datos (Glusker 1994). El R_{merge} y el R_{sym} son parámetros estadísticos que miden las desviaciones estándar entre las intensidades observadas múltiples veces bien sea por repetición o por simetría. La medida de $I/\sigma(I)$ nos indica la intensidad promedio de los datos con respecto al promedio de error de los mismos y la completitud indica el porcentaje de las reflexiones únicas observadas con respecto a las posibles.

Todos los datos recogidos se procesaron con el programa XDS (Kabsch 2010). El archivo resultante conteniendo las reflexiones (en formato .HKL) fue transformado a un formato .mtz a través de COMBAT (Collaborative Computational Project 1994) para poder ser luego, escalados con el programa Scala (Evans 2006).

Para estimar el número de moléculas en la unidad asimétrica y el porcentaje de solvente de los cristales se utilizó el coeficiente de Matthews (Matthews 1968). Tanto en el caso de los cristales de DnaK Δ Ct como de DnaK-NBD solo existía una molécula por unidad asimétrica con un contenido de solvente de 65.5 Y 56, respectivamente.

- Resolución de la estructura tridimensional

Uno de los principales problemas en la cristalografía de proteínas es la pérdida de información relacionada con las fases, el llamado problema de las fases. Cuando se recoge un conjunto de datos de difracción de rayos-X de un cristal se miden las intensidades de las ondas dispersadas en todas las direcciones de los diferentes planos

del cristal. Es a partir de estas intensidades que se derivan las amplitudes de las ondas dispersadas pero en el experimento se pierde la información de las fases. Para solventar el problema y recuperar las informaciones contenidas en las fases, hay tres métodos principales:

i) Método del Reemplazo Isomorfo Múltiple (*Multiple Isomorphous Replacement, MIR*). Este método requiere la inserción de átomos pesados, que son grandes dispersores de los rayos X, en el cristal. Las diferencias en la intensidad de las reflexiones de cristales que contienen átomos pesados con las reflexiones de cristales nativos se pueden usar para calcular un mapa de Patterson del cual se derivan las coordenadas de los átomos pesados con las cuales se puede entonces calcular las fases iniciales de la proteína (Perutz, Kendrew, 1938).

ii) Método de Dispersión Anómala Múltiple (*Multiple-wavelength Anomalous Dispersion, (MAD)*): En este caso los cambios en las intensidades de la difracción son debidas a modificaciones de las propiedades físicas de los átomos. De este modo, si la longitud de onda de la radiación incidente tiene una frecuencia próxima a la frecuencia natural de oscilación de los electrones de un determinado átomo, que pueden tratarse de átomos abundantes en la proteína, como por ejemplo P o S, o bien, átomos pesados añadidos al cristal (como en el caso MIR), se produce la denominada dispersión anómala que causará un cambio en la amplitud y en la fase. Esta información nos permitirá resolver el problema de las fases. (Hendrickson, Smith et al. 1988).

iii) Método de Reemplazo Molecular (*Molecular Replacement, MR*). Esta técnica utiliza un modelo estructural de una proteína con una secuencia de aminoácidos homóloga a la proteína que se quiere determinar, utilizándolo como un primer modelo que

posteriormente será refinado. Este procedimiento está basado en la observación de que proteínas homólogas en su secuencia peptídica muestran un plegamiento muy similar. El método consiste, entonces, en transferir el modelo de un sistema cristalino conocido al del cristal de la proteína con estructura desconocida. El posicionamiento del modelo en la celdilla unidad de la proteína desconocida requiere determinar su correcta orientación y su posición precisa, ambas operaciones, rotación y translación, calculadas mediante las denominadas funciones de rotación y de translación (Rossmann 1990).

En el presente trabajo, la existencia de modelos homólogos permitió solventar el problema de las fases mediante MR utilizando los programas MolRep (Vagin and Teplyakov 2010) o Phaser (Adams, Afonine et al. 2010). Como modelos de búsqueda se utilizaron la DnaK de *G. kaustophilus* (código PDB: 2V7Y) para determinar la estructura de la DnaK Δ Ct en presencia de AMP-PNP y esta, una vez resuelta, se utilizó como modelo de búsqueda para determinar las estructuras de las DnaK Δ Ct y DnaK-NBD en presencia de ATP. Las estructura de DnaK-NBD en presencia de AMP-PNP, ADP y en su forma apo se resolvieron utilizando como modelo de búsqueda la DnaK-NBD:ATP.

- **Construcción y refinamiento de los modelos moleculares**

Los mapas de densidad electrónica calculados después de solventar el problema de las fases fueron utilizados para reconstruir los modelos iniciales de manera manual con el programa Coot (Emsley and Cowtan 2004). Los modelos fueron mejorados alternando ciclos de refinamiento con el programa Refmac (Murshudov, Vagin et al. 1997, Murshudov, Skubak et al. 2011) incorporando TLS y de construcción interactiva del modelo con el programa Coot.

- Validación de la estructura

La validación de una estructura cristalina consiste en examinar tanto los datos experimentales como los del modelo para detectar errores o desviaciones y confirmar la idoneidad del modelo. Los indicadores utilizados para medir el buen ajuste del modelo a los datos experimentales son el R_{factor} y el R_{free} . El R_{factor} estima el desacuerdo entre los factores de estructura observados y los calculados, indicando lo bien que el modelo predice las amplitudes de las reflexiones medidas. Un R_{factor} por debajo del 20% indica generalmente la buena calidad del modelo, aunque depende de la resolución final del modelo. El R_{free} mide de forma objetiva el grado en el que el modelo atómico construido predice las amplitudes observadas al ser calculado, empleando el mismo algoritmo utilizado para calcular el R_{factor} , pero a partir de un pequeño conjunto de reflexiones (normalmente un 5%) tomadas aleatoriamente y no utilizadas en el refinamiento (Brunger 1992).

La calidad del modelo también debe ser evaluada según parámetros estereoquímicos. Se analiza que haya ausencia de choques entre las cadenas laterales de los aminoácidos, que las distancias entre átomos y los ángulos de los enlaces sean geoméricamente adecuados y que los parámetros térmicos sean razonables. Además de estos parámetros es importante contrastar que las conformaciones que adoptan los enlaces N-C α y C α -C del enlace peptídico no dan lugar a aproximaciones interatómicas energéticamente improbables. Las torsiones del enlace peptídico (Φ y ψ) presentan en principio libertad de giro sin embargo, cada uno de estos ángulos está energéticamente limitado a intervalos definidos en el diagrama de *Ramachandran*, que muestra las regiones permitidas de estos ángulos para cada uno de los tipos de

estructura secundaria. Se utilizó el servidor Molprobitry para validar todos los modelos (Chen, Arendall et al. 2010).

4.5.4 Resonancia superficial por plasmón (SPR)

La técnica de SPR generalmente se utiliza para identificar y caracterizar interacciones dos moléculas una de las cuales se encuentra inmovilizada a la superficie de un biosensor y la otra, el analito, es inyectada a través del biosensor (Jonsson, Fagerstam et al. 1991, Wilson 2002). El fenómeno de SPR se produce cuando la luz polarizada, rebota en una placa de oro eléctricamente conductora con un cierto ángulo de incidencia. La luz polarizada, que cubre un rango de los ángulos incidentes, es dirigida a la superficie del sensor, en particular a la cara del cristal siendo la luz es detectada dentro del sistema. La intensidad en el campo eléctrico, conocido como onda evanescente, se genera cuando la luz incide sobre el cristal. Esta onda evanescente interactúa y es absorbida por nubes de electrones libres en la capa de oro, generando ondas de densidad de carga de electrones llamadas plasmones y causando una reducción en la intensidad de la luz reflejada. El ángulo de resonancia en el que se produce este mínimo de intensidad es función del índice de refracción de la solución cerca de la capa de oro, en la cara opuesta a la superficie del sensor (Figura 4.3, A). La interacción entre el analito y la proteína inmovilizada produce un cambio en el índice de refracción en la superficie del chip que se calcula en unidades de resonancia (RU), siendo que 1 RU equivale a la densidad de superficie de una proteína a 1 pg/mm^{-2} . Las interacciones son monitorizadas en un sensograma que representa la interacción en

tiempo real. Cuando no hay interacción, el ángulo de SPR no cambia de manera que no habrá alteraciones en las unidades de resonancia (Figura 4.3, B).

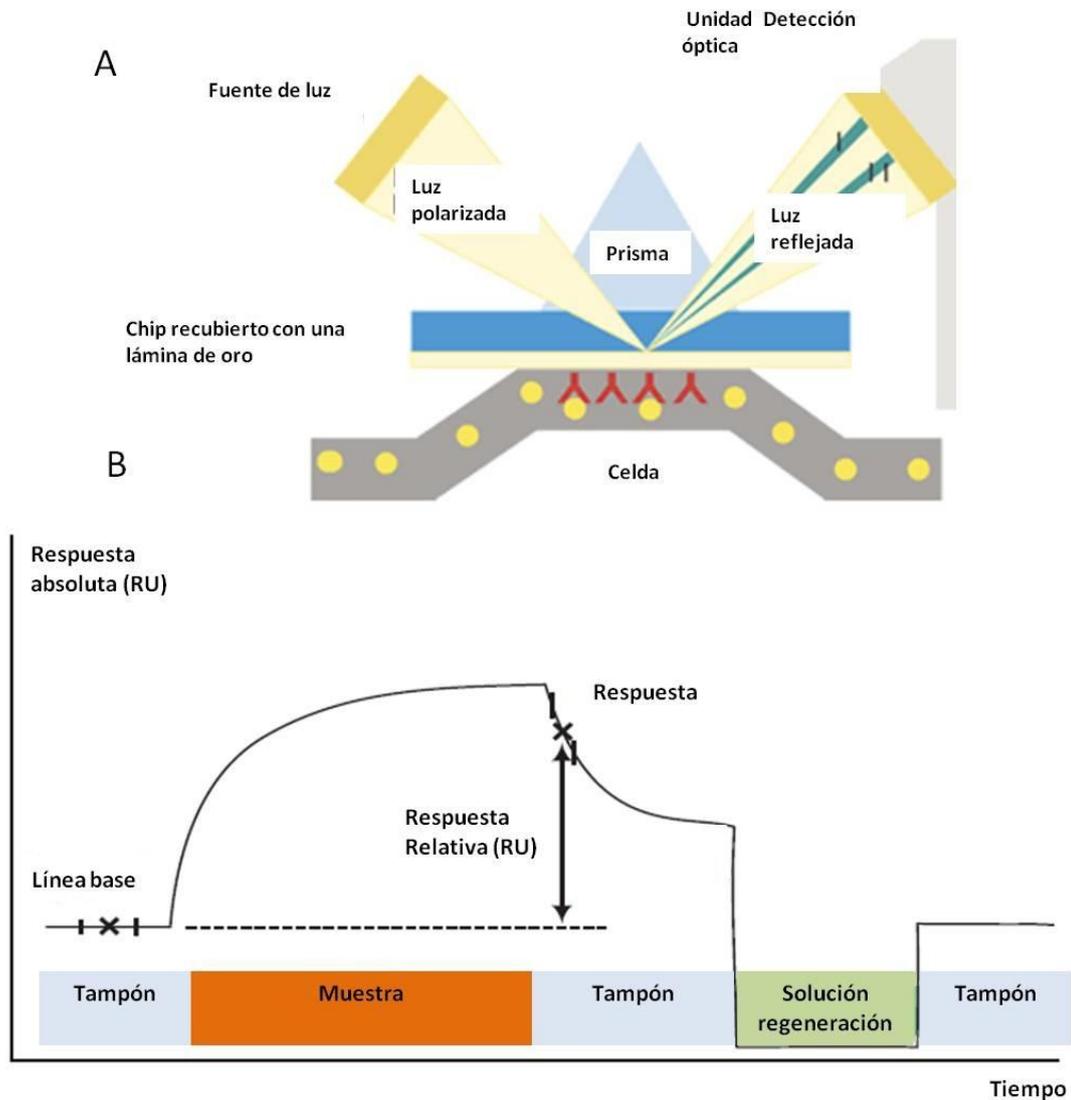


Figura 4.3 A) Esquema del fenómeno de SPR. Adaptación de la figura de <http://www.rci.rutgers.edu/~longhu/Biacore/>. B) Ilustración esquemática del sensograma obtenido en un experimento de SPR. Las barras bajo las curvas del sensograma indican las soluciones que pasan sobre la superficie del sensor. Figura adaptada de Biacore assay handbook (GE Healthcare, Uppsala, Sweden)

En el presente trabajo DnaKΔCt y los mutantes DnaKΔCt-L522D y DnaKΔCt-L522K fueron inmovilizados en la superficie de un biosensor CM5 (GE Healthcare, Uppsala, Sweden) previamente activado según el protocolo de GE Healthcare (GE Healthcare

Life science, Uppsala, Sweden). Este sensor es de los más versátiles que existen en el mercado, está formado por una matriz de dextran carboximetilada unido a la lámina de oro. La proteína se une a la matriz por grupos amina, tiol, aldehído o carboxi, de manera que hay una gran variabilidad para que se de la unión. La capacidad tan elevada de unión proporciona una elevada respuesta, ideal para ensayos de captura o en el que se encuentran involucradas moléculas pequeñas. Generalmente, se inyectó a 10 $\mu\text{L}/\text{min}$ 10 μL de proteína a 1 mg/mL hasta obtener RUs de $\sim 3000-4000$. El ensayo se realizó a temperatura ambiente. Las proteínas inmovilizadas en el biosensor fueron diluidas en tampón 0.01 M acetato de sodio pH 5 y el tampón utilizado para el ensayo fue 0.02 M fosfato de potasio pH 7.4. Se siguió un paso de desactivación del biosensor según el protocolo de GE Healthcare (GE Healthcare Life science, Uppsala, Sweden) y posteriormente se inyectaron consecutivamente 60 μL de los siguientes analitos a 20 $\mu\text{L}/\text{min}$, utilizando como tampón del experimento 0.15 M NaCl, 0.02 M Tris-HCl pH 8, 0.001 M MgCl_2 , 0.005% (v/v) surfactante P20: MG019-DnaJ (aminoácidos del 1 al 79), MG019_DnaJ-G/F (aminoácidos del 1 al 120), MG200-DnaJ (aminoácidos del 1 al 74) y MG200-DnaJ-G/F (aminoácidos del 1 al 120). La disociación fue monitorizada en cada medida durante 3 minutos después de las cuales el biosensor fue regenerado con 2M NaCl. La regeneración del chip se utiliza para eliminar los posibles analitos que han quedado fijados a la proteína, y poder repetir el experimento partiendo de una línea base óptima.

En el ensayo de interacción de las diferentes variantes de DnaK con GrpE la concentración máxima que se inyectó de esta fue de 10 nM, ya que se producía una interacción tan fuerte que obligaba a regenerar el chip de manera muy agresiva, porque la GrpE no se disociaba de DnaK. Se utilizaron dos tipos de protocolos de

inyección: inyecciones de 60 μL a un flujo de 20 $\mu\text{L}/\text{min}$ monitorizando la disociación durante 3 minutos o inyecciones de 120 μL de proteína a un flujo de 20 $\mu\text{L}/\text{min}$ monitorizando la disociación durante 6 minutos. El segundo protocolo permitió obtener una curva de respuesta más lenta, y detectar mejor el modelo bifásico entre la interacción de DnaK y GrpE . Para la regeneración del biosensor se utilizaron diferentes concentraciones de NaCl (hasta 2 M) y de SDS (hasta 0.08% (w/v)). Este protocolo de regeneración de la superficie del biosensor fue agresivo y provocó una desnaturalización y pérdida parcial de las proteínas inmovilizadas. En este ensayo se utilizó como tampón de corrida 0.15 M NaCl, 0.02 M Tris-HCl pH 8, 0.001 M MgCl_2 , 0.005% (v/v) surfactante P20, 0.5 M AMP-PNP.

4.7 Dispersión de rayos-X a bajo ángulo (SAXS)

La técnica de SAXS permite determinar el peso molecular y un modelo a baja resolución de la macromolécula en solución. La muestra en solución es cargada en un capilar que es irradiado con un haz de rayos-X. Estos son dispersados por los electrones de la capa electrónica de las macromoléculas en solución y esta señal es recogida por un detector. El patrón de dispersión recogido es transformado en una curva de dispersión (Baldon, 2014). La curva experimental obtenida se divide en tres regiones según la información que se puede extraer (Figura 4.4). En la región de Guinier el R_g experimental nos da información sobre el tamaño de las partículas pero hay que considerar que puede verse afectado por la agregación de las partículas en solución, por su polidispersidad y/o por la sustracción incorrecta de la curva

experimental del tampón. La región Fourier proporciona información sobre la forma de la partícula y la región Porod sobre su superficie. Todos los datos se analizaron con el paquete de programas ATSAS (Petoukhov, Franke et al. 2012).

Para utilizar la técnica de SAXS es necesario que las muestras sean homogéneas y monodispersas. Cuando, la muestra no es del todo homogénea o monodispersa, se puede acoplar a la técnica una columna de gel filtración en línea con el capilar donde se realizan las medidas de SAXS, permitiendo así, separar la posible mezcla de macromoléculas en solución (Pernot, Round et al. 2013). La técnica SEC-SAXS es particularmente apropiada para medir proteínas que tienen tendencia a la oligomerizar o a agregar.

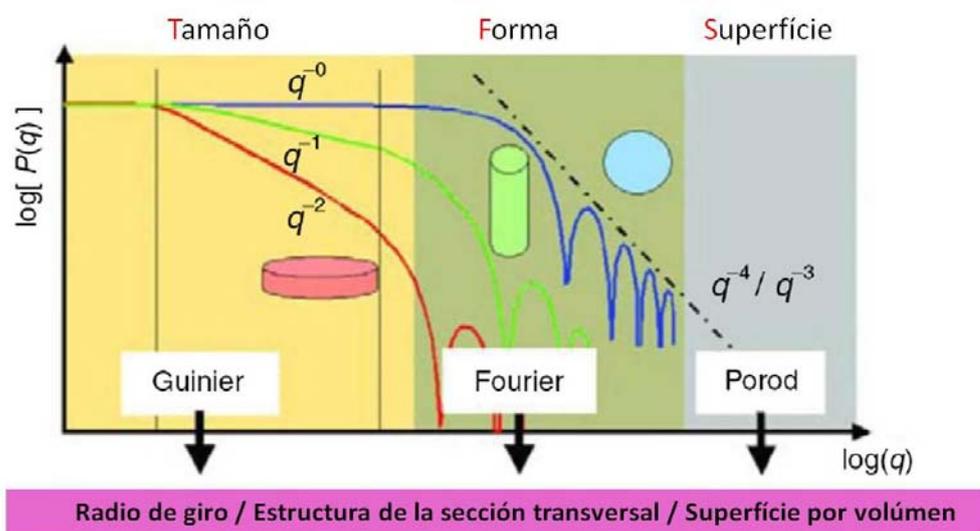


Figura 4.3 A) Esquema del fenómeno de SPR. Adaptación de la figura de <http://www.rci.rutgers.edu/~longhu/Biacore/>. B) Ilustración esquemática del sensograma obtenido en un experimento de SPR. Las barras bajo las curvas del sensograma indican las soluciones que pasan sobre la superficie del sensor. Figura adaptada de Biacore assay handbook (GE Healthcare, Uppsala, Sweden)

En el presente trabajo las muestras, medidas en la línea BM29 del ESRF (Grenoble, Francia), fueron inyectadas a través de una columna de gel filtración Superdex 200 increase 3.2/300 (GE Healthcare, Life Sciences, Uppsala, Sweden), equilibrada en 0.1 M NaCl, 0.02 M Tris-HCl pH 8, y inmediatamente irradiadas por rayos-X. Se han estudiado utilizando este protocolo soluciones de DnaK Δ Ct:AMP-PNP, DnaK-NBD:AMP-PNP y GrpE concentradas respectivamente a 15, 10, 7 y 15 mg/mL. Todas las mediciones fueron ejecutadas a 20°C, a un flujo de 0.075 mL/min y recogiendo 1 imagen cada segundo.

Bibliografía

Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 213-221.

Ahmad, A., A. Bhattacharya, R. A. McDonald, M. Cordes, B. Ellington, E. B. Bertelsen and E. R. Zuiderweg (2011). "Heat shock protein 70 kDa chaperone/DnaJ cochaperone complex employs an unusual dynamic interface." Proc Natl Acad Sci U S A **108**(47): 18966-18971.

Albanese, V., A. Y. Yam, J. Baughman, C. Parnot and J. Frydman (2006). "Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells." Cell **124**(1): 75-88.

Andreasson, C., J. Fiaux, H. Rampelt, S. Druffel-Augustin and B. Bukau (2008). "Insights into the structural dynamics of the Hsp110-Hsp70 interaction reveal the mechanism for nucleotide exchange activity." Proc Natl Acad Sci U S A **105**(43): 16519-16524.

Andreasson, C., J. Fiaux, H. Rampelt, M. P. Mayer and B. Bukau (2008). "Hsp110 is a nucleotide-activated exchange factor for Hsp70." J Biol Chem **283**(14): 8877-8884.

Bairoch, A., R. Apweiler, C. H. Wu, W. C. Barker, B. Boeckmann, S. Ferro, E. Gasteiger, H. Huang, R. Lopez, M. Magrane, M. J. Martin, D. A. Natale, C. O'Donovan, N. Redaschi and L. S. Yeh (2005). "The Universal Protein Resource (UniProt)." Nucleic Acids Res **33**(Database issue): D154-159.

Balish, M. F., T. W. Hahn, P. L. Popham and D. C. Krause (2001). "Stability of *Mycoplasma pneumoniae* cytoadherence-accessory protein HMW1 correlates with its association with the triton shell." J Bacteriol **183**(12): 3680-3688.

Bang, H., A. Pecht, G. Raddatz, T. Scior, W. Solbach, K. Brune and A. Pahl (2000). "Prolyl isomerases in a minimal cell. Catalysis of protein folding by trigger factor from *Mycoplasma genitalium*." Eur J Biochem **267**(11): 3270-3280.

Barraclough, R. and R. J. Ellis (1980). "Protein synthesis in chloroplasts. IX. Assembly of newly-synthesized large subunits into ribulose biphosphate carboxylase in isolated intact pea chloroplasts." Biochim Biophys Acta **608**(1): 19-31.

Bertelsen, E. B., L. Chang, J. E. Gestwicki and E. R. Zuiderweg (2009). "Solution conformation of wild-type *E. coli* Hsp70 (DnaK) chaperone complexed with ADP and substrate." Proc Natl Acad Sci U S A **106**(21): 8471-8476.

Bhattacharya, A., A. V. Kurochkin, G. N. Yip, Y. Zhang, E. B. Bertelsen and E. R. Zuiderweg (2009). "Allostery in Hsp70 chaperones is transduced by subdomain rotations." J Mol Biol **388**(3): 475-490.

Brunger, A. T. (1992). "Free R value: a novel statistical quantity for assessing the accuracy of crystal structures." Nature **355**(6359): 472-475.

Buchan, D. W., F. Minneci, T. C. Nugent, K. Bryson and D. T. Jones (2013). "Scalable web services for the PSIPRED Protein Analysis Workbench." Nucleic Acids Res **41**(Web Server issue): W349-357.

Buchberger, A., A. Valencia, R. McMacken, C. Sander and B. Bukau (1994). "The chaperone function of DnaK requires the coupling of ATPase activity with substrate binding through residue E171." EMBO J **13**(7): 1687-1695.

Bukau, B. and A. L. Horwich (1998). "The Hsp70 and Hsp60 chaperone machines." Cell **92**(3): 351-366.

Burgos, R., O. Q. Pich, M. Ferrer-Navarro, J. B. Baseman, E. Querol and J. Pinol (2006). "Mycoplasma genitalium P140 and P110 cytoadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development." J Bacteriol **188**(24): 8627-8637.

Burgos, R., O. Q. Pich, E. Querol and J. Pinol (2007). "Functional analysis of the Mycoplasma genitalium MG312 protein reveals a specific requirement of the MG312 N-terminal domain for gliding motility." J Bacteriol **189**(19): 7014-7023.

Burgos, R., O. Q. Pich, E. Querol and J. Pinol (2008). "Deletion of the Mycoplasma genitalium MG_217 gene modifies cell gliding behaviour by altering terminal organelle curvature." Mol Microbiol **69**(4): 1029-1040.

Calisto, B. M., A. Broto, L. Martinelli, E. Querol, J. Pinol and I. Fita (2012). "The EAGR box structure: a motif involved in mycoplasma motility." Mol Microbiol **86**(2): 382-393.

Calloni, G., T. Chen, S. M. Schermann, H. C. Chang, P. Genevaux, F. Agostini, G. G. Tartaglia, M. Hayer-Hartl and F. U. Hartl (2012). "DnaK functions as a central hub in the E. coli chaperone network." Cell Rep **1**(3): 251-264.

Cetinbas, M. and E. I. Shakhnovich (2013). "Catalysis of protein folding by chaperones accelerates evolutionary dynamics in adapting cell populations." PLoS Comput Biol **9**(11): e1003269.

Chambaud, I., R. Heilig, S. Ferris, V. Barbe, D. Samson, F. Galisson, I. Moszer, K. Dybvig, H. Wroblewski, A. Viari, E. P. Rocha and A. Blanchard (2001). "The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis." Nucleic Acids Res **29**(10): 2145-2153.

Chang, L. J., W. H. Chen, F. C. Minion and D. Shiuan (2008). "Mycoplasmas regulate the expression of heat-shock protein genes through CIRCE-HrcA interactions." Biochem Biophys Res Commun **367**(1): 213-218.

Chang, Y. W., Y. J. Sun, C. Wang and C. D. Hsiao (2008). "Crystal structures of the 70-kDa heat shock proteins in domain disjoining conformation." J Biol Chem **283**(22): 15502-15511.

Chayen, N. E. (1998). "Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques." Acta Crystallogr D Biol Crystallogr **54**(Pt 1): 8-15.

Cheetham, M. E. and A. J. Caplan (1998). "Structure, function and evolution of Dnal: conservation and adaptation of chaperone function." Cell Stress Chaperones **3**(1): 28-36.

Chen, V. B., W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson (2010). "MolProbity: all-atom structure validation for macromolecular crystallography." Acta Crystallogr D Biol Crystallogr **66**(Pt 1): 12-21.

Clark, G. W. and E. R. Tillier (2010). "Loss and gain of GroEL in the Mollicutes." Biochem Cell Biol **88**(2): 185-194.

Cloward, J. M. and D. C. Krause (2009). "Mycoplasma pneumoniae J-domain protein required for terminal organelle function." Mol Microbiol **71**(5): 1296-1307.

Collaborative Computational Project, N. (1994). "The CCP4 suite: programs for protein crystallography." Acta Crystallogr D Biol Crystallogr **50**(Pt 5): 760-763.

Csermely, P., T. Korcsmaros, I. A. Kovacs, M. S. Szalay and C. Soti (2008). "Systems biology of molecular chaperone networks." Novartis Found Symp **291**: 45-54; discussion 54-48, 137-140.

Dascher, C. C., S. K. Poddar and J. Maniloff (1990). "Heat shock response in mycoplasmas, genome-limited organisms." J Bacteriol **172**(4): 1823-1827.

Derre, I., G. Rapoport and T. Msadek (1999). "CtsR, a novel regulator of stress and heat shock response, controls clp and molecular chaperone gene expression in gram-positive bacteria." Mol Microbiol **31**(1): 117-131.

Dessau, M. A. and Y. Modis (2011). "Protein crystallization for X-ray crystallography." J Vis Exp(47).

Dhandayuthapani, S., W. G. Rasmussen and J. B. Baseman (1998). "Identification of mycoplasmal promoters in Escherichia coli using a promoter probe vector with Green Fluorescent Protein as reporter system." Gene **215**(1): 213-222.

Di Tommaso, P., S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J. M. Chang, J. F. Taly and C. Notredame (2011). "T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension." Nucleic Acids Res **39**(Web Server issue): W13-17.

Dominguez, R. (2004). "Actin-binding proteins--a unifying hypothesis." Trends Biochem Sci **29**(11): 572-578.

Dragovic, Z., S. A. Broadley, Y. Shomura, A. Bracher and F. U. Hartl (2006). "Molecular chaperones of the Hsp110 family act as nucleotide exchange factors of Hsp70s." EMBO J **25**(11): 2519-2528.

Ellis, J. (1987). "Proteins as molecular chaperones." Nature **328**(6129): 378-379.

Ellis, R. J. and S. M. Hemmingsen (1989). "Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures." Trends Biochem Sci **14**(8): 339-342.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-2132.

Evans, P. (2006). "Scaling and assessment of data quality." Acta Crystallogr D Biol Crystallogr **62**(Pt 1): 72-82.

Falah, M. and R. S. Gupta (1997). "Phylogenetic analysis of mycoplasmas based on Hsp70 sequences: cloning of the dnaK (hsp70) gene region of *Mycoplasma capricolum*." Int J Syst Bacteriol **47**(1): 38-45.

Fan, C. Y., S. Lee and D. M. Cyr (2003). "Mechanisms for regulation of Hsp70 function by Hsp40." Cell Stress Chaperones **8**(4): 309-316.

Flaherty, K. M., D. B. McKay, W. Kabsch and K. C. Holmes (1991). "Similarity of the three-dimensional structures of actin and the ATPase fragment of a 70-kDa heat shock cognate protein." Proc Natl Acad Sci U S A **88**(11): 5041-5045.

Fontana, A., G. Fassina, C. Vita, D. Dalzoppo, M. Zamai and M. Zambonin (1986). "Correlation between sites of limited proteolysis and segmental mobility in thermolysin." Biochemistry **25**(8): 1847-1851.

Fraser, C. M., J. D. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley, R. D. Fritchman, J. F. Weidman, K. V. Small, M. Sandusky, J. Fuhrmann, D. Nguyen, T. R. Utterback, D. M. Saudek, C. A. Phillips, J. M. Merrick, J. F. Tomb, B. A. Dougherty, K. F. Bott, P. C. Hu, T. S. Lucier, S. N. Peterson, H. O. Smith, C. A. Hutchison, 3rd and J. C. Venter (1995). "The minimal gene complement of *Mycoplasma genitalium*." Science **270**(5235): 397-403.

Frydman, J., E. Nimmesgern, K. Ohtsuka and F. U. Hartl (1994). "Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones." Nature **370**(6485): 111-117.

Fu, G. K., M. J. Smith and D. M. Markovitz (1997). "Bacterial protease Lon is a site-specific DNA-binding protein." J Biol Chem **272**(1): 534-538.

Garnier, M., X. Foissac, P. Gaurivaud, F. Laigret, J. Renaudin, C. Saillard and J. M. Bove (2001). "Mycoplasmas, plants, insect vectors: a matrimonial triangle." C R Acad Sci III **324**(10): 923-928.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch (2003). "ExpASY: The proteomics server for in-depth protein knowledge and analysis." Nucleic Acids Res **31**(13): 3784-3788.

Georgopoulos, C. and W. J. Welch (1993). "Role of the major heat shock proteins as molecular chaperones." Annu Rev Cell Biol **9**: 601-634.

Gibson, D. G., G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson, J. Zaveri, T. B. Stockwell, A. Brownley, D. W. Thomas, M. A. Algire, C. Merryman, L. Young, V. N. Noskov, J. I. Glass, J. C. Venter, C. A. Hutchison, 3rd and H. O. Smith (2008). "Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome." Science **319**(5867): 1215-1220.

Gibson, D. G., J. I. Glass, C. Lartigue, V. N. Noskov, R. Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z. Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison, 3rd, H. O. Smith and J. C. Venter (2010). "Creation of a bacterial cell controlled by a chemically synthesized genome." Science **329**(5987): 52-56.

Glass, J. I., N. Assad-Garcia, N. Alperovich, S. Yooseph, M. R. Lewis, M. Maruf, C. A. Hutchison, 3rd, H. O. Smith and J. C. Venter (2006). "Essential genes of a minimal bacterium." Proc Natl Acad Sci U S A **103**(2): 425-430.

Glusker, J. P. (1994). "X-ray crystallography of proteins." Methods Biochem Anal **37**: 1-72.

Greene, M. K., K. Maskos and S. J. Landry (1998). "Role of the J-domain in the cooperation of Hsp40 with Hsp70." Proc Natl Acad Sci U S A **95**(11): 6108-6113.

Guell, M., V. van Noort, E. Yus, W. H. Chen, J. Leigh-Bell, K. Michalodimitrakis, T. Yamada, M. Arumugam, T. Doerks, S. Kuhner, M. Rode, M. Suyama, S. Schmidt, A. C. Gavin, P. Bork and L. Serrano (2009). "Transcriptome complexity in a genome-reduced bacterium." Science **326**(5957): 1268-1271.

Han, W. and P. Christen (2001). "Mutations in the interdomain linker region of DnaK abolish the chaperone action of the DnaK/DnaJ/GrpE system." FEBS Lett **497**(1): 55-58.

Harding, M. M. (1999). "The geometry of metal-ligand interactions relevant to proteins." Acta Crystallogr D Biol Crystallogr **55**(Pt 8): 1432-1443.

Harrison, C. (2003). "GrpE, a nucleotide exchange factor for DnaK." Cell Stress Chaperones **8**(3): 218-224.

Harrison, C. J., M. Hayer-Hartl, M. Di Liberto, F. Hartl and J. Kuriyan (1997). "Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK." Science **276**(5311): 431-435.

Hartl, F. U. (1996). "Molecular chaperones in cellular protein folding." Nature **381**(6583): 571-579.

Hartl, F. U., A. Bracher and M. Hayer-Hartl (2011). "Molecular chaperones in protein folding and proteostasis." Nature **475**(7356): 324-332.

Hartl, F. U. and M. Hayer-Hartl (2002). "Molecular chaperones in the cytosol: from nascent chain to folded protein." Science **295**(5561): 1852-1858.

Hasselbring, B. M., J. L. Jordan, R. W. Krause and D. C. Krause (2006). "Terminal organelle development in the cell wall-less bacterium *Mycoplasma pneumoniae*." Proc Natl Acad Sci U S A **103**(44): 16478-16483.

Hasselbring, B. M. and D. C. Krause (2007). "Proteins P24 and P41 function in the regulation of terminal-organelle development and gliding motility in *Mycoplasma pneumoniae*." J Bacteriol **189**(20): 7442-7449.

Hasselbring, B. M., E. S. Sheppard and D. C. Krause (2012). "P65 truncation impacts P30 dynamics during *Mycoplasma pneumoniae* gliding." J Bacteriol **194**(11): 3000-3007.

Hatchel, J. M. and M. F. Balish (2008). "Attachment organelle ultrastructure correlates with phylogeny, not gliding motility properties, in *Mycoplasma pneumoniae* relatives." Microbiology **154**(Pt 1): 286-295.

Henderson, G. P. and G. J. Jensen (2006). "Three-dimensional structure of *Mycoplasma pneumoniae*'s attachment organelle and a model for its role in gliding motility." Mol Microbiol **60**(2): 376-385.

Hendrickson, W. A., J. L. Smith, R. P. Phizackerley and E. A. Merritt (1988). "Crystallographic structure analysis of lamprey hemoglobin from anomalous dispersion of synchrotron radiation." Proteins **4**(2): 77-88.

Himmelreich, R., H. Hilbert, H. Plagens, E. Pirkl, B. C. Li and R. Herrmann (1996). "Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*." Nucleic Acids Res **24**(22): 4420-4449.

Ho, C. S., C. W. Lam, M. H. Chan, R. C. Cheung, L. K. Law, L. C. Lit, K. F. Ng, M. W. Suen and H. L. Tai (2003). "Electrospray ionisation mass spectrometry: principles and clinical applications." Clin Biochem Rev **24**(1): 3-12.

Hoffmann, A., B. Bukau and G. Kramer (2010). "Structure and function of the molecular chaperone Trigger Factor." Biochim Biophys Acta **1803**(6): 650-661.

Huang, K., J. M. Flanagan and J. H. Prestegard (1999). "The influence of C-terminal extension on the structure of the "J-domain" in *E. coli* DnaJ." Protein Sci **8**(1): 203-214.

Hutchison, C. A., S. N. Peterson, S. R. Gill, R. T. Cline, O. White, C. M. Fraser, H. O. Smith and J. C. Venter (1999). "Global transposon mutagenesis and a minimal *Mycoplasma* genome." Science **286**(5447): 2165-2169.

- Inoue, M., N. Mitarai and A. Trusina (2012). "Circuit architecture explains functional similarity of bacterial heat shock responses." Phys Biol **9**(6): 066003.
- Jaffe, J. D., M. Miyata and H. C. Berg (2004). "Energetics of gliding motility in *Mycoplasma mobile*." J Bacteriol **186**(13): 4254-4261.
- Jiang, J., E. G. Maes, A. B. Taylor, L. Wang, A. P. Hinck, E. M. Lafer and R. Sousa (2007). "Structural basis of J cochaperone binding and regulation of Hsp70." Mol Cell **28**(3): 422-433.
- Jiang, J., A. B. Taylor, K. Prasad, Y. Ishikawa-Brush, P. J. Hart, E. M. Lafer and R. Sousa (2003). "Structure-function analysis of the auxilin J-domain reveals an extended Hsc70 interaction interface." Biochemistry **42**(19): 5748-5753.
- Johnson, K. A. and J. L. Rosenbaum (1992). "Polarity of flagellar assembly in *Chlamydomonas*." J Cell Biol **119**(6): 1605-1611.
- Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." J Mol Biol **292**(2): 195-202.
- Jonsson, U., L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg and et al. (1991). "Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology." Biotechniques **11**(5): 620-627.
- Jordan, J. L., K. M. Berry, M. F. Balish and D. C. Krause (2001). "Stability and subcellular localization of cytodherence-associated protein P65 in *Mycoplasma pneumoniae*." J Bacteriol **183**(24): 7387-7391.
- Kabsch, W. (2010). "Xds." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 125-132.
- Kenri, T., S. Seto, A. Horino, Y. Sasaki, T. Sasaki and M. Miyata (2004). "Use of fluorescent-protein tagging to determine the subcellular localization of mycoplasma pneumoniae proteins encoded by the cytodherence regulatory locus." J Bacteriol **186**(20): 6944-6955.
- Kerner, M. J., D. J. Naylor, Y. Ishihama, T. Maier, H. C. Chang, A. P. Stines, C. Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann and F. U. Hartl (2005). "Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*." Cell **122**(2): 209-220.
- Kim, Y. E., M. S. Hipp, A. Bracher, M. Hayer-Hartl and F. U. Hartl (2013). "Molecular chaperone functions in protein folding and proteostasis." Annu Rev Biochem **82**: 323-355.
- Koonin, E. V. (2000). "How many genes can make a cell: the minimal-gene-set concept." Annu Rev Genomics Hum Genet **1**: 99-116.
- Krause, D. C. (1996). "*Mycoplasma pneumoniae* cytodherence: unravelling the tie that binds." Mol Microbiol **20**(2): 247-253.

Krause, D. C. and M. F. Balish (2001). "Structure, function, and assembly of the terminal organelle of *Mycoplasma pneumoniae*." FEMS Microbiol Lett **198**(1): 1-7.

Kuhner, S., V. van Noort, M. J. Betts, A. Leo-Macias, C. Batisse, M. Rode, T. Yamada, T. Maier, S. Bader, P. Beltran-Alvarez, D. Castano-Diez, W. H. Chen, D. Devos, M. Guell, T. Norambuena, I. Racke, V. Rybin, A. Schmidt, E. Yus, R. Aebersold, R. Herrmann, B. Bottcher, A. S. Frangakis, R. B. Russell, L. Serrano, P. Bork and A. C. Gavin (2009). "Proteome organization in a genome-reduced bacterium." Science **326**(5957): 1235-1240.

Lartigue, C., S. Vashee, M. A. Algire, R. Y. Chuang, G. A. Benders, L. Ma, V. N. Noskov, E. A. Denisova, D. G. Gibson, N. Assad-Garcia, N. Alperovich, D. W. Thomas, C. Merryman, C. A. Hutchison, 3rd, H. O. Smith, J. C. Venter and J. I. Glass (2009). "Creating bacterial strains from genomes that have been cloned and engineered in yeast." Science **325**(5948): 1693-1696.

Laskey, R. A., B. M. Honda, A. D. Mills and J. T. Finch (1978). "Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA." Nature **275**(5679): 416-420.

Laufen, T., M. P. Mayer, C. Beisel, D. Klostermeier, A. Mogk, J. Reinstein and B. Bukau (1999). "Mechanism of regulation of hsp70 chaperones by DnaJ cochaperones." Proc Natl Acad Sci U S A **96**(10): 5452-5457.

Li, J., X. Qian and B. Sha (2003). "The crystal structure of the yeast Hsp40 Ydj1 complexed with its peptide substrate." Structure **11**(12): 1475-1483.

Lieutaud, P., B. Canard and S. Longhi (2008). "MeDor: a metasever for predicting protein disorder." BMC Genomics **9 Suppl 2**: S25.

Liu, Q. and W. A. Hendrickson (2007). "Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1." Cell **131**(1): 106-120.

Lluch-Senar, M., E. Querol and J. Pinol (2010). "Cell division in a minimal bacterium in the absence of ftsZ." Mol Microbiol **78**(2): 278-289.

Lluch-Senar, M., M. Vallmitjana, E. Querol and J. Pinol (2007). "A new promoterless reporter vector reveals antisense transcription in *Mycoplasma genitalium*." Microbiology **153**(Pt 8): 2743-2752.

Lorenzen, K. and E. van Duijn (2010). "Native mass spectrometry as a tool in structural biology." Curr Protoc Protein Sci **Chapter 17**: Unit17 12.

Maler, L., J. Blankenship, M. Rance and W. J. Chazin (2000). "Site-site communication in the EF-hand Ca²⁺-binding protein calbindin D9k." Nat Struct Biol **7**(3): 245-250.

Matthews, B. W. (1968). "Solvent content of protein crystals." J Mol Biol **33**(2): 491-497.

Mayer, M. P. (2013). "Hsp70 chaperone dynamics and molecular mechanism." Trends Biochem Sci **38**(10): 507-514.

McBride, M. J. (2001). "Bacterial gliding motility: multiple mechanisms for cell movement over surfaces." Annu Rev Microbiol **55**: 49-75.

McGuffin, L. J., J. D. Atkins, B. R. Salehe, A. N. Shuid and D. B. Roche (2015). "IntFOLD: an integrated server for modelling protein structures and functions from amino acid sequences." Nucleic Acids Res.

Melero, R., F. Moro, M. A. Perez-Calvo, J. Perales-Calvo, L. Quintana-Gallardo, O. Llorca, A. Muga and J. M. Valpuesta (2015). "Modulation of the Chaperone DnaK Allosterism by the Nucleotide Exchange Factor GrpE." J Biol Chem.

Miyata, M. (2008). "Centipede and inchworm models to explain Mycoplasma gliding." Trends Microbiol **16**(1): 6-12.

Miyata, M. (2008). "[Cytoskeleton of mycoplasma as gliding pathogenic bacteria: no tubulin, no actin]." Tanpakushitsu Kakusan Koso **53**(13): 1752-1758.

Miyata, M. and S. Seto (1999). "Cell reproduction cycle of mycoplasma." Biochimie **81**(8-9): 873-878.

Moran, C., G. K. Kinsella, Z. R. Zhang, S. Perrett and G. W. Jones (2013). "Mutational analysis of Sse1 (Hsp110) suggests an integral role for this chaperone in yeast prion propagation in vivo." G3 (Bethesda) **3**(8): 1409-1418.

Murshudov, G. N., P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin (2011). "REFMAC5 for the refinement of macromolecular crystal structures." Acta Crystallogr D Biol Crystallogr **67**(Pt 4): 355-367.

Murshudov, G. N., A. A. Vagin and E. J. Dodson (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallogr D Biol Crystallogr **53**(Pt 3): 240-255.

Musatovova, O., S. Dhandayuthapani and J. B. Baseman (2006). "Transcriptional heat shock response in the smallest known self-replicating cell, Mycoplasma genitalium." J Bacteriol **188**(8): 2845-2855.

Nair, S., I. Derre, T. Msadek, O. Gaillot and P. Berche (2000). "CtsR controls class III heat shock gene expression in the human pathogen *Listeria monocytogenes*." Mol Microbiol **35**(4): 800-811.

Nakamura, A., K. Takumi and K. Miki (2010). "Crystal structure of a thermophilic GrpE protein: insight into thermosensing function for the DnaK chaperone system." J Mol Biol **396**(4): 1000-1011.

Nakane, D. and M. Miyata (2007). "Cytoskeletal "jellyfish" structure of *Mycoplasma mobile*." Proc Natl Acad Sci U S A **104**(49): 19518-19523.

Narberhaus, F. (2002). "Alpha-crystallin-type heat shock proteins: socializing minichaperones in the context of a multichaperone network." Microbiol Mol Biol Rev **66**(1): 64-93; table of contents.

Nettleship, J. E., J. Brown, M. R. Groves and A. Geerlof (2008). "Methods for protein characterization by mass spectrometry, thermal shift (ThermoFluor) assay, and multiangle or static light scattering." Methods Mol Biol **426**: 299-318.

Niesen, F. H., H. Berglund and M. Vedadi (2007). "The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability." Nat Protoc **2**(9): 2212-2221.

O'Brien, M. C., K. M. Flaherty and D. B. McKay (1996). "Lysine 71 of the chaperone protein Hsc70 is essential for ATP hydrolysis." J Biol Chem **271**(27): 15874-15878.

Oshima, K., S. Kakizawa, H. Nishigawa, H. Y. Jung, W. Wei, S. Suzuki, R. Arashida, D. Nakata, S. Miyata, M. Ugaki and S. Namba (2004). "Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma." Nat Genet **36**(1): 27-29.

Papazisi, L., T. S. Gorton, G. Kutish, P. F. Markham, G. F. Browning, D. K. Nguyen, S. Swartzell, A. Madan, G. Mahairas and S. J. Geary (2003). "The complete genome sequence of the avian pathogen *Mycoplasma gallisepticum* strain R(low)." Microbiology **149**(Pt 9): 2307-2316.

Pelham, H. R. (1984). "Hsp70 accelerates the recovery of nucleolar morphology after heat shock." EMBO J **3**(13): 3095-3100.

Pelham, H. R. (1985). "Speculations on the functions of the major heat shock and glucose-regulated proteins."

Pelham, H. R. (1986). "Speculations on the functions of the major heat shock and glucose-regulated proteins." Cell **46**(7): 959-961.

Pellecchia, M., T. Szyperski, D. Wall, C. Georgopoulos and K. Wuthrich (1996). "NMR structure of the J-domain and the Gly/Phe-rich region of the *Escherichia coli* DnaJ chaperone." J Mol Biol **260**(2): 236-250.

Perales-Calvo, J., A. Muga and F. Moro (2010). "Role of DnaJ G/F-rich domain in conformational recognition and binding of protein substrates." J Biol Chem **285**(44): 34231-34239.

Pernot, P., A. Round, R. Barrett, A. De Maria Antolinos, A. Gobbo, E. Gordon, J. Huet, J. Kieffer, M. Lentini, M. Mattenet, C. Morawe, C. Mueller-Dieckmann, S. Ohlsson, W. Schmid, J. Surr, P. Theveneau, L. Zerrad and S. McSweeney (2013). "Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution." J Synchrotron Radiat **20**(Pt 4): 660-664.

Petoukhov, M. V., D. Franke, A. V. Shkumatov, G. Tria, A. G. Kikhney, M. Gajda, C. Gorba, H. D. Mertens, P. V. Konarev and D. I. Svergun (2012). "New developments in the program package for small-angle scattering data analysis." J Appl Crystallogr **45**(Pt 2): 342-350.

Pich, O. Q., R. Burgos, M. Ferrer-Navarro, E. Querol and J. Pinol (2006). "Mycoplasma genitalium mg200 and mg386 genes are involved in gliding motility but not in cytodherence." Mol Microbiol **60**(6): 1509-1519.

Pich, O. Q., R. Burgos, M. Ferrer-Navarro, E. Querol and J. Pinol (2008). "Role of Mycoplasma genitalium MG218 and MG317 cytoskeletal proteins in terminal organelle organization, gliding motility and cytodherence." Microbiology **154**(Pt 10): 3188-3198.

Pich, O. Q., R. Burgos, R. Planell, E. Querol and J. Pinol (2006). "Comparative analysis of antibiotic resistance gene markers in Mycoplasma genitalium: application to studies of the minimal gene complement." Microbiology **152**(Pt 2): 519-527.

Proft, T., H. Hilbert, G. Layh-Schmitt and R. Herrmann (1995). "The proline-rich P65 protein of Mycoplasma pneumoniae is a component of the Triton X-100-insoluble fraction and exhibits size polymorphism in the strains M129 and FH." J Bacteriol **177**(12): 3370-3378.

Razin, S. (1996). Mycoplasmas. Medical Microbiology. S. Baron. Galveston (TX).

Razin, S., M. Argaman and J. Avigan (1963). "Chemical Composition of Mycoplasma Cells and Membranes." J Gen Microbiol **33**: 477-487.

Razin, S. and E. Jacobs (1992). "Mycoplasma adhesion." J Gen Microbiol **138**(3): 407-422.

Razin, S., D. Yogev and Y. Naot (1998). "Molecular biology and pathogenicity of mycoplasmas." Microbiol Mol Biol Rev **62**(4): 1094-1156.

Regula, J. T., G. Boguth, A. Gorg, J. Hegermann, F. Mayer, R. Frank and R. Herrmann (2001). "Defining the mycoplasma 'cytoskeleton': the protein composition of the Triton X-100 insoluble fraction of the bacterium Mycoplasma pneumoniae determined by 2-D gel electrophoresis and mass spectrometry." Microbiology **147**(Pt 4): 1045-1057.

Reid, B. G., W. A. Fenton, A. L. Horwich and E. U. Weber-Ban (2001). "ClpA mediates directional translocation of substrate proteins into the ClpP protease." Proc Natl Acad Sci U S A **98**(7): 3768-3772.

Reinhard, L., H. Mayerhofer, A. Geerlof, J. Mueller-Dieckmann and M. S. Weiss (2013). "Optimization of protein buffer cocktails using Thermofluor." Acta Crystallogr Sect F Struct Biol Cryst Commun **69**(Pt 2): 209-214.

Relich, R. F. and M. F. Balish (2011). "Insights into the function of Mycoplasma pneumoniae protein P30 from orthologous gene replacement." Microbiology **157**(Pt 10): 2862-2870.

Rich, R. L. and D. G. Myszka (2007). "Survey of the year 2006 commercial optical biosensor literature." J Mol Recognit **20**(5): 300-366.

Richards, F. M. (1995). "Christian B. Anfinsen (1916-95)." Nature **376**(6535): 19.

Rossmann, M. G. (1990). "The molecular replacement method." Acta Crystallogr A **46** (Pt 2): 73-82.

Rottem, S. (2003). "Interaction of mycoplasmas with host cells." Physiol Rev **83**(2): 417-432.

Rottem, S. and D. Yogev (2000). "Mycoplasma interaction with eukaryotic cells." Subcell Biochem **33**: 199-227.

Saibil, H. (2013). "Chaperone machines for protein folding, unfolding and disaggregation." Nat Rev Mol Cell Biol **14**(10): 630-642.

Schmidt, B., S. Konig, D. Svergun, V. Volkov, G. Fischer and M. H. Koch (1995). "Small-angle X-ray solution scattering study on the dimerization of the FKBP25mem from *Legionella pneumophila*." FEBS Lett **372**(2-3): 169-172.

Seto, S., T. Kenri, T. Tomiyama and M. Miyata (2005). "Involvement of P1 adhesin in gliding motility of *Mycoplasma pneumoniae* as revealed by the inhibitory effects of antibody under optimized gliding conditions." J Bacteriol **187**(5): 1875-1877.

Seto, S., G. Layh-Schmitt, T. Kenri and M. Miyata (2001). "Visualization of the attachment organelle and cytoadherence proteins of *Mycoplasma pneumoniae* by immunofluorescence microscopy." J Bacteriol **183**(5): 1621-1630.

Seto, S. and M. Miyata (2003). "Attachment organelle formation represented by localization of cytoadherence proteins and formation of the electron-dense core in wild-type and mutant strains of *Mycoplasma pneumoniae*." J Bacteriol **185**(3): 1082-1091.

Seybert, A., R. Herrmann and A. S. Frangakis (2006). "Structural analysis of *Mycoplasma pneumoniae* by cryo-electron tomography." J Struct Biol **156**(2): 342-354.

Shapiro, A. L., E. Vinuela and J. V. Maizel, Jr. (1967). "Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels." Biochem Biophys Res Commun **28**(5): 815-820.

Shi, W., Y. Zhou, J. Wild, J. Adler and C. A. Gross (1992). "DnaK, DnaJ, and GrpE are required for flagellum synthesis in *Escherichia coli*." J Bacteriol **174**(19): 6256-6263.

Shi, X., L. Yan, H. Zhang, K. Sun, Z. Chang and X. Fu (2014). "Differential degradation for small heat shock proteins IbpA and IbpB is synchronized in *Escherichia coli*: implications for their functional cooperation in substrate refolding." Biochem Biophys Res Commun **452**(3): 402-407.

Spence, J., A. Cegielska and C. Georgopoulos (1990). "Role of Escherichia coli heat shock proteins DnaK and HtpG (C62.5) in response to nutritional deprivation." J Bacteriol **172**(12): 7157-7166.

Stivala, A., M. Wybrow, A. Wirth, J. C. Whisstock and P. J. Stuckey (2011). "Automatic generation of protein structure cartoons with Pro-origami." Bioinformatics **27**(23): 3315-3316.

Suarez, C. and F. Gudiol (2009). "[Beta-lactam antibiotics]." Enferm Infecc Microbiol Clin **27**(2): 116-129.

Suh, W. C., W. F. Burkholder, C. Z. Lu, X. Zhao, M. E. Gottesman and C. A. Gross (1998). "Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ." Proc Natl Acad Sci U S A **95**(26): 15223-15228.

Svergun, D. I. (1999). "Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing." Biophys J **76**(6): 2879-2886.

Swain, J. F., G. Dinler, R. Sivendran, D. L. Montgomery, M. Stotz and L. M. Gierasch (2007). "Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker." Mol Cell **26**(1): 27-39.

Taipale, M., D. F. Jarosz and S. Lindquist (2010). "HSP90 at the hub of protein homeostasis: emerging mechanistic insights." Nat Rev Mol Cell Biol **11**(7): 515-528.

Taylor-Robinson, D. and J. S. Jensen (2011). "Mycoplasma genitalium: from Chrysalis to multicolored butterfly." Clin Microbiol Rev **24**(3): 498-514.

Taylor-Robinson, D., J. G. Tully, P. M. Furr, R. M. Cole, D. L. Rose and N. F. Hanna (1981). "Urogenital mycoplasma infections of man: a review with observations on a recently discovered mycoplasma." Isr J Med Sci **17**(7): 524-530.

Taylor, R. C. and A. Dillin (2011). "Aging as an event of proteostasis collapse." Cold Spring Harb Perspect Biol **3**(5).

Thompson, J. D., D. G. Higgins and T. J. Gibson (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**(22): 4673-4680.

Tokuriki, N. and D. S. Tawfik (2009). "Chaperonin overexpression promotes genetic variation and enzyme evolution." Nature **459**(7247): 668-673.

Ueno, P. M., J. Timenetsky, V. E. Centonze, J. J. Wewer, M. Cagle, M. A. Stein, M. Krishnan and J. B. Baseman (2008). "Interaction of Mycoplasma genitalium with host cells: evidence for nuclear localization." Microbiology **154**(Pt 10): 3033-3041.

- Uenoyama, A., A. Kusumoto and M. Miyata (2004). "Identification of a 349-kilodalton protein (Gli349) responsible for cytoadherence and glass binding during gliding of *Mycoplasma mobile*." J Bacteriol **186**(5): 1537-1545.
- Uenoyama, A. and M. Miyata (2005). "Gliding ghosts of *Mycoplasma mobile*." Proc Natl Acad Sci U S A **102**(36): 12754-12758.
- Vagin, A. and A. Teplyakov (2010). "Molecular replacement with MOLREP." Acta Crystallogr D Biol Crystallogr **66**(Pt 1): 22-25.
- Vogel, M., M. P. Mayer and B. Bukau (2006). "Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker." J Biol Chem **281**(50): 38705-38711.
- Vogtherr, M., D. M. Jacobs, T. N. Parac, M. Maurer, A. Pahl, K. Saxena, H. Ruterjans, C. Griesinger and K. M. Fiebig (2002). "NMR solution structure and dynamics of the peptidyl-prolyl cis-trans isomerase domain of the trigger factor from *Mycoplasma genitalium* compared to FK506-binding protein." J Mol Biol **318**(4): 1097-1115.
- Vonskii, M. S., G. V. Astvatsaturians and S. N. Borkhsenius (1993). "[Expression of heat shock proteins in mycoplasma]." Dokl Akad Nauk **331**(1): 112-115.
- Waites, K. B., B. Katz and R. L. Schelonka (2005). "Mycoplasmas and ureaplasmas as neonatal pathogens." Clin Microbiol Rev **18**(4): 757-789.
- Wall, D., M. Zylicz and C. Georgopoulos (1994). "The NH₂-terminal 108 amino acids of the *Escherichia coli* DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication." J Biol Chem **269**(7): 5446-5451.
- Walsh, P., D. Bursac, Y. C. Law, D. Cyr and T. Lithgow (2004). "The J-protein family: modulating protein assembly, disassembly and translocation." EMBO Rep **5**(6): 567-571.
- Wawrzynow, A., B. Banecki, D. Wall, K. Liberek, C. Georgopoulos and M. Zylicz (1995). "ATP hydrolysis is required for the DnaJ-dependent activation of DnaK chaperone for binding to both native and denatured protein substrates." J Biol Chem **270**(33): 19307-19311.
- Weber, K. and M. Osborn (1969). "The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis." J Biol Chem **244**(16): 4406-4412.
- Wegrzyn, R. D. and E. Deuerling (2005). "Molecular guardians for newborn proteins: ribosome-associated chaperones and their role in protein folding." Cell Mol Life Sci **62**(23): 2727-2738.
- Wilbanks, S. M., C. DeLuca-Flaherty and D. B. McKay (1994). "Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. I. Kinetic analyses of active site mutants." J Biol Chem **269**(17): 12893-12898.

- Wilkins, M. R., E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel and D. F. Hochstrasser (1999). "Protein identification and analysis tools in the ExPASy server." Methods Mol Biol **112**: 531-552.
- Williams, T. A. and M. A. Fares (2010). "The effect of chaperonin buffering on protein evolution." Genome Biol Evol **2**: 609-619.
- Wilson, W. D. (2002). "Tech.Sight. Analyzing biomolecular interactions." Science **295**(5562): 2103-2105.
- Wittung-Stafshede, P., J. Guidry, B. E. Horne and S. J. Landry (2003). "The J-domain of Hsp40 couples ATP hydrolysis to substrate capture in Hsp70." Biochemistry **42**(17): 4937-4944.
- Wolgemuth, C. W., O. Igoshin and G. Oster (2003). "The motility of mollicutes." Biophys J **85**(2): 828-842.
- Wong, P. and W. A. Houry (2004). "Chaperone networks in bacteria: analysis of protein homeostasis in minimal cells." J Struct Biol **146**(1-2): 79-89.
- Wu, C. C., V. Naveen, C. H. Chien, Y. W. Chang and C. D. Hsiao (2012). "Crystal structure of DnaK protein complexed with nucleotide exchange factor GrpE in DnaK chaperone system: insight into intermolecular communication." J Biol Chem **287**(25): 21461-21470.
- Yang, Z., Y. Geng and W. Shi (1998). "A DnaK homolog in *Myxococcus xanthus* is involved in social motility and fruiting body formation." J Bacteriol **180**(2): 218-224.
- Young, J. C., V. R. Agashe, K. Siegers and F. U. Hartl (2004). "Pathways of chaperone-mediated protein folding in the cytosol." Nat Rev Mol Cell Biol **5**(10): 781-791.
- Yura, T. and K. Nakahigashi (1999). "Regulation of the heat-shock response." Curr Opin Microbiol **2**(2): 153-158.
- Yus, E., T. Maier, K. Michalodimitrakis, V. van Noort, T. Yamada, W. H. Chen, J. A. Wodke, M. Guell, S. Martinez, R. Bourgeois, S. Kuhner, E. Raineri, I. Letunic, O. V. Kalinina, M. Rode, R. Herrmann, R. Gutierrez-Gallego, R. B. Russell, A. C. Gavin, P. Bork and L. Serrano (2009). "Impact of genome reduction on bacterial metabolism and its regulation." Science **326**(5957): 1263-1268.
- Zahn, M., N. Berthold, B. Kieslich, D. Knappe, R. Hoffmann and N. Strater (2013). "Structural studies on the forward and reverse binding modes of peptides to the chaperone DnaK." J Mol Biol **425**(14): 2463-2479.
- Zhu, X., X. Zhao, W. F. Burkholder, A. Gragerov, C. M. Ogata, M. E. Gottesman and W. A. Hendrickson (1996). "Structural analysis of substrate binding by the molecular chaperone DnaK." Science **272**(5268): 1606-1614.

Zhuravleva, A., E. M. Clerico and L. M. Gierasch (2012). "An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones." Cell **151**(6): 1296-1307.

Zhuravleva, A. and L. M. Gierasch (2011). "Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones." Proc Natl Acad Sci U S A **108**(17): 6987-6992.

Zimm, B. H. (1948). "The dependence of the scattering of light on angle and concentration in linear polymer solutions." J Phys Colloid Chem **52**(1): 260-267.

Zuiderweg, E. R., E. B. Bertelsen, A. Rousaki, M. P. Mayer, J. E. Gestwicki and A. Ahmad (2013). "Allostery in the Hsp70 chaperone proteins." Top Curr Chem **328**: 99-153.