



Estudio de las proteínas de la partícula regulatoria del proteosoma y del ensamblaje de la lid y análisis estructural del dominio OBD de RepB y su unión al ADN

Juliana Amodio Debenjak

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

Facultad de Farmacia

Departamento de Bioquímica i Biología Molecular

**Estudio de las proteínas de la
partícula regulatoria del proteosoma
y del ensamblaje de la lid**

y

**Análisis estructural del dominio OBD
de RepB y su unión al ADN**

Juliana Amodio Debenjak

Barcelona, 2014



Memoria presentada por Juliana Amodio Debenjak, inscrita en el programa de doctorado de Biotecnología, para obtener el grado de Doctor.

El presente trabajo se realizó en el Institut de Biologia Molecular de Barcelona (IBMB-CSIC) y el Institute for Research in Biomedicine (IRB Barcelona) bajo la dirección del Prof. Miquel Coll.

Juliana Amodio
Debenjak

Miquel Coll

A Lu, el sol de mi luna
y a nuestro cachorrito Ivo

Agradecimientos

Esta tesis no habría sido posible sin la participación, apoyo y compañía de una amplia variedad de personas, cada una sumo su granito de arena ya sea científica o personalmente, a todos ellos mi más profundo agradecimiento:

Al Prof. Miquel Coll, mi director de tesis por darme la posibilidad de vivir esta experiencia.

A nuestros colaboradores:

- Prof. Montserrat Samsó y Tyler W. Steele del departamento de fisiología y biofísica de la Universidad de Virginia, por llevar a cabo los estudios de microscopia electrónica.
- El grupo de investigación de replicación y expresión del DNA en bacterias Gram-positivas del CIB de Madrid: Dra. Gloria del Solar, Dr. José Angel Ruíz Maso y Lorena Bordanaba Ruiseco por su participación en el proyecto de RepB.
- La Dra. Marta Vilaseca del IRB de Barcelona por los análisis de espectrometría de masas.
- Y el grupo del IBMB dirigido por el Dr. Bernat Crosas que además de cedernos la cepa de levadura y muchísimo apoyo, me convidaron con algún que otro trocito de pastel y mucha buena onda.

A los integrantes de mi comité de tesis: Patrick Aloy, Isabel Usón y Josefa Badia por sus aportaciones y sabios consejos.

A Roeland Boer por su apoyo científico y emocional y por las *brainstormings* acompañadas con cafés y chicles.

A Fabio Sessa con quien trabajamos codo a codo sin importar fines de semana ni horarios, por su cariño y sus enseñanzas.

A los chicos del lab: Zuza, Radek, Diana, Robert, Montse, Salva, Simo, Rosa, Albert y Cristina. A todos los cri!: Maria, Luca, Xavi, Mercé, Nere, Anna C, Anna R, Pablo, etc. A las técnicas: Esther, Leonor, Mailys, Nayibe, Queralt, Jenny e Isabel. Todos y cada uno de ellos me aportaron ideas, charlas científicas y no tan científicas, cafés y almuerzos, apoyo logístico y mucho cariño haciendo del lab mi segunda casa.

A los integrantes de la PAC quienes siempre estuvieron dispuestos a ayudar y a compartir un buen momento: Joan, Robert y especialmente Sonia que además se llevó un pedacito de mi corazón a Francia.

A mis amigos de allá: Martín, Vicky, Rober, Palo, Bele, Agos, tantos años cerca, tantos años lejos... pero siempre juntos. Y a mis amigos de acá:

Pablo, Adri, Malu (a quien le debo el arte de tapa) y Nury, por su compañía, paciencia y locura inherente necesaria para entenderme. Los quiero!!!

A mi familia, a mis padres quienes a pesar de la distancia siempre supieron estar presentes y a mis abuelos por ser mis modelos de vida y enseñarme que la vida debe recorrerse riendo y amando sinceramente.

Y por último pero no por ello menos importantes sino más bien todo lo contrario: a Lu, mi compañero de aventuras, por su amor, apoyo y comprensión, sin él no sería la misma persona y por supuesto no habría llegado hasta acá; por acompañarme a cada paso empujándome a sacar lo mejor de mí; por formar juntos esta pequeña gran familia de cinco. A nuestros mustélidos Messi y 13 por la apacible compañía y finalmente a Ivo nuestro cachorro, por traer a nuestra vida la felicidad más absoluta, por enseñarnos con su llegada nuestro lado más salvaje, primitivo y animal y demostrarnos finalmente que para ser felices *all we need is love*.

Índice

Abreviaturas	xiii
Listado de figuras	xvii
Listado de tablas	xix
Prefacio	xxi

Primera Parte: Estudio de las proteínas de la partícula regulatoria del proteosoma y del ensamblaje de la lid	1
1.1 Introducción	3
• 1.1.2 Ubiquitinación	6
• 1.1.3 Proteosoma 26S	7
• 1.1.4 La partícula central	8
• 1.1.5 La partícula regulatoria	9
• 1.1.6 La base	9
• 1.1.7 La lid	14
• 1.1.8 La biogénesis del proteosoma 26S	17
• 1.1.9 Objetivos	21
1.2 Materiales y métodos	23
• 1.2.1 Procedimientos generales	25
○ Purificación de ADN	25
○ Electroforesis en geles de agarosa	25
○ Determinación de concentración y pureza del ADN	25
○ Reacciones de digestión y ligación del ADN	25
○ Secuenciación de ADN	25
○ Medios de cultivos, cepas bacterianas y vectores	26
○ Transformación de bacterias	28
○ Electroforesis en geles de poliacrilamida SDS-PAGE	28
○ Columnas cromatográficas	29
○ Concentración y diálisis de proteínas	29
○ Western blots	29
○ Análisis por MS de proteínas	30
• 1.2.2 Proteínas recombinantes	30
○ Diseño de los constructos	30
○ PCR y clonado	30
○ Expresión y purificación	33
○ Ensayos cristalográficos	35
○ DLS	38
• 1.2.3 Lid endógena	39
○ Expresión y purificación de la lid endógena	39
○ Producción de anticuerpos monoclonales anti-Rpn11 y anti-Rpn8	40
○ Interacción de la lid endógena con los AcMo anti-Rpn11 y anti-Rpn8	41
• 1.2.4 Microscopia Electrónica	41
1.3 Resultados y discusión	43
• 1.3.1 Proteínas recombinantes	45
○ Rpn1 y Rpn2	45

○ Rpn10 y la interacción base-lid	45
○ Proteínas de la lid	49
▪ Proteínas individuales	49
▪ Complejos binarios	52
▪ Complejos triméricos	57
▪ Complejos tetraméricos	60
▪ Complejos pentaméricos	62
▪ Reconstitución heteróloga de la lid	64
• 1.3.2 Lid endógena	65
○ Expresión y purificación de la lid endógena	65
○ Interacción de la lid endógena con los AcMo anti-Rpn11 y anti-Rpn8	66
• 1.3.3 Microscopia electrónica	68
• 1.3.4 Biogénesis de la lid	71
1.4 Conclusiones	73
• Rpn1 y Rpn2	75
• Rpn10 y la interacción base-lid	75
• Proteínas de la lid	75
• Lid endógena	76
• Microscopía electrónica	76
Segunda Parte: Análisis estructural del dominio OBD de RepB y su unión al ADN	79
2.1 Introducción	81
• 2.1.1 Transferencia horizontal genética	84
• 2.1.2 Plásmidos bacterianos y su clasificación	86
• 2.1.3 Mecanismos de replicación	86
• 2.1.4 Plásmido pMV158	87
• 2.1.5 Replicación de círculo rodante (RCR)	88
• 2.1.6 Regulación de la RCR	89
• 2.1.7 Dso (double-strand origin)	90
• 2.1.8 Sso (single strand origin)	91
• 2.1.9 RepB	92
○ Superfamilia HUH	92
○ Proteínas relacionadas estructuralmente	94
○ RepB	95
• 2.1.10 Objetivos	100
2.2 Materiales y métodos	101
• 2.2.1 Ensayos cristalográficos	103
○ Preparación de la muestra de proteína	103
○ Preparación de los oligonucleótidos	103
○ Preparación de los complejos proteína-ADN	103
○ Cristalización	104
• 2.2.2 Resolución de la estructura	104
○ Recolección y procesamiento de los datos de difracción	104
○ Resolución y refinamiento estructural	104
○ Análisis de la estructura y creación de modelos	105
2.3 Resultados y discusión	107

• 2.3.1	Cristalización	109
• 2.3.2	Determinación estructural	109
○	Difracción y procesamiento de los datos	109
○	Resolución y refinamiento estructural	110
• 2.3.3	El dominio OBD de RepB y su unión a la DDR	113
• 2.3.4	Modelo de interacción hexamérico	120
• 2.3.5	RepB ₁₋₁₃₂ /23AB vs. proteínas relacionadas estructuralmente	121
2.4	Conclusiones	123
Bibliografía		127
•	Listado de webs	129
•	Bibliografía	131

Abreviaturas y anglicismos

#	Número
∞	Infinito
°C	Grado Celsius
Φ	Phi, ángulo
$\Delta\phi$	Delta phi, variación del ángulo
α	Alfa
Å	Ångstrom
Ac	Anticuerpo
AcMo	Anticuerpo monoclonal
Amp	Ampicilina
β	Beta
Beads	Perlas o microesferas
bp	Par de bases
BPV	Virus papiloma bovino
Cadena L	Cadena ligera o <i>light</i>
Cadena H	Cadena pesada o <i>heavy</i>
CC	<i>Coiled coil</i> o hélice superenrollada
COP9	<i>constitutive photomorphogenesis 9</i>
CP	<i>Core particle</i>
C-terminal	Carboxilo terminal
c/u	Cada uno
CV	Volumen de columna
d	Disponible
ddH ₂ O	Agua doblemente destilada
DDM	n-Dodecil- β -D-maltopiranosida
DDR	<i>Distal direct repeats</i>
DLS	<i>Dynamic light scattering</i>
DNasa	Desoxiribonucleasa
DO	Densidad óptica
dsADN	ADN doble cadena
dso	<i>Double-strand origin</i>
DTT	Ditiotreitol
EDTA	Ácido etilendiaminotetracético
eIF3	<i>Eukaryotic initiation factor 3</i>
EM	<i>Electron microscopy</i>
Fig.	Figura
FSC	<i>Fourier shell correlation</i>
Fwd	<i>Forward</i>
g	Constante de aceleración gravitatoria

gr	Gramo
GSH	Glutación
GST	Glutation S-transferasa
HGT	Horizontal gene transfer
i	Inmuno
Ii	Intercambio iónico
IPTG	Isopropil-β-D1-tiogalactopiranosido
JAMM	Jab1/Pad1/MPN metalo-enzima
kDa	Kilodalton
kV	Kilovoltio
l	Litro
LB	Medio Luria-Bertani
μ	Micro (10 ⁻⁶)
M	Molar
m	Mili (10 ⁻³)
mARN	ARN mensajero
MDa	Megadalton
min	Minuto
MPN	<i>Mpr1/Pad1 N-terminal</i>
MS	<i>Mass spectrometry</i>
MWCO	<i>Molecular weight cut off</i>
NCS	<i>Non-crystallographic symmetry</i>
Nº	Número
nm	Nanómetro
nt	Nucleótido
N-terminal	Amino terminal
o	Obtenido
OD	<i>Oligomerization domain</i>
OB	<i>Oligonucleotide/oligosaccharide binding</i>
OBD	<i>Origin binding domain</i>
p/p	Peso/peso
p/v	Peso/volumen
PAC	<i>Proteasome assembly chaperones</i>
PAC	Plataforma Automatizada de Cristalografía
paper	Publicación científica
Pba	Proteasome biogenesis-associated
PC	Proteosoma/ciclosoma
PCR	Reacción en cadena de la polimerasa
PCI	<i>Proteasome-COP9-IF3</i>
PDR	<i>Proximal direct repeats</i>
PVDF	Fluoruro de polivilideno

RBS	<i>Ribosome binding site</i>
RCR	Replicación de círculo rodante
Rev	<i>Reverse</i>
RMN	Resonancia magnética nuclear
R.m.s.	<i>Root mean square</i>
RNasa	Ribonucleasa
RP	<i>Regulatory particle</i>
rpm	Revoluciones por minuto
Rpn	<i>Regulatory particle non-ATPase</i>
Rpt	<i>Regulatory particle triple-A protein</i>
SDS	Dodecil sulfato de sodio
seg	Segundo
SOC	Medio super-óptimo con represión por catabolitos
SSB	<i>Single-stranded DNA binding</i>
ssso	<i>Single strand origin</i>
SV40	Virus simiano 40
t	Tiempo
t	Timo
TAE	Tris-acetato EDTA
Tag	Etiqueta
TBS	Buffer tris salino
TEM	Microscopía electrónica de transmisión
TEMED	N,N,N',N'-tetrametil etilendiamina
TPR	Repetición de tetratricopéptidos
U	Unidades
Ub	Ubiquitina
UE	Unión Europea
UMP	<i>Ubiquitin maturation protein</i>
UPS	<i>Ubiquitin-Proteasome System</i>
UV	Ultravioleta
V	Volt
v/v	Volumen/volumen
YPD	<i>Yeast extract peptone dextrose</i>

Listado de figuras:

- Fig. 1.1 Ubiquitinación.
- Fig. 1.2 Proteosoma 26S.
- Fig. 1.3 La partícula central.
- Fig. 1.4 Galería de imágenes de la partícula central del proteosoma.
- Fig. 1.5 Dominios HbYX.
- Fig. 1.6 Reorganización de la base inducida por el substrato.
- Fig. 1.7 Rpn10 y Rpn13.
- Fig. 1.8 Rpn1 y Rpn2.
- Fig. 1.9 Organización general de la base.
- Fig. 1.10 Proteínas PCI.
- Fig. 1.11 Dominio MPN.
- Fig. 1.12 Estructura de la lid.
- Fig. 1.13 Modelo de compromiso y degradación basado en la estructura.
- Fig. 1.14 Modelo esquemático del proceso de biogénesis de la partícula central en humanos.
- Fig. 1.15 Modelo esquemático de la biogénesis de la base.
- Fig. 1.16 Modelo esquemático del proceso de biogénesis de la lid.
- Fig. 1.17 El haz helicoidal.
- Fig. 1.18 Mapa del vector pGEX-6P-2RBS.
- Fig. 1.19 Esquema del vector pGEX-6P-2RBS.
- Fig. 1.20 Purificación de Rpn10.
- Fig. 1.21 *In vitro binding* de Rpn10.
- Fig. 1.22 Co-expresión de RpnX (3, 9, 11 y 12)/Rpn10.
- Fig. 1.23 Expresión y purificación de las proteínas Rpn individualmente.
- Fig. 1.24 Rpn12.
- Fig. 1.25 Rpn3/Rpn7.
- Fig. 1.26 Rpn8/Rpn11.
- Fig. 1.27 Versiones truncadas de Rpn11/Rpn8.
- Fig. 1.28 Rpn5/Rpn9 y versiones truncadas.
- Fig. 1.29 Rpn5/Rpn8 y versiones truncadas.
- Fig. 1.30 Análisis de trímeros.
- Fig. 1.31 Análisis de tetrámeros *full length*.
- Fig. 1.32 Análisis de tetrámeros truncados.
- Fig. 1.33 Análisis de pentámeros.
- Fig. 1.34 Esquema de las interacciones del complejo pentamérico.
- Fig. 1.35 Reconstrucción heteróloga de la lid.
- Fig. 1.36 Purificación de la lid endógena.
- Fig. 1.37 Western blot de los AcMo.
- Fig. 1.38 Interacción de la lid con los AcMo.
- Fig. 1.39 EM de la lid endógena.

- Fig. 1.40 EM de los subcomplejos.
- Fig. 1.41 Correlaciones por capas de Fourier (FSC).
- Fig. 1.42 Esquema de los eventos biogénicos de la formación del complejo pentamérico.
- Fig. 1.43 Esquema de la finalización del ensamblaje de la lid.
- Fig. 2.1 Mecanismos de transferencia genética horizontal.
- Fig. 2.2 Mapa del plásmido pMV158.
- Fig. 2.3 Modelo de la replicación de círculo rodante.
- Fig. 2.4 Regulación de la replicación.
- Fig. 2.5 Dso del plásmido pMV158 y las regiones de interacción con RepB.
- Fig. 2.6 Organización de las proteínas representativas de la superfamilia HUH.
- Fig. 2.7 Dominio HUH de clase Rep y Mob.
- Fig. 2.8 Comparación de estructuras relacionadas.
- Fig. 2.9 Secuencia primaria de RepB.
- Fig. 2.10 Estructura del monómero de RepB.
- Fig. 2.11 Estructura de los hexámeros de RepB.
- Fig. 2.12 Modelo del canal central del hexámero de RepB.
- Fig. 2.13 Cristalización de RepB₁₋₁₃₂/23AB.
- Fig. 2.14 Patrón de difracción del cristal.
- Fig. 2.15 Representación de Ramachandran para la estructura RepB₁₋₁₃₂-23AB.
- Fig. 2.16 Contenido de la unidad asimétrica.
- Fig. 2.17 Superposición del OBD sobre RepB *full length*.
- Fig. 2.18 Interacción de RepB₁₋₁₃₂ en el surco mayor de 23AB.
- Fig. 2.19 Esquema de las interacciones ADN/proteína.
- Fig. 2.20 Detalle de la interacción ADN/proteína.
- Fig. 2.21 Detalle de la interacción de D69 y S70 con la timina.
- Fig. 2.22 Análisis del *footprinting* de alta resolución en el contexto de la estructura de RepB₁₋₁₃₂/23AB.
- Fig. 2.23 Modelo de interacción de las DDR en el hexámero.
- Fig. 2.24 Modelo esquemático de la interacción de las PDR en el hexámero.
- Fig. 2.25 Superposición de RepB₁₋₁₃₂/23AB con proteínas relacionadas estructuralmente.

Listado de tablas:

Tabla 1.1: Cepas bacterianas

Tabla 1.2: Vectores

Tabla 1.3: Protocolo de PCR

Tabla 1.4: Primers

Tabla 1.5: Clones

Tabla 1.6: Buffers Rpn10

Tabla 1.7: Ensayos cristalográficos

Tabla 1.8: Código de los screens

Tabla 1.9: Buffers DLS

Tabla 1.10: Parámetros de los modelos 3D

Tabla 2.1: Estadísticas de la recolección de datos

Tabla 2.2: Estadísticas del refinamiento

Tabla 2.3: Parámetros locales de pares de bases

Tabla 2.4: Parámetros locales de pasos de pares de bases

Prefacio

Esta tesis, como muchas otras, no solamente es el recuento de resultados y conocimientos adquiridos sino más bien el viaje personal y científico que emprendí en septiembre de 2009 y me llevó a comprender algunos mecanismos biológicos y mucho de mi propia naturaleza como mujer científica.

Como la mayoría de los caminos que recorreremos en nuestras vidas pocos son lineales o continuos y muchos presentan más de una vuelta, con los proyectos de investigación muchas veces pasa lo mismo y éste es uno de esos casos. Esta tesis es la historia de dos proyectos que comenzaron en diferentes momentos y se llevaron a cabo independientemente uno del otro, cada uno me enfrentó a desafíos diferentes y me aportaron conocimientos en dos campos completamente distintos pero complementarios.

Por un lado el proyecto de las proteínas de la partícula reguladora del proteosoma me permitió mejorar mi habilidad para la expresión y purificación de proteínas, aprendiendo a integrar los resultados obtenidos y poniéndome a prueba con la publicación de más de un *paper* relacionado directamente con este proyecto durante su desarrollo. Por otro lado el proyecto de RepB me concedió la posibilidad de explorar el mundo de la cristalografía, desde la generación de cristales hasta la resolución de la estructura pasando por ese momento mágico en que con dos imágenes en estéreo se aprende a ver en 3D. Es por ello que esta tesis se divide en dos partes bien diferenciadas, los invito a recorrer los contenidos en el orden que ustedes prefieran y espero que disfruten de la lectura tanto como yo disfrute en desarrollarla.

Primera Parte:

Estudio de las proteínas de la
partícula reguladora del
proteosoma y del ensamblaje
de la lid

1.1 Introducción

Al igual que todos los sistemas de macromoléculas el proteoma se encuentra en un estado dinámico de síntesis y degradación, que asegura la homeostasis celular y la regulación de las proteínas y los procesos en los que éstas participan.

Debemos distinguir las proteínas "ajenas" de las proteínas "propias", las primeras, provenientes de la dieta, deben ser degradadas "fuera" del cuerpo es decir en el lumen del tracto digestivo. Éstas proteínas ajenas a nuestro cuerpo no pueden ser absorbidas directamente ya que dispararían la respuesta inmune, deben ser degradadas hasta aminoácidos que las células adquirirán para utilizar en la síntesis de proteínas propias. A su vez, las proteínas propias se pueden clasificar en proteínas extra e intracelulares que serán procesadas de manera diferencial. Las proteínas extracelulares entran en la célula por endocitosis, el endosoma resultante se fusiona con lisosomas portadores de enzimas proteolíticas que degradan el contenido de manera inespecífica. Durante el proceso las proteínas extracelulares nunca se exponen al citosol, asegurando de esta manera la integridad del medio intracelular. Evidentemente la degradación de proteínas intracelulares debe ser llevada a cabo por mecanismos completamente diferentes, ya que al convivir proteínas proteolíticas y sus substratos en el mismo espacio celular se requiere de un mecanismo específico y altamente regulado. Además el sistema debe poder distinguir entre componentes constitutivos y componentes efímeros, ya que por ejemplo la vida media de la hemoglobina es de 120 días mientras que la de la oncoproteína myc es de tan solo 10 minutos. La respuesta a la incógnita de cómo era regulada la degradación de proteínas intracelulares en eucariotas fue dada a finales de la década de los 70s por Hershko, Ciechanover y Rose al descubrir el sistema proteosoma-ubiquitina, descubrimiento por el que los tres recibieron el premio Nobel de química en el año 2004 [Schwartz *et al.*, 1999; Glickman *et al.*, 2002].

El sistema ubiquitina-proteosoma (o UPS, por sus siglas en inglés *Ubiquitin-Proteasome System*) debido a su función participa directa o indirectamente en casi todos los procesos celulares, como por ejemplo en la progresión a través del ciclo celular, transducción de señales, apoptosis, respuesta inmune, metabolismo, control de calidad proteico y desarrollo, entre otros. Por sus numerosos substratos y la cantidad de procesos en los que está involucrado no es sorprendente que aberraciones en el UPS devengan en patologías heredadas o adquiridas, incluyendo enfermedades autoinmunes, neurodegenerativas y reumatoides, cáncer, infecciones virales y caquexia. Estas patologías pueden agruparse en dos grupos: 1) las que se generan por una pérdida de función, por ejemplo por mutaciones que estabilicen el substrato y eviten su degradación; y las que aumentan la actividad del sistema, por ejemplo degradación acelerada o anormal del substrato [Schwartz *et al.*, 1999]. Por lo que, tanto la inhibición como el aumento de la actividad del UPS tienen valor clínico. La inhibición del UPS

ha sido aceptada como estrategia antineoplásica desde la aprobación del Bortezomib para el tratamiento del mieloma múltiple en 2003. En contraposición, estudios recientes han demostrado que el aumento de la actividad del UPS podría ser beneficioso en el tratamiento de enfermedades que involucran defectos en el plegamiento de proteínas, como algunas enfermedades neurodegenerativas y la replicación de algunos virus [Tomko *et al.*, 2013].

La degradación de proteínas a través del UPS consiste de dos pasos discretos: 1) la selección de los substratos por medio de la unión covalente de múltiples moléculas de ubiquitina (Ub), una pequeña proteína de 8kDa con un altísimo grado de conservación; y 2) la degradación a oligopéptidos de ese sustrato marcado por medio de la maquinaria proteolítica del sistema: el proteosoma 26S, que a su vez libera las ubiquitinas para su reciclaje [Glickman *et al.*, 2002].

1.1.2 Ubiquitinación:

La conjugación de la ubiquitina al sustrato es un proceso que requiere la acción de tres mecanismos sucesivos llevados a cabo por enzimas especializadas, inicialmente la ubiquitina es activada en una reacción ATP dependiente por la enzima activadora de ubiquitina o E1, en el segundo paso una enzima portadora de ubiquitina o E2 la transfiere a una ubiquitin-ligasa o E3 que finalmente cataliza la unión covalente de la ubiquitina al sustrato. Repitiendo la operación se sintetiza la cadena de poliubiquitinas, en la que el extremo C-terminal de cada ubiquitina se une a una lisina específica (Lys48) de la ubiquitina anterior. Usualmente existe solo una E1, mientras que existen múltiples especies de E2 y dos familias de E3, en las E3 de la familia HECT la ubiquitina se transfiere de la E2 a la E3 y finalmente al sustrato, mientras que en las de la familia RING la enzima cataliza la transferencia de la ubiquitina de la E2 directamente al sustrato (Fig. 1.1) [Hershko *et al.*, 1998].

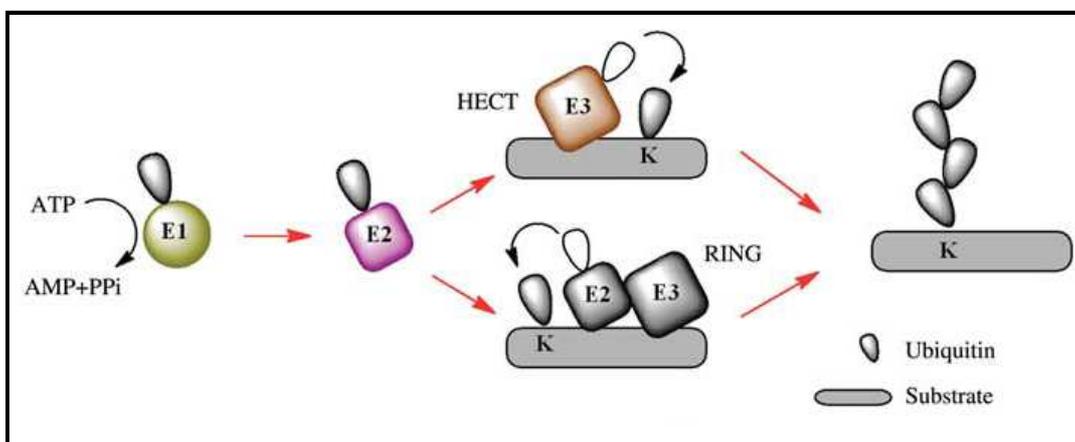


Fig. 1.1 Ubiquitinación. Representación esquemática de la acción secuencial de E1, E2 y E3 que ubiquitan al sustrato. Imagen tomada de Nagy *et al.*, 2010.

Normalmente una cadena de al menos 4 ubiquitinas es necesaria y suficiente para determinar el destino de ese sustrato al proteosoma 26S y su consecuente degradación, sin presentar aumento en la tasa de degradación para cadenas más largas. A su vez los sustratos deben presentar un sitio de iniciación desestructurado que les permita comprometerse con el proteosoma 26S [Lander *et al.*, 2013].

1.1.3 Proteosoma 26S:

El proteosoma 26S es el miembro de mayor tamaño y complejidad de la superfamilia de proteasas compartimentalizadas dependientes de ATP encontradas en todos los seres vivos. En eucariotas se encuentra evolutivamente muy conservado y se localiza tanto en el citoplasma como en el núcleo. Formas más simples de proteosomas han sido identificadas en arqueobacterias y actinobacterias, sin embargo la ubiquitina solo se encuentra en eucariotas [Tomko *et al.*, 2013].

Este complejo multicatalítico de 2.5 MDa está compuesto por al menos 33 subunidades diferentes organizadas en dos subcomplejos: la partícula central (o CP, por sus siglas en inglés *Core Particle*) o proteosoma 20S donde se produce la proteólisis y la partícula reguladora 19S (o RP, por sus siglas en inglés *Regulatory Particle*) localizada en uno o en ambos extremos y encargada de reconocer, desplegar y translocar los sustratos al interior de la CP [Tanaka, 2009]. La estructura general del proteosoma 26S de *Saccharomyces cerevisiae* ha sido resuelta por Beck *et al.*, 2012 mediante criomicroscopía electrónica a una resolución de 7.4Å (Fig. 1.2).

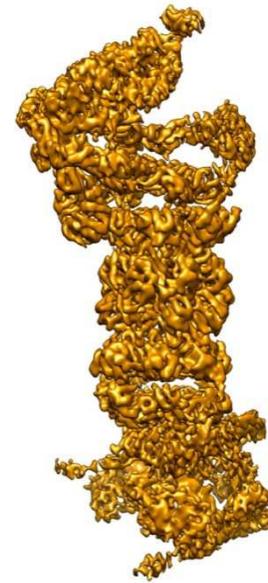


Fig. 1.2 Proteosoma 26S. Superficie del proteosoma 26S. Imagen modificada de Beck *et al.* 2012.

1.1.4 La partícula central:

La CP o proteosoma 20S (denominado así debido a su coeficiente de sedimentación) posee una masa aproximada de 730kDa y está compuesto por 28 subunidades. Cuando se analiza por microscopia electrónica tiene apariencia cilíndrica, tanto en levaduras como en células de mamíferos. Forma una partícula compacta que es el resultado del apilado axial de dos anillos α externos y dos anillos β internos, cada uno de ellos formado por siete subunidades α y β estructuralmente similares respectivamente;

formando una estructura α_1 - β_1 - β_7 - β_1 - β_7 - α_1 - β_7 - β_1 - β_7 - α_1 - β_7 . A diferencia del proteosoma de eucariotas, en procariontes los anillos están formados por subunidades idénticas por lo que su estructura es α - β - β - α (Fig. 1.3) y, siendo que la actividad proteolítica es llevada a cabo por las subunidades β , el proteosoma 20S procarionte tiene 14 sitios de corte idénticos. En cambio, en eucariotas, de las 7

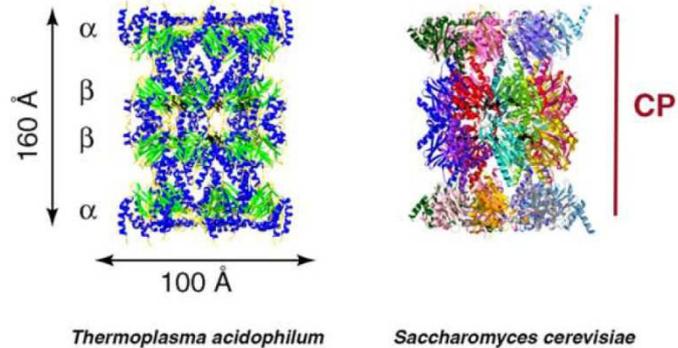


Fig. 1.3 La partícula central. Estructuras cristalográficas de las CP de procariontes (*T. acidophilum*) [Lowe *et al.*, 1995] y eucariotas (*S. cerevisiae*) [Groll *et al.*, 1997]. Imagen tomada de Gallastegui *et al.* 2010.

Cada una de ellas presenta un tipo de actividad proteolítica diferente: β_1 tiene actividad caspasa, β_2 actividad tipo tripsina y β_5 actividad tipo quimotripsina, lo que le confiere la capacidad de escindir los enlaces peptídicos en el extremo C-terminal de residuos ácidos, básicos e hidrofóbicos respectivamente. Dos pares de estos tres sitios activos quedan dentro de la cámara formada por la unión de los dos anillos β , donde los substratos son degradados generando oligopéptidos que varían en un rango de 3 a 15 aminoácidos (Fig. 1.4) [Tanaka, 2009].

En mamíferos se han descubierto cuatro subunidades β adicionales: β_{1i} , β_{2i} , β_{5i} y β_{5t} en las que "i" y "t" provienen de inmuno y timo respectivamente. Estas subunidades se encuentran sobre-expresadas en algunos tejidos del sistema inmune o bien su expresión está regulada por estímulos particulares como la exposición a interferón- γ , y reemplazan a las subunidades β tradicionales. La CP en la que se han intercambiado β_{1i} , β_{2i} y β_{5i} se denomina inmunoproteosoma y genera péptidos que presentan mayor afinidad por el complejo mayor de histocompatibilidad de tipo I; mientras que en la que se intercambian β_{1i} , β_{2i} y β_{5t} se denomina timoproteosoma y produce un repertorio más amplio de propio-péptidos para la selección positiva de células T durante su desarrollo en el timo [Tomko *et al.*, 2013].

Por otro lado, en las estructuras resueltas, el centro de los anillos α se encuentra bloqueado por los extremos N-terminales de las subunidades α , conformación que evitaría el paso de proteínas al interior de la cámara proteolítica. Además los anillos α forman compartimientos interiores adicionales que si bien su función no ha sido probada se especula que puedan retener substratos desplegados o bien productos parcialmente digeridos ofreciéndoles la oportunidad de volver a entrar en la cámara proteolítica para su completa degradación antes de ser liberados (Fig. 1.4) [Pickart *et al.*, 2004; Tanaka, 2009].

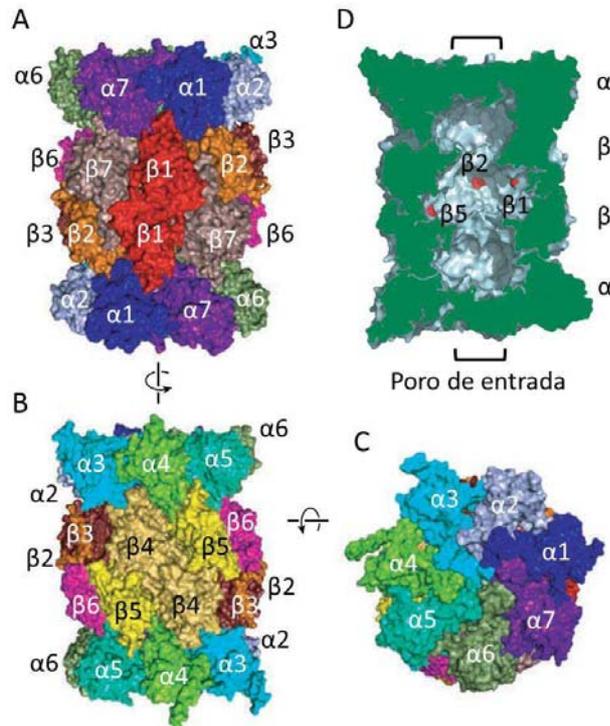


Fig. 1.4 Galería de imágenes de la partícula central del proteosoma. A,B) Vistas laterales de la superficie de la CP mostrando su organización general. C) Vista superior o inferior de la CP mostrando el poro de entrada cerrado por los extremos N-terminal de las subunidades α . D) Corte transversal de la CP que muestra la posición relativa de las subunidades proteolíticas, los corchetes muestran la localización del en poro de entrada. Imagen modificada de Murata *et al.* 2009 y Finley, 2009.

1.1.5 La partícula regulatoria:

Los proteosomas 20S aislados presentan una actividad proteolítica mínima. Para que sean eficientes y ubiquitin-dependientes deben asociarse a RPs, que se ubican sobre los anillos α en uno o en ambos extremos de la CP, en *S. cerevisiae* normalmente cada proteosoma 20S lleva dos RPs. La RP es un complejo de ~ 900 kDa compuesto por 19 subunidades que se organizan en dos subcomplejos: la base y la tapa o "lid" [Pickart *et al.*, 2004].

1.1.6 La base:

Los diez componentes de la base incluyen seis subunidades ATPasas AAA^+ homólogas denominadas Rpt₁₋₆ (por las siglas en inglés de *Regulatory*

particle triple-A protein) y cuatro subunidades sin actividad ATPasa: Rpn1, Rpn2, Rpn10 y Rpn13 (por las siglas en inglés de *Regulatory particle non-ATPase*).

Las Rpt forman un anillo hexamérico, organizado en el siguiente orden: Rpt1-Rpt2-Rpt6-Rpt3-Rpt4-Rpt5, que contacta con el anillo α de la CP y su función es permitir la apertura del anillo α , la translocación y el desdoblamiento del substrato [Murata *et al.*, 2009]. Cada Rpt está formada por una región N-terminal que contiene un dominio *coiled-coil* (CC), un dominio de unión a oligosacáridos u oligonucleótidos (OB, oligonucleotide/oligosaccharide binding), un dominio AAA+ ATPasa y un dominio C-terminal α -hélice [Zhang *et al.*, 2009].

La apertura del anillo α es llevada a cabo por tres de las seis subunidades: Rpt2, Rpt3 y Rpt5, éstas presentan un motivo C-terminal hidrofóbico-tirosina-X (HbYX) que al insertarse en los espacios presentes entre las subunidades α desencadenan el movimiento de sus extremos N-terminales abriendo el poro y permitiendo el paso de substratos (Fig. 1.5) [Saeki *et al.*, 2007; Smith *et al.*, 2007]. Si bien hay un desajuste de simetría (siete subunidades α versus seis Rpt) se cree que esta discrepancia podría permitir múltiples conformaciones y por ende proporciona flexibilidad estructural para la reorganización necesaria para la apertura del anillo α [Gallastegui *et al.*, 2010].

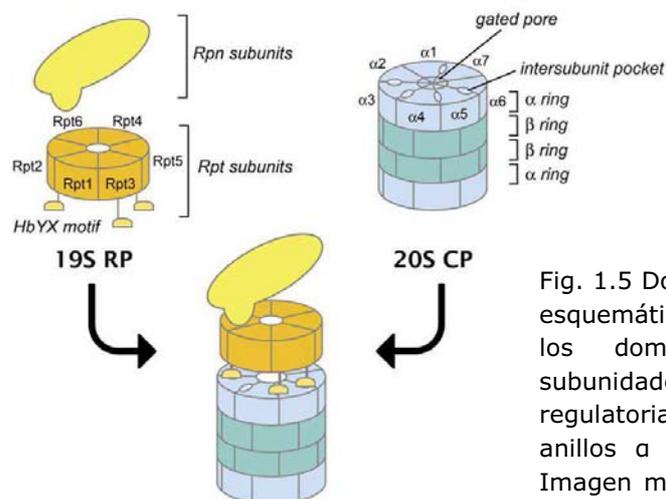


Fig. 1.5 Dominios HbYX. Ilustración esquemática de la interacción de los dominios HbYX de las subunidades Rpt de la partícula reguladora con los bolsillos de anillos α de la partícula central. Imagen modificada de Saeki *et al.* 2007.

La resolución de la estructura del proteosoma 26S durante la degradación de un substrato [Matyskiela, Lander *et al.*, 2013] demostró que se producen una amplia variedad de cambios conformacionales en cuanto a la RP y sugiere un mecanismo para la translocación del substrato. Cuando un substrato interactúa con el poro central del anillo de ATPasas se produce un movimiento global de la base que la coloca desde una posición donde los poros centrales de los anillos Rpt y α se encuentran desviados 10\AA a una posición axial perfectamente alineada (Fig. 1.6). A su vez, en

estado libre de sustrato los dominios ATPasa de las Rpt se encuentran organizados como una pronunciada escalera de caracol con Rpt3 arriba de todo, descendiendo a través de Rpt4, Rpt5 y Rpt1, con Rpt2 debajo de todo y Rpt6 situado entre Rpt3 y Rpt2 a una altura intermedia; cuando entra el sustrato la organización de las ATPasas varía notablemente hasta colocarse todas en un mismo plano aunque cada una de ellas se encuentra inclinada en diferente grado (Fig. 1.6), esto supone que para la translocación del sustrato la hidrólisis coordinada de ATP promovería la inclinación individual de cada subunidad siguiendo el espiral y generando la energía suficiente para translocar una porción del sustrato a través del poro, repitiendo este ciclo se impulsaría la translocación del sustrato por etapas hacia el interior de la CP.

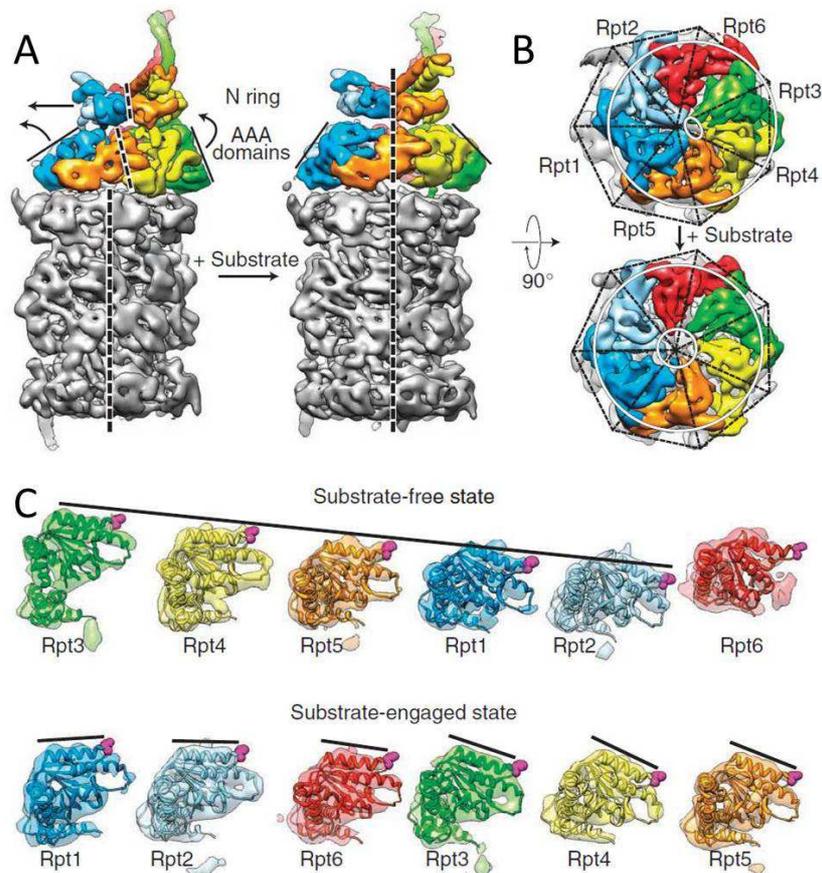
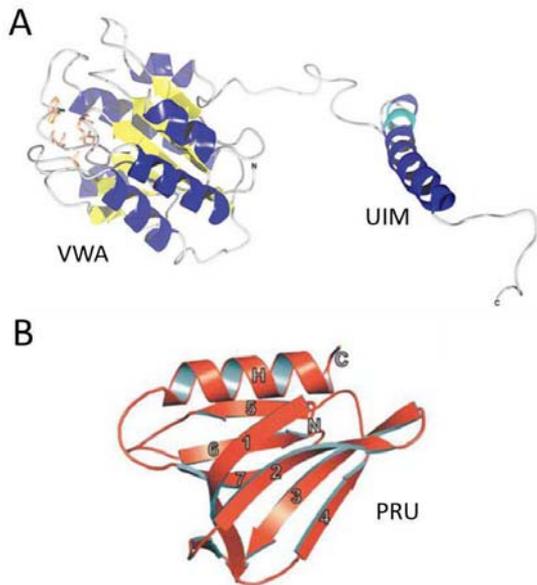


Fig. 1.6 Reorganización de la base inducida por el sustrato. A) La densidad segmentada de las subunidades Rpt1-6 (arcoíris) y la CP (gris) en ausencia (izquierda) y presencia (derecha) del sustrato, la línea de puntos indica el eje del canal central. La presencia del sustrato genera la alineación de los subcomplejos y forma un canal continuo. B) La CP (gris) y las ATPasas (arcoíris) vistas desde arriba en ausencia (arriba) y presencia (abajo) del sustrato, la línea de puntos negra marca la simetría heptamérica del anillo α y los círculos blancos el grado de alineamiento entre el anillo α y las Rpt. C) Posición de los dominios ATPasa de las Rpt en ausencia del sustrato formando una pronunciada escalera de caracol y en presencia del sustrato acomodadas en un mismo plano variando su alineación. Imagen modificada de Matyskiela, Lander *et al.*, 2013.

En cuando a las Rpns de la base, Rpn10 y Rpn13 cumplen la función de receptores primarios del sustrato poliubiquitinado, reconociéndolo por medio de sus dominios UIM (ubiquitin-interacting motif) y PRU (pleckstrin-like receptor for ubiquitin) respectivamente. La estructura tridimensional de estas subunidades fueron descritas, Schreiner *et al.*, 2008 resolvieron el dominio PRU de la Rpn13 murina por cristalografía y Rpn10 de *Schizosaccharomyces pombe* fue resuelta por una combinación de cristalografía y RMN por Riedinger *et al.*, 2010 (Fig. 1.7).



Por otro lado, las subunidades más grandes del proteosoma Rpn1 y Rpn2 (110 y 104kDa respectivamente) presentan en su porción central 11 repeticiones proteosoma/ciclosoma (PC) de α -hélices, cuyo plegamiento es toroidal. Cumplen la función de mediar la interacción entre el proteosoma y una amplia variedad de factores, entre ellos proteínas proteosomales asociadas como proteínas de unión a ubiquitina y factores auxiliares, y también la unión de Rpn13 que es mediada por Rpn2 [Rosenzweig *et al.*, 2012]. La estructura cristalográfica de Rpn2 ha sido resuelta por He *et al.*, 2012, presenta el aspecto de una pipa de tabaco y sus repeticiones PC están densamente empaquetadas aportando rigidez a la estructura. Comparaciones de las estructuras de Rpn1 y Rpn2 por EM revelan un aspecto similar, consecuente con el $\sim 20\%$ de identidad de secuencia que comparten (Fig. 1.8).

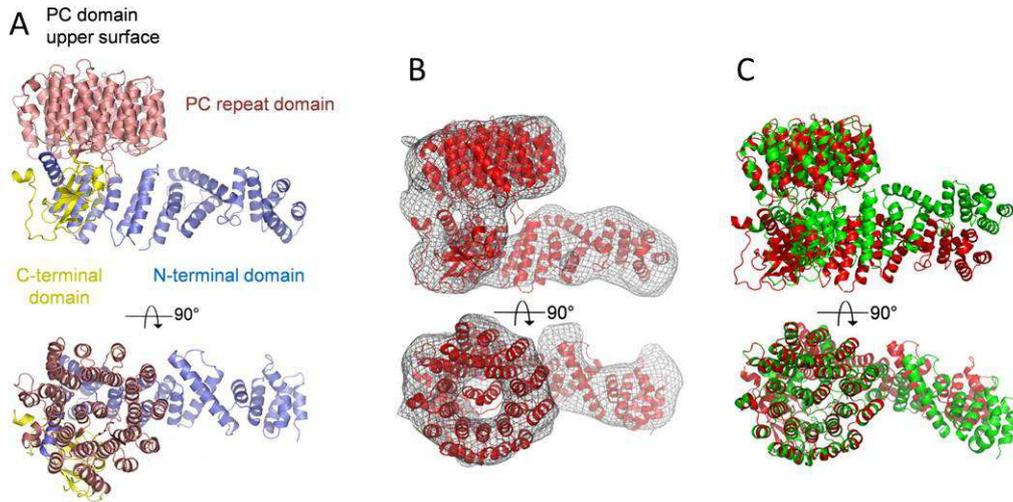


Fig. 1.8 Rpn1 y Rpn2. A) Estructura cristalográfica de Rpn2; B) Mapa 3D de Rpn1 obtenido por EM con la superposición del modelo de Rpn1 (en rojo) basado en la estructura de Rpn2; C) Superposición de la estructura de Rpn2 (en verde) y el modelo de Rpn1 (en rojo). Imagen modificada de He *et al.*, 2012.

La organización estructural de las subunidades de la base se puede observar en la siguiente Fig. 1.9:

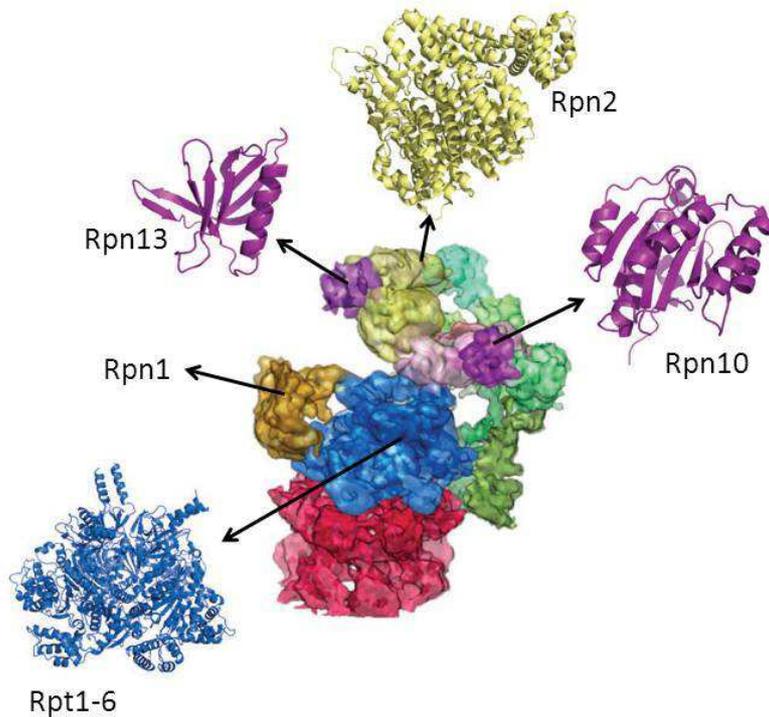


Fig. 1.9 Organización general de la base, incluyendo modelos atómicos y pseudoatómicos de las subunidades de la base. Imagen tomada de Tomko *et al.*, 2010.

1.1.7 La lid:

La lid esta estructuralmente emparentada con lo que se conoce como "complejos PCI" debido a un motivo no catalítico encontrado en algunas de sus subunidades; éstos son el signalosoma COP9 (*constitutive photomorphogenesis 9*) y el factor de transcripción eIF3 (*eukaryotic initiation factor 3*) [Finley, 2009]. Si bien entre eucariotas la composición del CSN y el eIF3 varían, su organización general se mantiene, todos presentan seis subunidades con dominio PCI (*Proteasome-COP9-IF3*) y dos con dominio MPN (*Mpr1/Pad1 N-terminal*). La composición de la lid permanece invariable en eucariotas, consiste de nueve subunidades no ATPasas (Rpn), de las cuales seis son proteínas PCI: Rpn3, Rpn5, Rpn6, Rpn7, Rpn9 y Rpn12; dos subunidades MPN: Rpn8 y Rpn11 y una pequeña proteína ácida Sem1 o Rpn15 [Tomko *et al.*, 2013].

Las proteínas PCI de N- a C-terminal están compuestas de repeticiones helicoidales tipo TPR (Repetición de tetratricopéptidos), el dominio PCI y al menos una hélice C-terminal [Forster *et al.*, 2013]. El dominio PCI presenta un plegamiento bipartido compuesto por un dominio N-terminal de repeticiones helicoidales seguido de un dominio de tipo hélice alada (*winged helix*), que presenta un plegamiento de tipo hélice-giro-hélice con una topología α - β - α - α - β - β donde las cadenas β forman una pequeña lámina β antiparalela en la superficie de la proteína, los bucles formados entre la segunda y la tercera cadena β y entre la tercera y el extremo C-terminal de la molécula forman las "alas" del dominio [Ellisdon *et al.*, 2012]. Recientemente las estructuras cristalográficas de Rpn6 y Rpn12 fueron descritas por [Pathare *et al.*, 2012] y [Boehringer *et al.*, 2012] respectivamente (Fig. 1.10).

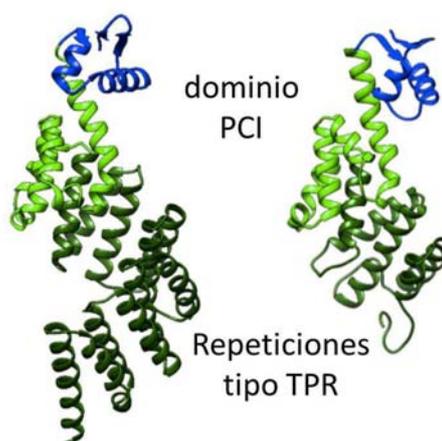


Fig. 1.10 Proteínas PCI. Estructura cristalográfica de Rpn6 (izquierda) y Rpn12 (derecha) donde podemos observar el dominio PCI, en azul la hélice alada y en verde claro las repeticiones helicoidales y el dominio N-terminal de repeticiones helicoidales tipo TPR en verde oscuro. Imagen modificada de Förster *et al.*, 2013.

El dominio MPN consiste de una secuencia de ~ 120 aminoácidos con estructura α/β en el extremo N-terminal de la proteína. La estructura del dominio MPN de Rpn8 humano fue resuelta por [Sanchez *et al.*, 2007] (Fig. 1.11), podemos observar que en el centro de la estructura está contenida una lámina β formada por 9 cadenas β rodeada por 3 hélices α , una cuarta hélice en el extremo C-terminal emerge del centro de la estructura y se intercambia con la hélice α de otra subunidad relacionada simétricamente,

sugiriendo que en solución el dominio es dimérico. A diferencia de Rpn8 que no es catalítica, Rpn11 contiene dos histidinas y un aspartato precedido de un glutamato conservado pero con espaciamiento variable entre ellos ($EX_nHXHX_{10}D$). Este motivo se denomina JAMM (Jab1/Pad1/MPN metalo-enzima), puede unir un ión de zinc y le confiere a la proteína actividad isopeptidasa [Verma *et al.*, 2002]. Si bien Rpn11 es la única deubiquitinasa integral del proteosoma su actividad se limita a cortar la cadena de ubiquitina por su base, separándola del sustrato; para obtener la ubiquitina monomérica es necesaria la acción de otra enzima deubiquitinante Ubp6 que interacciona de forma transitoria con el proteosoma a través de Rpn1 [Tanaka, 2009].

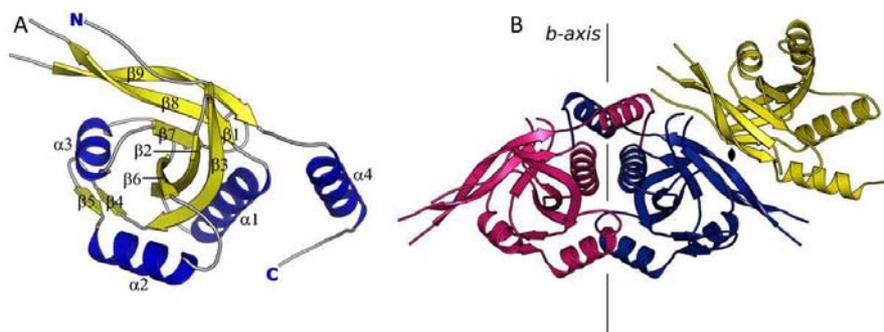


Fig. 1.11 Dominio MPN. A) Estructura cristalográfica del dominio MPN de Rpn8 humana; B) Representación de los dos posibles estados diméricos, la subunidades en azul y amarillo representan la posición de las cadenas en la unidad asimétrica, mientras que las cadenas en azul y rojo representan el dímero cristalográfico el cual basados en cálculos energéticos es que probablemente se encuentre en solución. Imagen tomada de Sanches *et al.*, 2007.

La estructura global de la lid ha sido determinada por EM por Lander *et al.*, 2012, los dominios PCI interactúan lateralmente formando una estructura con forma de herradura organizados en el siguiente orden Rpn9/Rpn5/Rpn6/Rpn7/Rpn3/Rpn12, desde donde los extremos N-terminales se extienden radialmente, mientras que Rpn8 cierra la herradura e interactúa con Rpn11 (Fig. 1.12), y Sem1 se ubica en la hendidura entre los dominios PCI de Rpn3 y Rpn7 [Bohn *et al.*, 2013]. Las hélices C-terminales de las proteínas PCI junto con las hélices C-terminales de Rpn8 y Rpn11 se unen formando lo que se conoce como haz helicoidal o *helical bundle*, esta estructura recientemente descubierta por Estrin *et al.*, 2013, funciona como un dispositivo que asegura el correcto ensamblaje de la lid.

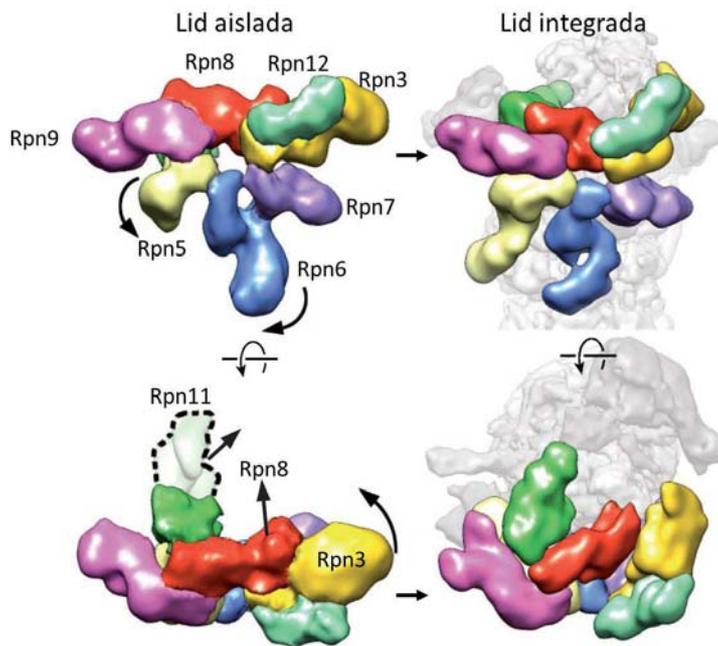


Fig. 1.12 Estructura de la lid resuelta por EM aislada e integrada al proteosoma 26S, los rearrreglos estructurales se encuentran marcados por flechas. Imagen modificada de Lander *et al.*, 2012.

Cuando se analiza la densidad del proteosoma 26S se observa que la lid se posiciona a un lado de la partícula reguladora y que posee numerosos puntos de contacto con la base y también con la CP. La comparación entre la lid aislada y la lid en el contexto del proteosoma demuestra que hay cambios conformacionales importantes entre una y otra (Fig. 1.12); en la lid aislada el dominio MPN de Rpn11 interacciona con Rpn9 y curva hacia arriba el extremo N-terminal de Rpn5, cuando la lid se une al proteosoma Rpn5 es liberada y su N-terminal contacta con la CP mientras que Rpn11 se extiende hacia Rpn2 [Lander *et al.*, 2012]. Además, cuando el sustrato interacciona con el anillo de ATPasas Rpn11 se desplaza 18\AA hacia el centro de la RP colocándose directamente encima del poro del anillo α . Probablemente las interacciones con Rpn5 y Rpn9 restringen el acceso al sitio activo de Rpn11 previniendo la deubiquitinación del sustrato en ausencia de la base y la CP, y su reubicación una vez que el sustrato está interaccionando con las ATPasas asegura que el mismo no será liberado antes de tiempo.

La actividad deubiquitinasa de Rpn11 es fundamental para el correcto desarrollo del proceso de degradación. Debido a que Rpn11 no presenta afinidad por la Ub para poder deubiquitarlo el sustrato debe permanecer unido al receptor de ubiquitinas Rpn10 o Rpn13. Ambos receptores se encuentran a unos $70\text{-}80\text{\AA}$ del sitio activo de Rpn11 por lo que se requieren 3 ubiquitinas para abarcar la distancia entre ellos. Siendo que el reconocimiento de la cadena de ubiquitinas se produce en la interfaz de unión de dos ubiquitinas consecutivas, al menos una cadena de tetra-ubiquitinas sería necesaria para el correcto procesamiento del sustrato. Esto explicaría por qué la degradación de proteínas depende de que la cadena de ubiquitina posea al menos 4 unidades.

Además, a fin de evitar la disociación del sustrato al ser deubiquitinado, el sustrato debe primero interactuar con la maquinaria translocadora de la base. Esta interacción requiere que el sustrato presente un sitio de iniciación desestructurado que debe ser lo suficientemente largo como para alcanzar el anillo de ATPasas. Se ha demostrado que la deubiquitinación es dependiente de ATP, lo que indica que, además de ser necesario que el sustrato este unido a un receptor, se requiere de la translocación del mismo a fin de "tirar" de la ubiquitina proximal hacia el sitio activo de Rpn11. A medida que el sustrato se despliega y va recorriendo el canal central Rpn11 lo escanea eliminando las ubiquitinas, facilitando la translocación y la degradación del mismo (Fig. 1.13) [Lander *et al.*, 2013; Matyskiela and Martin, 2013; Tomko *et al.*, 2013]

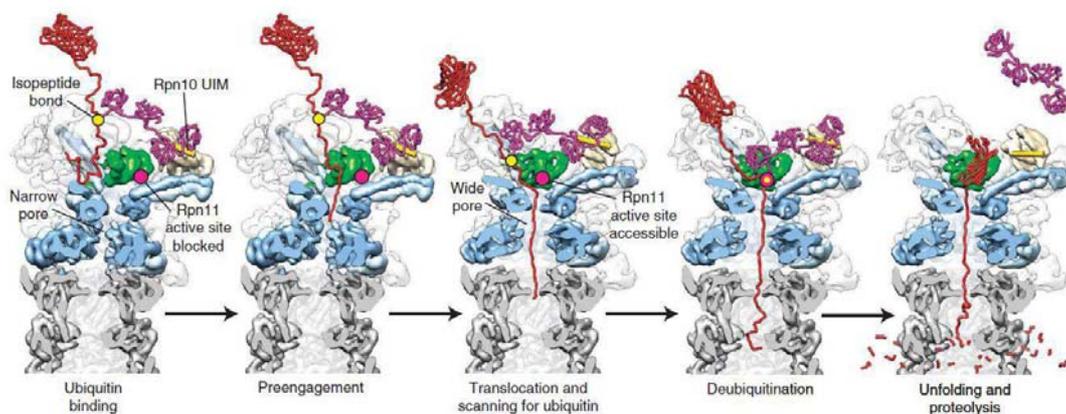


Fig. 1.13 Modelo de interacción y degradación basado en la estructura. En el primer paso, el sustrato (rojo) es reconocido a través de su cadena de ubiquitinas (violeta) por el dominio UIM de Rpn10 (cilindro amarillo). En este estado previo a la interacción el sitio de iniciación desestructurado del sustrato contacta con Rpt3 la subunidad superior de la escalera de caracol formada por las ATPasas. Una vez establecida la interacción las Rpts se reacomodan alineando el canal central y abriendo el anillo α de la CP (gris). Rpn11 (verde) se ubica directamente encima del anillo de ATPasas exponiendo su sitio activo (punto rosa) para escanear el sustrato a medida que es translocado. Todas las cadenas de ubiquitinas son separadas cuando el sitio del enlace isopeptídico pasa por Rpn11. Imagen tomada de Matyskiela, Lander *et al.*, 2013.

1.1.8 Biogénesis del proteosoma 26S

A diferencia del proteosoma 20S de procariontes que se ensambla de manera automática sin la participación de chaperonas, en eucariotes la biogénesis de la CP es mucho más compleja y requiere de la participación de chaperonas extrínsecas como PAC1-4 (*proteasome assembly chaperones*) y UMP1 (*ubiquitin maturation protein*) en humanos o sus ortólogos en levaduras Pba1-4 (*proteasome biogenesis-associated*) y Ump1; e intrínsecas como propéptidos y las colas C-terminales de las subunidades β .

PAC1-PAC2 y PAC3-PAC4 son los dos complejos heterodiméricos involucrados en la formación de los anillos α . Mientras PAC1-2 acompañan a la CP durante todo el proceso de biosíntesis evitando la dimerización espontánea de las subunidades α , PAC3-4 participa tempranamente favoreciendo la correcta colocación de las subunidades α y se libera cuando el anillo α está listo. El anillo α funciona como molde para la incorporación de las subunidades β en el siguiente orden: β_2 , β_3 , β_4 , β_5 , β_6 , β_1 y finalmente β_7 ; los extremos C-terminales y los propéptidos de los extremos N-terminales de las subunidades β son esenciales para el ensamblaje correcto del anillo β . Ump1 se integra al complejo posteriormente a β_2 , β_3 y β_4 y participa en la dimerización de los dos hemi-proteosomas. Finalmente la maduración del proteosoma se completa por autoproteólisis de los propéptidos β , activando los sitios activos de β_1 , β_2 y β_5 y liberando las cavidades proteosomales mediante la degradación de PAC1-2 y Ump1 (Fig. 1.14) [Murata *et al.*, 2009; Tanaka, 2009; Gallastegui *et al.*, 2010].

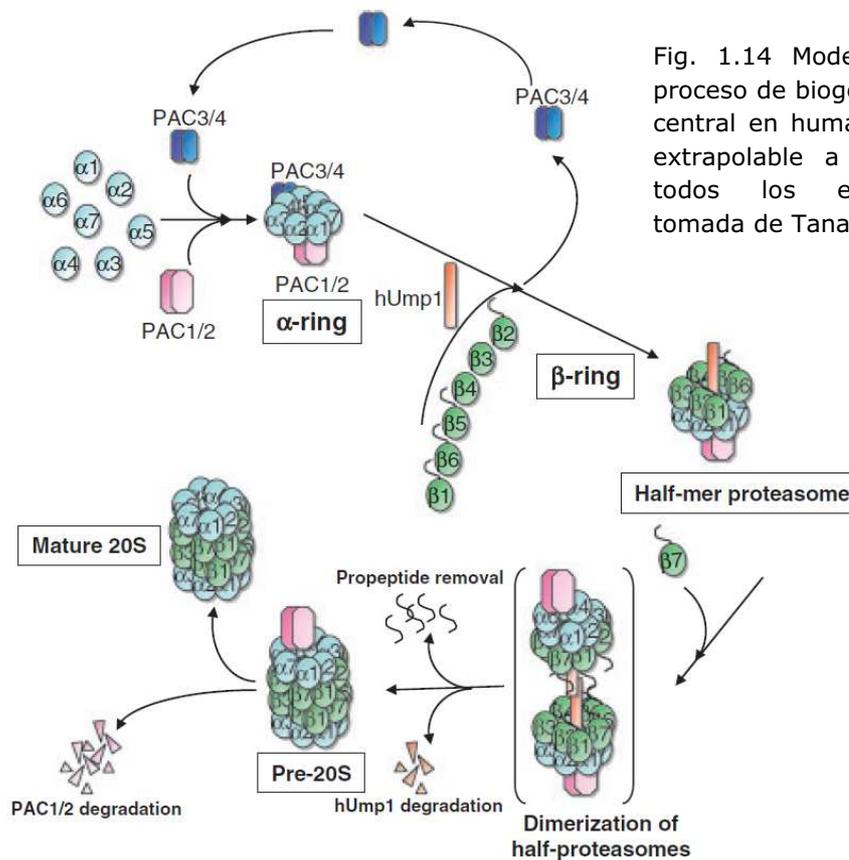


Fig. 1.14 Modelo esquemático del proceso de biogénesis de la partícula central en humanos. El esquema es extrapolable a lo que sucede en todos los eucariotas. Imagen tomada de Tanaka, 2009.

En cuanto a la biogénesis de la partícula reguladora podemos decir que el ensamblaje de la base y de la lid son eventos independientes y que una vez formados se unen junto con la CP para formar finalmente el proteosoma 26S.

Hasta la fecha cuatro chaperonas diferentes han sido identificadas en el proceso de biogénesis de la base: Nas2, Nas6, Hsm3 y Rpn14, en

conjunto se las denomina factores específicos de ensamblaje del complejo 19S. Cada una de estas chaperonas se une específicamente a una o varias subunidades Rpt, resultando en la formación de los siguientes complejos iniciales: Rpn14-Rpt6-Rpt3-Nas6, Rpt4-Rpt5-Nas2 y Rpt1-Rpt2-Rpn1-Hsm3, conocidos como módulos Rpn14/Nas6, Nas2 y Hsm3 respectivamente. Estos complejos se asocian unos con otros junto con Rpn2 y Rpn13 para formar un complejo precursor de la base, al que finalmente se le adicionan Rpn10 y la lid preformada para obtener la partícula regulatoria completa (Fig. 1.15) [Funakoshi *et al.*, 2009; Murata *et al.*, 2009; Gallastegui *et al.*, 2010]. En levaduras los módulos Rpn14/Nas6 y Nas2 se asocian primero y durante la incorporación del módulo Hsm3 se libera Nas2. En células de mamíferos en el complejo entre Rpn14/Nas6 y Hsm3 se observó q faltan las subunidades Rpt4 y Rpt5, indicando que el orden en que estos complejos se asocian varía entre especies. Alternativamente, los módulos podrían asociarse de diversas maneras ya que cada uno de ellos ha sido encontrado unido a la RP completa pero no al proteosoma 26S, indicando que se liberan antes o durante la unión de la RP a la CP [Tomko *et al.*, 2013].

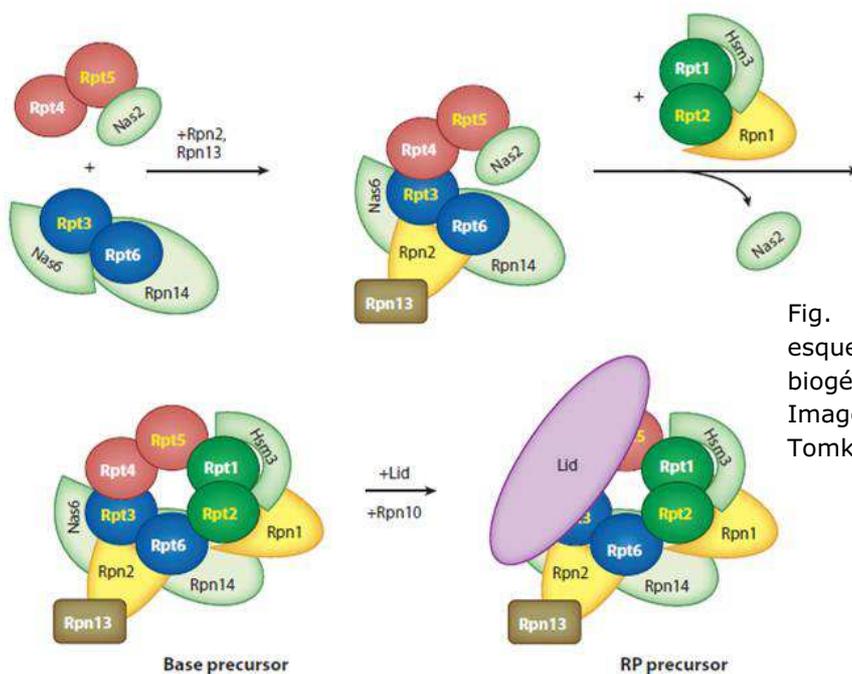


Fig. 1.15 Modelo esquemático de la biogénesis de la base. Imagen tomada de Tomko *et al.*, 2013.

En el caso de la biogénesis de la lid, si bien se han descrito diversos subcomplejos, el proceso de cómo se ensambla la lid o si requiere de enzimas chaperonas no ha sido comprobado aún. Basándonos en las interacciones observadas por espectrometría de masas (MS *mass spectrometry*) [Sharon *et al.*, 2006] y los subcomplejos observados in vivo [Fukunaga *et al.*, 2010; Tomko *et al.*, 2011] podemos delimitar dos subcomplejos el módulo 1 o complejo central (*core complex*) compuesto por Rpn5, Rpn6, Rpn8, Rpn9 y Rpn11 y el módulo 2 formado por Rpn3, Rpn7 y Sem1, que al unirse permiten la adición de Rpn12 completando la formación de la lid, y la posterior unión de la lid a la base (Fig. 1.16).

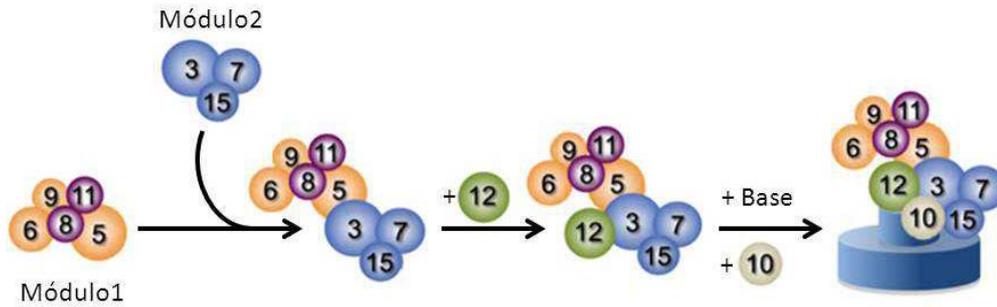


Fig. 1.16 Modelo esquemático del proceso de biogénesis de la lid, las subunidades del módulo 1 representadas en naranja son las que poseen el dominio PCI y en violeta las que presentan el dominio MPN. Imagen modificada de Tomko *et al.*, 2011.

Muy recientemente [Estrin *et al.*, 2013] describieron el haz helicoidal que se forma por la unión de las hélices C-terminales de todas las subunidades de la lid. Este paquete helicoidal está formado por 12 hélices, siete de ellas forman una estructura que recuerda el cilindro de un revolver, dentro del cual una hélice central (Rpn8) se encuentra rodeada de otras seis que se encuentran ligeramente inclinadas con respecto a la central (Rpn12, Rpn3, Rpn7, Rpn6 y dos hélices de Rpn11), las cinco hélices restantes se colocan junto a este cilindro unidas entre ellas (Fig. 1.17).

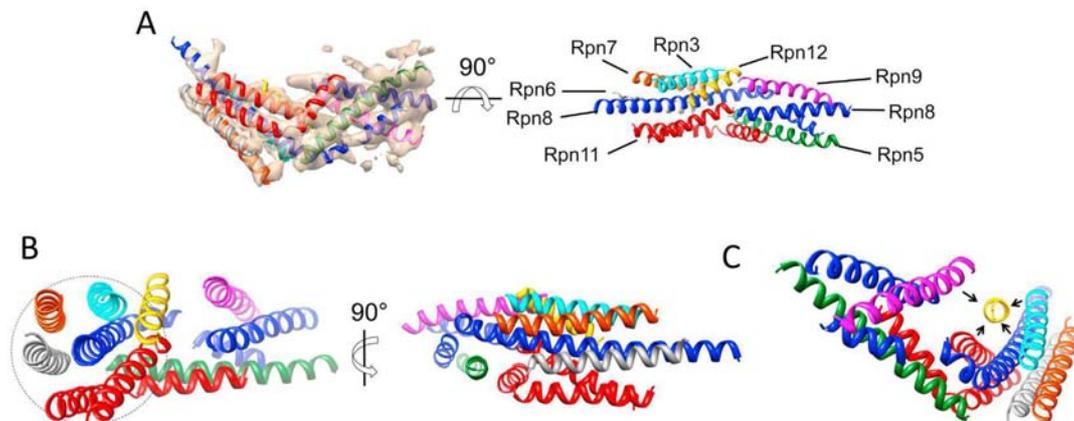


Fig. 1.17 El haz helicoidal. A-C) modelo del haz helicoidal desde cinco vistas diferentes. A) inclusión del haz en la densidad de la EM. B) Las hélices de Rpn3, Rpn6, Rpn7, Rpn8, Rpn11 y Rpn12 forman un cilindro (indicado por un círculo) con Rpn8 en posición central. C) La hélice de Rpn12 contacta con las hélices de Rpn3, Rpn8, Rpn9 y Rpn11 (indicado con flechas). Imagen tomada de Estrin *et al.*, 2013.

La topología del haz helicoidal sugiere un elegante mecanismo por el cual la incorporación de Rpn12 depende de la presencia de todas las otras hélices. La hélice de 15 residuos de Rpn12 contacta únicamente con Rpn3, Rpn8, Rpn9 y Rpn11, a su vez estas hélices mantienen importantes contactos con las otras hélices de haz helicoidal, esto sugiere que, ya sea directa o indirectamente a través de la conformación global del haz, Rpn12 controla la correcta formación del haz y asegura la asociación eficiente de la

lid con la base para completar la RP [Tomko *et al.*, 2011; Estrin *et al.*, 2013].

1.1.9 Objetivos

El objetivo de este trabajo es estudiar las proteínas que componen la partícula reguladora del proteosoma, las interacciones que se establecen entre ellas y analizar la biogénesis de la lid.

Por lo que definimos los siguientes objetivos parciales:

- Clonado, expresión/co-expresión y purificación de las proteínas de la base Rpn1 y Rpn2
- Clonado, expresión y purificación del receptor de ubiquitina Rpn10
- Analizar las interacciones que establece Rpn10 con las proteínas de la lid.
- Clonado, expresión y purificación de las proteínas de la lid de forma individual.
- Clonado, co-expresión/co-lisado y purificación de diversas combinaciones de las proteínas de la lid y versiones truncadas de las mismas.
- Purificación del complejo de la lid endógeno de *S. cerevisiae*.
- Realizar ensayos cristalográficos de todos los candidatos adecuados.
- Análisis de la lid endógena y subcomplejos de la misma mediante EM, con el fin de determinar el proceso de ensamblaje de la misma.

1.2 Materiales y métodos

1.2.1 Procedimientos generales

Purificación de ADN

La purificación de plásmidos, productos de PCR y fragmentos obtenidos de la digestión con enzimas de restricción se procesaron en una minicentrífuga de sobremesa (Eppendorf) con los kits de Qiagen miniprep y PCR purification, siguiendo las recomendaciones del fabricante.

Electroforesis en geles de agarosa

Se utilizó electroforesis en geles de agarosa para evaluar el tamaño de los amplicones, plásmidos y fragmentos resultantes de digestión con enzimas de restricción.

La concentración de los geles fue del 1% y fueron preparados y usados en buffer TAE (20mM Tris-Acetato, 0,5mM EDTA). Para sembrar las muestras se utilizó el buffer de carga 6X de Fermentas. Los geles se corrieron a 80-100V utilizando una fuente de poder Power Pac Basic (Biorad). Se usó el Lambda DNA/Eco91I (Fermentas) como estándar de peso molecular. Los geles se tiñeron con SYBER safe (Invitrogen) y fueron visualizados y fotografiados con un transiluminador U:genius (Syngene).

Determinación de concentración y pureza del ADN

La concentración y calidad del ADN se midió utilizando un nanodrop 1000 (Thermo Scientific).

Reacciones de digestión y ligación del ADN

Todas las reacciones de ligación o digestión se llevaron a cabo con enzimas de New England Biolabs o Fermentas siguiendo el protocolo recomendado por los fabricantes.

Secuenciación de ADN

Todos los clonajes realizados fueron verificados mediante secuenciación, las muestras fueron procesadas por Macrogen Korea.

Medios de cultivo, cepas bacterianas y vectores

La composición de los medios de cultivos y los aditivos utilizados se detallan en la siguiente lista:

- Medio Luria-Bertani (LB): triptona 1%, NaCl 1%, extracto de levadura 0,5%.
- Placas de LB: triptona 1%, NaCl 1%, extracto de levadura 0,5%, bacto-agar 1,5%.
- Medio super-óptimo con represión por catabolitos (SOC): triptona 2%, Extracto de levadura 0,5%, 10mM NaCl, 2,5mM KCl, 12mM MgCl₂, 10mM MgSO₄, 2,5mM glucosa.
- Antibióticos: ampicilina (solución stock: 100mg/ml en ddH₂O, concentración final 100µg/ml); kanamicina (solución stock: 50mg/ml en ddH₂O, concentración final: 50µg/ml); cloranfenicol (solución stock: 35mg/ml en etanol 100%, concentración final: 35µg/ml).
- IPTG: solución stock 1M.

Las cepas bacterianas utilizadas en este trabajo se describen en la siguiente tabla:

Tabla 1.1: Cepas bacterianas

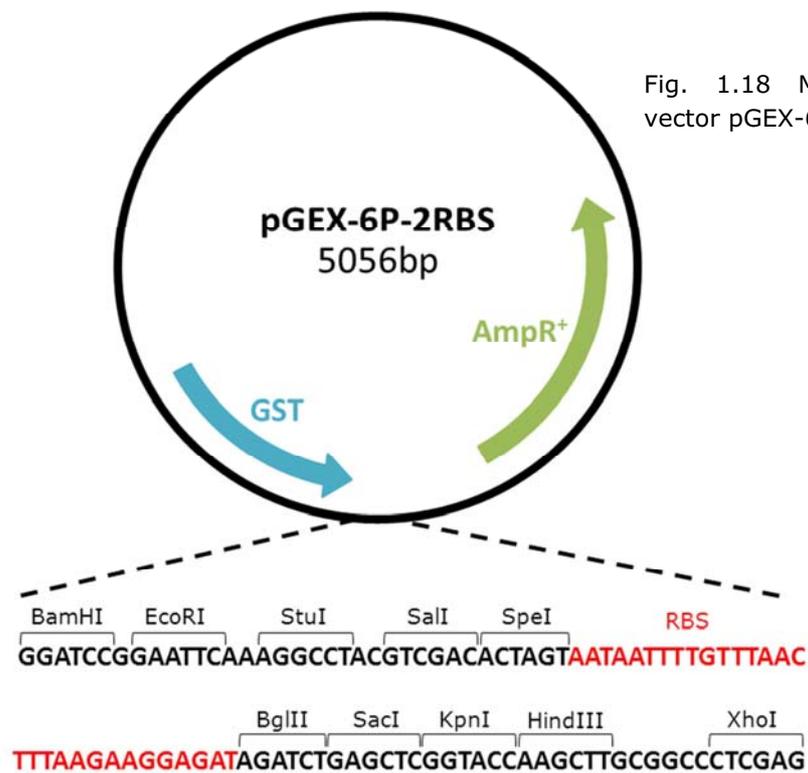
Cepa	Genotipo	Características
XL1 Blue	endA1 gyrA96(nal ^R) thi-1 recA1 relA1 lac glnV44 F' Tn10 proAB ⁺ lacI ^q Δ(lacZ) M15 hsdR17(r _K ⁻ m _K ⁺)	Cepa de clonado
DH5α	F ⁻ endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80d/lacZΔM15 Δ(lacZYA- argF)U169, hsdR17(r _K ⁻ m _K ⁺), λ ⁻	Cepa de clonado
BI21 (DE3)	<i>E. coli</i> B F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3 lacI lacUV5-T7 gene 1 ind1 sam7 nin5)	Cepa de expresión con el lisógeno del fago λ DE3 que expresa el gen de la RNA polimerasa del fago T7 bajo el control del promotor <i>lacUV5</i>
BI21 (DE3) Codon-plus RIL	<i>E. coli</i> B F ⁻ ompT hsdS (r _B -m _B -) dcm ⁺ Tet ^R <i>E. coli</i> gal λ (DE3) endA Hte argU ileY leuW Cam ^R	Cepa de expresión con DE3, esta cepa codifica tRNAs para codones de arginina AGA, AGG; isoleucina AUA; y leucina CUA

Los vectores utilizados en este trabajo se describen en la siguiente tabla:

Tabla 1.2: Vectores

Vector	Tamaño	Antibiótico	Promotor	N-ter	Proteasa	Referencia
pGEX-4T	4969bp	Amp	pTac/LacI	GST	Trombina	Amersham
pGEX-6P-2RBS	5056bp	Amp	pTac/LacI	GST	PreScission	[Sessa <i>et al.</i> , 2005]

El vector pGEX-6P-2RBS es un vector de expresión que ha sido modificado a partir de un pGEX-6P al que se le añadió un segundo sitio de unión al ribosoma (RBS *ribosome binding site*) (Fig. 1.18) [Sessa *et al.*, 2005].



Este vector permite la expresión de dos proteínas donde la primera es expresada fusionada a GST mientras que la segunda solo puede ser purificada si interacciona con la primera (Fig. 1.19).

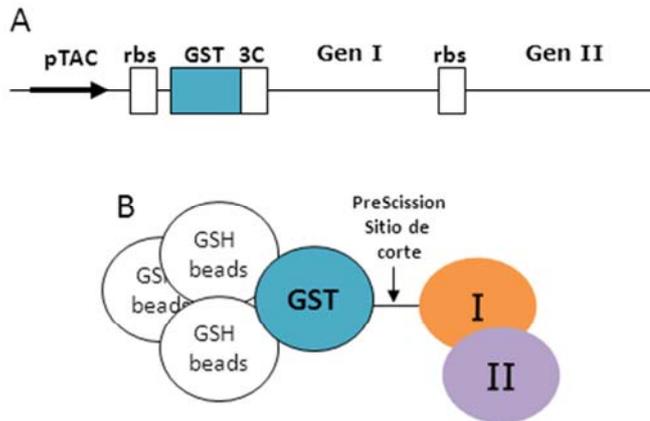


Fig. 1.19 Esquema del vector pGEX-6P-2RBS.

Representación esquemática de A) la región de expresión del vector y B) la purificación de las proteínas de interés, con la matriz de glutatión (GSH).

Transformación de bacterias

El protocolo de transformación usado fue el siguiente: a 50µl de bacterias competentes, obtenidas por el método del CaCl₂ [Sambrook *et al.*, 2006], se le agregó 1µl de ADN. La mezcla se incubó durante 30 minutos en hielo y luego durante 45 segundos a 42°C. Las bacterias se enfriaron durante 2 minutos en hielo y fueron colocadas en medio SOC para crecerlas durante una hora antes de plaquearlas.

Electroforesis en geles de poliacrilamida SDS-PAGE

Las proteínas y los complejos recombinantes fueron analizados por SDS-PAGE [Laemmli, 1970]. La composición del gel separador fue: acrilamida 12% p/v, bis-acrilamida 0,27% p/v, SDS 0,1% en buffer Tris 0,36M pH 8.8 mientras que la del gel concentrador fue: acrilamida 5% p/v, bis-acrilamida 0,13% p/v, SDS 0,1% p/v en Tris 0,12M pH 6.8. La polimerización se realizó usando N,N,N',N'-tetrametil etilendiamina (TEMED) 0,06% v/v como iniciador y persulfato de amonio 0,05% p/v como catalizador.

Las muestras se prepararon hirviéndolas durante 5 minutos a 95°C en buffer de siembra (55mM Tris pH 6,8; SDS 1,6%; glicerol 8,3%; azul de bromofenol 0,002%, 2- β-mercaptoetanol 5%). Las electroforesis se realizaron con el sistema Mini Protean II (BioRad) usando como buffer de corrida 25mM Tris pH 8.8, 192mM glicina, SDS 0,1%, a 150V-250V, usando como estándar de peso molecular el SeeBlue[®] preteñido (Invitrogen). A continuación, los geles fueron teñidos con una solución de Coomassie Blue G-250 (Sigma) 0,5 % p/v, metanol 50% v/v y ácido acético 10% v/v.

Columnas cromatográficas

Todas las columnas se compraron a GE Healthcare y fueron utilizadas en máquinas Äkta explorer o purifier 10, a temperatura ambiente o a 4°C dependiendo el caso. Las columnas que se usaron en este trabajo se detallan en la siguiente lista:

- GST trap FF (1ml y 5ml)
- Superdex 200 (10/300 y 16/60)
- Superosa 6 10/300
- Mono Q 5/50

En las purificaciones por afinidad con GST (Glutathion S-transferasa) de menor volumen se utilizó GST Sepharose Fast Flow (Amersham Biosciences).

Concentración y diálisis de proteínas

La concentración de proteínas se hizo con dispositivos Vivaspin (Sartorius) de diferentes volúmenes (0,5ml, 2ml, 6ml y 20ml) y membranas de diferentes tamaños de poro (3000, 5000, 10000 y 30000 MWCO). Y la diálisis se realizó utilizando "Slide-A-Lyzer Dialysis Cassettes" (Pierce) para volúmenes de 0,1-0,5µl y 1-3ml o los botones de diálisis (Hampton Research) de 50µl en el caso de volúmenes menores.

La concentración de las diferentes proteínas se cuantificaron mediante el método de Bradford [Bradford, 1976] con el reactivo de BioRad y la pureza de las muestras se evaluó por SDS-PAGE.

Western blots

La transferencia se realizó con el sistema Trans-blot® semi-seco (BioRad) siguiendo las recomendaciones del fabricante y se utilizaron los siguientes reactivos: buffer de transferencia (200ml de metanol, 3gr de Tris, 3gr de glicina, 3ml de SDS al 10%, en un volumen final de 1l), papel extragrueso Whatmann 3MM (GE Healthcare) y membrana de nitrocelulosa o PVDF (GE Healthcare).

El bloqueo se realizó durante toda la noche a 4°C en una solución de TBS-Tween (Tris 50mM, NaCl 150mM y Tween 20 0.05%) con 5% de leche descremada. La incubación con el anticuerpo primario y secundario se realizaron a temperatura ambiente, con 3 lavados de 8min cada uno con TBS-Tween entre la incubación de cada anticuerpo. El revelado fue llevado a cabo por quimioluminiscencia por lo que se incubó la membrana con el sustrato ECL (GE Healthcare) o ECL plus (GE Healthcare) y revelado en

película fotográfica Kodak con una maquina reveladora Hyper processor (Amersham farmacia biotech).

Análisis por MS de proteínas

Todos los estudios de MS fueron llevados a cabo en la plataforma científica de espectrometría de masa del IRB Barcelona.

1.2.2 Proteínas recombinantes

Diseño de los constructos

El diseño de los constructos truncados fue definido mediante el análisis de la secuencia de cada proteína con una variedad de herramientas bioinformáticas, detalladas en la siguiente lista:

- SMART (Web1) [Schultz *et al.*, 1998; Letunic *et al.*, 2012]
- PrDOS (Web2) [Ishida *et al.*, 2007]
- COILS (Web3) [Lupas *et al.*, 1991]
- Protein CCD at the NKI (Web4) [Mooij *et al.*, 2009]
- PSIPRED (Web5) [Jones, 1999; Buchan *et al.*, 2010]

También se tomaron en cuenta los resultados obtenidos en el trabajo realizado por Fu *et al.*, 2001.

PCR y clonado

Las PCRs se realizaron con el kit KOD DNA polymerase (Novagen) siguiendo las recomendaciones del fabricante. Como templado se utilizaron los genes completos (*full-length*) de todas las Rpns insertados en un vector pGEX-4T, cedidos amablemente por el Dr. Bernat Crosas del Instituto de Biología Molecular de Barcelona (CSIC). En la siguiente tabla se especifican los detalles de la reacción:

Tabla 1.3: Protocolo de PCR		
Componentes	Volumen	Concentración final
Primer Fwd	4µl	0,4µM
Primer Rev	4µl	0,4µM
Templado	1µl	
Buffer 10X	5µl	1X
MgCl ₂ 25mM	3µl	1,5mM

dNTPs (2mM c/u)	5µl	0,2mM c/u
H ₂ O ultrapura	27,6µl	
KOD DNA polymerase	0,4µl	0,02U/µl
Volumen total	50µl	
# de ciclo	Tiempo	Temperatura
1 ciclo	30 seg	98°C
30 ciclos	10 seg	98°C
	30 seg	58°C
	60 seg	72°C
Elongación final	10 min	72°C
Espera	∞	4°C

Las PCRs se realizaron en un termociclador Eppendorf mastercycle gradient y los resultados se analizaron mediante geles de agarosa. Los primers utilizados para el clonaje de los diversos constructos se detallan en la siguiente tabla:

Nombre	Secuencia
Rpn1 F BamHI	CGC-GGATCC-ATGGTAGACGAAAGTGATAAGAAACA
Rpn1 F BglII	CGC-AGATCT-ATGGTAGACGAAAGTGATAAGAAACA
Rpn1 R StuI	CGC-AGGCCT-TTACTCCTCTTCACGATAGTCAGG
Rpn1 R SacI	CGC-GAGCTC-TTACTCCTCTTCACGATAGTCAGG
Rpn1 ^{PC} K438 F BamHI	CGC-GGATCC-ATGAAGTATCTCTACGTTGATGAGCC
Rpn1 ^{PC} K438 F BglII	CGC-AGATCT-ATGAAGTATCTCTACGTTGATGAGCC
Rpn1 ^{PC} V876 R SalI	CGC-GTCGAC-TTACACCTTATTCAAACATGTGCATC
Rpn1 ^{PC} V876 R SacI	CGC-GAGCTC-TTACACCTTATTCAAACATGTGCATC
Rpn1 ^{PC} V876 R XhoI	CGC-CTCGAG-TTACACCTTATTCAAACATGTGCATC
Rpn2 ^{PC} K350 F BamHI	CGC-GGATCC-ATGAAGGCGAACTTGCCA
Rpn2 ^{PC} K350 F BglII	CGC-AGATCT-ATGAAGGCGAACTTGCCA
Rpn2 ^{PC} V730 R SalI	CGC-GTCGAC-TTATACGGACTTGGTATCCAATGTG
Rpn2 ^{PC} V730 R XhoI	CGC-CTCGAG-TTATACGGACTTGGTATCCAATGTG
Rpn5 F BamHI	CGC-GGATCC-ATGTCAAGAGATGCACCAATTAAG
Rpn5 F EcoRI+2	CGC-GAATTC-AA-ATGTCAAGAGATGCACCAATTAAG
Rpn5 F ΔN G169 F BamHI	CGC-GGATCC-ATGGGCTCCATGGAAATGT
Rpn5 F ΔN A249 F BamHI	CGC-GGATCC-ATGGCCATTAAATCAGATGAGG
Rpn5 F ΔN I313 F BamHI	CGC-GGATCC-ATGATTGTTCAAAAAACCTATGAGC
Rpn5 R SalI	CGC-GTCGAC-TCATTTAGCTTGCAAACCGTG
Rpn9 F ΔN K160 F BglII	CGC-AGATCT-ATGAAAAGGATTCCATTCCATT
Rpn9 F ΔN G261 F BglII	CGC-GGATCC-ATGGGTGATTTTGATAAATTCGACTCA
Rpn9 R XhoI	CGC-CTCGAG-CTAAACCCAGATGGATTGGC

Muchos genes estaban ya clonados por lo que para la producción de muchas de las combinaciones se realizaron mediante subclonajes, lo que eliminó la necesidad de repetir PCRs. La digestión con enzimas de restricción, purificación de los amplicones y vectores, y la ligación fueron realizados siguiendo la metodología anteriormente explicada. Los clones obtenidos (o) y los clones disponibles en el laboratorio (d), organizados en las diversas líneas de investigación llevadas a cabo se detallan en la siguiente tabla:

Tabla 1.5: Clones			
Proyecto	Gen/genes	Vector	o/d
Rpn1 y Rpn2	Rpn1 fl	pGEX-4T	d
	Rpn2	pGEX-4T	d
	Rpn1 ^{PC}	pGEX-6P-2RBS	o
	Rpn2 ^{PC}	pGEX-6P-2RBS	o
	Rpn1/Rpn2 ^{PC}	pGEX-6P-2RBS	o
	Rpn2 ^{PC} /Rpn1	pGEX-6P-2RBS	o
	Rpn1 ^{PC} /Rpn2 ^{PC}	pGEX-6P-2RBS	o
	Rpn2 ^{PC} /Rpn1 ^{PC}	pGEX-6P-2RBS	o
Rpn10 y la interacción base/lid	Rpn10	pGEX-4T	d
	Rpn3/Rpn10	pGEX-6P-2RBS	o
	Rpn5/Rpn10	pGEX-6P-2RBS	o
	Rpn9/Rpn10	pGEX-6P-2RBS	o
	Rpn11/Rpn10	pGEX-6P-2RBS	o
	Rpn12/Rpn10	pGEX-6P-2RBS	o
Proteínas de la lid	Rpn3	pGEX-4T	d
	Rpn5	pGEX-4T	d
	Rpn6	pGEX-4T	d
	Rpn7	pGEX-4T	d
	Rpn8	pGEX-4T	d
	Rpn9	pGEX-4T	d
	Rpn11	pGEX-4T	d
	Rpn12	pGEX-4T	d
	Rpn5 IC	pGEX-6P-2RBS	d
	Rpn5 _{169-f} IC	pGEX-6P-2RBS	o
	Rpn5 _{249-f} IC	pGEX-6P-2RBS	o
	Rpn5 _{313-f} IC	pGEX-6P-2RBS	o
	Rpn6 IC	pGEX-6P-2RBS	d
	Rpn6 _{110-f} IC	pGEX-6P-2RBS	d
Rpn8 IC	pGEX-6P-2RBS	d	

Rpn9 IC	pGEX-6P-2RBS	d
Rpn11 IC	pGEX-6P-2RBS	d
Rpn11 ₁₋₂₀₅ IC	pGEX-6P-2RBS	d
Rpn3/Rpn7	pGEX-6P-2RBS	d
Rpn5/Rpn9	pGEX-6P-2RBS	d
Rpn5/Rpn9 _{261-f}	pGEX-6P-2RBS	o
Rpn5 _{169-f} /Rpn9 _{160-f}	pGEX-6P-2RBS	o
Rpn5 _{313-f} /Rpn9	pGEX-6P-2RBS	o
Rpn5 _{313-f} /Rpn9 _{261-f}	pGEX-6P-2RBS	o
Rpn5/Rpn8	pGEX-6P-2RBS	d
Rpn5 _{169-f} /Rpn8	pGEX-6P-2RBS	o
Rpn5 _{249-f} /Rpn8	pGEX-6P-2RBS	o
Rpn5 _{313-f} /Rpn8	pGEX-6P-2RBS	o
Rpn11/Rpn8	pGEX-6P-2RBS	d
Rpn11/Rpn8 ₁₋₁₈₆	pGEX-6P-2RBS	d
Rpn11/Rpn8 ₁₋₂₄₀	pGEX-6P-2RBS	d
Rpn11 ₁₋₂₀₅ /Rpn8	pGEX-6P-2RBS	d
Rpn11 ₁₋₂₀₅ /Rpn8 ₁₋₁₈₆	pGEX-6P-2RBS	d
Rpn11 ₁₋₂₀₅ /Rpn8 ₁₋₂₄₀	pGEX-6P-2RBS	d

Expresión y purificación

La expresión heteróloga de las proteínas de interés fusionadas a GST se llevó a cabo en *E. coli* BI21 o BI21 codon-plus RIL según cada caso. Cada partida de expresión se empezó de placas con bacterias transformadas frescas (de no más de una semana) a partir de las que se lanzó un precultivo en LB con los antibióticos correspondientes que se creció durante toda la noche a 37°C. Al día siguiente este precultivo se utilizó como inóculo y se diluyó 1/100 en LB con antibióticos y se creció a 37°C en agitación (180rpm) en un agitador Innova 44 (New Brunswick Scientific). Al alcanzar una $DO_{600nm} \sim 0.9-1.0$ se le agregó IPTG (isopropil- β -D1-tiogalactopiranósido) a una concentración final de 0,3mM, la inducción se realizó durante toda la noche a 18°C en agitación (180rpm). Dependiendo el caso y de la cantidad de proteína que se esperaba obtener los cultivos se realizaron en dos volúmenes diferentes: cultivos de 500ml en Erlenmeyers de 2l o 2l de cultivo en Erlenmeyers de 6l.

Para obtener la masa bacteriana los cultivos inducidos se centrifugaron a 4000g por 15 min en una centrífuga Beckman Coulter Avanti J-26 XPI en diferentes rotores dependiendo del volumen. El pellet de bacterias obtenido se resuspendió en 30ml por litro de cultivo de buffer de lisis (50mM Tris pH 7.6, 300mM NaCl, 1mM DTT, 1mM EDTA, 5% glicerol)

con inhibidores de proteasas (Complete Protease Inhibitor Cocktail Tablets de Roche), se realizó la lisis por sonicación y se centrifugó a 12000g por 45min. Para realizar la purificación por afinidad el sobrenadante obtenido se incubó con 300µl por litro de cultivo de GST Sepharose Fast Flow o se pasó a través de una GST trap (1ml o 5ml) previamente equilibradas. A continuación la matriz fue lavada con 10 volúmenes de columna (CV) de buffer de lisis y luego se equilibró con buffer de corte (50mM Tris pH 7.6, 150mM NaCl, 1mM DTT, 1mM EDTA). La escisión de la proteína de fusión se realizó sobre la matriz por lo que una vez equilibrada se adicionaron 10U de proteasa PreScission (Amersham) o de trombina (Amersham) por miligramo de proteína según corresponda y se incubó durante toda la noche a 4°C, el producto escindido se recuperó y se concentró hasta llegar a un volumen adecuado para cargar en las columnas de exclusión molecular Superdex200 (500µl para la 10/300 o 2ml para la 16/60).

En caso de ser necesaria una mayor pureza se adicionó un paso de purificación con una columna de intercambio iónico o monoQ realizada con los siguientes buffers: buffer A (50mM Tris pH 7.6, 50mM NaCl, 1mM DTT, 1mM EDTA) y Buffer B (50mM Tris pH 7.6, 50mM NaCl, 1mM DTT, 1mM EDTA) con un gradiente de elución al 40% en 20CV. Por otro lado, en los casos en que las proteínas o complejos purificados eran para realizar ensayos de cristalización en la última columna de exclusión molecular se intercambié el buffer por el que sería utilizado en la producción de gotas.

Para optimizar la purificación de Rpn10 siguiendo el mismo protocolo pero modificando los buffers se evaluaron tres condiciones diferentes: alta sal, con detergente Tween20 y en presencia de aminoácidos estabilizadores [Golovanov *et al.*, 2004]. La composición de dichos buffers se muestra en la siguiente tabla:

Tabla 1.6: Buffers Rpn10					
Alta sal		Tween 20		Aminoácidos	
Buffer de Lisis					
Tris	50mM	Tris	50mM	Tris	50mM
NaCl	1,5M	NaCl	500mM	NaCl	500mM
EDTA	1mM	EDTA	1mM	EDTA	1mM
Glicerol	5%	Tween20	0,01%	L-arginina	50mM
pH	7.6	pH	7.6	L-acido glutámico	50mM
				pH	7.6
Buffer de Corte					
Tris	50mM	Tris	50mM	Tris	50mM
NaCl	1M	NaCl	300mM	NaCl	300mM
CaCl ₂	2,5mM	CaCl ₂	2,5mM	CaCl ₂	2,5mM
Glicerol	5%	Tween20	0,01%	L-arginina	50mM
pH	7.6	pH	7.6	L-acido glutámico	50mM
				pH	7.6

Para generar muchos de los complejos se combinaron diferentes clones mono o bicistrónicos, cada población se creció de manera independiente y juntaron para ser co-lisadas por sonicación y siguiendo el mismo protocolo a partir de este punto.

Ensayos cristalográficos

Todos los ensayos cristalográficos se llevaron a cabo en la Plataforma Automatizada de Cristalografía (PAC) del Parc Científic de Barcelona con reactivos de Hampton Research y Molecular Dimensions.

Los *screenings* iniciales se realizaron en placas MRC de 96 pocillos (Swissci AG). Las condiciones de cristalización fueron preparadas por un robot Freedom Evo (Tecan) y las gotas de 200nl (100nl de proteína + 100nl de buffer) fueron dispensadas utilizando los robots Cartesian (Genomics Solutions) y Pheonix (Art Robbins). Las optimizaciones se realizaron a mano en placas Linbro de 24 pocillos (Hampton Research). Todos los ensayos se hicieron con el método de difusión de vapor en gota sentada.

El seguimiento de las gotas se efectuó mediante imágenes tomadas por un robot Crystal Farm (Bruker AXS) para las placas de 96 pocillos y mediante observación por microscopio óptico SMZ1000 (Nikon) para las de 24 pocillos.

Para discriminar cristales de sal de cristales proteicos se utilizó el Crystalmation (Rigaku) que permite tomar fotografías de las gotas con luz UV.

Las proteínas y complejos purificados, con las condiciones iniciales y los screenings a los que fueron sometidos se encuentran detallados en la siguiente tabla:

Proteína/complejo	Concen-tración	Temp	Buffer inicial	Screen
Rpn5	7.7mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 2, 3, 8, 11, 15, 16, 17, 21, 22 y 23
	7.8mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1 y 4
	10mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1 y 4
			α -quimotripsina 1/5000 p/p 10mM Tris, 50mM NaCl pH	PAC 1 y 4

Materiales y métodos

			7.6 Papaina 1/5000 p/p	
Rpn9	8mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 4, 11, 21 y 22
Rpn12	7.4mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 2, 3, 5, 6, 8, 12, 13, 15, 16, 17, 18, 19, 20 y 23
	5mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 4, 11, 21 y 22
Rpn5/Rpn9/Rpn8		4°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 2, 3, 4, 5, 21, 22 y 23
	9.4mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 5, 7, 8, 9, 11, 12, 13, 14, 21, 22 y 23 Midas Morpheus
	9mg/ml	20°C	20mM Tris, 150mM NaCl pH7.6	PAC 1, 2, 3, 4, 15, 16 y 17
	15mg/ml	20°C	20mM Tris, 150mM NaCl pH7.6 Purificado por Ii	PAC 1, 2, 3, 4, 15, 16 y 17
	9.9mg/ml	4°C	20mM Citrato de Na, 50mM NaCl pH 6	PAC 1, 4, 5, 9, 10, 11 y 14 Clear Strategy I y II (pH 5.5 y 6.5) Midas
	9.8mg/ml	20°C	20mM Citrato de Na, 50mM NaCl pH 6 α -quimotripsina 1/5000 p/p	PAC 1, 4, 5, 11 y 23
				20mM Citrato de Na, 50mM NaCl pH 6 Papaina 1/5000 p/p
Rpn5 _{169-f} /Rpn9 _{160-f} / Rpn8	11.7mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 2, 21, 23 y Morpheus
Rpn5 _{169-f} /Rpn9 _{160-f} / Rpn8/Rpn6	11.8mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 2, 3, 4, 5, 15, 16, 17, 21, 22, 23 y Morpheus
Rpn5 _{169-f} /Rpn9 _{160-f}	8.8mg/ml	20°C	10mM Tris,	PAC 1 y 2

Rpn8/Rpn6 _{110-f}			50mM NaCl pH 7.6	
Rpn5 _{169-f} /Rpn9 _{160-f} /Rpn8/Rpn11	11mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 6, 8, 14 y 20
	7mg/ml	20°C	50mM Tris, 150mM NaCl pH7.6	PAC 1, 2, 3, 4, 15, 16 y 17
Rpn5/Rpn9/Rpn8/Rpn11/Rpn6	15mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 2, 3 y 4
			α-quimotripsina 1/5000 p/p	
Rpn5 _{169-f} /Rpn9 _{160-f} /Rpn8/Rpn11/Rpn6	15mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 2, 3 y 4
			α-quimotripsina 1/5000 p/p	
	10 y 20mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 2, 3, 4, 15, 16 y 17
Rpn5 _{169-f} /Rpn9 _{160-f} /Rpn8/Rpn11/Rpn6 _{110-f}	7.5mg/ml	20°C	10mM Tris, 50mM NaCl pH 7.6	PAC 1, 2, 3, 4, 15, 16, 17, 21, 22 y 23

Los ensayos con la metodología de proteólisis *in situ* se realizaron con el kit Floppy-Choppy de Jena Bioscience siguiendo las recomendaciones del fabricante. Por otro lado, los screen de la plataforma están organizados según el siguiente código:

Código	Screen	# de condiciones	Referencia
PAC1	Crystal Screen I	48	Hampton Research
	Crystal Screen II	48	
PAC2	Wizard I	48	Emerald Bio
	Wizard II	48	
PAC3	Index	96	Hampton Research
PAC4	Salt RX	96	Hampton Research
PAC5	A/S Ion Screen	48	Hampton Research
	Ammonium sulphate	24	
	Quick phosphate	24	
PAC6	PEG 6000	24	Hampton Research

Materiales y métodos

	PEG 6000/LiCl	24	
	PEG 400	24	
	PEG 4000/LiCl	24	
PAC7	Crystal Screen Lite	48	Hampton Research
	Crystal Screen Cryo	48	
PAC8	PEG-Ion	48	Hampton Research
	MPD	24	
	NaCl	24	
PAC9	Natrix	48	Hampton Research
	Complex screen	40	
PAC10	Protein-dna Screen	96	Diseñado por la PAC
PAC11	Memfac	48	Hampton Research
	Memplus	48	Molecular Dimensions
PAC12	MemStart	48	Molecular Dimensions
	MemSys	48	
PAC13	Memgold	96	Molecular Dimensions
PAC14	Wizard Cryo 1	48	Emerald Bio
	Wizard Cryo 2	48	
PAC15	JBS1	24	Jena Bioscience
	JBS2	24	
	JBS3	24	
	JBS4	24	
PAC16	JBS5	24	Jena Bioscience
	JBS6	24	
	JBS7	24	
	JBS8	24	
PAC17	JBS9	24	Jena Bioscience
	JBS10	24	Molecular Dimensions
	Clear Strategy I - 8.5	24	
	Clear Strategy II - 8.5	24	
PAC18	Clear Strategy I - 4.5	24	Molecular Dimensions
	Clear Strategy I - 5.5	24	
	Clear Strategy II - 4.5	24	
	Clear Strategy II - 5.5	24	
PAC19	Clear Strategy I - 6.5	24	Molecular Dimensions
	Clear Strategy I - 7.5	24	
	Clear Strategy II - 6.5	24	
	Clear Strategy II - 7.5	24	
PAC20	Citrat Screen	24	Diseñado por la PAC
	Mc Pherson Screen	24	Molecular Dimensions
	Clear Strategy I	24	
	Clear Strategy II	24	
PAC21	PACT premier HT-96	96	Molecular Dimensions
PAC22	Pi-PEG Screen	96	Jena Bioscience
PAC23	Pi - minimal Screen	96	Jena Bioscience
	Midas	96	Molecular Dimensions
	Morpheus	96	Molecular Dimensions

DLS

Con el fin de testear diversos buffers para Rpn5, Rpn12 y para el complejo Rpn5/Rpn9/Rpn8, las proteína y el trímero a una concentración

aproximada de 1mg/ml se dializaron con una batería de buffers durante toda la noche [Proteau *et al.*, 2010], la composición de los mismos se detalla en la siguiente tabla:

Tabla 1.9: Buffers DLS	
Buffer #	Composición
Buffer 0	150mM NaCl, 50mM Tris pH 7.6
Buffer 1	50mM NaCl, 10mM Tris pH 7.6
Buffer 2	50mM NaCl, 100mM Citrato de Na pH 5
Buffer 3	50mM NaCl, 100mM Citrato de Na pH 6
Buffer 4	50mM NaCl, 100mM PIPES pH 6.5
Buffer 5	50mM NaCl, 100mM Fosfato de Na pH 7
Buffer 6	50mM NaCl, 100mM HEPES pH 7.5
Buffer 7	50mM NaCl, 100mM HEPES pH 8
Buffer 8	50mM NaCl, 100mM Tris pH 8.5
Buffer 9	50mM NaCl, 100mM Bicina pH 9

Al día siguiente cada muestra se colocó en cubetas de cuarzo y se analizó con el DLS Zetasizer nanoseries (Malvern). Se realizó una medición estándar en la que se realizan tres series de diez mediciones cada una y se promedian los resultados.

1.2.3 Lid endógena

Expresión y purificación de la lid endógena

Para obtener la lid endógena de *S. cerevisiae* se siguió el protocolo descrito por [Leggett *et al.*, 2005]. Se utilizó una cepa de *S. cerevisiae* que presenta un *tag* de proteína A y un sitio de corte para la proteasa TEV en el extremo C-terminal de Rpn11, cedida amablemente por el Dr. Bernat Crosas. A partir de una placa con medio YPD (1,5% extracto de levadura, 2% peptona y 2% glucosa) y 1,5% agar, se lanzó un inóculo de 40ml de medio YPD líquido con antibióticos (ampicilina y tetraciclina a una concentración final de 50mg/l y 10mg/l respectivamente) que se creció durante toda la noche a 30°C y 220rpm, al día siguiente con el inóculo densamente crecido se realizó un precultivo de 2l y se volvió a incubar durante toda la noche. Este precultivo se escaló hasta 50l de cultivo en el servicio de fermentación de la Universidad de Barcelona, donde se fermentó durante 30hs totales, a las 24hs se agregó glucosa adicional a fin de obtener mayor biomasa. Las células se recuperaron mediante centrifugación a 4°C, 3000g por 15min, para conservarlas se resuspendieron en buffer de

lisis y congelación (50mM Tris pH 7.5, 100mM NaCl, 10mM MgCl₂, 0,18mM DDM, 1mM DTT e inhibidores de proteasas Complete de Roche) en una relación 350ml/kg, se congelaron en nitrógeno líquido y se almacenaron a -80°C.

Para lisar las levaduras se utilizaron Bead Beaters (Biospec Products) con contenedores de metal, en los que se colocaron 200ml de la suspensión de células previamente descongeladas, 200ml de bolitas de vidrio (0,5mm de diámetro), DNAsas (Roche) y RNAsas (Roche) a una concentración final de 0,4µg/ml o 1U/ml y 5µg/ml respectivamente y se procesaron durante una hora en ciclos de 30seg encendido, 90seg apagado. El lisado se centrifugó a 50.000g durante dos horas y se filtró mediante gasas y papel de filtro. El lisado clarificado fue llevado a pH neutro con Tris 1M y se incubó durante toda la noche con IgG sepharose 6 fast flow (GE Healthcare), que al día siguiente fueron lavadas con 20CV de buffer de lisis y congelación, seguido de 30CV de buffer con alta concentración de sal (50mM Tris pH 7.5, 1,5mM NaCl, 10mM MgCl₂, 0,18mM DDM, 1mM DTT e inhibidores de proteasas Complete de Roche) que separa la lid de la base y la CP, finalmente se realizó un último lavado con buffer A (50mM Tris pH 7.6, 50mM NaCl, 1mM DTT). A fin de eliminar el *tag* de proteína A se adicionaron 1000U de proteasa TEV que se dejaron actuar durante toda la noche a 4°C. Al día siguiente se recuperó el complejo escindido y se cargó en una columna monoQ usando los buffers A y B (50mM Tris pH 7.6, 50mM NaCl, 1mM DTT) en un gradiente del 40% en 20CV. Las fracciones del pico obtenido se concentraron y se pasaron por una columna de Superosa 6 10/300 con el siguiente buffer: 50mM Tris pH 7.6, 150mM NaCl, 1mM DTT.

Producción de anticuerpos monoclonales anti-Rpn11 y anti-Rpn8

Los anticuerpos monoclonales (AcMo) anti-Rpn11 y anti-Rpn8 fueron solicitados a la Unidad de Producción de Anticuerpos (CAbs) del CIBER-BBN/CSI. Para su obtención se utilizaron los siguientes péptidos.

- Rpn11: DPKKHLSETADETL

10	20	30	40	50	60	70
MERLQRLMMN	SKVGSADTGR	DDTKETVYIS	SIALLKMLKH	GRAGVPMEM	GLMLGEFVDD	YTVNVVDVFA
80	90	100	110	120	130	140
MPQSGTGVSV	EAVDDVFQAK	MMDMLKQTGR	DQMVVGWYHS	HPGFGCWLSS	VDVNTQKSFE	QLNSRAVAVV
150	160	170	180	190	200	210
VDPIQSVK GK	VVIDAFRLID	TGALINNLEP	RQTTSENTGLL	NKANIQALIH	GLNRHYYSLN	IDYHKTAKET
220	230	240	250	260	270	280
KMLMNLHKEQ	WQSGLKMYDY	EEKEESNLAA	TKSMVKIAEQ	YSKRIEEKE	LTEEELKTRY	VGRQDPKKHL
290	300	307				
<u>SETADETL</u> EN	NIVSVLTAGV	NSVAIK*				

- Rpn8: DSESESGDKEATAP

10	20	30	40	50	60	70
MSLQHEKVTI	APLVLLSALD	HYERTQTKEN	KRCVGVILGD	ANSSTIRVTN	SFALPFEEDE	KNSDVWFLDH
80	90	100	110	120	130	140
NYIENMNEC	KKINAKEKLI	GWYHSGPKLR	ASDLKINELF	KKYTQNNPLL	LIVDVKQQGV	GLPTDAYVAI
150	160	170	180	190	200	210
EQVKDDGTST	EKTFLHLPCT	IEABEABEIG	VEHLLRDVRD	QAAGGLSIRL	TNQLKSLKGL	QSKLKDVVEY
220	230	240	250	260	270	280
LDKVINKEP	INHITLGLKQ	DVFNLLPNLG	TPDDDEIDVE	NHDRINISNN	LQKALTVKTN	DELMVIYISN
290	300	310	320	330	339	
LVRSIIAFDD	LIENKIQNKK	IQEQRVKDKQ	SKVSD	<u>DSESE</u>	<u>SGDKEATAP</u>	IQRKNKKN*

La avidéz y especificidad de los anticuerpos monoclonales fue evaluada por western blot.

Interacción de la lid endógena con los AcMo anti-Rpn11 y anti-Rpn8

Se combinaron 300µg de AcMo con 200µg de lid endógena durante 2hs a temperatura ambiente en buffer 50mM Tris pH 7.6, 150mM NaCl y 1mM EDTA y DTT. Finalizado el tiempo de incubación se pasaron las muestras a través de una columna de Superosa 6 y las fracciones se evaluaron por SDS-PAGE.

1.2.4 Microscopía electrónica:

Los complejos Rpn5/Rpn9/Rpn8 (o "trímero"), Rpn5/Rpn9/Rpn8/Rpn6 (o "tetramero") y Rpn5/Rpn9/Rpn8/Rpn11/Rpn6 (o "pentámero") fueron analizados por microscopía electrónica de transmisión con tinción negativa por nuestros colaboradores del departamento de fisiología y biofísica de la Universidad Commonwealth de Virginia, Richmond en Estados Unidos: Tyler W. Steele y la Prof. Montserrat Samsó.

La tinción negativa de las muestras se llevó a cabo sobre rejillas de 400 de malla recubiertas de carbono y las muestras en un rango de concentración entre 0,02 y 0,05mg/ml y se fijaron mediante su incubación con 0,75% formato de uranilo. Se utilizó un microscopio Tecnai F20 G2 funcionando a 120kV bajo condiciones de baja dosis y las imágenes se capturaron a 80.000x en una cámara Gatan 4k x 4k CCD, utilizando para la lid endógena un valor de desenfoque de 0.6-1.6µm y 2.5-4µm para los subcomplejos.

Para el procesamiento de las imágenes se utilizaron los programas SPIDER [Frank *et al.*, 1996] y EMAN2 [Ludtke, 2010]. El análisis 2D se realizó mediante un proceso de alineación de referencia libre, seguido de un proceso iterativo de análisis estadístico multivariante, clasificación y alineación con referencias múltiples [Samsó *et al.*, 1998] en SPIDER; y las reconstrucciones 3D se construyeron usando un algoritmo de líneas

comunes seguido por el refinamiento iterativo utilizando EMAN2. Se confirmó para todos los complejos una alta correspondencia entre las proyecciones 2D de las reconstrucciones 3D y orientaciones 2D promedio, que es un indicador robusto de la validez de la reconstrucción 3D. El umbral del volumen de las reconstrucciones 3D se calculó con la masa molecular de cada complejo y según el valor promedio de densidad proteica ($0.87\text{Da}/\text{\AA}^3$) [Quillin *et al.*, 2000], mientras que para calcular la resolución se dividió cada conjunto de datos en dos (partículas pares e impares) para calcular la curva de la correlación por capas en el espacio de Fourier (*Fourier Shell Correlation*), las reconstrucciones finales se exponen e interpretan con la resolución límite obtenida utilizando un *cutoff* del 0.5.

Las galerías de imágenes 2D se generaron con el programa WEB [Frank *et al.*, 1996] y las representaciones de la superficie de los modelos 3D con el programa UCSF Chimera [Pettersen *et al.*, 2004]. La segmentación de la lid endógena se realizó con el algoritmo Segger incluido en Chimera [Pintilie *et al.*, 2010].

1.3 Resultados y discusión

3.1.1 Proteínas recombinantes

Rpn1 y Rpn2

Las proteínas Rpn1 *full-length* y Rpn2₁₋₈₅₅ a la que le faltaban los últimos 90 aminoácidos del extremo C-terminal por un error de clonado original, se expresaron correctamente y con el peso molecular esperado, sin embargo al escindir la GST ambas proteínas precipitaron y la pequeña fracción que permaneció soluble al ser analizada por cromatografía de exclusión molecular eluía en el volumen muerto de la columna, indicando que la proteína se encontraba agregada.

Se decidió entonces, intentar expresar y purificar versiones truncadas de ambas proteínas, que contuviesen el dominio PC, eliminando los extremos flexibles. Se diseñaron los siguientes constructos: Rpn1₄₃₈₋₈₇₆ (Rpn1^{PC}) y Rpn2₃₅₀₋₇₃₀ (Rpn2^{PC}) siguiendo la estrategia utilizada por Effantin *et al.*, 2009, lamentablemente tampoco de este modo fue posible obtener proteínas solubles ya que presentaron un comportamiento similar al de las proteínas *full length*.

Al comienzo de este proyecto en la literatura se proponía que Rpn1 y Rpn2 interaccionaban formando un heterodímero, en el que las regiones PC de forma toroidal se encontraban apiladas axialmente [Rosenzweig *et al.*, 2008; Effantin *et al.*, 2009], por este motivo se decidió co-expresar Rpn1 y Rpn2 en el contexto del vector pGEX-6P-2RBS. La co-expresión, ya sea de las proteínas *full length*, de los truncados o la combinación de ambas no resolvió el conflicto de la insolubilidad de estas proteínas ni tampoco demostró la interacción entre ellas. Tiempo después [Lander *et al.*, 2012] demostró mediante EM y marcando las subunidades con anticuerpos la posición de ambas proteínas en lados opuestos de la RP.

Rpn10 y la interacción base-lid

Inicialmente se expresó y purificó Rpn10 siguiendo el protocolo estándar (descrito en 2.1.1) resultando en una buena expresión y cantidad de proteína recuperada en la purificación por afinidad pero al escindir la GST en la cromatografía de exclusión molecular una parte importante de la muestra salía en el volumen muerto de la columna indicando su estado oligomérico y presentaba múltiples contaminantes. A fin de encontrar un protocolo más propicio para la proteína se evaluó su purificación utilizando 3 condiciones diferentes: alta sal, con detergente tween20 y en presencia de aminoácidos estabilizadores. La purificación con alta sal fue con la que se obtuvieron los mejores resultados: buena cantidad de proteína monomérica y la presencia de un solo contaminante que pudo ser eliminado por

intercambio iónico, resultando en Rpn10 con un 95% de pureza al ser analizada por SDS-PAGE (Fig. 1.20). Desafortunadamente cuando el protocolo de purificación fue puesto a punto y me disponía a realizar ensayos cristalográficos la estructura de Rpn10 fue publicada por Riedinger *et al.*, 2010, por lo que se decidió no continuar con esta línea de investigación.

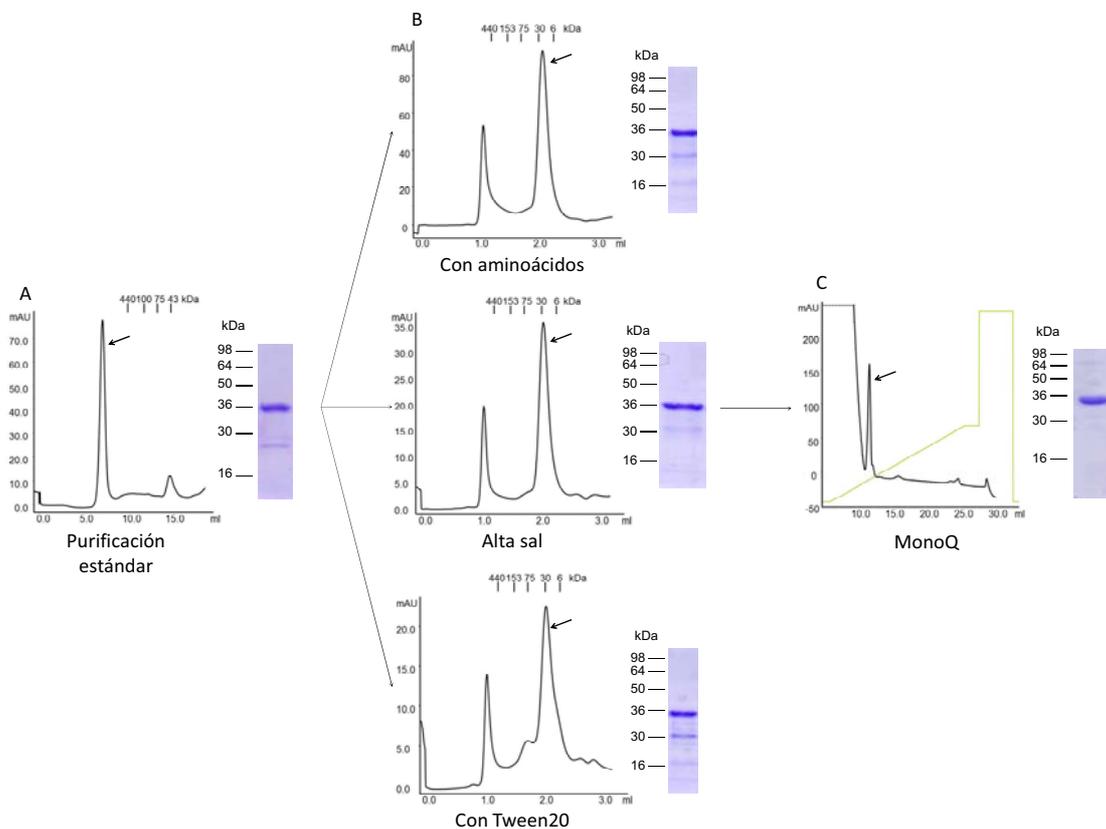


Fig. 1.20 Purificación de Rpn10. A) Superdex200 10/300 de Rpn10 purificada en condición estándar, se observa el pico por fuera del rango resolutive de la columna indicando que la proteína esta agregada. B) Superdex200 5/150 de Rpn10 purificada en presencia de aminoácidos, alta sal y tween20, se observa que en las 3 condiciones Rpn10 eluye con el peso molecular esperado siendo el mejor resultado el obtenido en alta sal ya que presenta menor cantidad de contaminantes. C) MonoQ de Rpn10 purificada con alta sal, con este paso extra de purificación se logran eliminar los contaminantes resultando en una muestra con un 95% de pureza. A la derecha de cada perfil de elución se muestra el resultado del SDS-PAGE de los picos señalados con una flecha.

Por otro lado, en la literatura del momento se hipotetizaba que Rpn10 funcionaba como bisagra conectando y estabilizando la unión entre la lid y la base, ya que en mutantes Δ Rpn10 la lid se disociaba del conjunto base-CP [Glickman *et al.*, 1998; Gallastegui *et al.*, 2010]. Por esta razón se decidió indagar en las posibles interacciones entre Rpn10 y las proteínas de la lid, inicialmente se realizó un *in vitro binding* en el que se inmovilizaron todas las subunidades de la lid mediante la GST a la matriz de purificación y

se incubaron con Rpn10 purificada, este experimento dio como resultado que Rpn10 interaccionaba con todas ellas en mayor o menor medida incluyendo la GST sola, indicando la naturaleza "pegajosa" de Rpn10 (Fig. 1.21).

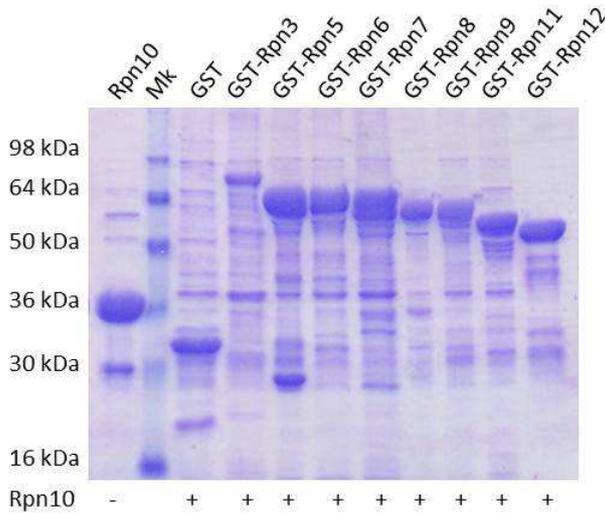


Fig. 1.21 *In vitro binding* de Rpn10. En la primer calle se observa a Rpn10, seguido del marcador de peso molecular (Mk) y de las *beads* de glutatión con las subunidades de la lid unidas a través de la GST que se incubaron con Rpn10 a fin de observar la afinidad entre ellas. Rpn10 interacciona con todas ellas incluyendo el control negativo (GST sola)

A pesar de la interacción inespecífica de Rpn10 con GST, se decidió analizar la co-expresión bicistrónica de Rpn3, Rpn5, Rpn9, Rpn11 y Rpn12 con Rpn10 a fin de verificar si la interacción observada en el *in vitro binding* al escindir la GST se mantenía. Mediante co-expresión y previamente a escindir la GST se observa que en todos los casos Rpn10 interactúa en mayor o menor medida, sin embargo cuando se escinde la GST y se evalúa el complejo por exclusión molecular solo en el caso de Rpn3 la interacción se mantiene, mientras que para Rpn9, Rpn11 y Rpn12 la asociación con Rpn10 no es detectable en SDS-PAGE, al analizar las fracciones de la elución de Rpn9/Rpn10 y Rpn12/Rpn10 mediante western blot con un anticuerpo anti-Rpn10 se observa que estos complejos mantienen una relación sub-estequiométrica (Fig. 1.22). Rpn5/Rpn10 no fue evaluada tras el corte de la GST.

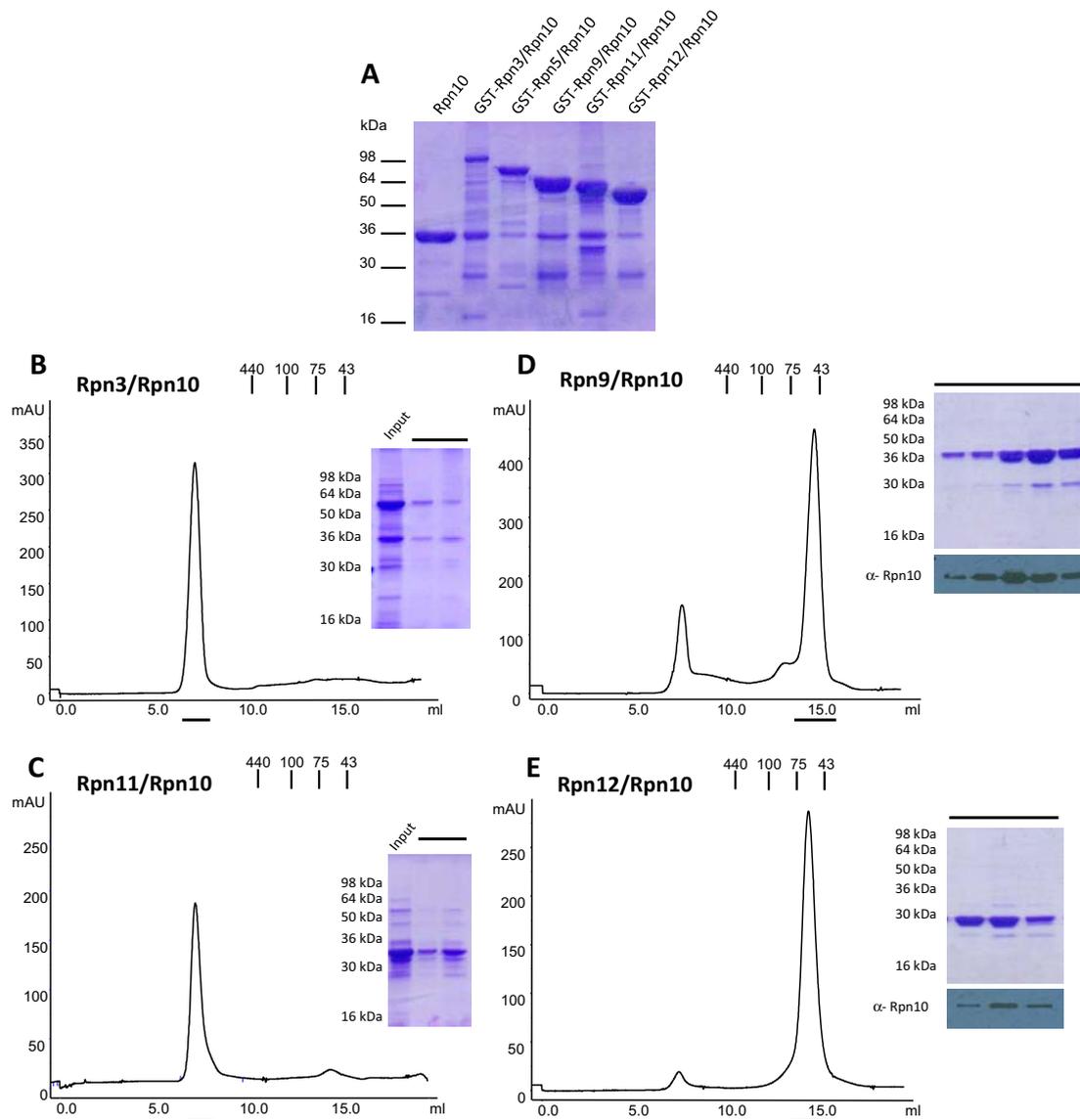


Fig. 1.22 Co-expresión de RpnX (3, 9, 11 y 12)/Rpn10. A) SDS-PAGE de la expresión de los clones bicistrónicos aún fusionados a GST. B y C) Superdex200 10/300 y SDS-PAGE de Rpn3/Rpn10 y Rpn11/Rpn10, los picos salen en el volumen muerto indicando que están agregadas, Rpn11 eluye sola sin Rpn10 mientras que para Rpn3/Rpn10 se observa interacción. D y E) Superdex200 10/300 y SDS-PAGE de Rpn9/Rpn10 y Rpn12/Rpn10, ambos eluyen dentro del rango de resolución de la columna, al ser analizados por SDS-PAGE no se detecta la presencia de Rpn10 que solo puede ser observada mediante western blot indicando que existe una interacción subestequiométrica.

Tomando en cuenta estos resultados no podemos inferir que las asociaciones que establece Rpn10 con otras subunidades de la lid sean específicas, ya que el comportamiento promiscuo de Rpn10 no permite diferenciar a ciencia cierta que interacciones son reales y cuales son artefactos. Si bien Rpn3/Rpn10 parece que mantienen una relación estequiométrica luego del corte de la GST la relación entre estas proteínas no es clara, ya que en las estructuras de la lid se observa que Rpn10 se

encuentra ubicada del lado opuesto a Rpn3, dándonos a entender que esta interacción podría ser un artefacto.

Proteínas de la lid

A fin de caracterizar la lid, encontrar proteínas o complejos adecuados para la cristalización y analizar las interacciones intermoleculares que mantienen este complejo se decidió evaluar las subunidades de manera individual así también como combinaciones de 2 o más proteínas co-expresandolas o co-lisandolas. A su vez se analizaron fragmentos de algunas de ellas de modo de obtener información relevante sobre la participación de dominios discretos en el ensamblaje de la lid.

Proteínas individuales

Inicialmente Fabio Sessa quien arrancó el proyecto antes de mi llegada al laboratorio, expresó las subunidades individualmente siguiendo el protocolo descrito en el apartado 1.2.2, los resultados obtenidos fueron los siguientes (Fig. 1.23):

- Rpn3: completamente insoluble
- Rpn5: soluble, a pesar de que una fracción precipita en las *beads* al escindir la GST, en la Superdex200 eluye en un pico simétrico con el peso molecular esperado y al ser analizada por SDS-PAGE se observa una pureza del 95%.
- Rpn6: soluble, al igual que Rpn5 una pequeña fracción de proteína escindida queda en las *beads*, en la Superdex200 y en su respectivo gel podemos observar que presenta el peso molecular esperado aunque la proteína está acompañada de un contaminante de aproximadamente 30kDa.
- Rpn7: insoluble, si bien una pequeña fracción permanece soluble y puede ser recuperada mediante la purificación por afinidad, al escindir la GST Rpn7 precipita dejando solo una cantidad mínima de proteína soluble insuficiente para realizar su análisis mediante exclusión molecular.
- Rpn8: soluble, aunque gran parte de la muestra precipita al eliminar el *tag* de GST, al ser analizada por Superdex200 se observa como la proteína eluye a todo lo largo del perfil indicando que la misma se encuentra en diversos estados oligoméricos, además de presentar varios contaminantes.
- Rpn9: soluble, se puede observar una doble banda que indica algún tipo de degradación, eluye de la columna en un pico simétrico con el peso molecular esperado y con una pureza del 95%.

- Rpn11: insoluble, una pequeña fracción puede ser recuperada por afinidad pero al escindir el *tag* de GST precipita resultando en una cantidad insuficiente para su análisis.
- Rpn12: soluble, luego del corte de la GST prácticamente toda la muestra permanece soluble, al ser analizada por Superdex200 eluye en un pico simétrico en el peso molecular correspondiente a un dímero, aunque presenta una triple banda indicando una posible degradación de la proteína.

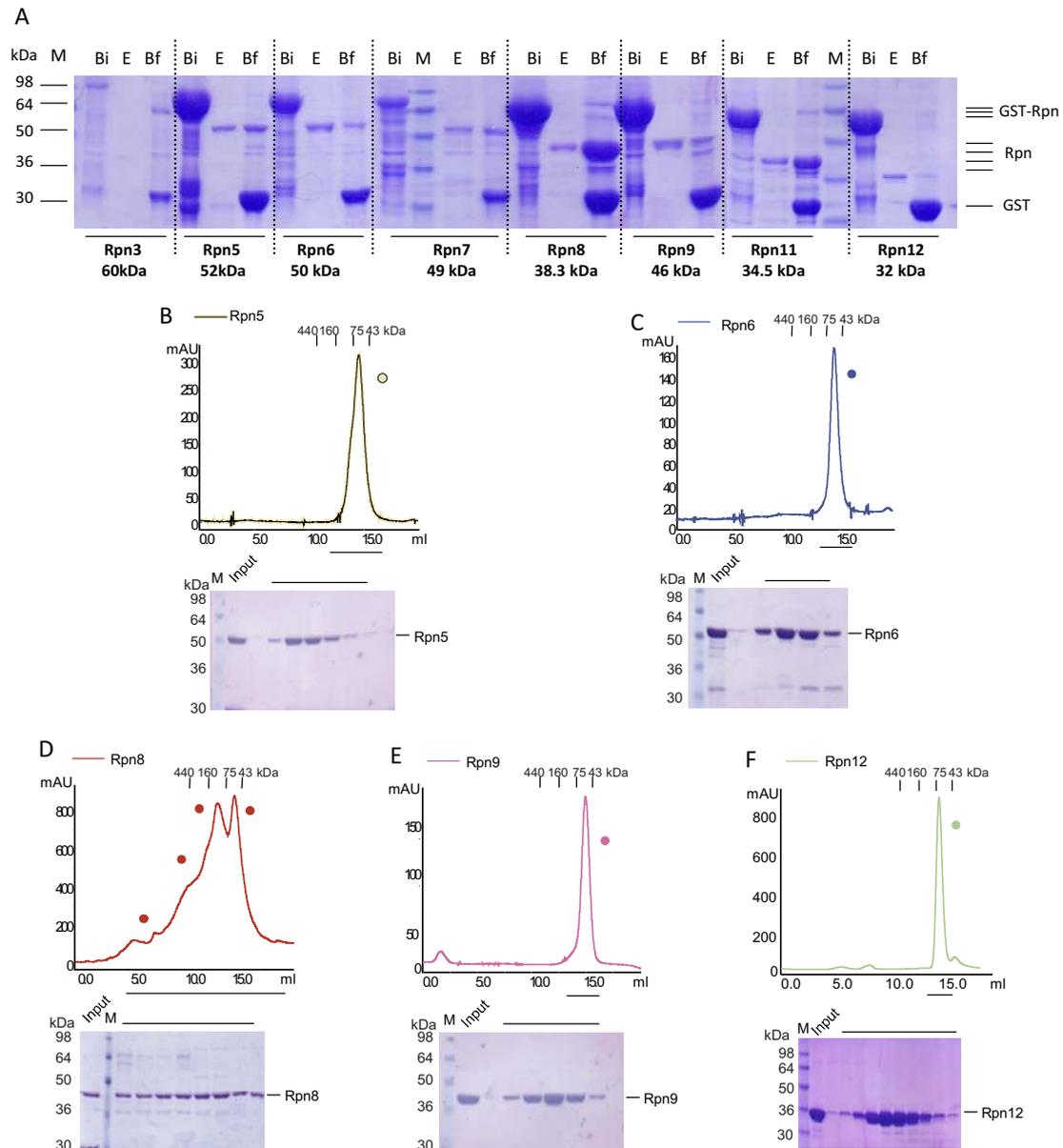


Fig. 1.23 Expresión y purificación de las proteínas Rpn individualmente. A) SDS-PAGE tras la escisión de la GST con las muestras: M marcador de peso molecular, Bi *beads* iniciales, E eluido, Bf *beads* finales para cada una de las subunidades de la lid. B-F) Superdex200 10/300 y SDS-PAGE de las proteínas B) Rpn5 C) Rpn6 D) Rpn8 E) Rpn9 y F) Rpn12.

Teniendo en cuenta estos resultados se seleccionaron los candidatos más interesantes para proseguir con la optimización del protocolo de purificación a fin de testarlos en ensayos cristalográficos. Los posibles candidatos por sus buenos resultados iniciales fueron Rpn5, Rpn6, Rpn9 y Rpn12. Sin embargo, como la estructura homóloga de Rpn6 de *Drosophila melanogaster* fue descrita [Pathare *et al.*, 2012] se descartó del conjunto.

Se realizaron ensayos cristalográficos para las proteínas **Rpn5** y **Rpn9** (ver tabla 1.7: Ensayos cristalográficos) siguiendo el protocolo estándar sin obtener resultados positivos para ninguna de las dos. Como Rpn9 presenta algún tipo de degradación se decidió continuar solo con Rpn5, con la que se realizaron ensayos cristalográficos utilizando la técnica de proteólisis *in situ* sin obtener cristales. Además se buscó un buffer inicial que diera mayor estabilidad a la proteína mediante DLS siguiendo el protocolo detallado en el apartado 1.2.2, si bien los resultados obtenidos por este método sugerían que el buffer 2 (50mM NaCl, 100mM Citrato de Na pH 5) confería una mejor monodispersión de la proteína en solución al intentar purificar Rpn5 en esta condición la proteína precipitó. Probablemente esta diferencia de comportamiento se deba a que este buffer solo aporta estabilidad a bajas concentraciones y no a la alta concentración necesaria para realizar los ensayos.

En el caso de **Rpn12** se realizaron ensayos cristalográficos iniciales que resultaron infructuosos (ver tabla 1.7: Ensayos cristalográficos), mientras que por otro lado se intentó optimizar el protocolo de purificación adicionando un paso de intercambio iónico con una columna MonoQ, a fin de separar la triple banda obtenida tras la purificación por afinidad, sin embargo el resultado obtenido no fue satisfactorio ya que la columna no discriminó las diferentes especies de Rpn12 en la solución (Fig. 1.24A). También se analizó la proteína por DLS y se seleccionó el buffer más prometedor, lamentablemente la estructura de Rpn12 de *S. pombe* fue publicada [Boehringer *et al.*, 2012]; la misma comparte una identidad de secuencia del 34% con nuestra proteína de interés (Fig. 1.24B) por lo que se decidió no continuar con su estudio. El estado dimérico y la problemática de la degradación de esta proteína también fue observada en su homóloga donde el extremo C-terminal era propenso a degradación, finalmente la estructura del dímero fue resuelta utilizando versiones truncadas.

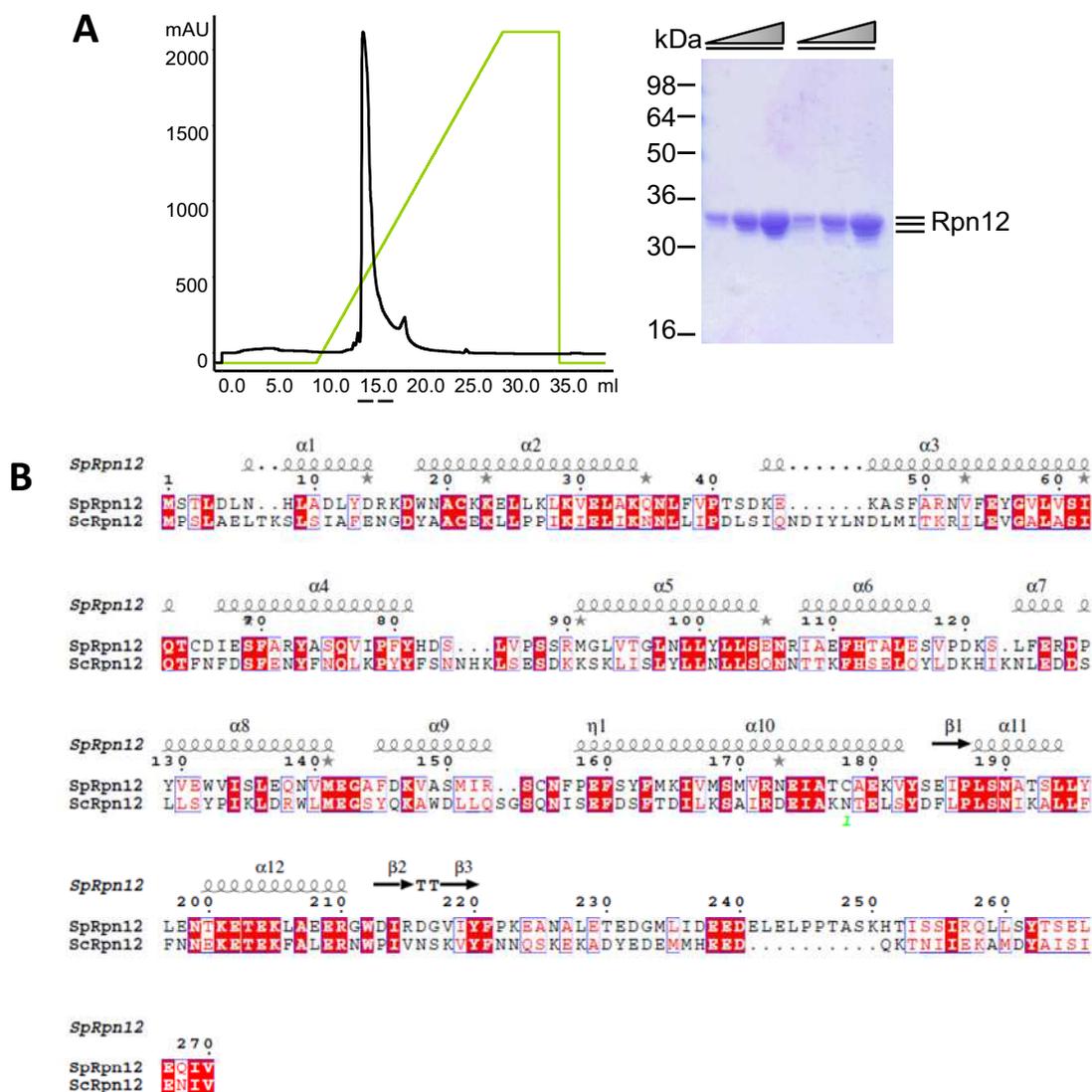


Fig. 1.24 Rpn12. A) Perfil de elución de la MonoQ y análisis por SDS-PAGE de las fracciones del pico en concentraciones crecientes. B) Alineamiento de las secuencias y estructuras secundarias de Rpn12 de *S. cerevisiae* (Sc) y *S. pombe* (Sp).

Complejos binarios

Tomando en consideración los resultados obtenidos por Fu *et al.*, 2001. Se clonaron y expresaron los siguientes complejos bicistrónicos: Rpn3/Rpn7, Rpn5/Rpn9, Rpn5/Rpn8 y Rpn11/Rpn8, mediante su expresión en el vector pGEX-6P-2RBS. Los complejos binarios Rpn3/Rpn7 y Rpn11/Rpn8 fueron analizados por el Dr. Fabio Sessa, se añaden sus resultados en esta tesis a fin de comprender el panorama completo de las interrelaciones establecidas entre las proteínas de la lid.

Se comprobó que el complejo **Rpn3/Rpn7** interactúa en una relación estequiométrica de 1:1, sin embargo el problema de insolubilidad y agregación que presentaban las proteínas de manera individual se mantiene. Es decir que gran parte de la muestra precipita luego del corte de

la GST y la poca cantidad que se obtiene soluble al ser analizada mediante la columna de exclusión molecular eluye en el volumen muerto de la misma evidenciando que el complejo está agregado (Fig. 1.25).

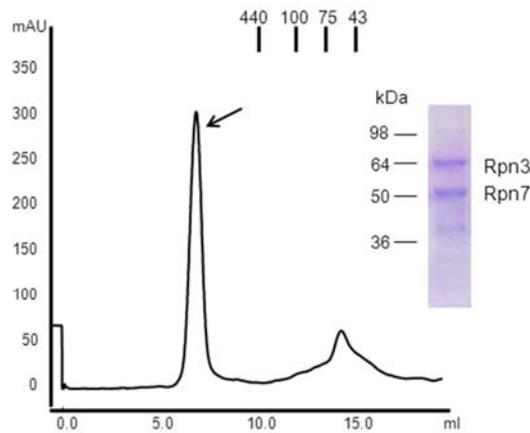


Fig. 1.25 Rpn3/Rpn7. Perfil de elución del complejo Rpn3/Rpn7 en la columna Superdex200 y SDS-PAGE del pico obtenido.

La co-expresión de **Rpn11/Rpn8** dio como resultado un complejo estequiométrico, sorprendentemente el problema de insolubilidad presentado por Rpn11 al ser expresada individualmente es superado al co-expresarla con Rpn8. Sin embargo, los diversos estados de oligomerización de Rpn8 se mantienen en el complejo y Rpn11 presenta degradación (Fig. 1.26A). Con el fin de delimitar los dominios discretos de estas proteínas que mantienen unido al complejo se crearon constructos racionales de los dominios MPN y PCI, diseñados teniendo en cuenta los resultados del análisis de la secuencia primaria de ambas proteínas mediante una variedad de herramientas bioinformáticas (apartado 1.2.2). Mediante la predicción de la estructura secundaria de Rpn8 se identificaron dos regiones conservadas *coiled coil* separadas por una corta secuencia desordenada, por lo que los constructos resultantes fueron: Rpn8₁₋₁₈₆ que contiene estrictamente el dominio MPN y Rpn8₁₋₂₄₀ que contiene el dominio MPN y la región *coiled coil* proximal (Fig. 1.26B). En el caso de Rpn11 se diseñó un solo constructo que abarca el dominio JAMM/MPN: Rpn11₁₋₂₀₅.

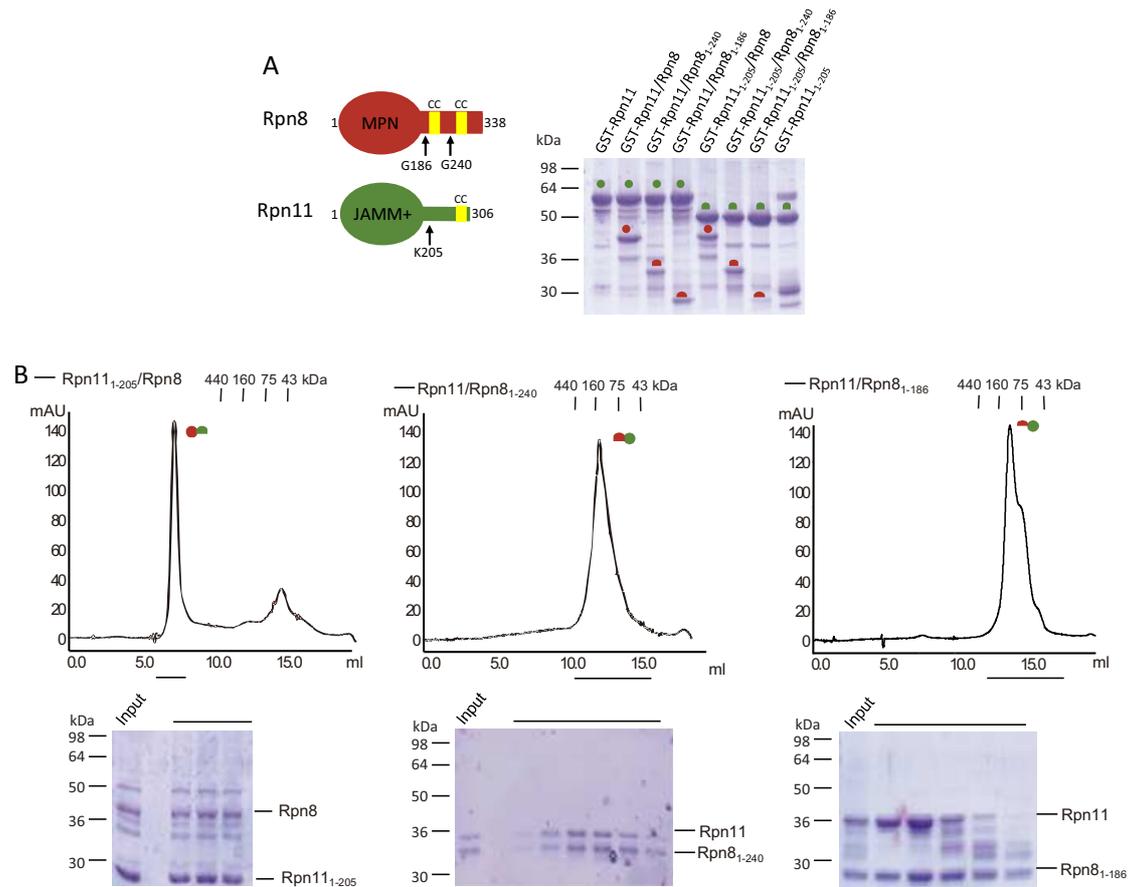


Fig. 1.27 Versiones truncadas de Rpn11/Rpn8. A) Representación esquemática de los dominios de Rpn8 y Rpn11, en amarillo se muestran las regiones *coiled coil* (CC) y con flechas los límites de los constructos truncados en el extremo C-terminal (izquierda), y el análisis por SDS-PAGE de su co-expresión (derecha). B) Perfiles de elución y análisis por SDS-PAGE de las fracciones de los complejos Rpn11₁₋₂₀₅/Rpn8 (izquierda), Rpn11/Rpn8₁₋₂₄₀ (centro) y Rpn11/Rpn8₁₋₁₈₆ (derecha).

La expresión bicistrónica de **Rpn5/Rpn9** forma un heterodímero estequiométrico, sin embargo Rpn9 presenta indicios de degradación (Fig. 1.28A) por lo que se descartó la posibilidad de realizar ensayos de cristalización con este complejo. Al igual que para el complejo Rpn11/Rpn8 se decidió analizar el comportamiento de versiones truncadas, la predicción de la estructura secundaria de estas proteínas revela que el dominio PCI, localizado en el extremo C-terminal está formado por aproximadamente 100 residuos organizados en hélices α y que Rpn9 presenta además dos regiones *coiled coil* a cada lado del dominio PCI. Basándonos en estos datos y en la literatura se clonaron dos constructos previamente descritos por Fu *et al.*, 2001: Rpn5_{169-f} y Rpn9_{160-f} y dos versiones que contienen estrictamente el dominio PCI: Rpn5_{313-f} y Rpn9_{261-f}.

Mediante co-expresión se identificó para el complejo Rpn5_{169-f}/Rpn9_{160-f} una interacción comparable a la establecida entre las versiones *full length*. Sin embargo, se observó una reducción de la afinidad al co-

expresar los complejo Rpn5/Rpn9_{261-f} y Rpn5_{313-f}/Rpn9, en el último caso como las proteínas GST-Rpn5_{313-f} y Rpn9 presentan el mismo peso molecular fue necesario eliminar el *tag* a fin de poder apreciar la interacción entre las proteínas. A su vez cuando se expresaron los dominios PCI estrictamente la afinidad disminuyó aún más (Fig. 1.28B), demostrando que la unión de Rpn5 y Rpn9 está mediada, no solo por el dominio PCI, sino también por las regiones medias de ambas proteínas que proporcionan estabilidad al complejo.

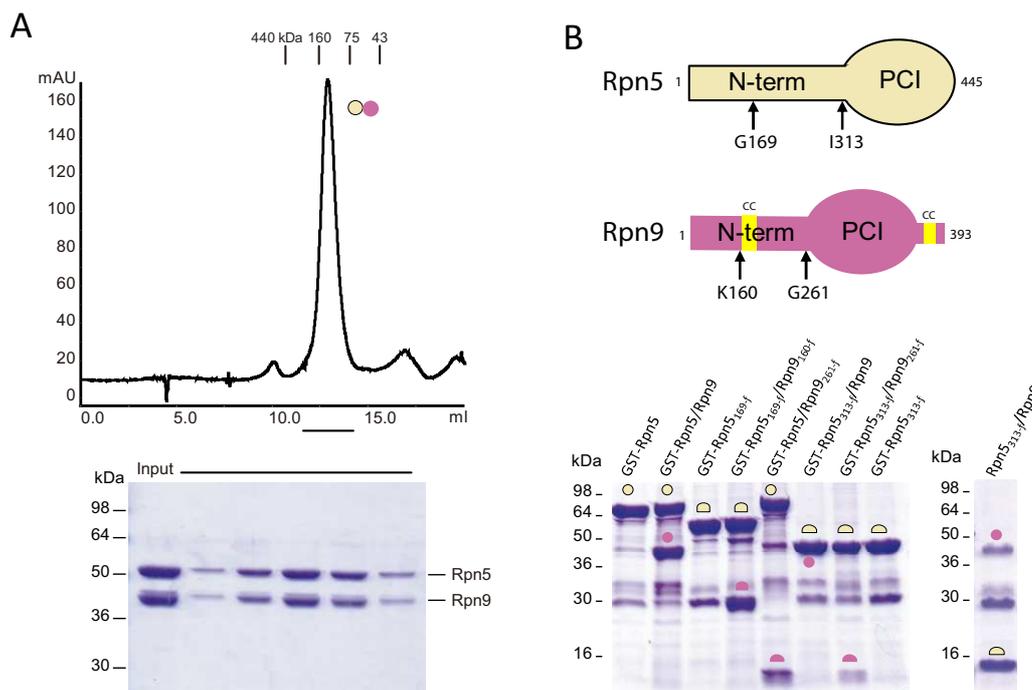


Fig. 1.28 Rpn5/Rpn9 y versiones truncadas. A) Perfil de elución y SDS-PAGE de las fracciones del complejo Rpn5/Rpn9. B) Arriba se encuentra la representación esquemática de los dominios de Rpn5 y Rpn9, en amarillo se muestran las regiones *coiled coil* (CC) y con flechas los límites de los constructos truncados en el extremo N-terminal; abajo el análisis por SDS-PAGE de la co-expresión de los diferentes complejos truncados.

Sorprendentemente la co-expresión del complejo **Rpn5/Rpn8** resultó en un complejo estequiométrico y que eluye de la Superdex200 como un pico simétrico y con peso molecular esperado, lo que revela que Rpn5 ocupa la región de Rpn8 responsable del proceso de oligomerización observado cuando se la expresa individualmente. Desafortunadamente su estudio cristalográfico no se llevó a cabo porque la integridad del complejo no se mantenía y se disgregaba con el paso del tiempo. En vista de este resultado se decidió testear por co-expresión la capacidad de Rpn8 de unir tres constructos de Rpn5 truncados en el extremo N-terminal: Rpn5_{169-f}, Rpn5_{249-f} y Rpn5_{313-f}. Se observó que Rpn8 une con igual afinidad los tres constructos (Fig. 1.29), indicando que el dominio PCI es la región mínima capaz de interactuar con Rpn8. Al analizar estos complejos binarios mediante columna de exclusión molecular se observa que mientras que

Rpn5_{249-f}/Rpn8 y Rpn5_{313-f}/Rpn8 eluyen siguiendo el mismo perfil de Rpn8 sola, el complejo formado por Rpn5_{169-f}/Rpn8 eluye como un pico simétrico, demostrando que la región de Rpn5 entre los residuos 169 a 249 es la responsable de capturar el *coiled coil* distal de Rpn8 responsable de los diferentes estados oligoméricos de la proteína (Fig. 1.29).

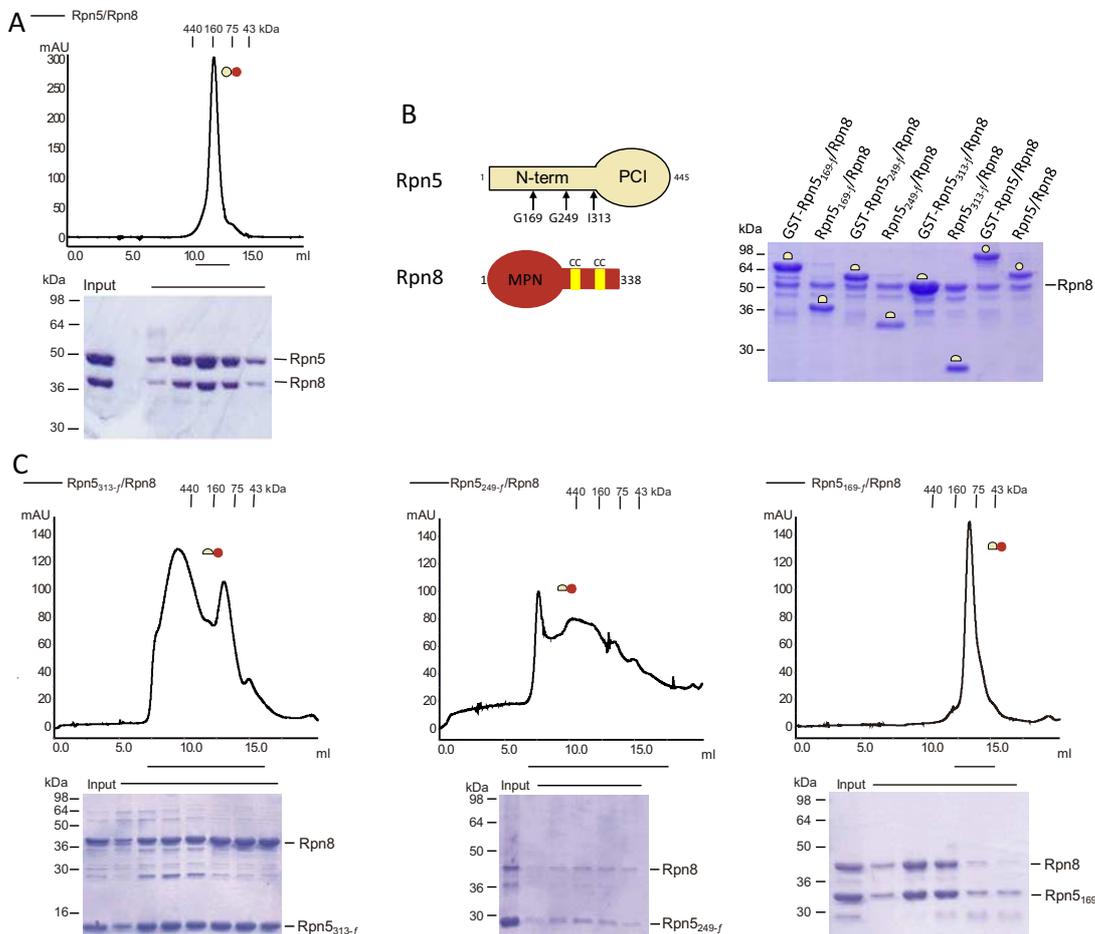


Fig. 1.29 Rpn5/Rpn8 y versiones truncadas. A) Perfil de elución y análisis por SDS-PAGE de las fracciones del complejo Rpn5/Rpn8. B) Representación esquemática de los dominios de Rpn5 y Rpn8, en amarillo se muestran las regiones *coiled coil* (CC) y con flechas los límites de los constructos truncados (izquierda) y el análisis por SDS-PAGE de su co-expresión (derecha). C) Perfiles de elución y análisis por SDS-PAGE de las fracciones de los complejos Rpn5_{313-f}/Rpn8 (izquierda), Rpn5_{249-f}/Rpn8 (centro) y Rpn5_{169-f}/Rpn8 (derecha).

Complejos triméricos

Para analizar interacciones ternarias se co-expresaron complejos binarios por un lado y proteínas individuales por otro y se co-lisaron ambos cultivos conjuntamente para purificarlos siguiendo el protocolo estándar. Se realizaron las siguientes combinaciones: Rpn11/Rpn8+Rpn5, Rpn11/Rpn8+Rpn9, Rpn5/Rpn9+Rpn8, Rpn5/Rpn9+Rpn11 y además se decidió comprobar si la proteína Rpn6 era capaz de unirse a alguno de los 3 complejos binarios.

Al combinar el complejo Rpn11/Rpn8 con Rpn9 no se observan cambios en los perfiles de elución, pero al expresar este complejo con Rpn5 se genera un nuevo complejo estequiométrico: **Rpn11/Rpn8/Rpn5**, que como era de esperarse no presenta los diferentes estados oligoméricos de Rpn8 ya que la presencia de Rpn5 lo evita, aunque Rpn11 sigue siendo inestable y se degrada (Fig. 1.30). En cuanto a Rpn5/Rpn9 con Rpn11 tampoco se observa interacción, mientras que con Rpn8 se obtiene un complejo estequiométrico y estable: **Rpn5/Rpn9/Rpn8** (Fig. 1.30). Este heterotrímero y el mismo complejo pero con las versiones truncadas de Rpn5 y Rpn9 (**Rpn5_{169-f}/Rpn9_{160-f}/Rpn8**) se sometieron a una amplia variedad de ensayos cristalográficos a 4°C y a 20°C e inclusive se intentaron cristalizar con la técnica de proteólisis *in situ*; los detalles se especifican en la tabla 1.7, sin obtener resultados positivos. En relación a la interacción de Rpn6 con los complejos binarios no se observaron variaciones en los perfiles de elución (Fig. 1.30).

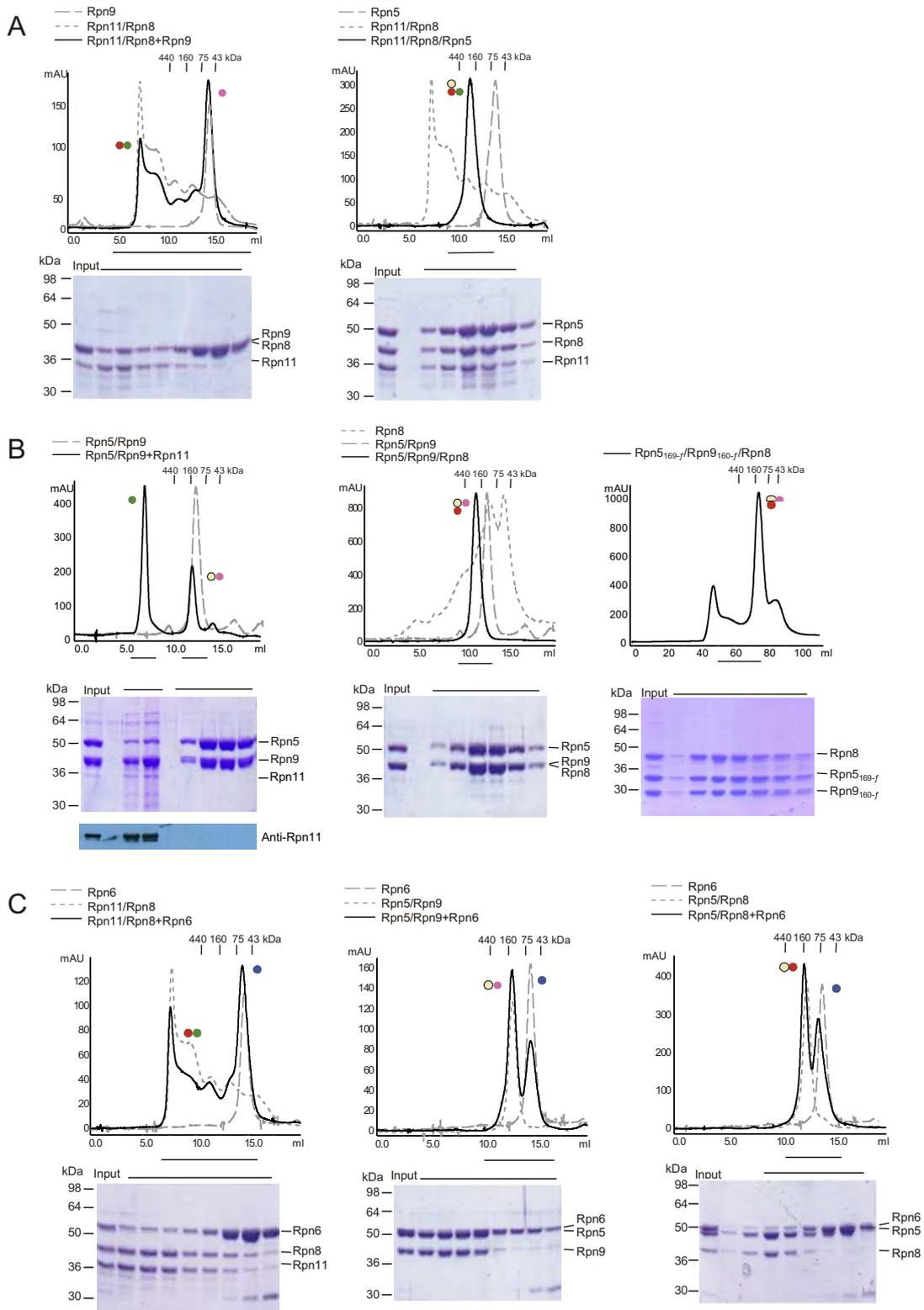


Fig. 1.30 Análisis de trímeros. Perfiles de elución y análisis por SDS-PAGE de las fracciones de A) Rpn11/Rpn8 con Rpn9 (derecha) y Rpn11/Rpn8/Rpn5 (izquierda). B) Rpn5/Rpn9 con Rpn11, Rpn11 no se aprecia en el SDS-PAGE por lo que se comprobó su presencia por western blot (derecha), Rpn5/Rpn9/Rpn8 (centro) y Rpn5_{169-f}/Rpn9_{160-f}/Rpn8 (izquierda). C) Rpn11/Rpn8 con Rpn6 (derecha), Rpn5/Rpn9 con Rpn6 (centro) y Rpn5/Rpn8 con Rpn6 (izquierda).

Complejos tetraméricos

A continuación se analizó la capacidad de interacción entre los dos complejos binarios Rpn11/Rpn8 y Rpn5/Rpn9 y de los complejos ternarios Rpn11/Rpn8/Rpn5 y Rpn5/Rpn9/Rpn8 con Rpn6. En el primer caso, se realizó la co-expresión de los bicomplejos por separado para luego co-lisar ambos cultivos, mientras para los otros dos se co-expresó el bicomplejo y las otras dos proteínas por separado para finalmente co-lisar los tres cultivos conjuntamente.

Después de la eliminación del *tag* de GST y de la columna de exclusión molecular se obtuvieron dos complejos heterotetraméricos estables estequiométricos y monodispersos, formados por Rpn11/Rpn8/Rpn5/Rpn9 y Rpn5/Rpn9/Rpn8/Rpn6. Mientras que para la combinación de Rpn11/Rpn8+Rpn5+Rpn6 no se observaron cambios en los perfiles de elución (Fig. 1.31). Ambos complejos cuaternarios fueron sometidos a ensayos cristalográficos en todas las condiciones disponibles en la PAC y a dos temperaturas 4°C y 20°C que fueron realizados por el Dr. Fabio Sessa sin obtener resultados positivos.

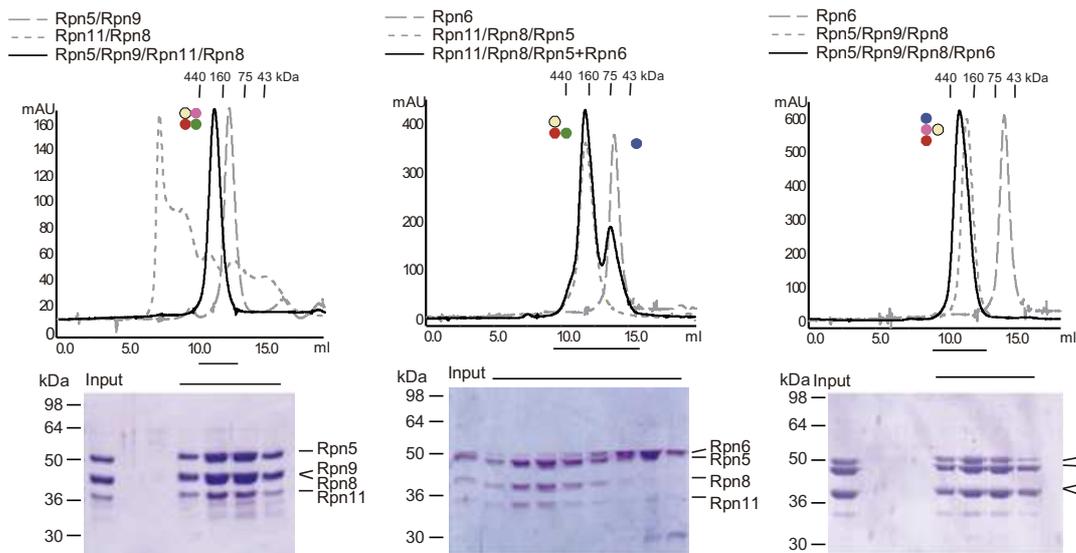


Fig. 1.31 Análisis de tetrámeros *full length*. Perfiles de elución y análisis por SDS-PAGE de las fracciones de Rpn5/Rpn9/Rpn11/Rpn8 (derecha), Rpn11/Rpn8/Rpn5+Rpn6 (centro) y Rpn5/Rpn9/Rpn8/Rpn6 (izquierda).

Tomando en cuenta los resultados previos de los complejos binarios, ternarios y cuaternarios podemos inferir que la interacción principal que mantiene unido a estos complejos es la establecida entre Rpn5 y Rpn8, siendo Rpn5 la responsable de unir a Rpn9 y Rpn8 a Rpn11. También se deduce que Rpn11 probablemente no establece interacciones directas ni con Rpn9, ni con Rpn6. A su vez, podemos apreciar que es necesaria la formación de un complejo trimérico compuesto por Rpn5/Rpn9/Rpn8 para

que Rpn6 forme parte del complejo, siendo insuficiente la presencia de solo dos de estas proteínas.

Con la intención de analizar en detalle las áreas de interacción y de encontrar posibles candidatos para ensayos cristalográficos se expresaron y analizaron los complejos cuaternarios reemplazando algunas proteínas *full length* por versiones truncadas disponibles: Rpn11₁₋₂₀₅, Rpn8₁₋₂₄₀, Rpn8₁₋₁₈₆, Rpn5_{169-f}, Rpn9_{160-f} y Rpn6_{110-f}. Cuando se co-purifica Rpn5/Rpn9 con Rpn11₁₋₂₀₅/Rpn8 o Rpn11/Rpn8₁₋₂₄₀ se obtienen dos complejos tetraméricos, indicando que ni la eliminación del extremo C-terminal de Rpn11, ni del *coiled coil* distal de Rpn8 perturban la asociación del complejo. Sin embargo cuando Rpn5/Rpn9 se combinó con Rpn11/Rpn8₁₋₁₈₆ no se observaron cambios en los perfiles de elución de los heterodímeros, demostrando que el *coiled coil* proximal es esencial para la unión de Rpn5 y por lo tanto para la formación de tetracomplejo. Por otro lado la eliminación de los extremos N-terminales de Rpn5, Rpn9 y Rpn6 tampoco afecta a la conformación de los complejos (Fig. 1.32), por lo que, con la intención de obtener cristales, se purificaron los siguientes complejos: Rpn5_{169-f}/Rpn9_{160-f}/Rpn11/Rpn8, Rpn5_{169-f}/Rpn9_{160-f}/Rpn8/Rpn6 y Rpn5_{169-f}/Rpn9_{160-f}/Rpn8/Rpn6_{110-f}. Lamentablemente no se obtuvieron resultados positivos (tabla 1.7).

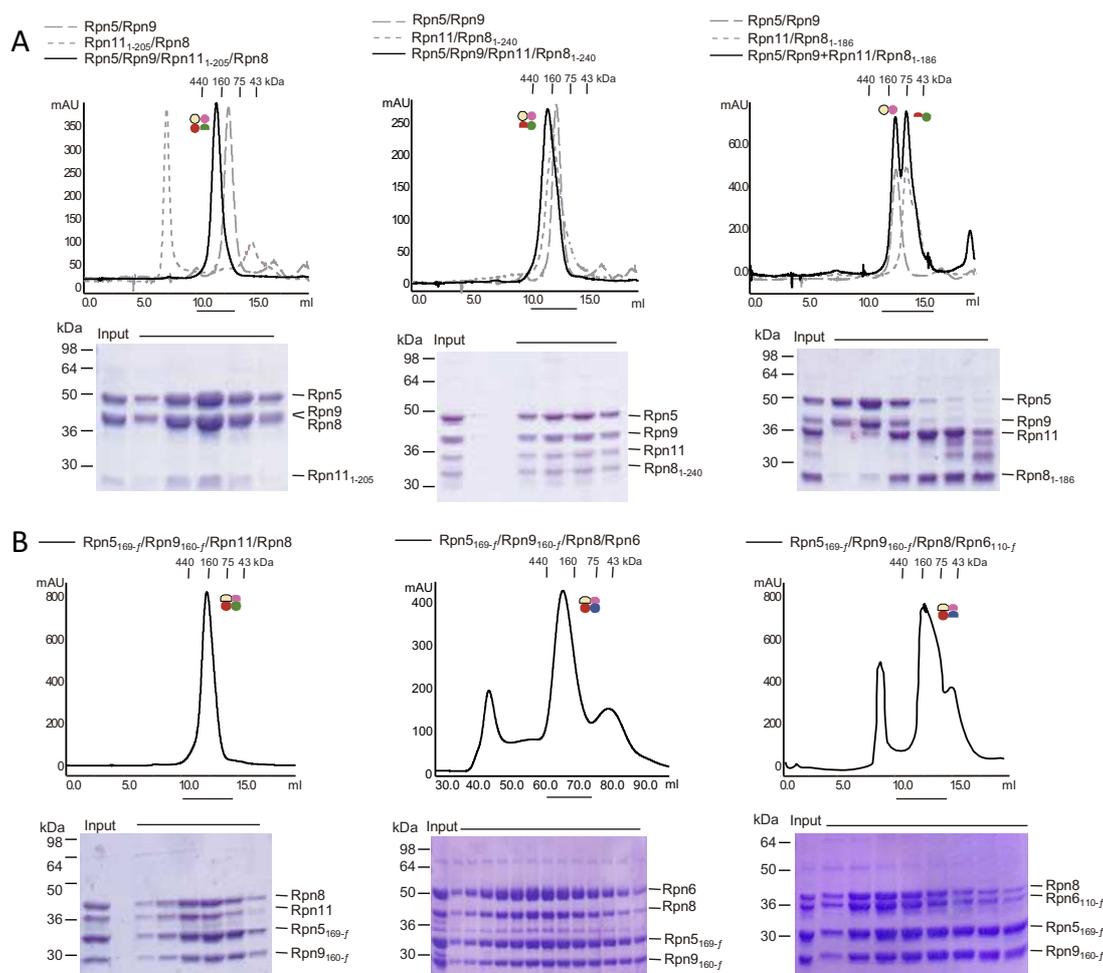


Fig. 1.32 Análisis de tetrámeros truncados. Perfiles de elución y análisis por SDS-PAGE de las fracciones A) Rpn5/Rpn9/Rpn11₁₋₂₀₅/Rpn8 (derecha), Rpn5/Rpn9/Rpn11/Rpn8₁₋₂₄₀ (centro) y Rpn5/Rpn9 con Rpn11/Rpn8₁₋₁₈₆ (izquierda). B) Rpn5_{169-f}/Rpn9_{160-f}/Rpn11/Rpn8 (derecha), Rpn5_{169-f}/Rpn9_{160-f}/Rpn8/Rpn6 (centro) y Rpn5_{169-f}/Rpn9_{160-f}/Rpn8/Rpn6_{110-f} (izquierda).

Complejos pentaméricos

Finalmente teniendo en cuenta que los dos complejos tetramérico comparten 3 de las 4 proteínas que los componen se decidió integrarlos, por lo que se co-expresaron los bicomplejos Rpn5/Rpn9 y Rpn11/Rpn8 junto con Rpn6, de este modo se obtuvo un pentacomplejo estable y estequiométrico. Siguiendo la misma estrategia que para los tetrámeros, se analizó el efecto de las versiones truncadas sobre el ensamblaje. Se observó que ni la eliminación del extremo C-terminal de Rpn11 ni de los extremos N-terminales de Rpn5, Rpn9 o Rpn6 afectaron a la conformación del pentámero mientras que, al analizar el extremo C-terminal de Rpn8, se mostró claramente que, como era de esperar, la eliminación de todo el C-terminal (Rpn8₁₋₁₈₆) impide la formación del complejo. Interesantemente, al retirar el *coiled coil* distal, se obtiene el complejo tetramérico previamente

descrito pero Rpn6 no es capaz de unirse a él, demostrando que este dominio es esencial para la incorporación de Rpn6 al pentámero (Fig. 1.33).

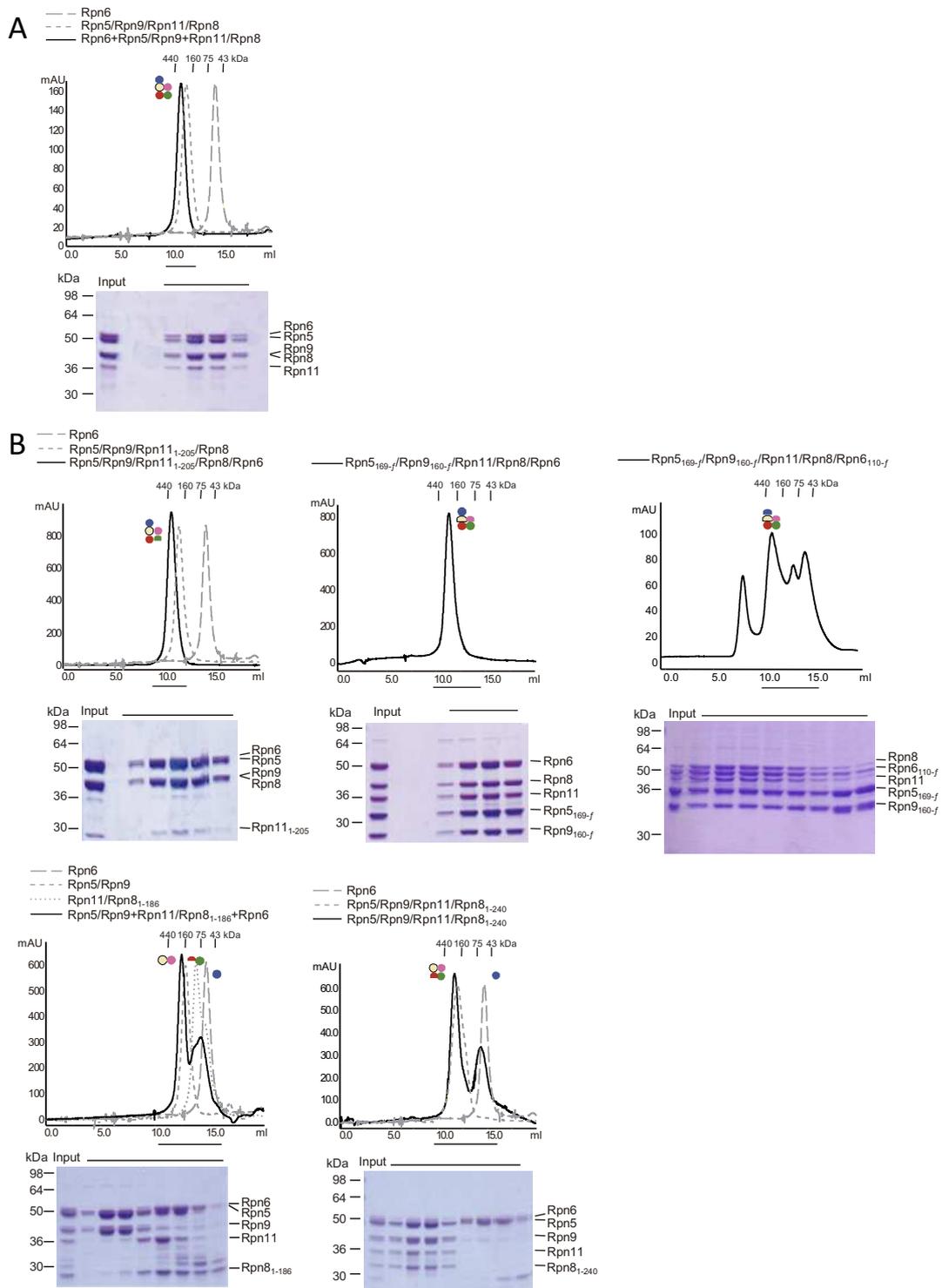


Fig. 1.33 Análisis de pentámeros. Perfiles de elución y análisis por SDS-PAGE de las fracciones de A) Rpn5/Rpn9/Rpn11/Rpn8/Rpn6. B) Rpn5/Rpn9/Rpn11₁₋₂₀₅/Rpn8/Rpn6 (derecha), Rpn5_{169-f}/Rpn9_{160-f}/Rpn11/Rpn8/Rpn6 (centro) y Rpn5_{169-f}/Rpn9_{160-f}/Rpn11/Rpn8/Rpn6_{110-f} (izquierda). C) Rpn5/Rpn9+Rpn11/Rpn8₁₋₁₈₆+Rpn6 (derecha) y Rpn5/Rpn9/Rpn11/Rpn8₁₋₂₄₀+Rpn6 (izquierda).

Se realizaron todos los ensayos cristalográficos a 4°C y 20°C para el complejo pentamérico *full length* (Dr. Fabio Sessa) y también se aplicó la técnica de proteólisis *in situ*. Lamentablemente, la búsqueda de un cristal fue infructuosa. También se intentó cristalizar en vano los complejos Rpn5_{169-f}/Rpn9_{160-f}/Rpn8/Rpn11/Rpn6 y Rpn5_{169-f}/Rpn9_{160-f}/Rpn8/Rpn11/Rpn6_{110-f} (tabla 1.7).

Con el fin de integrar los resultados obtenidos en esta serie de experimentos, generamos el siguiente esquema (Fig. 1.34) en el que se resumen y detallan las regiones discretas de interacción las proteínas.

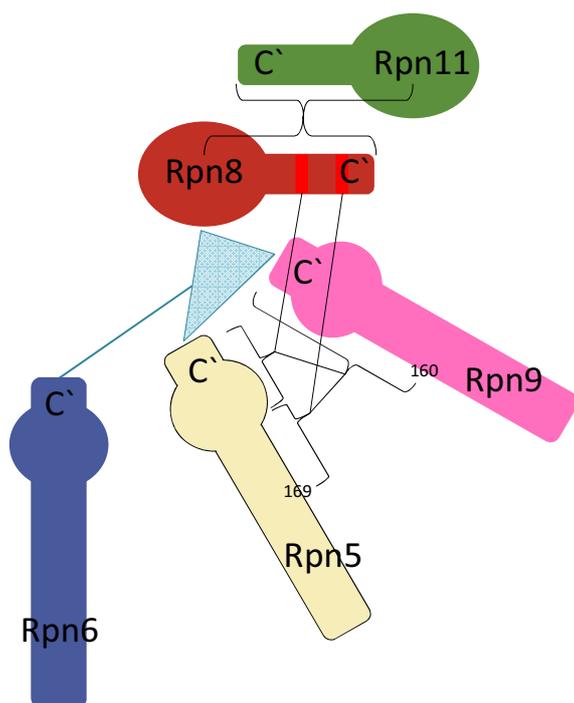


Fig. 1.34 Esquema de las interacciones del complejo pentamérico. Las interacciones binarias entre áreas discretas de las proteínas se muestran con corchetes unidos por líneas negras, mientras que el triángulo azul conectado por una línea hace referencia a la necesidad de un complejo trimérico pre-ensamblado para la incorporación de Rpn6. Los cuadrados en tonos más oscuros en Rpn8 muestran la presencia relativa de las secuencias *coiled coil* y las áreas abultadas la posición de los dominios PCI o MPN según corresponda.

Reconstitución heteróloga de la lid

Como paso final del análisis bioquímico se intentó reconstituir el complejo de la lid al completo, para ello se co-expresaron y co-lisaron las cinco proteínas que componen el pentámero, junto con el bicomplejo Rpn3/Rpn7 y Rpn12. Desafortunadamente no se observaron cambios en los perfiles de elución; posiblemente el estado de agregación observado para el bicomplejo Rpn3/Rpn7 impide el ensamblaje del complejo (Fig. 1.35). Si bien podemos asegurar que el pentacomplejo se forma independientemente y sin la necesidad de chaperonas, no podemos afirmar esto mismo para todo el complejo de la lid. De hecho, a pesar de estar utilizando una estrategia similar a la aplicada por Lander *et al.*, 2012 no obtenemos el complejo íntegro, por lo que se podría pensar que la presencia de la chaperona Hsp90 es la clave para el ensamblaje completo de la lid.

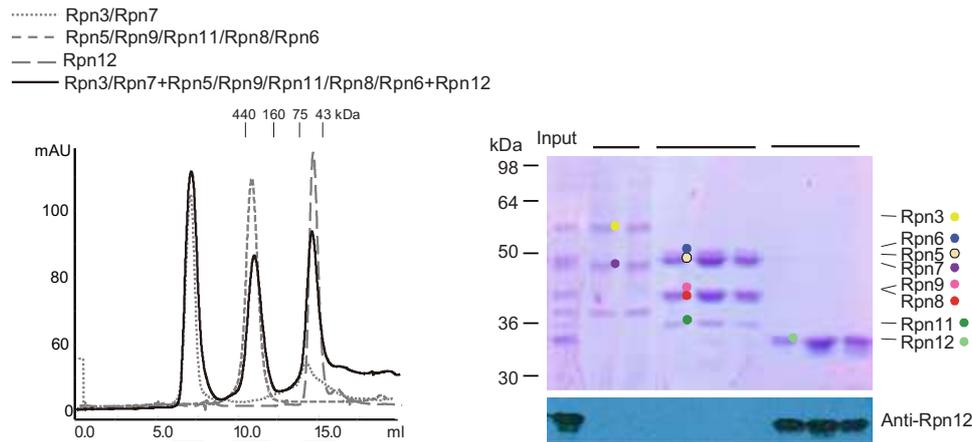


Fig. 1.35 Reconstrucción heteróloga de la lid. Perfil de elución y análisis de las fracciones por SDS-PAGE de la combinación de Rpn3/Rpn7, el pentacomplejo y Rpn12.

1.3.2 Lid endógena

Expresión y purificación de la lid endógena

Del cultivo de levaduras en el fermentador de 50l se obtuvo un rendimiento promedio de 1,7kg de biomasa total seca.

El complejo obtenido luego de la purificación por afinidad y la eliminación del *tag* de proteína A fue sometido a una cromatografía de intercambio iónico y a continuación las fracciones del pico obtenido fueron concentradas e inyectadas en una columna de exclusión molecular superosa 6. Los perfiles de elución y el correspondiente análisis por SDS-PAGE se muestran en la Fig. 1.36.

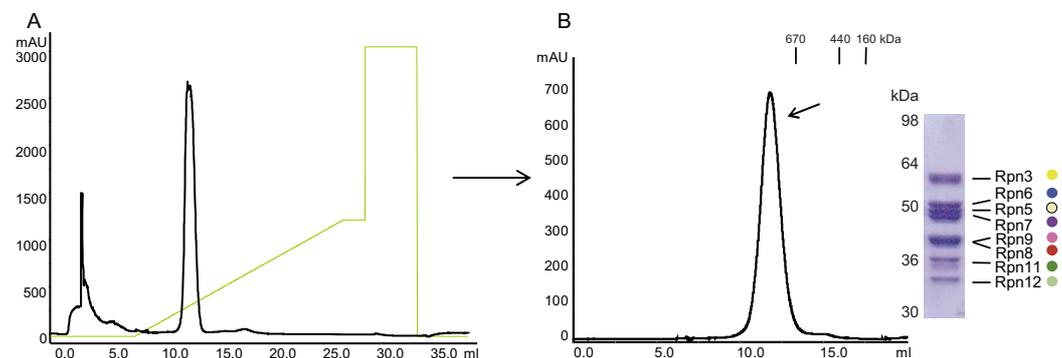


Fig. 1.36 Purificación de la lid endógena. A) Cromatografía monoQ del complejo de la lid obtenido luego de la purificación por afinidad y la remoción del tag de proteína A. B) Cromatografía superosa 6 del pico obtenido de la monoQ y análisis por SDS-PAGE del pico señalado con una flecha.

La purificación resultó en un complejo homogéneo que fue analizado por MS, demostrando la presencia estequiométrica de todas las subunidades

Rpn de la lid, excepto Sem1. La ausencia o bien la imposibilidad de detectar esta subunidad también ha sido observada por Lander *et al.*, 2012 que no la detectan ni en la lid endógena ni en la reconstrucción de la lid de forma recombinante, y por Fukunaga *et al.*, 2010 quienes la detectaron mediante la utilización de un Ac anti-Sem1 pero fueron incapaces de obtener resultados por MS, probablemente debido a la pobre ionización de los péptidos de esta subunidad.

Interacción de la lid endógena con los AcMo anti-Rpn11 y anti-Rpn8

Inicialmente se comprobó la avidéz y especificidad de ambos AcMo por western blot. El resultado fue que si bien la especificidad del anti-Rpn8 era correcta, la avidéz era realmente baja obligándonos a utilizar títulos altísimos (1/50) para obtener una banda de intensidad baja a pesar de utilizar sustrato ECL plus y tras un largo periodo de exposición de la lámina fotográfica. Por el contrario, en el caso del anti-Rpn11 tanto la especificidad como la avidéz eran correctas ya que, utilizando títulos de hasta 1/6000, los resultados eran claros y las bandas resultantes intensas, utilizando sustrato ECL normal y periodos de exposición de pocos segundos (Fig. 1.37).

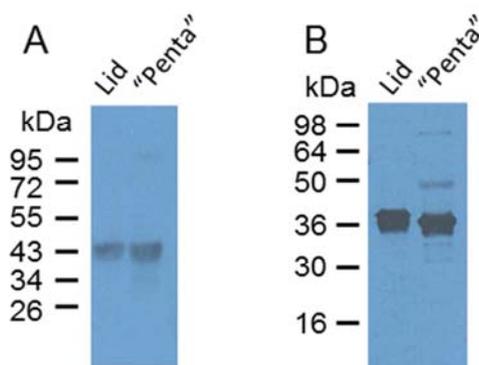


Fig. 1.37 Western blot de los AcMo A) anti-Rpn8 y B) anti-Rpn11

Se incubaron ambos AcMo con la lid en experimentos aislados y en ambos casos se obtuvieron como resultado dos picos bien diferenciados en sus perfiles de elución. Sin embargo, en el caso del AcMo anti-Rpn8 cada pico correspondía a la lid endógena y al Ac respectivamente, sin interacción entre ambas entidades (Fig. 1.38A); mientras que en el caso del anti-Rpn11 en el primer pico se detectó la presencia de un complejo formado por Rpn5/Rpn9/Rpn8/Rpn11/Rpn6 y el Ac y en el segundo pico un complejo compuesto por Rpn3/Rpn5/Rpn6/Rpn7 cuyo peso es similar al del Ac y por lo tanto co-eluyen en el mismo pico (Fig. 1.38B). Por otro lado, Rpn12 no fue detectado por SDS-PAGE pero, al analizar las fracciones por western blot, pudimos observar que esta subunidad se había disociado y se encontraba libre en la mezcla (Fig. 1.38B).

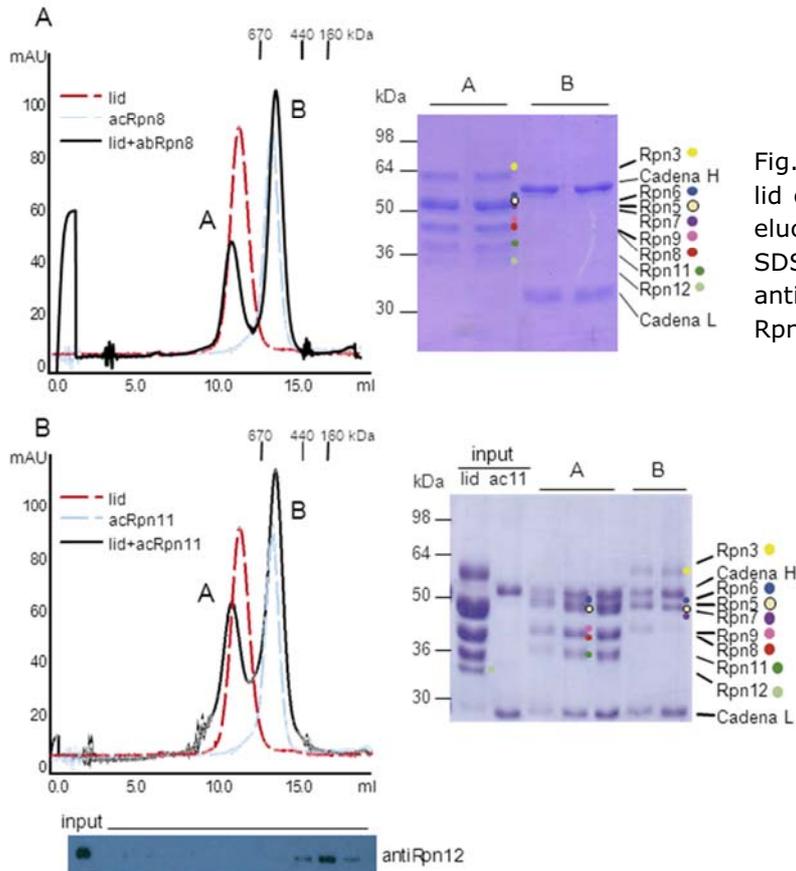


Fig. 1.38 Interacción de la lid con los AcMo. Perfiles de elución en superosa 6 y SDS-PAGE con A) el AcMo anti-Rpn8 y B) el AcMo anti-Rpn11.

El resultado obtenido de la interacción de la lid endógena con el AcMo anti-Rpn11 nos provee información indirecta de la conformación de la lid y su estructura modular, donde se observa un penta-complejo definido (Rpn5/Rpn9/Rpn8/Rpn11/Rpn6) al igual que en nuestros ensayos recombinantes y en los estudios realizados por Sharon *et al.*, 2006; Fukunaga *et al.*, 2010 y Tomko *et al.*, 2011. A su vez se confirma que Rpn12 se mantiene monomérico si la lid no está completa, tal y como lo expone Tomko *et al.*, 2011. El reciente descubrimiento del haz helicoidal que mantiene unida a la lid [Estrin *et al.*, 2013], explicaría el efecto del Ac sobre el complejo. El epítipo reconocido por el mismo se localiza en el extremo C-terminal de Rpn11 por lo que su interacción con esta fracción de la proteína fácilmente podría interferir sobre la topología del haz helicoidal y por ende disociaría el complejo.

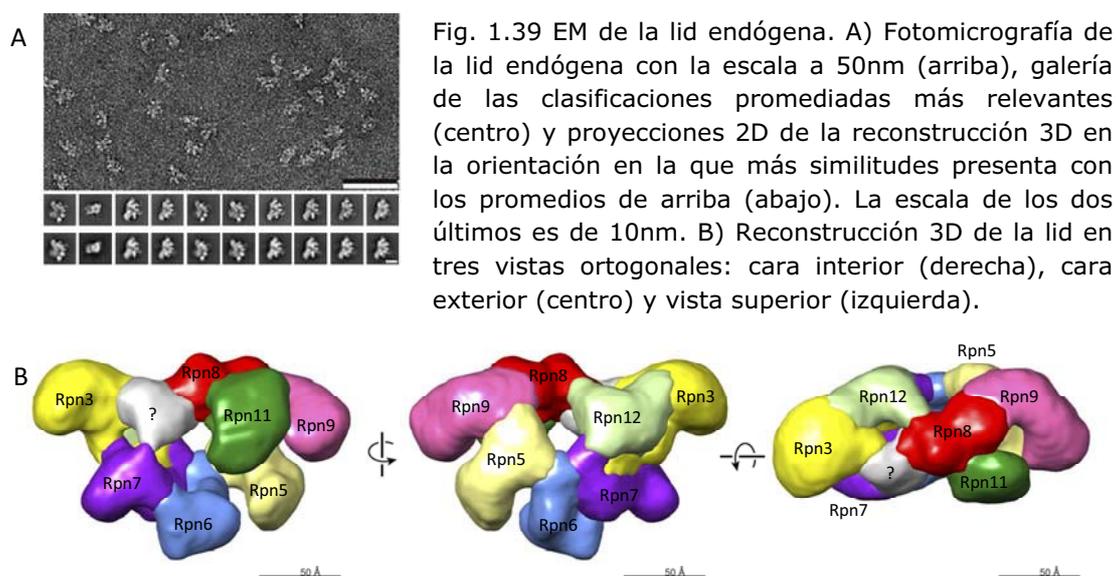
Como se comentó en la introducción tanto al incremento como la inhibición de la actividad proteosomal pueden ser de utilidad para el tratamiento de una amplia variedad de enfermedades, por lo que el anticuerpo monoclonal anti-Rpn11 podría poseer algún tipo de valor clínico. A día de hoy la inmunoterapia tiene una aplicación limitada en el campo de la oncología y solo doce anticuerpos han sido aprobados por la FDA. Todos ellos actúan sobre dianas extracelulares sobretodo antígenos de superficie [Scott *et al.*, 2012], por lo que el hecho de que la diana de nuestro anticuerpo sea citosólica podría ser una limitación importante. Sin embargo,

si bien el anticuerpo *per se* podría no ser aplicable en el ámbito clínico, es interesante notar que su actividad podría ser reinterpretada para desarrollar drogas específicas. Es decir, el hecho de que el anticuerpo al reconocer el extremo C-terminal de Rpn11 produzca la disgregación del complejo nos lleva a pensar en la posibilidad diseñar compuestos que, de una u otra manera, interfieran en esta región.

1.3.3 Microscopía electrónica

Para obtener información estructural sobre la lid se realizó microscopía electrónica de transmisión (TEM) con tinción negativa, análisis 2D y reconstrucciones 3D de la lid endógena y de tres de los subcomplejos de la lid generados mediante expresión heteróloga. Estos subcomplejos representan la adición consecutiva de subunidades de una en una: (Rpn5/Rpn9/Rpn8) + Rpn6 → (Rpn5/Rpn9/Rpn8/Rpn6) + Rpn11 → (Rpn5/Rpn9/Rpn11/Rpn8/Rpn6), lo que nos permite obtener una visión dinámica del proceso de biogénesis de este complejo.

La TEM de la lid revela una distribución homogénea de partículas con una estructura compacta de forma triangular con un lado largo y recto y otros dos más cortos e irregulares. Dependiendo de la orientación de las partículas se observa una acumulación de tintura a través de la región central del complejo. La reconstrucción 3D resultante de 22Å tiene unas dimensiones de ~18x12x10nm³. Es una estructura con cinco dominios ensamblados en forma de abanico (cuyas puntas forman los lados irregulares) y una densidad extra que recorre la región recta. De los cinco dominios el central es el que parece tener límites más difusos, sugiriendo que este dominio es flexible (Fig. 1.39).



La organización cuaternaria del complejo se determinó tomando en consideración el mapa de interacciones obtenido en este trabajo, la información proporcionada por los subcomplejos de los que hablaremos más adelante y trabajos publicados recientemente [Lander *et al.*, 2012; Estrin *et al.*, 2013]. El área coloreada en gris representa la masa del complejo a la que no se le asigna ninguna subunidad en particular, por lo que hipotetizamos que corresponde a Sem1 o al haz helicodal. La diferencia fundamental entre la lid aislada de Lander *et al.*, 2012 y la nuestra es la conformación de los extremos N-terminales de las subunidades PCI, que en el primer caso se observan más extendidas, de hecho este mismo grupo detecta importantes cambios conformacionales de estos extremos al comparar la lid aislada de la lid unida al proteosoma indicando la flexibilidad inherente del complejo.

El trímero Rpn5/Rpn9/Rpn8 muestra partículas monodispersas caracterizadas por una forma de "V" con dos brazos de $\sim 10\text{nm}$ de largo cuyos extremos están separados por $\sim 2\text{nm}$, uno de los brazos presenta una torsión en el extremo. Se asignó que los brazos corresponden a Rpn5 y Rpn9 y la zona voluminosa del vértice a Rpn8. Al superponer este complejo sobre la estructura de la lid completa nos permite establecer con mayor definición los límites de Rpn8 (Fig. 1.40).

El tetrámero Rpn5/Rpn9/Rpn8/Rpn6 muestra partículas individuales con tres brazos. La reconstrucción 3D exhibe tres estructuras globulares de tamaños similares $\sim 3\text{nm}$ que emergen de una base común desestructurada y con dimensiones máximas totales de 13nm . Estos resultados sugieren que el tercer brazo del complejo corresponde a Rpn6 (Fig. 1.40). A pesar de poseer más masa, este complejo presenta menor resolución y mayor variabilidad conformacional que el trímero. Ambos fenómenos pueden explicarse debido a la flexibilidad previamente descrita de Rpn6 en la lid aislada [Lander *et al.*, 2012] y a las limitaciones que esto impone para la reconstrucción 3D del modelo.

El pentámero Rpn5/Rpn9/Rpn11/Rpn8/Rpn6 presenta partículas individuales con forma de tipo "W", en la reconstrucción 3D también se observan tres brazos que emergen de una estructura sólida central con dimensiones máximas de 11nm , aunque en este caso la estructura es mucho más compacta que en el caso del tetrámero, indicando que la adición de Rpn11 proporciona rigidez estructural al complejo (Fig. 1.40).

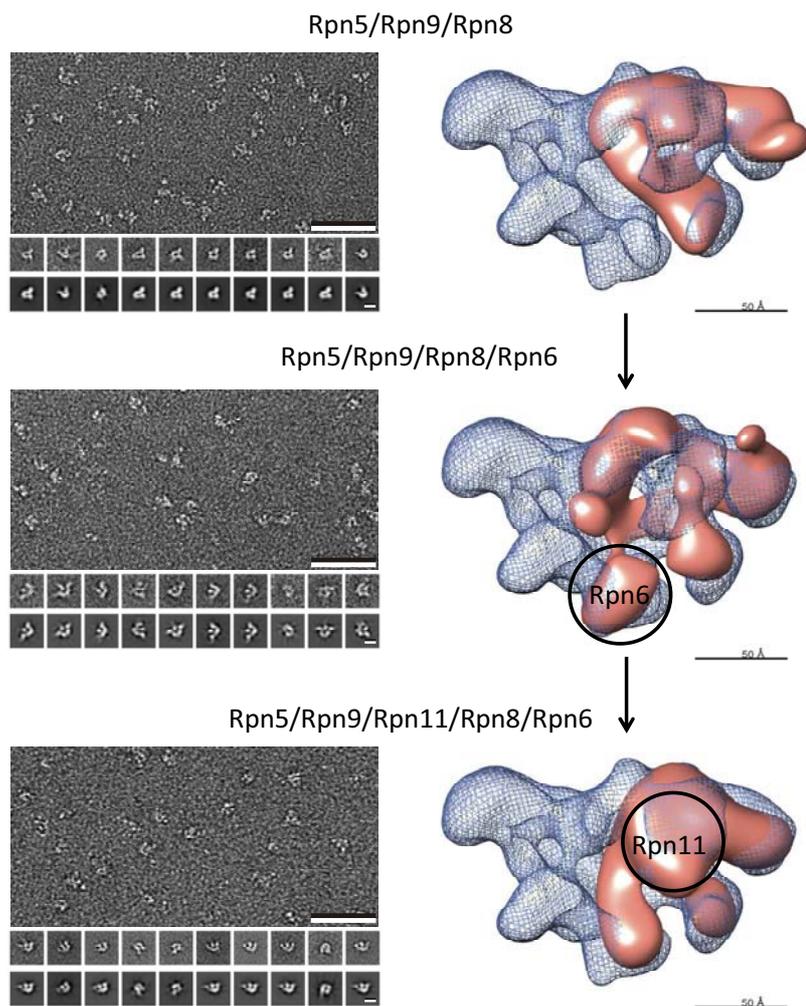


Fig. 1.40 EM de los subcomplejos. A la derecha de la imagen podemos observar las fotomicrografías de los subcomplejos con la escala a 50nm, galerías de las clasificaciones promediadas más relevantes y proyecciones 2D de las reconstrucciones 3D en la orientación en la que más similitudes presentan con los promedios de arriba. A la izquierda se encuentran las reconstrucciones 3D superpuestas sobre la estructura de la lid endógena, los círculos muestran las masas extras que aparecen luego de la adición de cada subunidad.

La resolución y la cantidad de partículas analizadas para la reconstrucción de estos complejos se detallan en la siguiente tabla y las correlaciones por capas en el espacio de Fourier (FSC, *Fourier Shell Correlations*) en la Fig. 1.41:

Tabla 1.10: Parámetros de los modelos 3D

Complejo	Nº de partículas	Resolución (Å)
Rpn5/Rpn9/Rpn8	9.176	31,1
Rpn5/Rpn9/Rpn8/Rpn6	5.757	32,9
Rpn5/Rpn9/Rpn11/Rpn8/Rpn6	6.121	31,1
Lid endógena	15.598	22,0

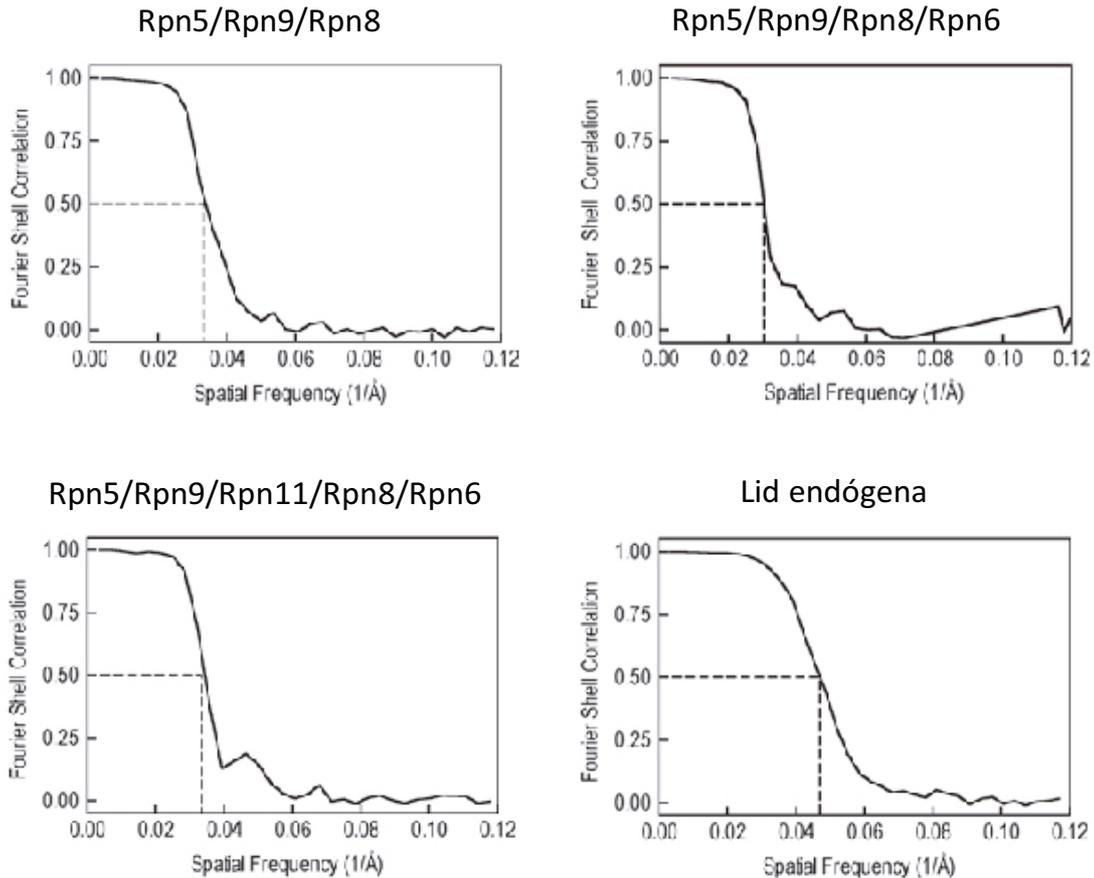


Fig. 1.41 Correlaciones por capas de Fourier (FSC). Los valores de resolución se determinaron de acuerdo a las correlaciones de Fourier utilizando el criterio del umbral 0,5 (línea de puntos).

1.3.4 Biogénesis de la lid

Estos resultados nos permiten generar un modelo sobre los eventos biogénicos que se producen durante en el proceso de ensamblaje de la lid. En el siguiente esquema (Fig. 1.42) se observa un camino principal (líneas continuas) representado por los subcomplejos que en nuestros experimentos resultaron ser los más estables y homogéneos, que al ser analizados por EM muestran claramente la adición consecutiva de subunidades al subcomplejo pentamérico de la lid. Sin embargo, y debido a que en durante el análisis bioquímico de estas proteínas se observaron muchas combinaciones posibles no podemos descartar otros caminos biogénicos alternativos (líneas discontinuas). Es interesante notar que Rpn8 cumple un rol central, funcionando como centro de nucleación para las proteínas que conforman el complejo pentamérico.

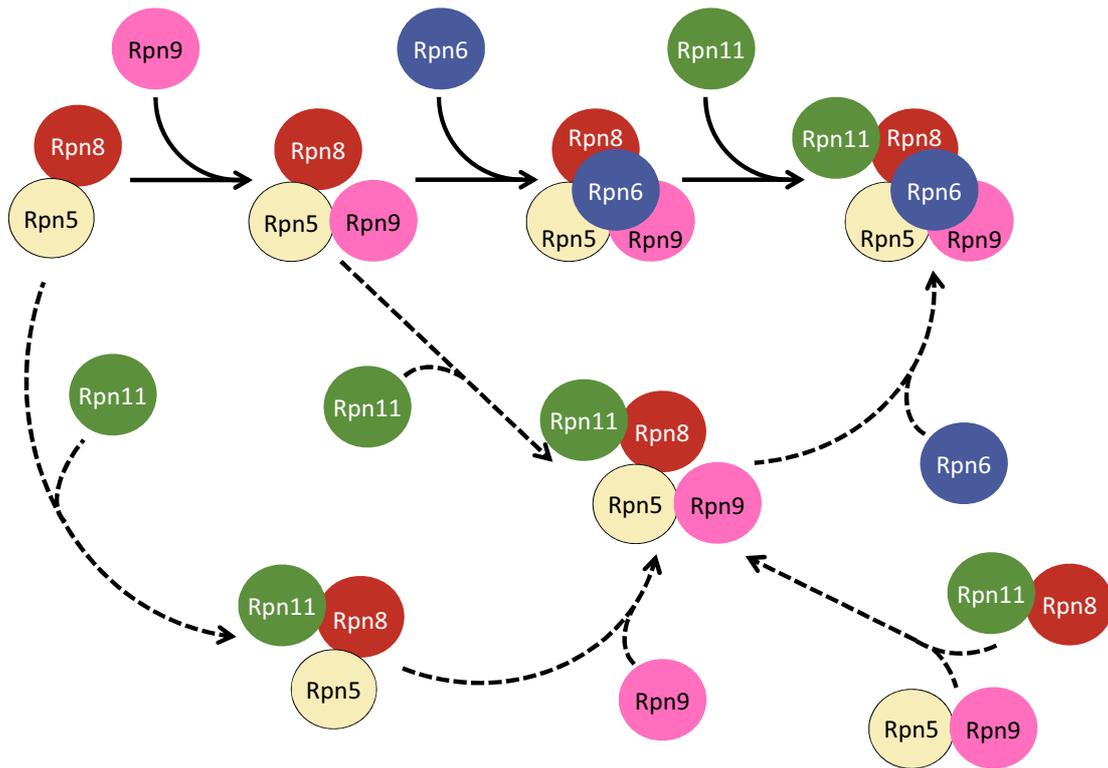


Fig. 1.42 Esquema de los eventos biogénicos de la formación del complejo pentamérico. Las líneas continuas muestran el camino metabólico más probable, mientras que las líneas discontinuas proponen todos los caminos alternativos posibles observados en este estudio.

El ensamblaje de la lid se completa mediante la adición del complejo binario formado por Rpn3/Rpn7, aunque no conocemos la naturaleza de las interacciones que se establecen con el complejo pentamérico y finalmente, como se describe en la literatura [Tomko *et al.*, 2011], Rpn12 concluye el proceso (Fig 1.43).

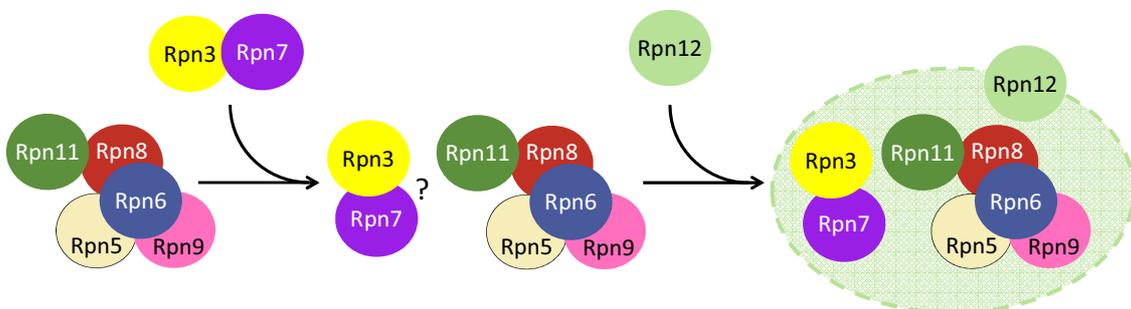


Fig 1.43 Esquema de la finalización del ensamblaje de la lid. Una vez formado el complejo pentamérico se produce la adición del complejo binario Rpn3/Rpn7 y finalmente Rpn12 completa el complejo de la lid.

1.4 Conclusiones

Rpn1 y Rpn2:

- Rpn1 y Rpn2 no pudieron ser purificadas de manera soluble, ya sea individualmente, en complejo o versiones truncadas de las mismas.

Rpn10 y la interacción base-lid:

- Rpn10 pudo ser purificada con un 95% de pureza utilizando un protocolo de purificación en condiciones de alta sal.
- Las interacciones que establece Rpn10 con otras subunidades de la lid ya sea por *in vitro binding* o por co-expresión son artefactos biológicos, debido probablemente a la naturaleza "pegajosa" de Rpn10.

Proteínas de la lid:

- En cuanto a las proteínas individuales Rpn3, Rpn7 y Rpn11 son insolubles, Rpn8 se obtiene soluble pero presenta diferentes estados oligoméricos, y Rpn5, Rpn6, Rpn9 y Rpn12 son solubles aunque las dos últimas presentan degradación.
- Al co-expresar Rpn3/Rpn7 se obtiene un complejo estequiométrico que se encuentra agregado en solución.
- La insolubilidad de Rpn11 es superada al ser co-expresada con Rpn8, generando un complejo heterodimérico que presenta diversos estados de agregación al igual que Rpn8 sola. La interacción entre estas proteínas probablemente sea de modo antiparalelo, donde el dominio MPN de una contacta con el extremo C-terminal de la otra y viceversa.
- La co-expresión de Rpn5 y Rpn9 genera un complejo estequiométrico, estas proteínas interactúan a través no solo de sus dominios PCI sino también de las regiones medias de ambas proteínas que estabilizan el heterodímero.
- Rpn5 y Rpn8 generan un complejo heterodimérico estequiométrico, donde el dominio PCI de Rpn5 es la región mínima necesaria capaz de interactuar con Rpn8 y la región de Rpn5 entre los residuos 169 a 249 captura el *coiled coil* distal de Rpn8 que es el responsable de los diferentes estados oligoméricos observados al purificar esta proteína individualmente o acompañada de Rpn11.
- Es posible obtener dos complejos triméricos Rpn5/Rpn9/Rpn8 y Rpn11/Rpn8/Rpn5, que apuntan al rol central de la interacción Rpn5/Rpn8, donde Rpn5 recluta a Rpn9 y Rpn8 a Rpn11.
- Mediante co-expresión y co-lisado es posible obtener dos tetracomplejos: Rpn11/Rpn8/Rpn5/Rpn9 y Rpn5/Rpn9/Rpn8/Rpn6.

Del análisis de los mismos podemos inferir que Rpn11 no establece contactos ni con Rpn 9 ni con Rpn6 y que Rpn6 requiere de la presencia conjunta de Rpn5/Rpn9/Rpn8 para formar parte del complejo. El análisis de versiones truncadas demuestra que el *coiled coil* proximal de Rpn8 es esencial para la interacción con Rpn5 y por ende para la integridad del complejo, mientras que los extremos N-terminales de Rpn5, Rpn9 y Rpn6 no afectan las asociaciones interprotéicas.

- Es posible obtener un pentacomplejo formado por Rpn5/Rpn9/Rpn11/Rpn8/Rpn6, y se demostró que Rpn6 no solo requiere la presencia de Rpn5, Rpn9 y Rpn8 sino que además el *coiled coil* distal de Rpn8 es clave para su unión al complejo.
- Podemos concluir que la formación del complejo pentamérico es un proceso que se lleva a cabo sin la participación de chaperonas.
- La reconstrucción de la lid completa no fue posible, probablemente debido al estado de agregación del complejo Rpn3/Rpn7 o tal vez debido a la falta de alguna chaperona como por ejemplo Hsp90.
- No fue posible obtener cristales de ninguno de los complejos o proteínas individuales sometidos a ensayos cristalográficos.

Lid endógena:

- Mediante la purificación por afinidad es posible obtener en complejo de la lid a partir de cultivos de *S. cerevisiae*, el complejo presenta todas las subunidades Rpn en cantidades estequiométricas. Sem1 no fue detectada.
- La interacción del anticuerpo monoclonal anti-Rpn11 produce la fragmentación de la lid, probablemente debido a que interfiere con el haz helicoidal.

Microscopía electrónica:

- Se obtuvieron modelos 3D para la lid endógena, el trímero Rpn5/Rpn9/Rpn8, el tetrámero Rpn5/Rpn9/Rpn8/Rpn6 y el pentámero Rpn5/Rpn9/Rpn11/Rpn8/Rpn6, que aportan evidencia estructural de que los subcomplejos están correctamente formados ya que conservan similitudes con la lid completa.

Biogénesis de la lid:

- El estudio de las interacciones y los subcomplejos nos aportan una visión dinámica del proceso temprano de biogénesis de la lid de forma secuencial. Rpn8 funciona como centro de nucleación donde se van adicionando consecutivamente las subunidades en el siguiente orden: Rpn5, Rpn9, Rpn6 y Rpn11, para formar el complejo

pentamérico. A este complejo se une el complejo binario Rpn3/Rpn7 y finalmente Rpn12 para completar la lid.

Segunda Parte:

Análisis estructural del
dominio OBD de RepB y su
unión al ADN

2.1 Introducción

La teoría microbiana de la enfermedad también conocida como teoría germinal fue enunciada en la segunda mitad del siglo XIX, la cual fue estructurada y demostrada por Koch, quien publicó sus postulados en 1890 [Castillo, 2007]. Esta investigación sentó las bases de la microbiología actual y del estudio de la etiología de las enfermedades infecciosas.

Sin embargo no fue hasta 40 años después que comenzaron a surgir compuestos capaces de combatir estas enfermedades. Ehrlich fue quien comenzó la búsqueda de antimicrobianos a principios del 1900 y encontró compuestos derivados del arsénico que cumplían dicha función, también Fleming en el año 1929 descubre un compuesto producido por el hongo *Penicillium notatum* con propiedades antibacterianas al que denominó penicilina. Aunque estos compuestos fueron descubiertos primero no fueron los primeros en ser utilizados masivamente, los compuestos de arsénico son tóxicos para el ser humano y la penicilina no pudo ser obtenida en grandes cantidades hasta 1940. Fueron las sulfamidas, descubiertas por Domagk en 1932, los primeros antimicrobianos utilizados masivamente para el tratamiento de un amplio espectro de infecciones bacterianas [Bauman, 2007].

Selman Waksman utilizó por primera vez en 1941 la palabra "antibiótico" para describir cualquier compuesto que antagoniza el crecimiento de microorganismos. A partir de 1945-1955 el desarrollo de la penicilina, junto con la estreptomocina, el cloranfenicol y la tetraciclina marcaron el comienzo de la era de los antibióticos. Los antibióticos son probablemente la forma de quimioterapia más exitosa de la historia de la medicina; no es necesario reiterar cuántas vidas han salvado y cuan significativamente han contribuido a la lucha contra las enfermedades infecciosas que eran las principales causas de morbilidad y mortalidad humana [Clardy *et al.*, 2009; Aminov, 2010].

Lamentablemente uno de los efectos colaterales del mal uso o abuso de los antibióticos es que las bacterias se vuelvan resistentes a sus efectos. A día de hoy la evolución de la resistencia a los antibióticos es un grave problema sanitario, el surgimiento de patógenos resistentes a múltiples fármacos es el causante de elevadas tasas de mortalidad. Cada año alrededor de 25.000 pacientes en la Unión Europea (UE) mueren por una infección bacteriana multi-resistente, los costos económicos estimados para dichas infecciones representan costes y pérdidas de productividad de 1.500 millones de euros como mínimo cada año en asistencia sanitaria adicional [Aminov, 2010].

Los patógenos se vuelven resistentes a los compuestos antimicrobianos por al menos uno los siguientes mecanismos: 1) produciendo enzimas que destruyan o desactiven al compuesto; 2) disminuyendo o impidiendo por completo la entrada del compuesto al interior de la célula; 3) modificando el receptor diana impidiendo o

disminuyendo la efectividad de unión del compuesto. 4) alterando la química metabólica o abandonando los pasos metabólicos sensibles al compuesto; 5) bombeando al compuesto fuera de la célula antes de que pueda cumplir su objetivo (eflujo). Entre las bacterias, los organismos individuales pueden adquirir resistencia de dos maneras posibles: a través de nuevas mutaciones en los genes cromosomales o adquiriendo plásmidos extracromosomales, denominados plásmidos R, por medio de la transferencia horizontal de ADN [Bauman, 2007; Madigan, 2009].

2.1.1 Transferencia horizontal genética

La transferencia horizontal de genes (HGT *horizontal gene transfer*) también conocida como transferencia lateral es un mecanismo natural por el cual un organismo es capaz de transferir directamente fragmentos de su ADN a otro organismo que no es su descendiente. El nombre de transferencia horizontal o lateral se acuñó en contraposición a la transferencia vertical que sucede cuando los organismos reciben el material genético de sus progenitores.

La transferencia horizontal de ADN tiene un efecto profundo en la evolución genética de las bacterias, permitiendo que individuos contemporáneos de diferentes especies o inclusive lejanos filogenéticamente compartan genes. Una vez que una bacteria adquiere material genético foráneo puede o no estabilizarse. Si se estabiliza, este material será heredado por su progenie, de esta manera la adquisición de ADN foráneo mediante transferencia horizontal puede pasar a la descendencia por transferencia vertical [Schaechter *et al.*, 2006].

La HGT es responsable de la rápida y amplia propagación de los determinantes genéticos de resistencia a diversos antibióticos. Este hecho sumado a la subsecuente selección impuesta por el mal uso de antibióticos amenaza con socavar la efectividad de los mismos [Gogarten *et al.*, 2009]. La HGT puede ocurrir mediante tres mecanismos fundamentales: transformación, transducción y conjugación (Fig. 2.1) [Schaechter *et al.*, 2006].

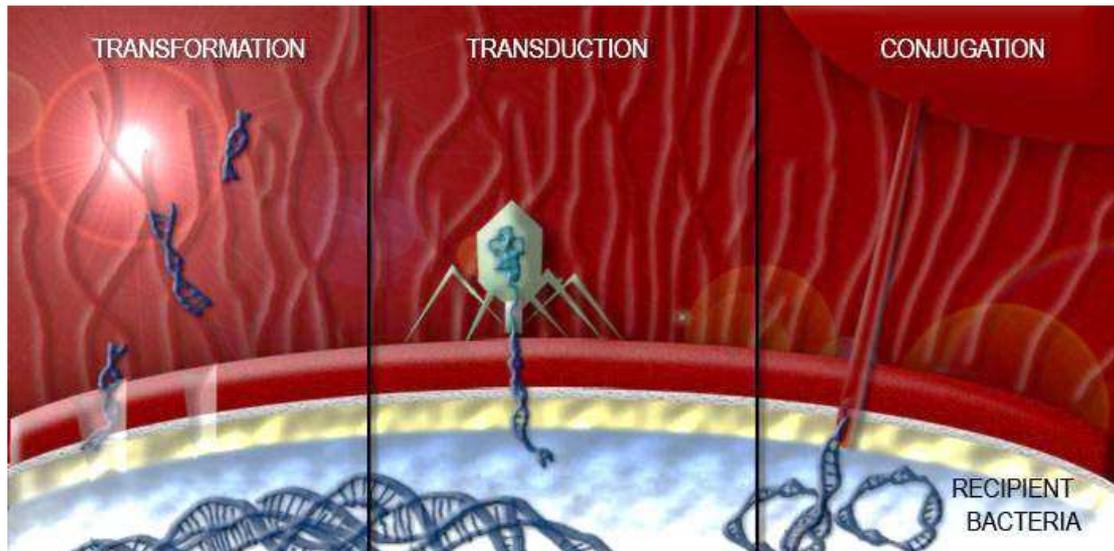


Fig. 2.1 Mecanismos de transferencia genética horizontal. Transformación mediante la adquisición de ADN desnudo del medio, transducción a través de un bacteriófago que actúa como vector de transmisión y conjugación mediada por un plásmido. Tomada de la Web6

La transformación involucra la introducción de ADN desnudo (sin proteínas asociadas) desde el exterior de la célula. Las bacterias receptoras capaces de ser transformadas, natural o artificialmente, se denominan competentes. El ADN libre puede ser el resultado de la muerte celular o de la secreción activa de la célula donante. Si el ADN transformado no posee características autoreplicativas podría integrarse mediante recombinación homóloga al genoma huésped, dependiendo de la similitud de secuencia; esto supone que la transformación efectiva queda limitada a especies bacterianas estrechamente relacionadas.

La transducción es la transferencia de material genético de una bacteria a otra mediada por bacteriófagos (virus bacteriano). Este mecanismo es la consecuencia de errores que ocurren regularmente durante el ensamblaje de los viriones. Las partículas defectivas resultantes se denominan partículas transductoras y contienen ADN de la célula huésped que reemplaza total o parcialmente el ADN del bacteriófago. Al igual que en la transformación, el ADN transducido debe ser integrado en el genoma de la célula receptora limitando este mecanismo a especies filogenéticamente cercanas.

La conjugación es la transferencia unidireccional de material genético a través del contacto directo entre la bacteria donante y la bacteria receptora. Este mecanismo depende de una molécula de ADN circular extracromosomal llamada plásmido que tiene la capacidad de autoreplicarse. Algunos de ellos, llamados plásmidos de amplio espectro o promiscuos, pueden ser transferidos entre individuos de diferentes especies

o incluso entre individuos de diferentes reinos como es el caso del plásmido Ti de *Agrobacterium* que se transfiere a células vegetales.

2.1.2 Plásmidos bacterianos y su clasificación

Los plásmidos bacterianos son moléculas de ADN extracromosomales, que se definen como elementos que codifican para funciones prescindibles dentro de la célula, si bien pueden proporcionar ventajas evolutivas como por ejemplo capacidad para fijar nitrógeno o resistencia a metales pesados, radiación o antibióticos. No son esenciales para la viabilidad del huésped.

Poseen la capacidad de autoreplicarse de forma controlada y permanecer establemente en las poblaciones bacterianas mediante la transferencia vertical y horizontal. Suelen ser circulares y bicatenarios, variando en su tamaño desde menos de 1kb hasta más de 250kb y en el número de copias en el huésped desde una sola copia hasta varios cientos por célula, pudiendo constituir hasta el 25% del genoma bacteriano. La supervivencia de los plásmidos depende en gran medida de su capacidad para mantener un número de copias constantes, que se consigue mediante circuitos de regulación negativa del inicio de la replicación [del Solar *et al.*, 1998].

En cuanto a su clasificación se pueden organizar según su capacidad de movilización, pudiendo ser conjugativos, es decir que se auto-transmiten ya que contienen toda la información genética necesaria implicada en la transferencia por conjugación; o movilizables, que carecen de dicha información y por ende solo pueden ser transferidos cuando coexisten en una misma célula con plásmidos conjugables que aportan la maquinaria necesaria para su transmisión. Otra forma de clasificar los plásmidos bacterianos es según su mecanismo de replicación [del Solar *et al.*, 1998; Francia *et al.*, 2004; Espinosa, 2013].

2.1.3 Mecanismos de replicación

Existen tres tipos de mecanismos de replicación denominados de tipo theta, desplazamiento de cadena y de círculo rodante [del Solar *et al.*, 1998].

El mecanismo de tipo theta lleva este nombre por la apariencia característica (θ) que adoptan los intermediarios de replicación al ser observados por microscopía electrónica. Este método de replicación requiere la separación de las hebras parentales, la síntesis de un *primer* de ARN y la iniciación de la síntesis de ADN mediante la extensión covalente del *primer*

de ARN. La síntesis es continua en una de las hebras y discontinua en la otra, aunque la síntesis de ambas parece ser un evento acoplado.

La replicación por desplazamiento de cadena requiere de la acción de tres proteínas codificadas por el propio plásmido que promueven la apertura del origen de replicación y sintetizan *primers* en las dos hebras, desde donde comienza la polimerización de ADN en ambas direcciones de manera continua. Este mecanismo de replicación se asocia con vectores de amplio espectro posiblemente debido a que no requiere de la maquinaria celular del huésped.

El tercer mecanismo de replicación de plásmidos se denomina de círculo rodante, es ubicuo en bacterias Gram-positivas y ha sido descrito también en bacterias Gram-negativas, archaea y algunos virus. La replicación de círculo rodante será explicada en detalle para el plásmido pMV158 en los siguientes apartados.

2.1.4 Plásmido pMV158

El plásmido pMV158 aislado por primera vez de *Streptococcus agalactiae* por Burdett, 1980 es el plásmido prototípico de la super-familia pMV158. Tiene un tamaño de 5536bp y puede establecerse en una amplia variedad de huéspedes, tanto en bacterias Gram-positivas como Gram-negativas. Actualmente ha sido identificado en firmicutes (*S. mutants*, *S. thermophilus*, *S. oralis*, *S. sanguis*, *S. gordonii*, *Bacillus subtilis*, *B. stearothermophilus*, *B. cereus*, *Listeria*, *Staphylococcus aureus*, *Lactococcus lactis*, *Enterococcus faecalis* y *Clostridium*), actino-bacterias (*Corynebacterium glutamicum*, *Brevibacterium*) y en la gammaproteobacteria *E. coli*.

Es un plásmido movilizable por lo que puede ser transmitido con la ayuda de plásmidos conjugativos, como los de la familia pIP501/pAM β 1 en Gram-positivas o los plásmidos RP4 (IncP) y R388 (incW) en Gram-negativas, o bien por el mecanismo de transformación y se replica mediante el mecanismo del círculo rodante [Francia *et al.*, 2004; Espinosa, 2013].

La organización genética del plásmido pMV158 está compuesta por tres módulos: el módulo DET (*antibiotic resistance determinant*) que determina la resistencia al antibiótico tetraciclina, el módulo MOV (*mobilization*) compuesto por el origen de transferencia (oriT) y los genes *mob* involucrados en la movilización conjugativa, y el módulo LIC (*leading strand initiation and control*) que contiene el origen de replicación de la cadena líder (*dso*, *double-strand origin*) y los genes de replicación (RepB) y su control. Este plásmido cuenta además con dos orígenes de replicación de cadena retrasada (*sso*, *single strand origin*) (Fig. 2.2) [del Solar *et al.*, 1993].

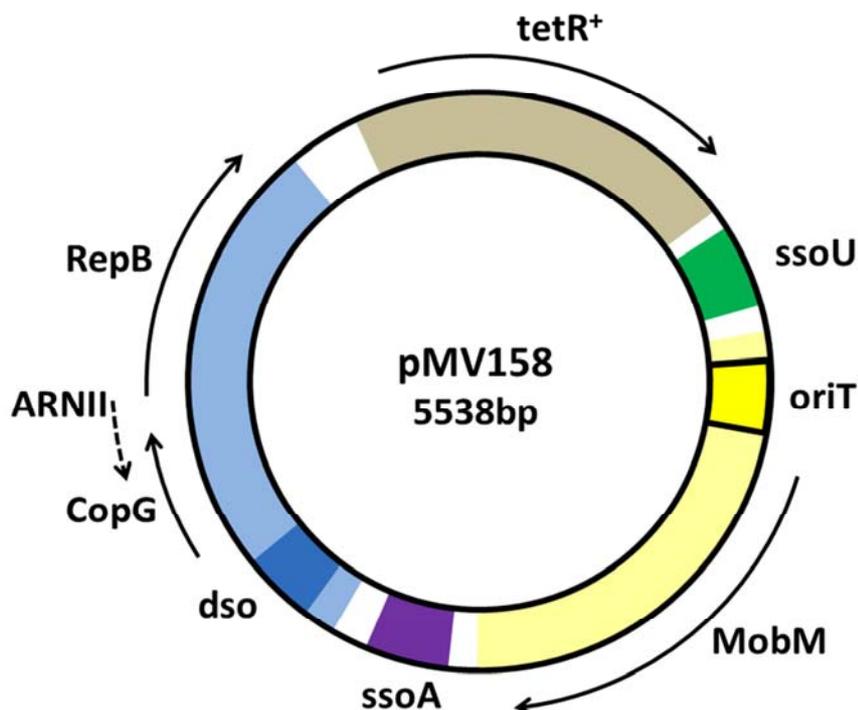


Fig. 2.2: Mapa del plásmido pMV158. En el esquema se muestran el módulo DET (canela) que codifica para la resistencia a tetraciclina, el módulo MOB (amarillo claro) que codifica para la proteína responsable de la movilización MobM y que aloja también al oriT (amarillo), y el módulo LIC (azul) con el operón CopG-RepB. Las flechas señalan la dirección de la transcripción y el ARNII (flecha discontinua). También se muestran los orígenes de cadena doble, dso en azul oscuro y de cadena simple, ssoA en violeta y ssoU en verde.

2.1.5 Replicación de círculo rodante (RCR)

La replicación de círculo rodante (RCR) es unidireccional y se la considera asimétrica ya que la síntesis de la cadena líder y de la cadena retrasada no está acoplada (Fig. 2.3) [del Solar *et al.*, 1993; del Solar *et al.*, 1998; Khan, 2000; Espinosa, 2013].

El primer paso, es decir la síntesis de la cadena líder, comienza cuando la proteína iniciadora de la replicación (RepB) interactúa con el origen de replicación de doble cadena (dso) del plásmido a través de un reconocimiento específico de secuencia. La interacción de la RepB con el dso promueve la curvatura del ADN permitiendo la extrusión de una horquilla (*hairpin*) que expone el sitio de corte o *nick site* localizado en el *loop* de simple cadena de la horquilla. RepB presenta actividad tipo topoisomerasa que le permite escindir al ADN generando un aducto inestable entre la proteína y el ADN. A su vez, carece de actividad helicasa por lo que se reclutan de la maquinaria celular del huésped proteínas de unión al ADN de simple cadena (SSB, *single-stranded DNA binding*) y una helicasa (probablemente PcrA) para estabilizar y desenrollar al ADN a medida que la replicación avanza. El corte en el ADN realizado por RepB genera un

extremo 3'-OH libre que actúa como *primer*, desde donde la ADN polimerasa del huésped comienza la síntesis de la cadena líder, esta elongación (acompañada por el desplazamiento de la cadena parental no molde) continúa hasta que se alcanza el dso reconstituido, donde culmina la síntesis de la cadena líder. Es en este punto donde RepB, mediante reacciones de transferencia, liga la nueva cadena y corta su unión con la cadena desplazada ligándola también. El producto final de este paso es una molécula de ADN de doble cadena formada por la recientemente sintetizada cadena líder y la hebra parental y un intermediario de cadena simple correspondiente a la otra hebra parental.

Por lo tanto, el segundo paso de la RCR es la conversión a doble cadena del intermediario antes mencionado. Este proceso comienza a través del reconocimiento del sso por parte de la ARN polimerasa del huésped responsable de la síntesis de un *primer* de ARN desde donde la ADN polimerasa cataliza la síntesis de la cadena retardada. Finalmente, el último paso corresponde al superenrollado de los productos de replicación por parte de la ADN girasa del huésped.

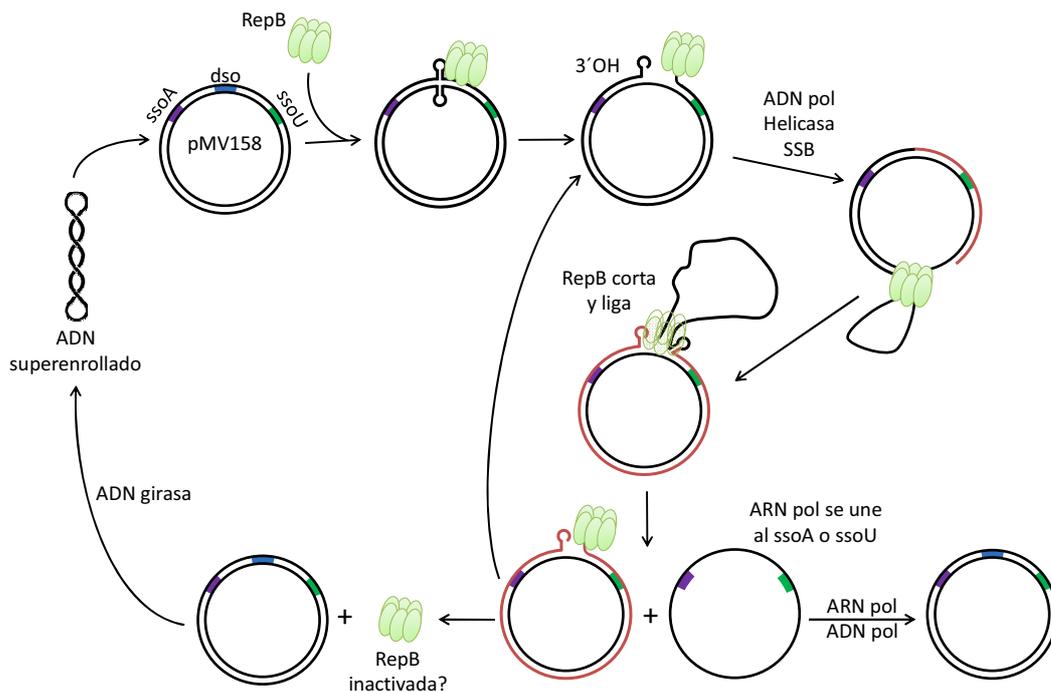


Fig. 2.3 Modelo de la replicación de círculo rodante. Esquema de la replicación del plásmido pMV158 llevada a cabo por la proteína hexamérica RepB, los detalles se explican en texto principal.

2.1.6 Regulación de la RCR

Para co-existir establemente dentro del hospedador los plásmidos bacterianos deben mantener un número de copias constante por lo que ejercen un estricto control sobre los eventos replicativos. En el caso del

plásmido pMV158 el factor limitante es la disponibilidad de la proteína iniciadora RepB, cuya expresión esta modulada por la acción de los productos de dos genes: un ARN antisentido denominado ARNII y la proteína represora CopG (Fig. 2.4).

CopG es una proteína de tan solo 45 aminoácidos que ejerce control a nivel de la transcripción de RepB, los genes *copG* y *repB* se co-transcriben a partir de un mismo promotor sobre el que opera CopG reprimiendo la síntesis de RepB así también como la suya propia. En cambio, ARNII es un ARN antisentido de 48nt de corta vida (vida media de 1min) que actúa a nivel de la traducción, el gen *arnII* se superpone en la región intergénica entre *copG* y *repB* y se transcribe en dirección opuesta siendo su secuencia complementaria al ARN mensajero (mARN) *copG-repB* por lo que inhibe su traducción [del Solar *et al.*, 1995; Gomis-Ruth *et al.*, 1998; Espinosa, 2013].

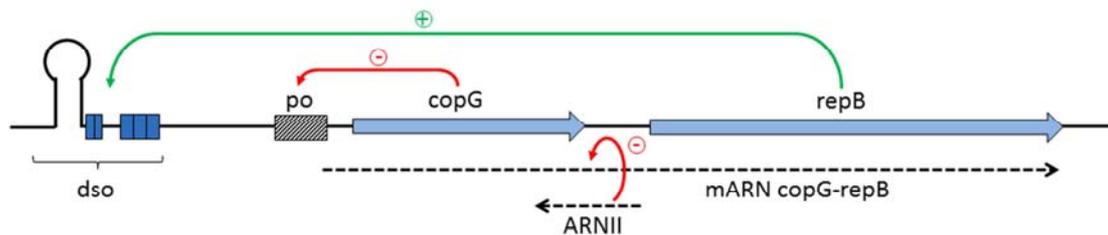


Fig. 2.4 Regulación de la replicación. Esquema de los circuitos de control que operan en la replicación del plásmido pMV158, se representan: los genes *copG* y *repB* (flechas celestes), el promotor que controla su expresión (*po*, cuadro rayado), los ARNs (flechas discontinuas) y las regiones *bind* (cuadros azules) y *nic* (horquilla) del *dso*. La actividad positiva (+ en verde) y negativa (- en rojo) de los productos génicos sobre sus dianas se esquematiza con flechas.

2.1.7 Dso (double-strand origin)

El *dso* es el origen de replicación de la cadena líder y puede ser dividido física y funcionalmente en dos loci denominados *bind*, donde se une RepB y *nic*, donde RepB cataliza el corte de la cadena en el sitio específico de corte o *nick site*.

El locus *bind* se compone de 3 repeticiones directas de 11bp cada una localizadas 84bp corriente abajo (*downstream*) del sitio de corte o *nick site*. Durante mucho tiempo fue un misterio como RepB se une en un lugar y ejerce su actividad en otra región alejada, a pesar de que la unión de RepB al ADN produce una curvatura en el mismo, no era suficiente para explicarlo. No fue hasta 2007 cuando Ruiz-Masó y su equipo [Ruiz-Masó *et al.*, 2007] descubrieron un sitio de unión secundario que presenta menor afinidad formado por dos repeticiones directas de 7bp cada una localizadas justo al lado del sitio *nic*. A partir de este descubrimiento las repeticiones

localizadas en el locus *bind* se denominan DDR (*distal direct repeats*) y las del sitio *nic* PDR (*proximal direct repeats*) y se postula el siguiente modelo: RepB se une a las DDR aumentando la concentración local de la misma, como consecuencia la proteína se une ahora a las PDR y promueve el cambio conformacional del ADN generando la extrusión cruciforme de la horquilla en el sitio *nic*, exponiendo en el *loop* de dicha horquilla el sitio de corte en configuración de simple cadena, permitiéndole a RepB cortar en el enlace fosfodiéster específico (5'-TACTACG/AC-3') mediante el ataque nucleofílico de la tirosina catalítica (Y99) (Fig. 2.5) [del Solar *et al.*, 1993; Moscoso *et al.*, 1997; Espinosa, 2013].

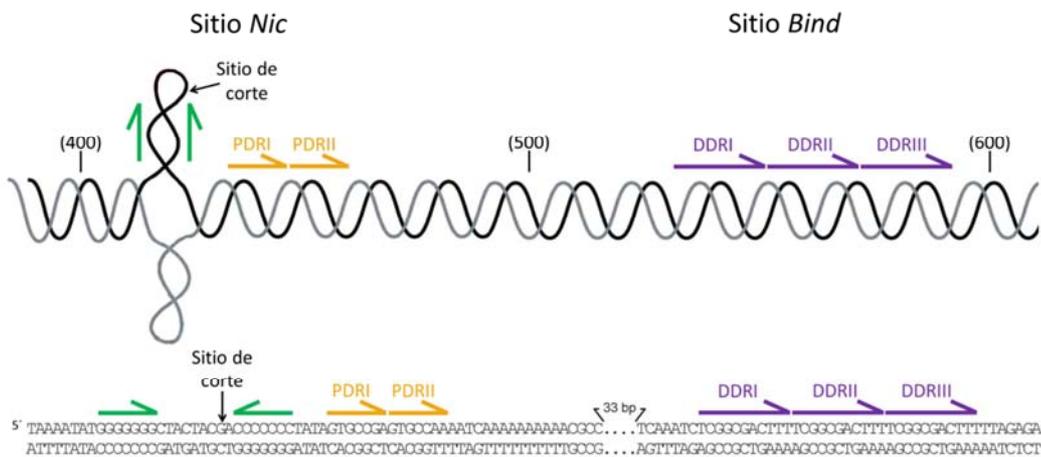


Fig. 2.5 Dso del plásmido pMV158 y las regiones de interacción con RepB. Representación esquemática (arriba) y secuencia nucleotídica (abajo) del dso del plásmido pMV158, se muestran los componentes del sitio *nic*: sitio de corte, repeticiones invertidas que forman la horquilla y las repeticiones directas proximales (PDR); y las repeticiones directas distales (DDR) del sitio *bind*. Como referencia entre paréntesis se expresan las coordenadas de la secuencia del plásmido. Imagen modificada de Boer *et al.*, 2009 y Ruiz-Maso *et al.*, 2004.

2.1.8 Sso (*single strand origin*)

La cadena parental que es desplazada al avanzar la síntesis de la nueva hebra de ADN es convertida a doble cadena utilizando el origen de replicación de la cadena retrasada o sso, que es reconocido por la ARN polimerasa del huésped. La tasa de conversión del intermediario de simple cadena a doble cadena depende de la eficiencia con que las proteínas del huésped reconocen un determinado sso. Se han descrito 4 tipos diferentes de sso: ssoA, ssoW, ssoT y ssoU, los dos primeros funcionan eficientemente solo en los huéspedes nativos mientras que los dos últimos son eficientes en diferentes huéspedes. El plásmido pMV158 es portador de dos sso uno de tipo A y otro tipo U, lo que determina en parte el amplio espectro del mismo [Khan, 2005; Espinosa, 2013].

2.1.9 RepB

La proteína iniciadora RepB pertenece a una superfamilia de proteínas que comparte el mecanismo de actividad endonucleolítica basado en una tirosina catalítica y un metal divalente, siendo capaces de cortar y unir al ADN de forma similar a una topoisomerasa tipo I. Esta superfamilia se caracteriza por la presencia de dos dominios particulares: el motivo HUH, formado dos residuos de histidina que rodean a un residuo hidrofóbico voluminoso en la secuencia primaria de la proteína (His-hidrofóbico-His, HUH) cuya función es la de modular la unión del catión divalente necesario para la actividad; y el motivo Y que contiene una o dos tirosinas separadas por un número variable de residuos, responsable de la unión al ADN.

Superfamilia HUH

La superfamilia HUH incluye proteínas iniciadoras de la replicación de plásmidos de eubacterias, de colifagos, de virus de ADN de simple cadena, de geminivirus y de parvovirus de mamíferos, denominadas en su conjunto como proteínas de clase Rep. A su vez, si modificamos la orientación relativa entre los dominios esta superfamilia abarca también a las proteínas de clase Mob, donde el dominio Y se encuentra corriente arriba (upstream) del dominio HUH y viceversa para la clase Rep. Esto se debe a eventos evolutivos de permutación circular, lo sorprendente de este fenómeno es que no solo los dominios son superponibles sino también que la posición relativa de la tirosina catalítica es prácticamente idéntica (Fig. 2.6) [Dyda *et al.*, 2003; Guasch *et al.*, 2003].

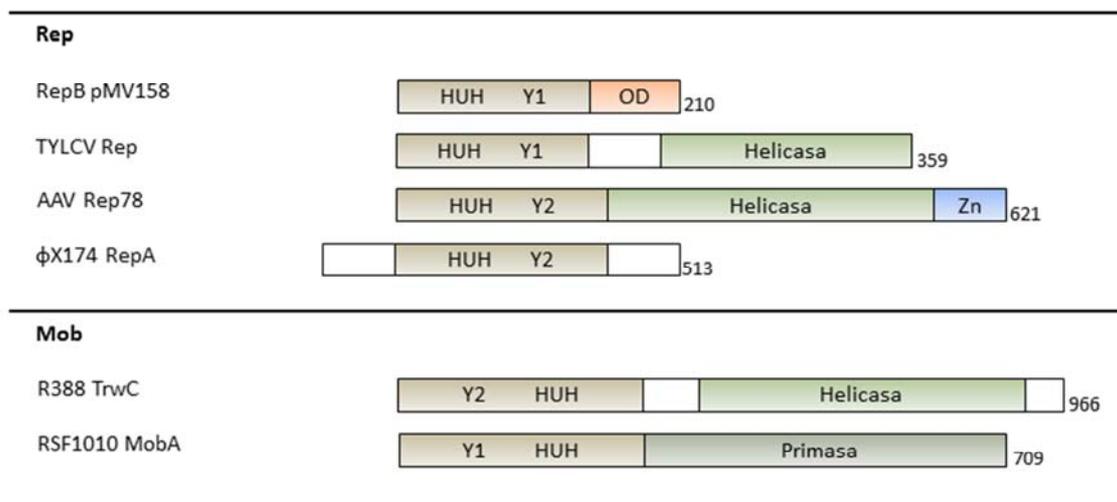


Fig. 2.6 Organización de las proteínas representativas de la superfamilia HUH. Contienen los dominios HUH, helicasa, primasa, dominio de oligomerización (OD) y de unión al Zn (Zn), el tamaño de cada proteína se expresa en número de aminoácidos. Imagen modificada de Chandler *et al.*, 2013.

Debido a la gran variedad de procesos en los que están involucradas las proteínas de esta superfamilia no es sorprendente la presencia de dominios anexos al HUH, en muchos casos con función desconocida aunque los dominios helicasa, primasa, de unión al Zn o de multimerización son recurrentes [Chandler *et al.*, 2013].

Las endonucleasas HUH catalizan la ruptura y la unión de ADN de simple utilizando la(s) tirosina(s) del motivo Y para crear un intermediario fosfotirosina 5' y un OH libre 3' en el sitio de corte. La unión covalente entre el ADN y la tirosina catalítica es tradicional para esta familia, si bien no está descrito en la literatura comunicaciones personales con el grupo de Gloria del Solar, José Ruiz Maso y Lorena Bordanaba del CIB de Madrid, con quienes colaboramos, nos confirman la existencia de un intermediario fosfotirosina estable. Por otro lado, el extremo OH 3' puede servir como *primer* para la replicación, así como también puede actuar como nucleófilo para la transferencia de cadena para resolver el intermediario fosfotirosina en la etapa de terminación de la replicación. Estas enzimas requieren metales divalentes para facilitar el corte localizando y polarizando el enlace fosfodiéster escindible. Dependiendo de la enzima Mg^{2+} , Mn^{2+} u otros iones pueden ser utilizados *in vitro*, aunque normalmente Mg^{2+} , Mn^{2+} son los cofactores fisiológicos. El par de histidinas del dominio HUH proporciona dos de los tres ligandos necesarios para la coordinación del catión, la ubicación y la identidad del tercer ligando (invariablemente un residuo polar) puede ser ácido glutámico, ácido aspártico, histidina o glutamina variando a través de la superfamilia [del Solar *et al.*, 1998; Chandler *et al.*, 2013].

Estructuralmente el dominio HUH se caracteriza por cinco hebras antiparalelas que forman una lámina β flanqueada por un número variable de hélices α , la tirosina activa se localiza en una de estas hélices α . Las tres hebras centrales de la lámina β tienen la misma disposición y topología en ambas clases de proteínas, con el motivo HUH localizado en el medio. En cambio las hebras flanqueantes de la lámina β son proporcionadas por el extremo C-terminal del dominio en las proteínas de clase Rep, y por el extremo N-terminal en las de clase Mob (Fig. 2.7) [Gomis-Ruth *et al.*, 2006; Boer *et al.*, 2009; Chandler *et al.*, 2013].

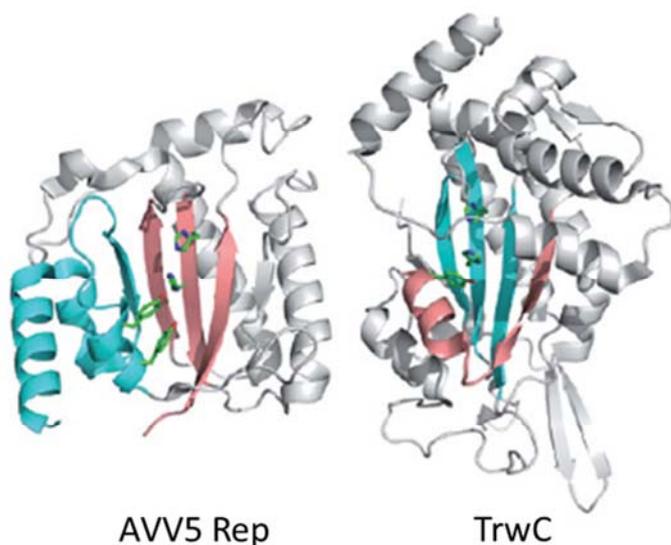


Fig. 2.7 Dominio HUH de clase Rep y Mob. La clase Rep representada por el dominio del virus adeno-asociado 5 (AAV5) (PDB 1M55) y la clase Mob por el dominio de la proteína TrwC del plásmido R388 (PDB 1OMH) ambas presentan una lámina β formada por cinco hebras con un centro $\beta\alpha\beta$ que lleva el motivo HUH en la hebra central. Debido a la organización del dominio característica de cada clase la posición del centro difiere ($\beta\alpha\beta\beta\alpha\beta$ para Rep y $\beta\alpha\beta\alpha\beta\beta$ para Mob). En rosa se indican los residuos amino-terminales y en celeste los carboxi-terminales. Imagen tomada de Chandler *et al.*, 2013.

Las proteínas de esta familia deben, además de catalizar el corte inicial en el dso, participar en el proceso de terminación, realizando el corte de la cadena líder una vez que ha sido completamente replicada y la ligación de la cadena desplazada para su liberación. De este modo la proteína A de ϕ X174 que actúa como un monómero se vale de las dos tirosinas catalíticas que presenta en el sitio activo, mientras que las proteínas Rep del plásmido pT181 actúan como dímero y la única tirosina catalítica de cada monómero participa en la replicación, se ha propuesto un mecanismo similar para las proteínas Rep de la familia de plásmidos pMV158 que actúan como hexámeros [Khan, 1997; Khan, 2000]

Proteínas relacionadas estructuralmente

Es interesante que estructuras homólogas al dominio de esta superfamilia se encuentren en los dominios de unión al ADN de proteínas virales no relacionadas con la replicación de círculo rodante, como es el caso de la proteína E1 del virus papiloma bovino (BPV) [Enemark *et al.*, 2000; Enemark *et al.*, 2002] y la proteína antígeno-T del virus simiano 40 (SV40) [Luo *et al.*, 1996]. El dominio de unión al ADN de estas proteínas virales también exhibe un plegamiento de tres capas α - β - α que contiene en su centro una lámina β antiparalela de cinco hebras.

La homología estructural que se observa no se ve reflejada a nivel aminoacídico ya que entre ellas no comparten una identidad de secuencia primaria significativa, algunas de las proteínas virales no presentan el dominio HUH y ni las tirosinas del sitio activo y por ende carecen de la actividad nucleasa que caracteriza a las proteínas iniciadoras de la superfamilia RCR. Otra distinción estructural es que en las proteínas Rep el

sitio activo se localiza en el interior de una hendidura, mientras que en las proteínas virales que carecen del mismo la hendidura se pierde debido a que las extensiones N-terminales ocupan este espacio (Fig. 2.8) [Hickman *et al.*, 2002].

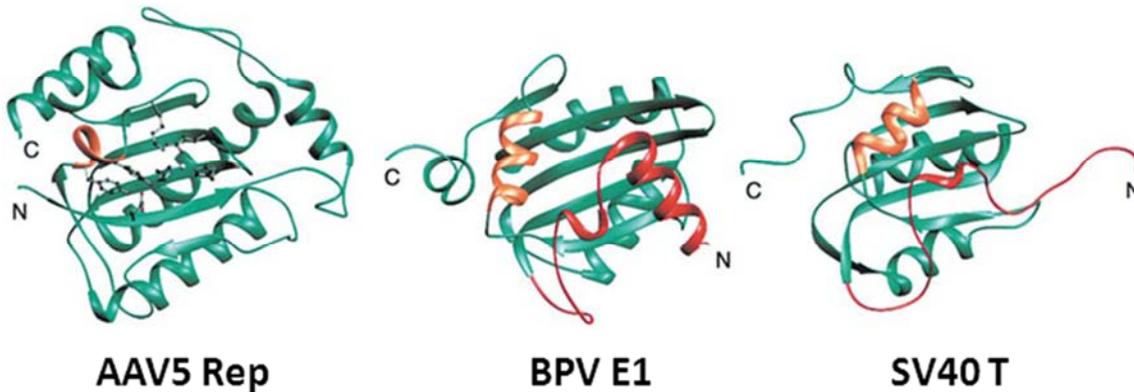


Fig. 2.8 Comparación de estructuras relacionadas. En la estructura del dominio de la proteína Rep del virus adeno-asociado 5 (AAV5) se encuentran representados los residuos que forman parte del sitio activo, mientras que en los dominios de unión al ADN de la proteína E1 del virus papiloma bovino y el antígeno T del virus simiano 40 los extremos N-terminales ocupan esta región. Imagen tomada de Hickman *et al.*, 2002.

RepB

RepB presenta actividad nucleasa, ligasa y transferasa *in vitro* tanto en ADN monocatenario como superenrollado. La actividad nucleasa está mediada por la tirosina 99. Si bien presenta otra tirosina cercana (Fig. 2.9) la misma no se encuentra correctamente localizada para la catálisis y solo la Y99 es esencial para la actividad, ya que su reemplazo por otro residuo produce la pérdida de la misma [Moscoso *et al.*, 1997; Chandler *et al.*, 2013]. En cuanto al uso de iones metálicos se comprobó que RepB solo es activa en presencia de Mn^{2+} o Co^{2+} *in vitro* [Boer *et al.*, 2009]. Por otro lado, en lo que respecta a los modelos de terminación de la replicación se demostró que la transferencia de cadena ocurre en dos pasos o en un número par de eventos de transesterificación [Moscoso *et al.*, 1997; del Solar *et al.*, 1998].

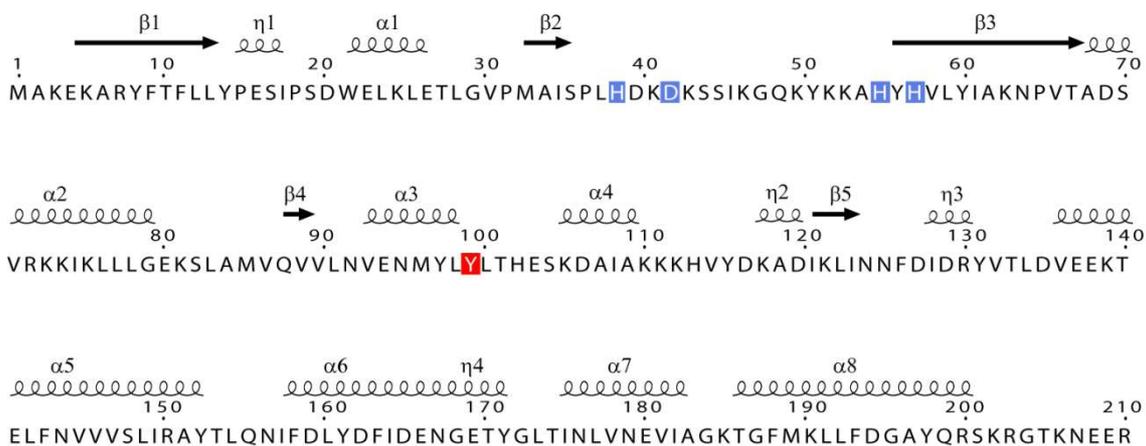


Fig. 2.9 Secuencia primaria de RepB. Se incluye la anotación de la estructura secundaria, η hace referencia a las hélices 3_{10} . Encuadrados en azul se encuentran los residuos que participan en la coordinación del Mn^{2+} y en rojo la tirosina catalítica. Imagen generada mediante el servidor ESPript [Gouet *et al.*, 1999].

La estructura de RepB fue resuelta en nuestro laboratorio en 2009 [Boer *et al.*, 2009], se obtuvieron dos estructuras a partir de cristales con forma trigonal y tetragonal respectivamente (Fig. 2.11), ambas presentan una estructura cuaternaria hexamérica con forma de copa. Cada monómero del hexámero consiste en dos dominios conectados por una corta región “bisagra”, donde el dominio N-terminal de unión al origen (OBD, *origin-binding domain*) porta la actividad catalítica y el dominio C-terminal es responsable de la hexamerización del complejo (OD, *oligomerization domain*).

El dominio OD (T133-R210) está formado completamente por hélices α , las hélices $\alpha 5$ y $\alpha 8$ se colocan en paralelo mientras que las hélices más cortas $\alpha 6$ y $\alpha 7$, conectadas por una hélice 3_{10} , se encuentran perpendicularmente a las dos primeras. Los OD forman un anillo toroidal de simetría séxtuple.

Los OBD en cambio, consisten de una lámina β formada por cinco hebras antiparalelas rodeada por hélices a ambos lados. Las hélices $\alpha 1$ y $\alpha 2$, en una de las caras de la lámina β , se acomodan diagonalmente con respecto a la misma. El sitio activo se encuentra al otro lado de la lámina en una cavidad rodeada por las hélices $\alpha 3$ y $\alpha 4$, la hélice $\alpha 3$ porta la tirosina catalítica (Y99), y los residuos que coordinan el ion metálico (H39, D42, H55 y H57) se ubican en las hebras $\beta 2$ y $\beta 3$ y en un loop largo y flexible que las conecta (Fig. 2.10).

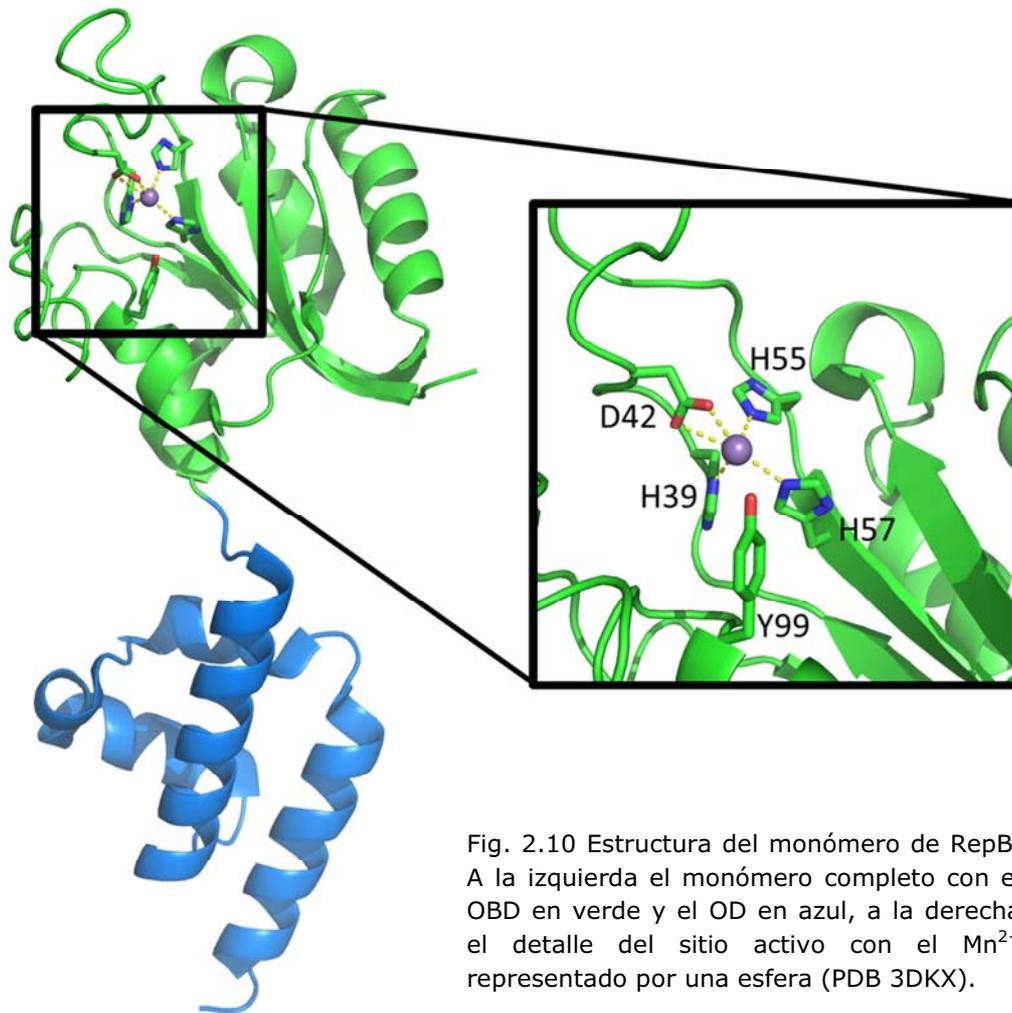
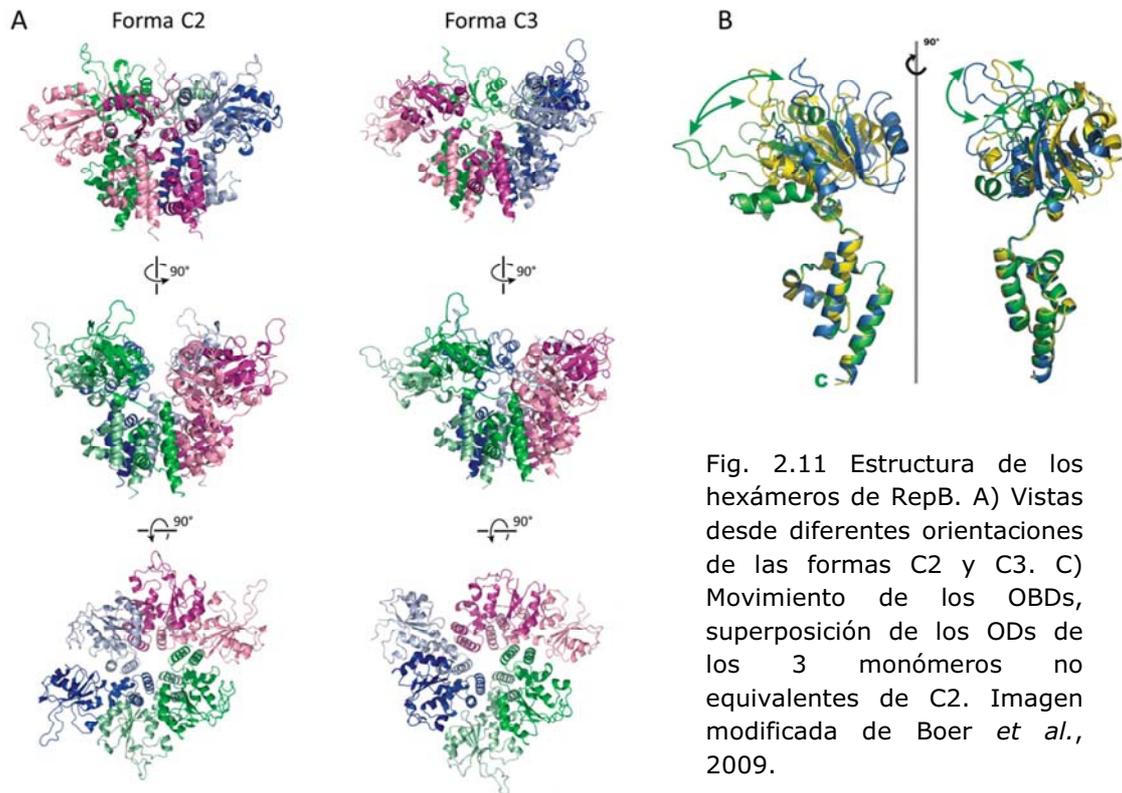


Fig. 2.10 Estructura del monómero de RepB. A la izquierda el monómero completo con el OBD en verde y el OD en azul, a la derecha el detalle del sitio activo con el Mn^{2+} representado por una esfera (PDB 3DKX).

Los OBD no siguen la simetría hexagonal de los OD, en la forma trigonal dos sets de tres OBDs se relacionan mediante simetría cristalográfica binaria (C_2), pudiendo considerarse un dímero de trímeros; mientras que en la forma tetragonal los OBDs muestran simetría ternaria (C_3) considerándose un trímero de dímeros. Esta variabilidad es el resultado de las diferentes orientaciones y localizaciones de los OBDs con respecto al anillo formado por los ODs, los residuos alrededor de la V132 actúan como una bisagra entre ambos dominios permitiendo un rango de movimiento de hasta 55° (Fig. 2.11). Esta flexibilidad favorecería la reorientación de los OBDs involucrados en la catálisis, permitiendo que se aproximen para llevar a cabo las reacciones de corte y ligación necesarias para la terminación del proceso.



Teniendo en cuenta que cada OBD puede unirse a una de las tres repeticiones directas del sitio *bind* [Ruiz-Maso *et al.*, 2007], RepB₆ podría ser capaz de albergar hasta dos moléculas de ADN que contengan esta secuencia; sin embargo se demostró que RepB₆ puede acomodar solo una molécula de ADN que contenga el sitio *bind* [Boer *et al.*, 2009].

En base a los resultados obtenidos al analizar el hexámero de RepB mediante microscopía electrónica en presencia y ausencia del sustrato, donde se identifica una masa extra en el centro del complejo (Fig. 2.12A), se propone un modelo en el que el ADN de simple cadena pasa por el centro del hexámero. La similitud estructural con otros sistemas como el la helicasa E1 del BPV [Enemark *et al.*, 2006] o el antígeno T del SV40 [VanLoock *et al.*, 2002] en los que el ADN ocupa el poro central respaldan este modelo de acción (Fig. 2.12B). A su vez, dado que *in vivo* el hexámero co-existe con formas monoméricas y dimericas de RepB se postula que el ensamblado del complejo en el origen del plásmido se produce por la adición consecutiva de monómeros o dímeros alrededor del sitio de unión [Boer *et al.*, 2009].

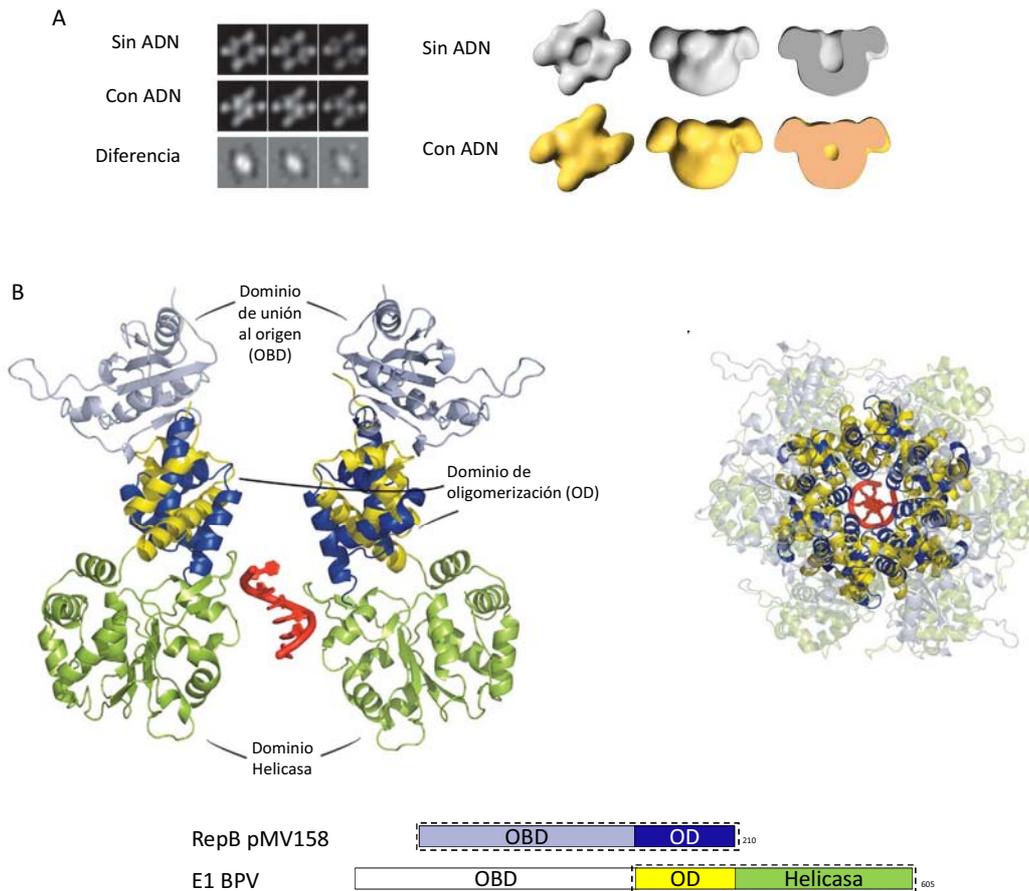


Fig. 2.12 Modelo del canal central del hexámero de RepB. A) Resultados del análisis de RepB en presencia y ausencia de sustrato. A la derecha se muestran las fotomicrográficas y a la izquierda los modelos 3D. B) Comparación estructural de E1 BPV y RepB. El diagrama de abajo representa la organización de dominios de ambas proteínas, remarcado con un cuadro discontinuo se muestran los dominios representados en el esquema siguiendo el mismo código de colores. A la derecha se muestra la superposición de dos monómeros que flanquean el canal central de RepB y E1, en rojo se muestra el ADN en el centro de E1, a la izquierda la vista superior de la superposición de todos los monómeros de los hexámeros de ambas proteínas, con los dominios helicasa y OBD en un segundo plano. Imagen modificada de Boer *et al.*, 2009.

En el trabajo de la estructura *full length* de RepB [Boer *et al.*, 2009] se caracterizó la proteína demostrando su naturaleza hexamérica mediada por los dominios OD, se describió la arquitectura del sitio activo situado en el dominio OBD y se observó la flexibilidad de los dominios OBD con respecto a los OD. Sin embargo algunos interrogantes quedaron pendientes para posteriores estudios: como es la unión de esta proteína al ADN diana y como se ajusta esto al contexto hexamérico de la proteína. Con el fin de analizar y responder a estos interrogantes es que, en el presente trabajo se decidió resolver la estructura del dominio OBD en presencia su ADN diana, la DDR.

2.1.10 Objetivos:

- Obtener cristales de RepB₁₋₁₃₂/23AB con una adecuada calidad de difracción.
- Determinar la estructura cristalográfica de RepB₁₋₁₃₂/23AB.
- Analizar en detalle la unión del dominio OBD de RepB a las DDR.
- Generar un modelo que explique el funcionamiento de la proteína en el contexto del complejo hexamérico.

2.2 Materiales y métodos

2.2.1 Ensayos cristalográficos

Preparación de la muestra de proteína

El constructo de RepB₁₋₁₃₂ purificado según el protocolo descrito por Ruiz-Maso *et al.*, 2004 fue amablemente cedido por el Dr. José L. Ruiz-Masó y la Dra. Gloria del Solar. La concentración de proteína se determinó mediante la medición de la absorbancia a 280nm con un nanodrop 1000 (Thermo Scientific), para lo que se estimó el coeficiente de extinción molar a partir de la secuencia de la proteína utilizando la herramienta ProtParam del servidor web ExpASY bioinformatics resource portal (Web7).

La muestra de RepB utilizada se encontraba en el siguiente buffer: 20mM Tris pH 8, 1mM EDTA, 5mM DTT, 5% etilenglicol y 1M NaCl a una concentración de 9,8mg/ml.

Preparación de los oligonucleótidos

Los oligonucleótidos se adquirieron en la casa comercial Biomers, las muestras, con un grado de pureza HPLC según el fabricante, se utilizaron sin mediar pasos extra de purificación, sometiéndolas a un proceso de hibridación o *annealing*.

Los oligonucleótidos 23A (TCGGCGACTTTTCGGCGACTTTT) y 23B (AAAAGTCGCCGAAAAGTCGCCGA) a 7,5mM y 14,6mM respectivamente se combinaron en una relación 1:1 y se hibridaron colocando la mezcla en un baño a 80°C durante 5 minutos para luego apagar el baño y dejar que se enfriara lentamente hasta llegar a temperatura ambiente.

El dsADN de 23bp corresponde a un fragmento del locus de unión de 42bp del plasmido pMV158 (coordenadas 533-555).

Preparación de los complejos proteína-ADN

Los complejos se prepararon, inmediatamente después del tratamiento al ADN, mezclando la proteína con el ADN en relación molar 2:1. Se diluyó la mezcla en el mismo buffer de la proteína libre de sal a fin de disminuir la concentración de la misma 4 veces y se reconcentró al volumen inicial utilizando dispositivos de filtrado Vivaspin (Sartorius) con membrana de corte de peso molecular de 3000MWCO, a 4500g y a 4°C en una centrifuga MULTIFUGE 3 L-R Heraeus.

Cristalización

Los ensayos iniciales se realizaron siguiendo la metodología descrita en el apartado 1.2.2 Ensayos cristalográficos.

Una vez encontradas las condiciones iniciales de cristalización, se procedió al escalado de los ensayos, pasando de placas de 96 pocillos a placas de 24 pocillos y gotas de 200nl (100nl de complejo + 100nl de condición) a 2 μ l (1 μ l de complejo + 1 μ l de condición) y a la optimización de las condiciones, teniendo en cuenta: el pH del buffer, la concentración del agente precipitante, la concentración de las sales y/o de los aditivos.

Los cristales obtenidos fueron retirados (pescados) con *loops* de nylon de 0,05-0,1mm o 0,1-0,2mm (Molecular Dimensions) según el tamaño de los cristales. Se crio-protegieron pasando el cristal montado en el *loop* por una mezcla de la solución madre con la adición de 10% de glicerol, previamente a sumergir el mismo en nitrógeno líquido para su almacenamiento.

2.2.2 Resolución de la estructura

Recolección y procesamiento de los datos de difracción

La difracción de los cristales se llevó a cabo en la línea XALOC del sincrotrón ALBA, Cerdanyola del Vallès (Barcelona). La recolección de datos se realizó a la longitud de onda del pico de absorción del selenio (0,97949Å) y se obtuvieron 3 conjuntos de datos:

- Set 1: $\varphi=0^{\circ}$ $\Delta\varphi=0,1$ $t= 0,4$ seg 851 imágenes
- Set 2: $\varphi=75^{\circ}$ $\Delta\varphi=0,1$ $t= 0,4$ seg 648 imágenes
- Set 3: $\varphi=139^{\circ}$ $\Delta\varphi=0,1$ $t= 0,4$ seg 400 imágenes

Los datos de la difracción fueron indexados e integrados con MOSFLM [Leslie *et al.*, 2007] y se escalaron con el programa Scala [Evans, 2006] de la suite de programas de CCP4 [Collaborative Computational Project, 1994].

Resolución y refinamiento estructural

La resolución de la estructura se llevó a cabo mediante reemplazo molecular utilizando el programa PHASER a través del CCP4, utilizando como modelo de búsqueda el fragmento OBD de RepB (código PDB 3DKX) y el ADN fue modelado a mano base a base en Coot [Emsley *et al.*, 2010].

La construcción del modelo se hizo en Coot, alternando con ciclos de refinamiento en Refmac5 [Murshudov *et al.*, 1997] utilizando restricciones de simetría no cristalográfica (NCS, por sus siglas en inglés *non-crystallographic symmetry*) siendo estrictas para cadena principal y permisivas para las cadenas laterales; así también como parámetros de TLS obtenidos mediante el servidor online TLSMD [Painter *et al.*, 2006]. La validación del modelo se realizó mediante las herramientas del programa Coot y el servidor MolProbity [Davis *et al.*, 2007] (Web8), se analizaron los parámetros geométricos más relevantes como son las distancias y los ángulos de enlace, las conformaciones de los posibles rotámeros y los diagramas de Ramachandran.

Análisis de la estructura y creación de modelos

Los parámetros geométricos del ADN se obtuvieron con el programa 3DNA [Lu *et al.*, 2003] y las interacciones proteína-ADN se analizaron utilizando el programa Nucplot [Luscombe *et al.*, 1997]. Todas las figuras fueron preparadas con PyMOL [Schrödinger].

El modelo de interacción del hexámero con las DDR se construyó a partir de la estructura resuelta en este trabajo; se aisló un OBD con su DDR correspondiente y esta unidad OBD/DDR se superpuso sobre las nueve posiciones relativas de los OBD con respecto de los OD (observadas en los dos hexámeros: 3 orientaciones del complejo C2 y seis del C3) [Boer *et al.*, 2009]. Mediante el análisis visual directo se seleccionaron las orientaciones relativas que permitían unir las tres DDR consecutivas en una misma molécula, el resultado de este montaje son cada OBD, OD y DDR ubicados espacialmente pero disociados entre sí, por lo que se realizó un proceso de minimización geométrica, refinando por Refmac5 sin datos cristalográficos.

La superposición de las proteínas virales Rep de AAV5 (1RZ9) [Hickman *et al.*, 2004] y E1 de BPV (1KS4) [Enemark *et al.*, 2002] sobre nuestra estructura se realizó utilizando el programa Coot. Se alineo la cadena principal del motivo HUH (H55-H57) con los aminoácidos correspondientes de las proteínas homólogas, siendo éstos los residuos 89-91 de la proteína Rep de AAV5 y 230-232 de la proteína E1 de BPV.

2.3 Resultados y discusión

2.3.1 Cristalización

De los ensayos iniciales con el complejo RepB₁₋₁₃₂/23AB se obtuvieron cristales múltiples con forma de aguja en la condición E9 del PAC10 (50mM MgCl₂, 50mM TRIS pH 8.0, 20% PEG3350), que fueron optimizados variando la concentración de PEG3350 al 23%, dando como resultado agujas individuales y de buen tamaño (Fig. 2.13). Los cristales se recogieron y congelaron en nitrógeno líquido, previa crioprotección del cristal mediante su inmersión en una solución crioprotectante con la misma condición donde se generó el cristal más la adición de un 10% de glicerol.

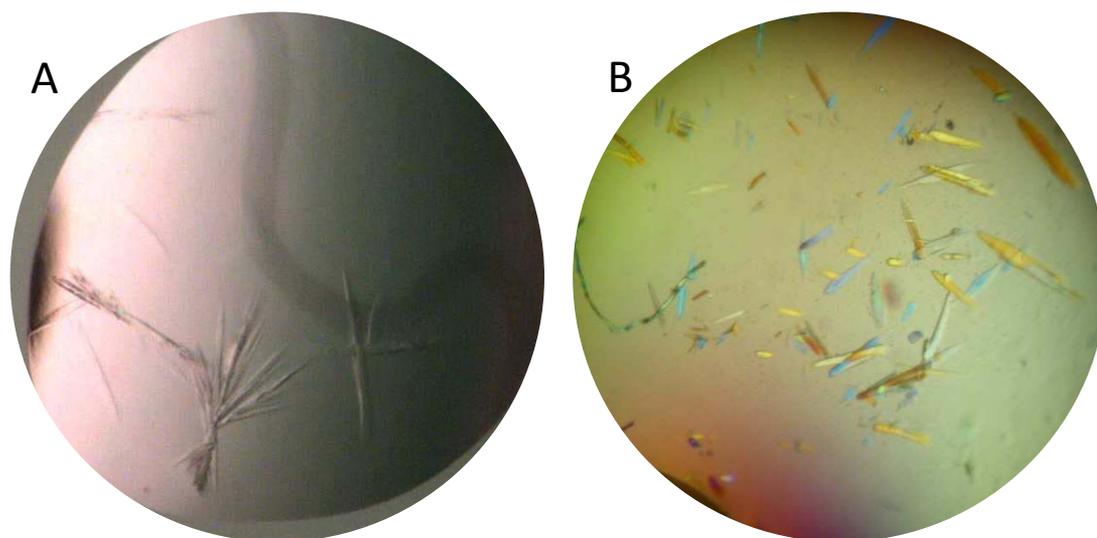


Fig. 2.13 Cristalización de RepB₁₋₁₃₂/23AB. A) Cristales obtenidos en la condición E9 del PAC10. B) Optimización

2.3.2 Determinación estructural

Difracción y procesamiento de los datos

Los datos de difracción se obtuvieron en el sincrotrón ALBA, en la línea XALOC. En total se obtuvieron 3 conjuntos de datos que se combinaron.

El cristal es monocíclico primitivo, pertenece al grupo espacial P2₁. La Fig. 2.14 muestra un ejemplo del patrón de difracción del cristal, que difracta hasta una resolución de 3Å y las estadísticas de la recolección de datos del cristal se detallan en la siguiente tabla:

Tabla 2.1: Estadísticas de la recolección de datos

Grupo espacial	P2 ₁
Dimensiones de la celda	
a b c (Å)	49.23 33.61 289.67
α β γ (°)	90.00 93.50 90.00
Longitud de onda (Å)	0.979490
Rango de resolución (Å)	25 - 3
Nº de reflexiones	
Totales (outershell)	41499 (4032)
Únicas (outershell)	17963 (2051)
Completitud (outershell) (%)	91.3 (73.9)
Rmerge	0.164 (0.530)
Intensidad promedio $\langle I/\sigma I \rangle$ (outershell)	3.7 (1.4)
Multiplicidad (outershell) (%)	2.3 (2.0)
Completitud anómala (outershell) (%)	62.5 (36.8)
Multiplicidad anómala (outershell) (%)	1.4 (1.3)

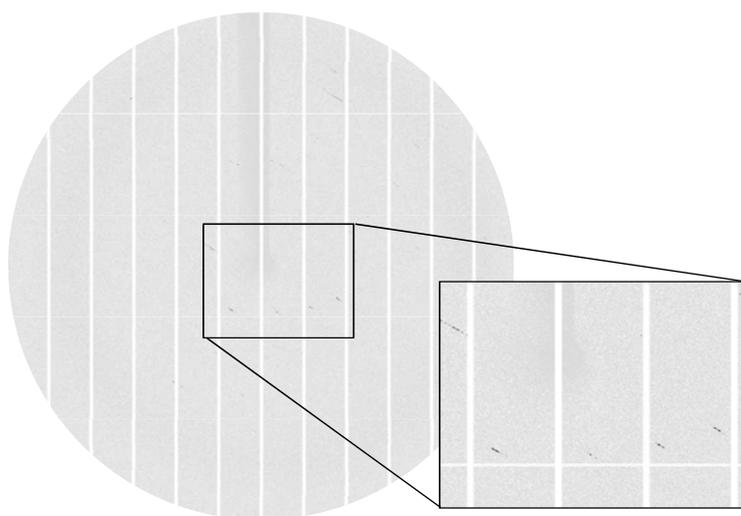


Fig. 2.14 Patrón de difracción del cristal

Resolución y refinamiento estructural

Considerando que la unidad asimétrica contiene 4 moléculas de proteína y 2 moléculas de oligonucleótidos (doble cadena) el porcentaje de solvente es de 58.86% y el coeficiente de Matthews 2.67 Å³/Da. La estructura se resolvió por reemplazo molecular utilizando como modelo inicial el OBD de la estructura de RepB resuelta por [Boer *et al.*, 2009] (código PDB 3DKX). El refinamiento se realizó iterativamente hasta que Rwork y Rfree convergieron en 0.228 y 0.296 respectivamente. Los

principales parámetros y las estadísticas finales de todos los refinamientos se muestran en la siguiente tabla:

Tabla 2.2: Estadísticas del refinamiento

Nº de reflexiones	17036
Rwork/Rfree	0.228/0.296
Nº de átomos (factor B)	6168 (84.71)
Proteína (factor B)	4292 (84.40)
ADN (factor B)	1874 (85.43)
Ligando Ca ²⁺ (factor B)	2 (72.02)
Aguas	0
Factor B Wilson (Å ²)	78.011
Desviaciones R.m.s	
Enlaces (Å)	0.010
Ángulos (°)	1.393
Ramachandran plot*	
Outliers, n (%)	13 (2.51)
Favorecidos, n (%)	443 (85.69)
Conformaciones no favorables de las cadenas laterales, n (%)*	28 (5.98)

*según Molprobit

En la estructura se identificaron perfectamente el oligonucleótido y la cadena principal de la proteína; en cuanto a las cadenas laterales de los aminoácidos no todas pudieron ser resueltas debido a la baja resolución de los datos, esto podría explicar los valores relativamente altos de Rwork y Rfree. En la Fig. 2.15 podemos observar los Ramachandran plots [Lovell *et al.*, 2003] generados por Molprobit.

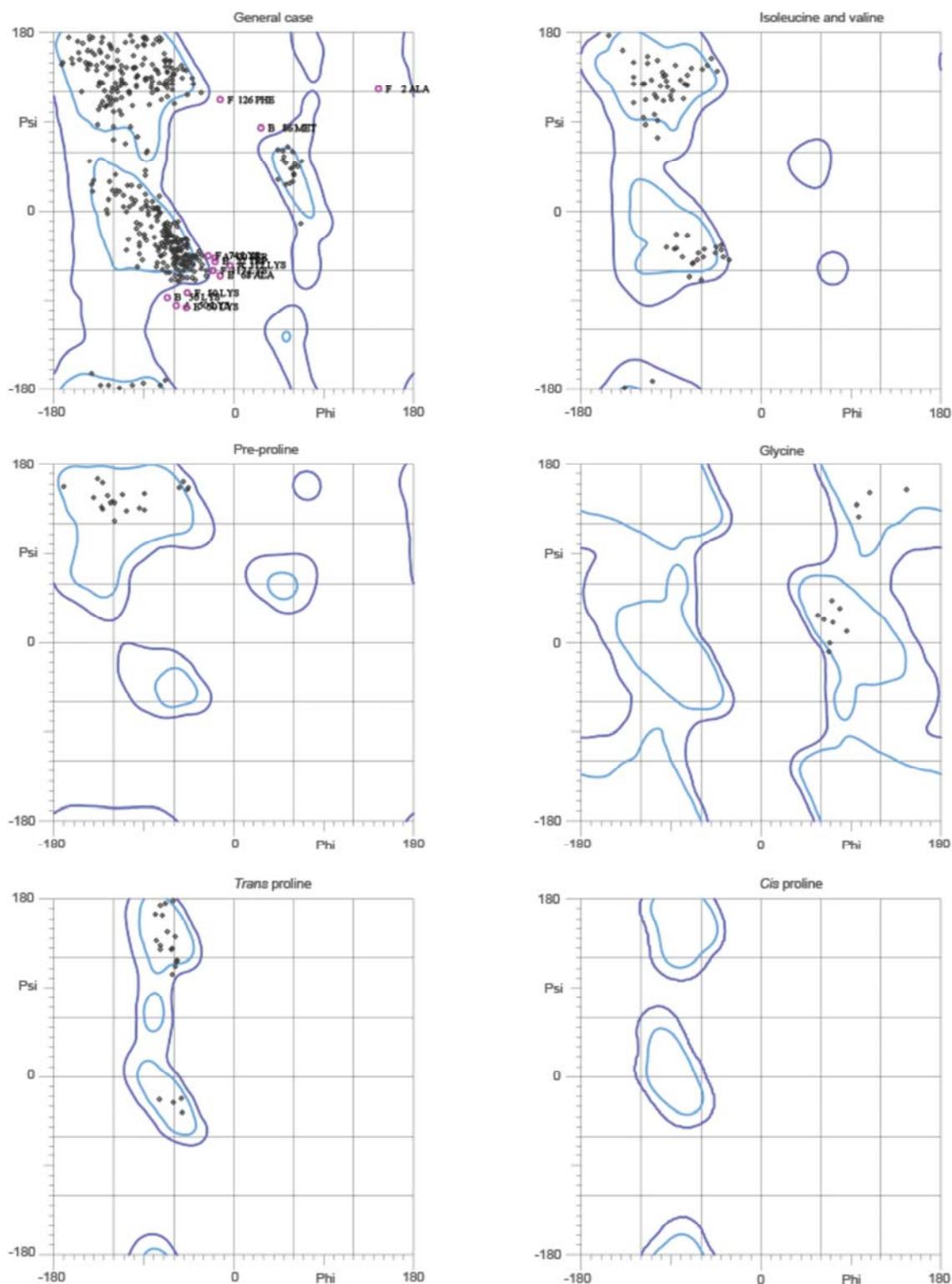


Fig. 2.15 Representación de Ramachandran para la estructura RepB₁₋₁₃₂-23AB.

La unidad asimétrica contiene 2 moléculas de ADN, localizadas longitudinalmente (es decir extremo romo con extremo romo) y en direcciones opuestas en cuanto a su secuencia, y 4 de RepB₁₋₁₃₂; cada molécula de ADN contiene dos repeticiones directas por lo que con cada dúplex interactúan dos moléculas de RepB₁₋₁₃₂ (Fig. 2.16).

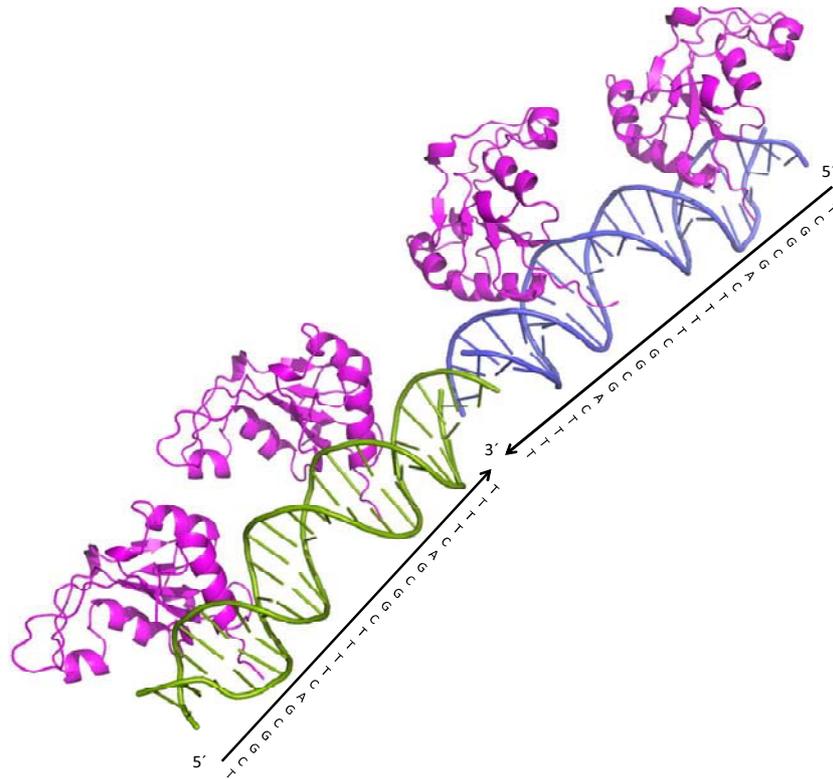


Fig. 2.16 Contenido de la unidad asimétrica

2.3.3 El dominio OBD de RepB y su unión a la DDR

El plegamiento global del OBD es prácticamente idéntico al observado en la estructura *full length* (Fig. 2.17) (al superponer ambas estructuras presenta una desviación r.m.s. de tan solo 0.8561) excepto por la zona del *loop* que se encuentra entre las hebras $\beta 2$ y $\beta 3$ debido muy probablemente a la flexibilidad de esta región. Por lo tanto presenta el plegamiento característico de esta familia de proteínas con una lámina β central formada por cinco hebras β antiparalelas flanqueada a ambos lados por hélices α .

Por otro lado el análisis del ADN presente en la estructura mediante el programa 3DNA [Lu *et al.*, 2003] demuestra que el mismo es un típico ADN de tipo B, prácticamente recto ya que solo presenta una curvatura de $\sim 4^\circ$, que se encuentra levemente comprimido y que el ángulo de giro

Fig. 2.17 Superposición del OBD (magenta) sobre RepB *full length* (celeste).

de las bases es un poco menor, siendo el clásico de 34,6° cuando en nuestra estructura es de 33,02°. Los parámetros geométricos del mismo se muestran en las siguientes tablas:

Tabla 2.3: Parámetros locales de pares de bases

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 T-A	0.13	-0.42	0.69	-29.00	-9.21	-0.32
2 C-G	-0.66	0.38	0.18	-17.09	9.93	2.24
3 G-C	0.26	-0.05	-0.25	-8.69	-12.62	-5.56
4 G-C	0.45	0.09	0.00	-12.23	-2.06	-0.01
5 C-G	0.17	-0.24	0.18	11.32	3.04	1.59
6 G-C	0.05	-0.45	0.66	20.29	-12.31	-8.33
7 A-T	0.87	-0.07	0.11	4.64	-13.97	-9.64
8 C-G	-0.04	-0.22	0.55	3.88	-13.78	-4.52
9 T-A	0.23	-0.50	0.85	-5.34	-18.90	-4.28
10 T-A	0.10	-0.06	0.24	-4.14	-20.56	5.27
11 T-A	-0.15	0.32	-0.23	-8.71	-8.91	1.44
12 T-A	-0.79	0.76	0.40	-19.89	-6.85	15.90
13 C-G	-0.63	0.49	-0.28	8.20	-1.50	12.67
14 G-C	1.52	-0.15	-0.58	2.37	-11.03	-12.98
15 G-C	-1.18	0.13	0.09	-9.98	-12.93	2.40
16 C-G	0.49	-0.28	0.25	-4.13	1.40	-6.07
17 G-C	-0.08	-0.53	0.55	8.59	-12.79	-6.16
18 A-T	-0.25	-0.23	0.90	10.44	-10.36	-6.67
19 C-G	-0.47	0.00	0.69	-5.41	-7.34	1.89
20 T-A	-0.10	-0.00	0.02	-6.19	-13.92	-0.22
21 T-A	-0.61	-0.01	-0.28	-9.07	-31.14	-4.26
22 T-A	-0.01	0.12	0.34	-9.35	-20.01	0.98
23 T-A	-1.53	0.53	0.12	-5.38	-10.41	9.00
24 A-T	1.75	0.56	0.16	5.80	-10.37	6.19
25 A-T	0.28	0.06	0.18	10.48	-16.66	-2.01
26 A-T	0.65	-0.09	-0.36	8.14	-29.48	-5.44
27 A-T	0.06	-0.19	-0.08	4.52	-12.73	-1.73
28 G-C	0.20	-0.09	0.52	6.39	-6.91	-1.18
29 T-A	0.09	-0.25	0.79	-11.34	-12.34	-7.58
30 C-G	-0.25	-0.46	0.64	-11.17	-10.00	-8.21
31 G-C	-0.62	-0.13	-0.04	-0.67	-1.52	-8.03
32 C-G	1.12	0.31	-0.27	6.78	-10.26	0.57
33 C-G	-1.58	0.07	-0.52	-6.75	-7.63	-16.42
34 G-C	0.80	0.57	-0.22	-10.91	-2.50	11.31
35 A-T	0.89	0.74	0.31	19.36	-7.20	15.58
36 A-T	0.25	0.22	-0.15	10.74	-9.73	-1.66
37 A-T	0.09	-0.07	0.06	5.55	-21.95	5.82
38 A-T	-0.09	-0.62	0.52	6.20	-16.00	-7.48
39 G-C	-0.09	-0.26	0.18	-5.60	-13.74	-5.74
40 T-A	-1.09	-0.12	0.07	-6.46	-12.15	-12.09
41 C-G	-0.20	-0.43	0.50	-24.12	-9.57	-13.53
42 G-C	-0.14	-0.05	-0.29	-13.98	2.81	-2.53
43 C-G	-0.29	0.17	-0.26	19.33	-1.63	-1.35
44 C-G	-0.24	-0.06	-0.54	11.24	-11.67	-7.48
45 G-C	0.60	0.32	0.33	16.95	10.10	-0.28
46 A-T	-0.13	-0.46	0.84	28.18	-11.87	-3.18
Average	-0.00	-0.01	0.17	-0.35	-9.98	-1.78
SD	0.68	0.34	0.40	12.30	8.29	7.26

Tabla 2.4: Parámetros locales de pasos de pares de bases							
Step	Shift	Slide	Rise	Tilt	Roll	Twist	
1	TC/GA	0.38	1.10	3.14	0.39	6.77	26.51
2	CG/CG	-1.15	1.19	3.34	-1.07	-2.15	38.90
3	GG/CC	-0.33	-0.78	3.51	-3.30	-0.70	28.46
4	GC/GC	-0.16	-0.01	2.74	0.20	4.28	24.52
5	CG/CG	-0.03	0.39	3.11	-1.96	-3.54	40.23
6	GA/TC	0.44	-0.68	3.52	5.99	-1.62	43.23
7	AC/GT	0.59	-1.36	3.20	-6.10	5.31	31.65
8	CT/AG	-0.61	-1.15	3.47	-2.48	-3.12	36.05
9	TT/AA	-0.54	-1.25	2.88	4.66	-4.02	29.89
10	TT/AA	-0.13	-0.32	3.08	3.54	-1.66	35.09
11	TT/AA	1.06	-0.34	3.57	-1.42	3.65	28.70
12	TC/GA	0.59	0.59	2.82	3.77	7.27	29.69
13	CG/CG	-1.13	0.56	3.61	-1.02	3.85	43.49
14	GG/CC	1.17	-0.97	3.63	-0.10	4.07	25.20
15	GC/GC	-0.76	0.06	3.35	-0.73	8.12	35.21
16	CG/CG	-0.33	0.28	2.90	-2.48	-2.89	32.10
17	GA/TC	0.41	-0.65	2.98	0.79	-8.72	42.05
18	AC/GT	0.65	-0.98	3.85	-2.44	3.51	34.85
19	CT/AG	-0.69	-0.85	2.99	3.37	-0.11	35.69
20	TT/AA	-1.12	-0.06	3.14	-2.34	-8.02	34.94
21	TT/AA	0.15	-0.68	3.31	-4.51	-3.60	40.63
22	TT/AA	0.79	-0.19	3.14	2.94	-7.77	31.50
23	TA/TA	-1.81	2.02	3.17	-1.97	1.39	-10.93
24	AA/TT	-0.79	-0.20	3.12	-3.18	-10.20	31.88
25	AA/TT	-0.12	-0.74	3.35	4.55	-2.48	39.73
26	AA/TT	0.99	-0.10	3.18	1.58	-9.12	35.71
27	AG/CT	0.61	-0.93	2.95	-2.88	0.38	34.45
28	GT/AC	-0.58	-0.98	3.92	1.55	1.16	35.05
29	TC/GA	-0.39	-0.75	3.04	-2.19	-8.88	41.35
30	CG/CG	0.26	0.25	2.97	5.43	-0.61	32.60
31	GC/GC	0.76	0.13	3.31	-0.52	6.39	35.46
32	CC/GG	-1.31	-0.96	3.60	-2.48	5.97	24.63
33	CG/CG	1.11	0.51	3.61	1.07	-0.06	44.22
34	GA/TC	-0.50	0.54	2.74	-2.85	8.27	29.50
35	AA/TT	-1.12	-0.43	3.42	-0.20	2.97	28.23
36	AA/TT	0.27	-0.33	3.17	-0.79	-2.03	36.34
37	AA/TT	0.40	-1.20	2.89	-4.53	-2.60	29.73
38	AG/CT	0.67	-1.13	3.59	2.35	-2.48	35.47
39	GT/AC	-0.60	-1.27	3.20	3.75	5.45	31.11
40	TC/GA	-0.63	-0.78	3.50	-6.33	-1.91	43.40
41	CG/CG	0.03	0.35	3.04	2.81	-2.81	41.30
42	GC/GC	0.21	0.08	2.51	-2.64	2.14	24.38
43	CC/GG	0.28	-0.84	3.62	2.85	2.21	28.09
44	CG/CG	1.06	1.13	3.36	-0.70	-3.38	38.78
45	GA/TC	-0.61	1.03	3.33	-0.56	5.87	27.01
Average		-0.06	-0.24	3.24	-0.23	-0.12	33.02
SD		0.74	0.79	0.31	3.04	4.96	8.77

La interacción del OBD con el ADN se produce en el surco mayor, participan el extremo N-terminal, la hélice $\alpha 2$ y los loops a ambos lados de la misma, es decir los loops entre la hélice $\alpha 2$ y la hebra $\beta 4$ y entre la hélice

$\alpha 2$ y la hebra $\beta 3$. Es interesante notar que el sitio activo queda alejado con respecto a la molécula de ADN (Fig. 2.18). Otra particularidad que presenta es que los OBD interactúan en la misma cara de la molécula de ADN, debido a que lo hacen con las regiones ricas en G y C que se encuentran separadas por las regiones AT que tienen la longitud justa de una vuelta de hélice, dejando orientadas las zonas GC hacia un mismo lado (Fig. 2.16).

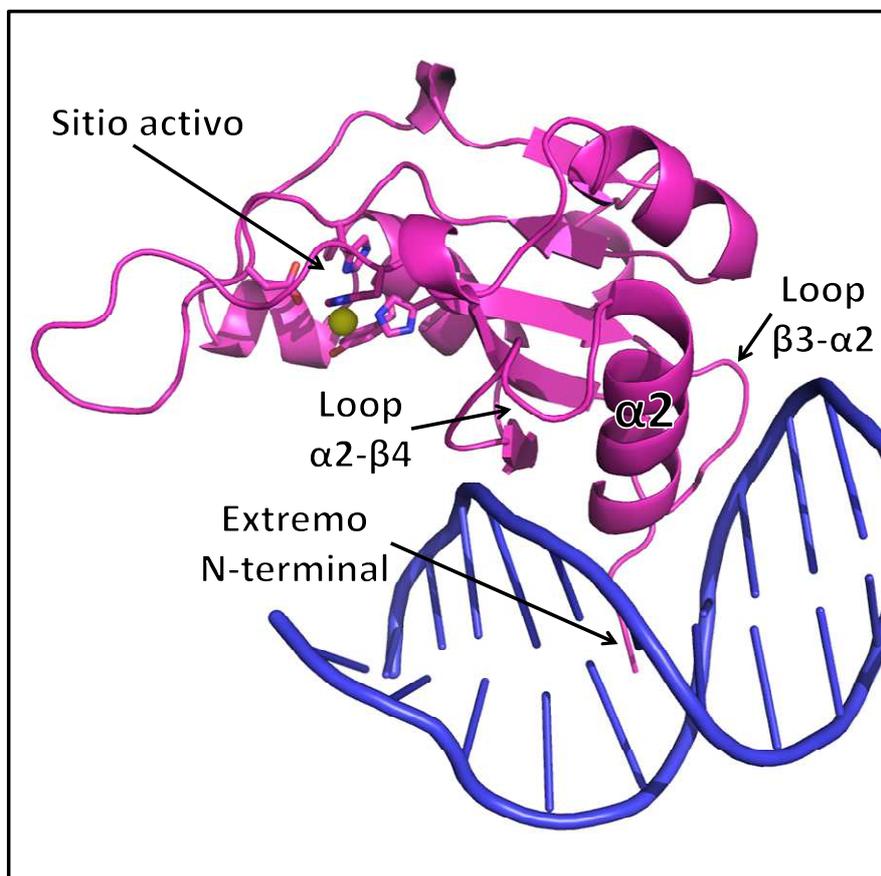


Fig. 2.18 Interacción de RepB₁₋₁₃₂ en el surco mayor de 23AB

Tomando como referencia los resultados obtenidos del análisis de la estructura mediante el programa Nucplot [Luscombe *et al.*, 1997] (Fig. 2.19), podemos identificar gráficamente los residuos que participan y las regiones del ADN que contactan. Si bien se observan algunas diferencias en cuanto a las áreas contactadas del ADN debido a la poca definición de algunas cadenas laterales por la baja resolución de la estructura, es claro cuáles son los residuos del OBD que participan: K3 y K5 localizados en el extremo N-terminal; T67 en $\beta 3$; D69, S70, R72, K73, K74 y K76 en $\alpha 2$; M86, V87 y Q88 en el loop entre $\alpha 2$ y $\beta 4$ (Fig. 2.20).

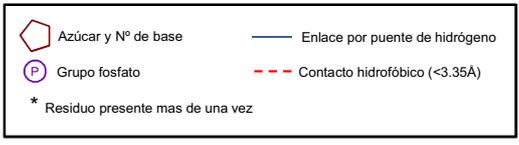
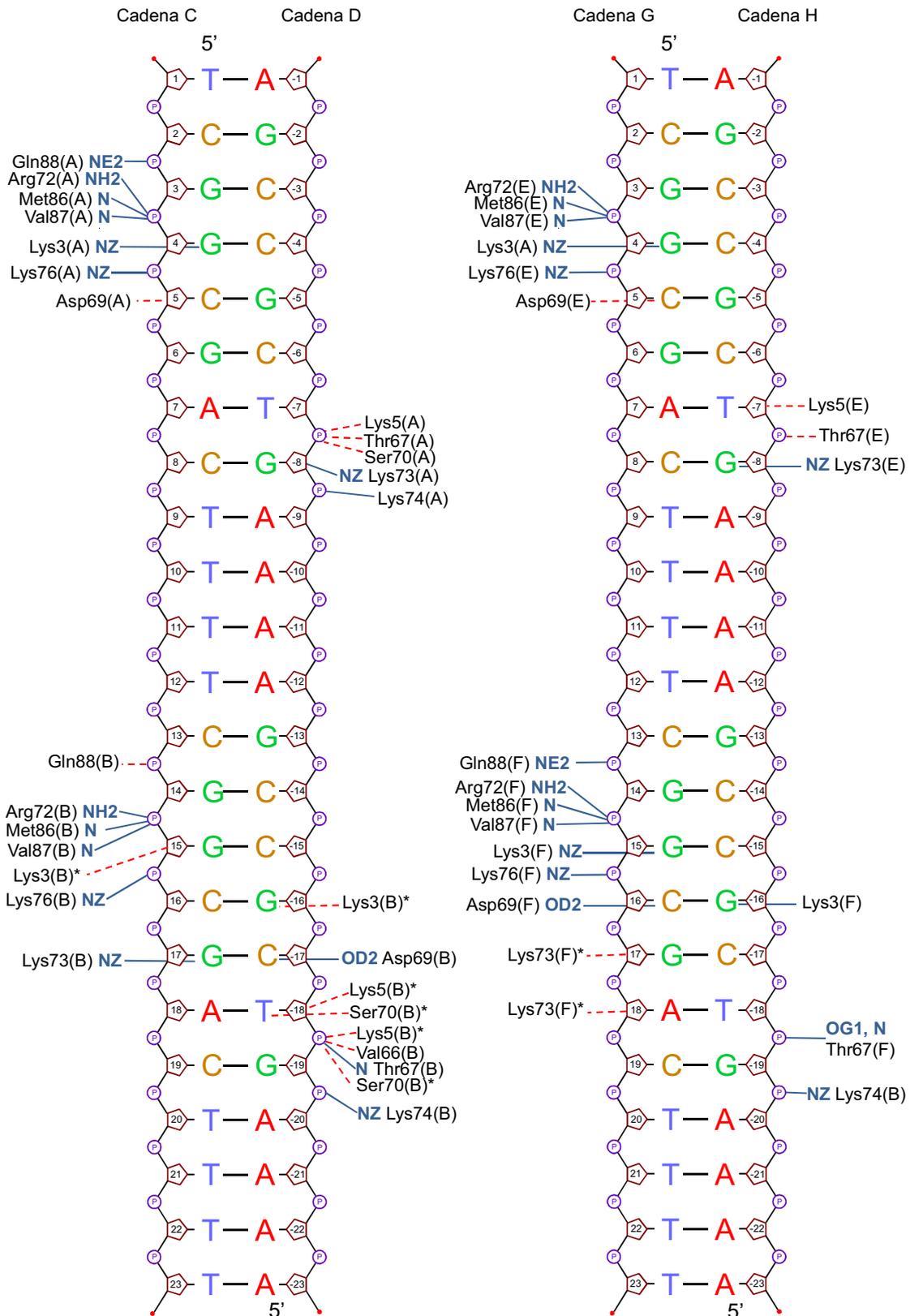


Fig. 2.19 Esquema de las interacciones ADN/proteína, generado mediante el programa Nucplot.

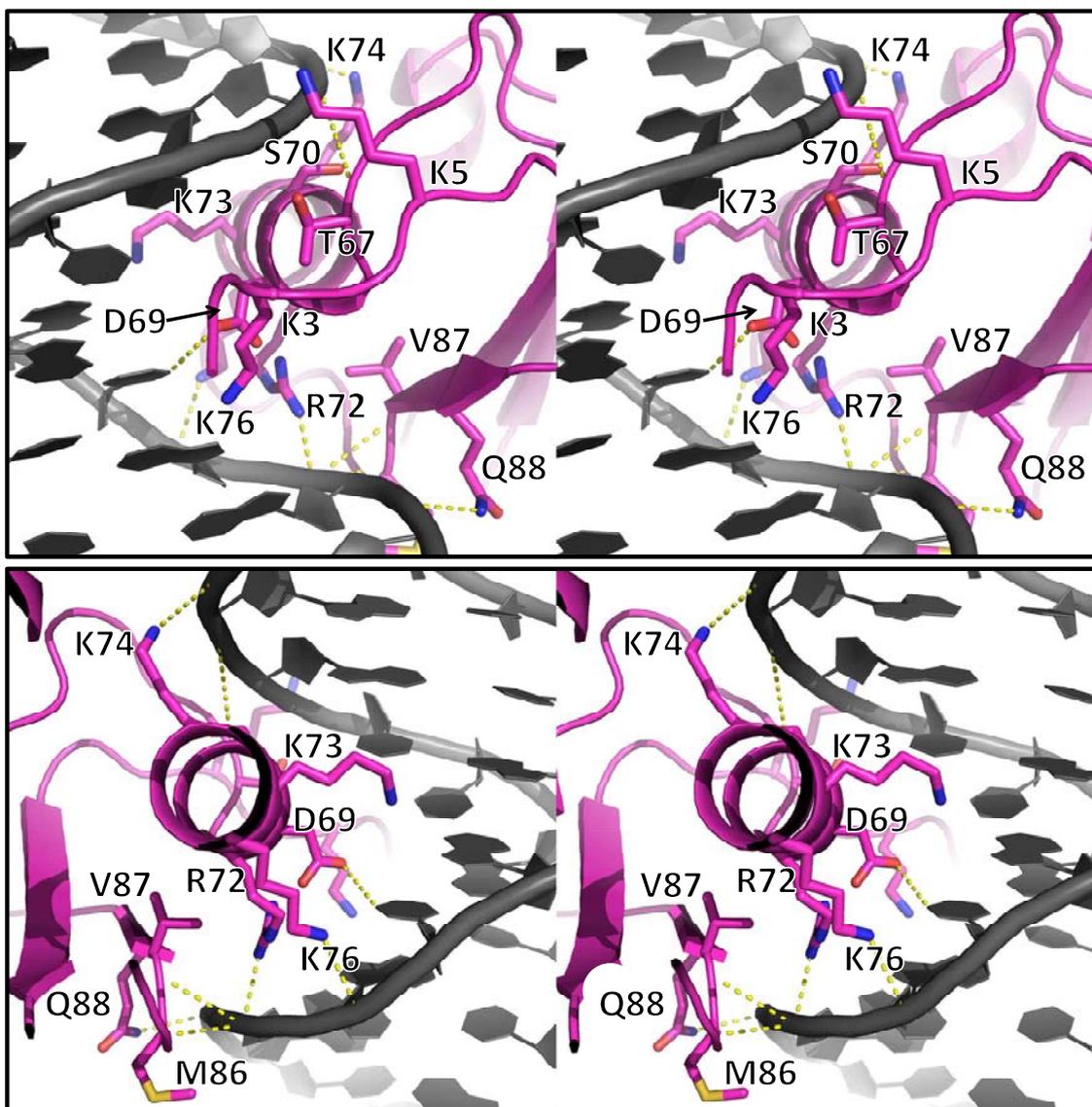


Fig. 2.20 Detalle de la interacción ADN/proteína. Se muestran en estéreo las vistas superior e inferior del OBD (magenta) con los aminoácidos que participan en la interacción (bastones) y el ADN (gris). Las interacciones por enlace de hidrógeno se representan con líneas punteadas amarillas.

Estos resultados son congruentes con el hecho de que al reemplazar los residuos R72 o K76 por alaninas se afecta gravemente la unión del OBD a las DDR y que la sustitución simultánea de los aminoácidos R72, K73 y K74 suprime la capacidad de unión del OBD por completo [Boer *et al.*, 2009]. La mayoría de los contactos que se observan en la estructura se establecen con la cadena de fosatos-azúcar del ADN. Se observan, sin embargo, dos interacciones específicas con las bases. Es el caso del D69 cuyo grupo carboxilo contacta el nitrógeno N4 de la citosina (5 o 16). Por otra parte, la K3 tiene la cadena lateral hacia el surco mayor, a distancia de puente de hidrógeno con las guaninas cercanas, aunque la falta de densidad impide precisar esta interacción. Por otro lado, existe una interacción que el nucplot no detecta en la que el carbono β del residuo D69 y el carbono α de

la S70 forman un bolsillo hidrofóbico que aloja el grupo metilo de la timina -7 y -18 (Fig21).

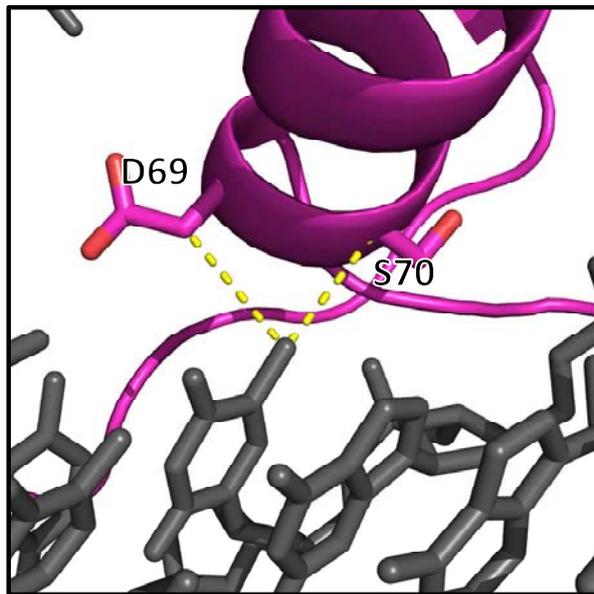


Fig. 21 Detalle de la interacción de D69 y S70 con la timina.

Al comparar nuestros resultados con los obtenidos en el análisis del *footprinting* de alta resolución realizado por Ruiz-Maso *et al.*, 2007 observamos que concuerdan, sobretodo en cuanto a la protección de la cadena principal del ADN en el surco mayor (Fig. 2.22).

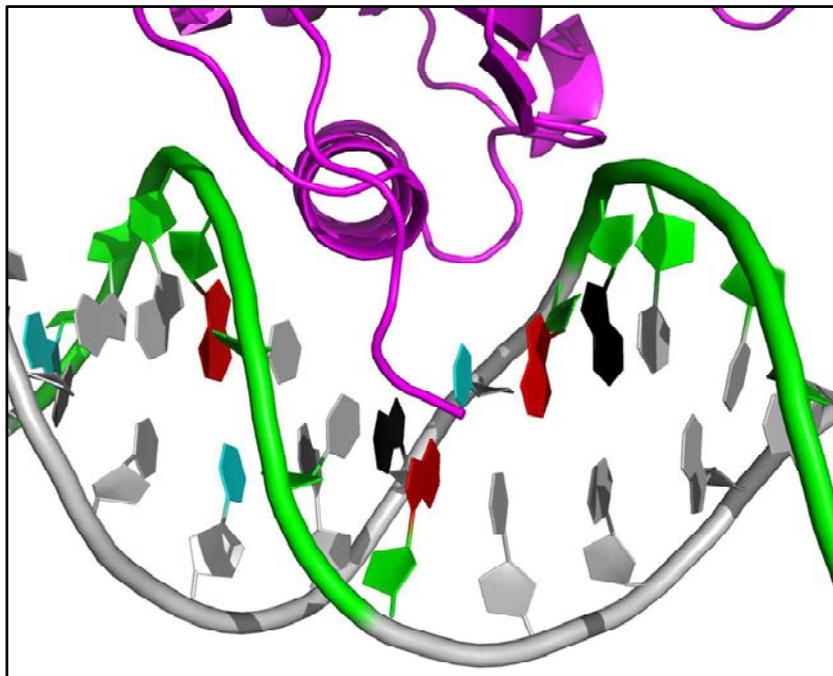


Fig. 2.22 Análisis del *footprinting* de alta resolución en el contexto de la estructura de RepB₁₋₁₃₂/23AB. En la imagen se muestran el OBD de RepB (magenta) y el ADN (gris) donde se resaltan en verde las zonas de la cadena principal que se encuentran protegidas al ataque de OH•, en rojo las bases protegidas de la metilación por DMS, en negro y cian aquellas bases que presentan sensibilidad aumentada al DMS en mayor o menor medida respectivamente.

2.3.4 Modelo de interacción hexamérico

Cabe destacar que RepB es capaz de interactuar con ADN que presenta diferentes secuencias, como es el caso de las DDR y las PDR en configuración de doble cadena, y que a su vez reconoce y lleva a cabo su actividad nucleolítica sobre ADN de simple cadena, como es la horquilla. Apuntando a que esta proteína presenta dos sitios de reconocimiento, uno (detallado en este trabajo) responsable de interactuar con ADN de doble cadena y reconocer secuencias con un cierto parecido aunque con importantes diferencias tanto de identidad como de largo; y otro responsable del reconocimiento de la horquilla portadora del sitio de corte donde además ejerce su actividad. Con el fin de comprender el modo en el que RepB actúa desarrollamos un modelo que integra la nueva información adquirida de la interacción del OBD con las DDR en el contexto del hexámero, el modelo obtenido se observa en la Fig. 2.23.

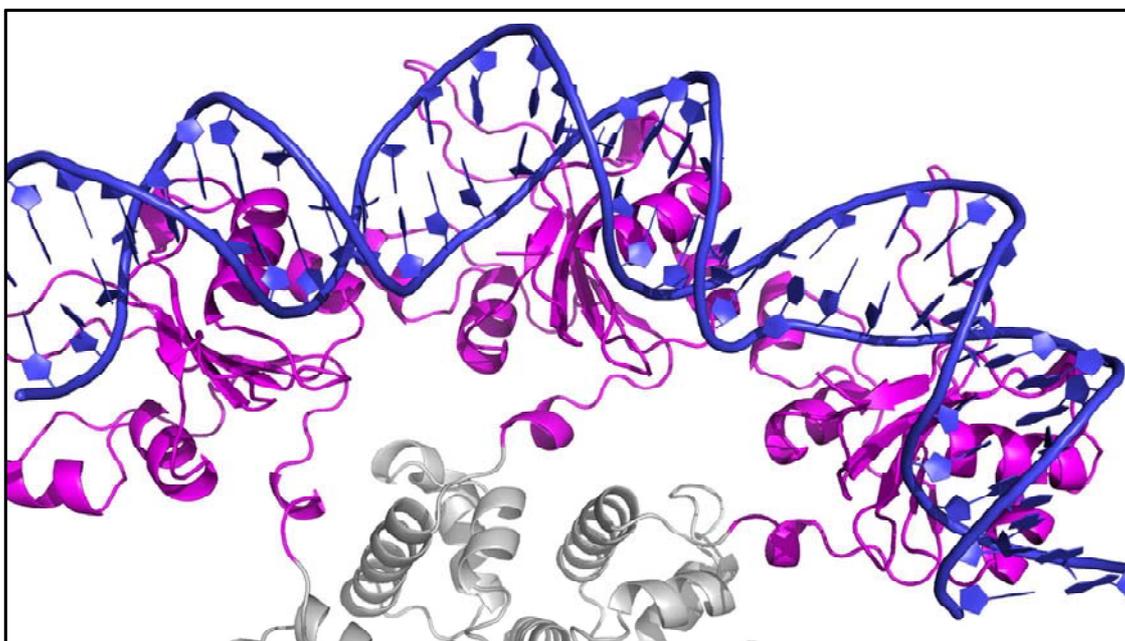


Fig. 2.23 Modelo de interacción de las DDR en el hexámero. Se muestran en gris los OD, en magenta los OBD y en azul el ADN.

En el modelo se observa claramente que la unión de RepB al ADN induce una curvatura importante en el mismo, los OBD interaccionan con las regiones ricas en GC y desde la misma cara de la molécula, generando posiblemente distorsiones en el ADN y dejando libres a su vez las regiones AT permitiendo que el ADN alcance una curvatura de 120° . Estos resultados son consistentes con los presentados por Ruiz-Maso *et al.*, 2007 quienes demuestran que el ADN unido a RepB presenta justamente una curvatura de 120° .

Por otro lado hemos querido construir un modelo hipotético de cómo actuaría RepB con las PDR y la horquilla que aloja el sitio de corte, proponemos que dos OBD consecutivos en el hexámero reconocen y unen

las PDR, posicionando la horquilla en el sitio activo del tercer OBD donde se ejerce la actividad endonucleolítica (Fig. 2.24). Este modelo explicaría porque el sitio activo queda alejado del sitio de unión al ADN de doble cadena, ya que de esta manera queda espacio libre para alojar la horquilla. A su vez, la estructura de la proteína Rep de AAV5, cuya posición del sitio activo e interacción con una horquilla similar coinciden con la de RepB, apoya esta propuesta [Hickman *et al.*, 2004].

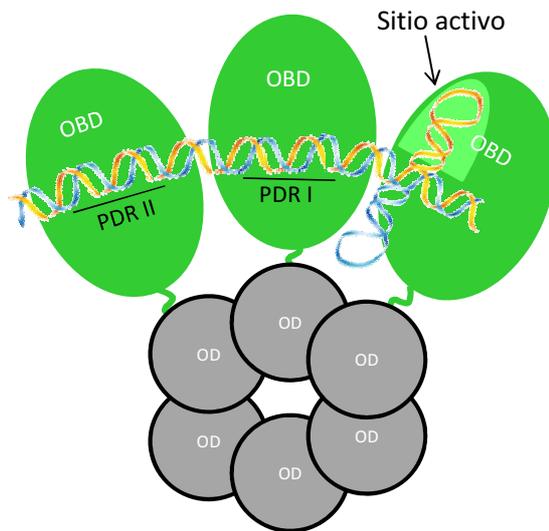


Fig. 2.24 Modelo esquemático de la interacción de las PDR en el hexámero. Los ovalos verdes representan los OBD con el sitio activo resaltado en verde claro y los círculos grises los OD. También se muestra la posición de las PDR con respecto a la horquilla de simple cadena con líneas negras.

Como perspectiva futura de este proyecto sería interesante intentar la cristalización del mismo constructo RepB₁₋₁₃₂ con un oligonucleótido que presente la secuencia de las PDR o de las PDR junto con la horquilla, ya que resolvería algunos de los interrogantes aun abiertos.

2.3.5 RepB₁₋₁₃₂/23AB vs. proteínas relacionadas estructuralmente

Decidimos comparar la estructura obtenida en este trabajo con proteínas relacionadas estructuralmente, aunque virales, cuyas estructuras han sido resueltas en presencia de su ADN diana, como es el caso de los dominios de unión al ADN de la proteína Rep de AAV5 [Hickman *et al.*, 2004] y la proteína E1 de BPV [Enemark *et al.*, 2002] (Fig. 2.25). Cabe notar que AAV5 se une al ADN tanto en el surco mayor como en el surco menor, mientras que la proteína de BPV al igual que la RepB lo hace solo en el surco mayor.

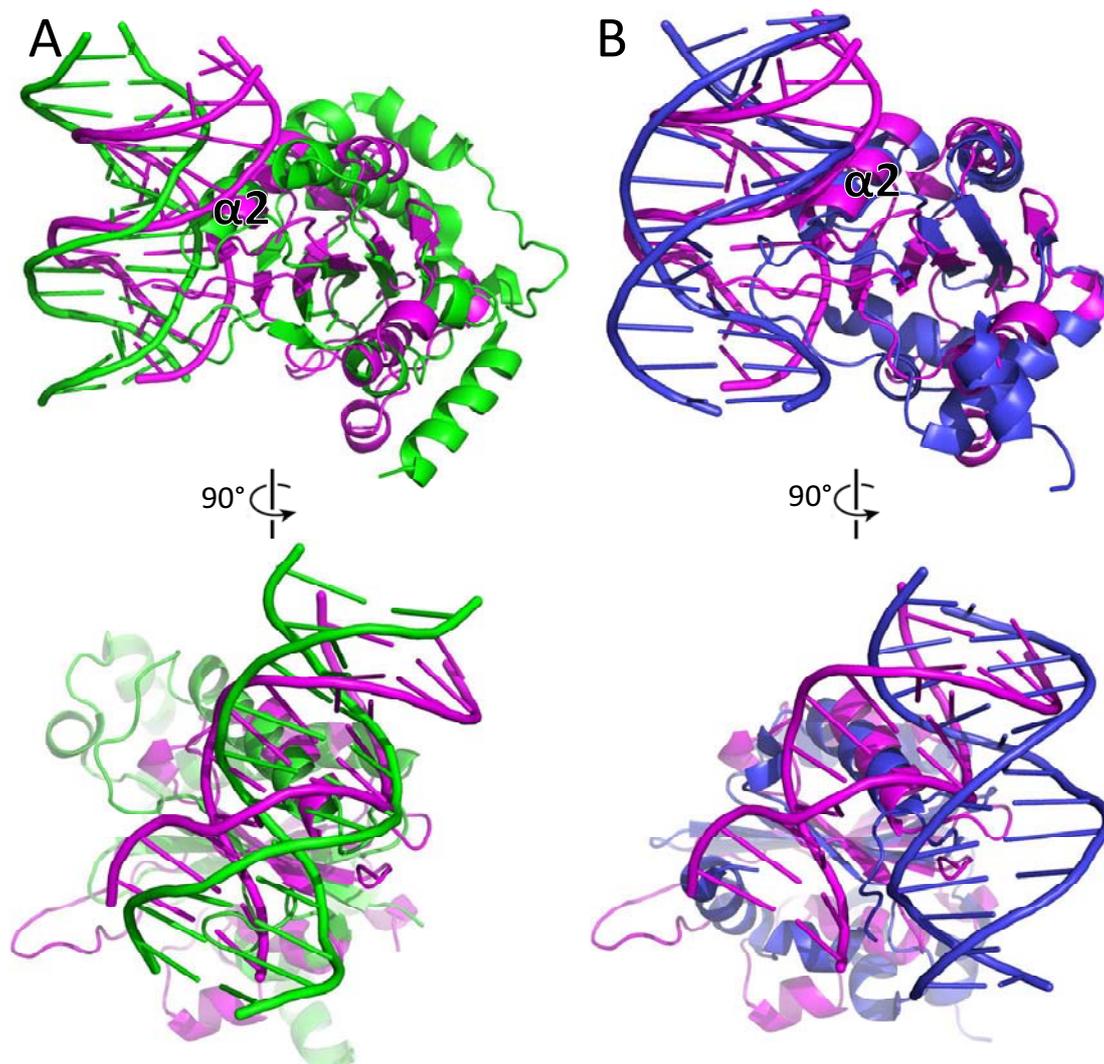


Fig. 2.25 Superposición de RepB₁₋₁₃₂/23AB con proteínas relacionadas estructuralmente. Se muestran dos vistas de RepB en magenta superpuesta sobre A) Rep de AAV5 en verde y B) E1 de BPV en azul.

Al analizar la superposición se observa que en ambos casos la interacción está mediada por una hélice α y, en el caso de AAV5, un *loop*, y en el de BPV, un *loop* y el N-terminal, coincidiendo claramente con nuestro resultado. A pesar de que, a diferencia de nuestro complejo, la interacción de AAV5 se establece mediante el contacto de la hélice en el surco menor y el *loop* en el surco mayor la posición relativa con respecto a RepB es mucho más parecida que en el caso de BPV. En esta última observamos un evidente cambio de orientación de la molécula de ADN con respecto a nuestra estructura al superponer ambas. A pesar de las diferencias podemos concluir que la interacción ADN/proteína de esta familia de proteínas es análoga.

2.3 Conclusiones

- La optimización de los cristales obtenidos en la condición E9 del PAC10 dio como resultado cristales de tamaño y características adecuadas que difractaron hasta 3Å de resolución.
- Utilizando estos datos se resolvió la estructura de RepB₁₋₁₃₂ unida a la DDR.
- Se comprobó que el OBD de RepB interactúa con el ADN diana en el surco mayor de la región rica en G y C, contactando con el hélice $\alpha 2$ y los *loops* adyacentes a la misma, dejando libres las zonas AT.
- Se determinaron los aminoácidos que participan en la unión del ADN, confirmando resultados de mutaciones puntuales previamente descritas.
- La mayoría de los contactos proteína-ADN son inespecíficos, pero tres de ellos son específicos de secuencia. En particular el puente de hidrógeno entre el D69 y la citosina, y el bolsillo hidrofóbico formado por D69 y S70 que aloja a la timina.
- Se determinaron las regiones del ADN contactadas directamente por la proteína, confirmando los resultados de *footprinting* previamente descritos.
- Se propone un modelo de unión al sitio *bind* en el contexto del hexámero, que además reproduce el ángulo de curvatura de las DDR descrito en la literatura.
- Se propone un modelo de unión al sitio *nic*.
- De la comparación de la estructura de RepB₁₋₁₃₂/23AB con otras estructuras relacionadas de proteínas virales concluimos que la interacción con el ADN se produce con la participación de regiones homólogas de las proteínas, aunque contactan con el ADN de forma diferente.

Bibliografía

Listado de webs:

Web1: <http://smart.embl-heidelberg.de/>

Web2: <http://prdos.hgc.jp/cgi-bin/top.cgi>

Web3: http://embnet.vital-it.ch/software/COILS_form.html

Web4: <http://xtal.nki.nl/ccd/Welcome.html>

Web5: <http://bioinf.cs.ucl.ac.uk/psipred/>

Web6: <http://amrls.cvm.msu.edu/microbiology/molecular-basis-for-antimicrobial-resistance/acquired-resistance/acquisition-of-antimicrobial-resistance-via-horizontal-gene-transfer>

Web7: <http://web.expasy.org/protparam/>

Web8: <http://molprobity.biochem.duke.edu/>

Bibliografía

Aminov, R. I. (2010). "A brief history of the antibiotic era: lessons learned and challenges for the future." *Front Microbiol* 1: 134.

Bauman, R. (2007). *Microbiology with diseases by taxonomy*, Pearson International

Beck, F., P. Unverdorben, S. Bohn, A. Schweitzer, G. Pfeifer, E. Sakata, S. Nickell, J. M. Pitzko, E. Villa, W. Baumeister and F. Forster (2012). "Near-atomic resolution structural model of the yeast 26S proteasome." *Proc Natl Acad Sci U S A* 109(37): 14870-5.

Boehringer, J., C. Riedinger, K. Paraskevopoulos, E. O. Johnson, E. D. Lowe, C. Khoudian, D. Smith, M. E. Noble, C. Gordon and J. A. Endicott (2012). "Structural and functional characterization of Rpn12 identifies residues required for Rpn10 proteasome incorporation." *Biochem J* 448(1): 55-65.

Boer, D. R., J. A. Ruiz-Maso, J. R. Lopez-Blanco, A. G. Blanco, M. Vives-Llacer, P. Chacon, I. Uson, F. X. Gomis-Ruth, M. Espinosa, O. Llorca, G. del Solar and M. Coll (2009). "Plasmid replication initiator RepB forms a hexamer reminiscent of ring helicases and has mobile nuclease domains." *EMBO J* 28(11): 1666-78.

Bohn, S., E. Sakata, F. Beck, G. R. Pathare, J. Schnitger, I. Nagy, W. Baumeister and F. Forster (2013). "Localization of the regulatory particle subunit Sem1 in the 26S proteasome." *Biochem Biophys Res Commun* 435(2): 250-4.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." *Anal Biochem* 72: 248-54.

Buchan, D. W., S. M. Ward, A. E. Lobley, T. C. Nugent, K. Bryson and D. T. Jones (2010). "Protein annotation and modelling servers at University College London." *Nucleic Acids Res* 38(Web Server issue): W563-8.

Burdett, V. (1980). "Identification of tetracycline-resistant R-plasmids in *Streptococcus agalactiae* (group B)." *Antimicrob Agents Chemother* 18(5): 753-60.

Castillo, C. F. (2007). "Los postulados de Koch: revisión histórica y perspectiva actual." *RCCV* 1(2): 262-266.

Clardy, J., M. A. Fischbach and C. R. Currie (2009). "The natural history of antibiotics." *Curr Biol* 19(11): R437-41.

Collaborative Computational Project, N. (1994). "The CCP4 suite: programs for protein crystallography." *Acta Crystallogr D Biol Crystallogr* 50(Pt 5): 760-3.

Chandler, M., F. de la Cruz, F. Dyda, A. B. Hickman, G. Moncalian and B. Ton-Hoang (2013). "Breaking and joining single-stranded DNA: the HUH endonuclease superfamily." *Nat Rev Microbiol* 11(8): 525-38.

Davis, I. W., A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson and D. C. Richardson (2007). "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids." *Nucleic Acids Res* 35(Web Server issue): W375-83.

del Solar, G., M. Moscoso and M. Espinosa (1993). "Rolling circle-replicating plasmids from gram-positive and gram-negative bacteria: a wall falls." *Mol Microbiol* 8(5): 789-96.

del Solar, G., P. Acebo and M. Espinosa (1995). "Replication control of plasmid pLS1: efficient regulation of plasmid copy number is exerted by the combined action of two plasmid components, CopG and RNA II." *Mol Microbiol* 18(5): 913-24.

del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa and R. Diaz-Orejas (1998). "Replication and control of circular bacterial plasmids." *Microbiol Mol Biol Rev* 62(2): 434-64.

Dyda, F. and A. B. Hickman (2003). "A mob of reps." *Structure* 11(11): 1310-1.

Effantin, G., R. Rosenzweig, M. H. Glickman and A. C. Steven (2009). "Electron microscopic evidence in support of alpha-solenoid models of proteasomal subunits Rpn1 and Rpn2." *J Mol Biol* 386(5): 1204-11.

Ellisdon, A. M. and M. Stewart (2012). "Structural biology of the PCI-protein fold." *Bioarchitecture* 2(4).

Emsley, P., B. Lohkamp, W. G. Scott and K. Cowtan (2010). "Features and development of Coot." *Acta Crystallogr D Biol Crystallogr* 66(Pt 4): 486-501.

Enemark, E. J., G. Chen, D. E. Vaughn, A. Stenlund and L. Joshua-Tor (2000). "Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus." *Mol Cell* 6(1): 149-58.

Enemark, E. J., A. Stenlund and L. Joshua-Tor (2002). "Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex." *EMBO J* 21(6): 1487-96.

Enemark, E. J. and L. Joshua-Tor (2006). "Mechanism of DNA translocation in a replicative hexameric helicase." *Nature* 442(7100): 270-5.

Espinosa, M. (2013). "Plasmids as models for studying macromolecular interactions: the pMV158 paradigm." *Res Microbiol* 164(3): 199-204.

Estrin, E., J. R. Lopez-Blanco, P. Chacon and A. Martin (2013). "Formation of an Intricate Helical Bundle Dictates the Assembly of the 26S Proteasome Lid." *Structure*.

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

Finley, D. (2009). "Recognition and processing of ubiquitin-protein conjugates by the proteasome." *Annu Rev Biochem* 78: 477-513.

Forster, F., P. Unverdorben, P. Sledz and W. Baumeister (2013). "Unveiling the long-held secrets of the 26S proteasome." *Structure* 21(9): 1551-62.

Francia, M. V., A. Varsaki, M. P. Garcillan-Barcia, A. Latorre, C. Drainas and F. de la Cruz (2004). "A classification scheme for mobilization regions of bacterial plasmids." *FEMS Microbiol Rev* 28(1): 79-100.

Frank, J., M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj and A. Leith (1996). "SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields." *J Struct Biol* 116(1): 190-9.

Fu, H., N. Reis, Y. Lee, M. H. Glickman and R. D. Vierstra (2001). "Subunit interaction maps for the regulatory particle of the 26S proteasome and the COP9 signalosome." *EMBO J* 20(24): 7096-107.

Fukunaga, K., T. Kudo, A. Toh-e, K. Tanaka and Y. Saeki (2010). "Dissection of the assembly pathway of the proteasome lid in *Saccharomyces cerevisiae*." *Biochem Biophys Res Commun* 396(4): 1048-53.

Funakoshi, M., R. J. Tomko, Jr., H. Kobayashi and M. Hochstrasser (2009). "Multiple assembly chaperones govern biogenesis of the proteasome regulatory particle base." *Cell* 137(5): 887-99.

Gallastegui, N. and M. Groll (2010). "The 26S proteasome: assembly and function of a destructive machine." *Trends Biochem Sci* 35(11): 634-42.

Glickman, M. H., D. M. Rubin, O. Coux, I. Wefes, G. Pfeifer, Z. Cjeka, W. Baumeister, V. A. Fried and D. Finley (1998). "A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3." *Cell* 94(5): 615-23.

Glickman, M. H. and A. Ciechanover (2002). "The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction." *Physiol Rev* 82(2): 373-428.

Gogarten, M. B., J. P. Gogarten and L. Olendzenski (2009). *Horizontal gene transfer : genomes in flux*. New York, Springer.

Golovanov, A. P., G. M. Hautbergue, S. A. Wilson and L. Y. Lian (2004). "A simple method for improving protein solubility and long-term stability." *J Am Chem Soc* 126(29): 8933-9.

Gomis-Ruth, F. X., M. Sola, P. Acebo, A. Parraga, A. Guasch, R. Eritja, A. Gonzalez, M. Espinosa, G. del Solar and M. Coll **(1998)**. "The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator." *EMBO J* 17(24): 7404-15.

Gomis-Ruth, F. X. and M. Coll **(2006)**. "Cut and move: protein machinery for DNA processing in bacterial conjugation." *Curr Opin Struct Biol* 16(6): 744-52.

Gouet, P., E. Courcelle, D. I. Stuart and F. Metoz **(1999)**. "ESPrInt: analysis of multiple sequence alignments in PostScript." *Bioinformatics* 15(4): 305-8.

Groll, M., L. Ditzel, J. Lowe, D. Stock, M. Bochtler, H. D. Bartunik and R. Huber **(1997)**. "Structure of 20S proteasome from yeast at 2.4 Å resolution." *Nature* 386(6624): 463-71.

Guasch, A., M. Lucas, G. Moncalian, M. Cabezas, R. Perez-Luque, F. X. Gomis-Ruth, F. de la Cruz and M. Coll **(2003)**. "Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC." *Nat Struct Biol* 10(12): 1002-10.

He, J., K. Kulkarni, P. C. da Fonseca, D. Krutauz, M. H. Glickman, D. Barford and E. P. Morris **(2012)**. "The structure of the 26S proteasome subunit Rpn2 reveals its PC repeat domain as a closed toroid of two concentric alpha-helical rings." *Structure* 20(3): 513-21.

Hershko, A. and A. Ciechanover **(1998)**. "The ubiquitin system." *Annu Rev Biochem* 67: 425-79.

Hickman, A. B., D. R. Ronning, R. M. Kotin and F. Dyda **(2002)**. "Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep." *Mol Cell* 10(2): 327-37.

Hickman, A. B., D. R. Ronning, Z. N. Perez, R. M. Kotin and F. Dyda **(2004)**. "The nuclease domain of adeno-associated virus rep coordinates replication initiation using two distinct DNA recognition interfaces." *Mol Cell* 13(3): 403-14.

Ishida, T. and K. Kinoshita **(2007)**. "PrDOS: prediction of disordered protein regions from amino acid sequence." *Nucleic Acids Res* 35(Web Server issue): W460-4.

Jones, D. T. **(1999)**. "Protein secondary structure prediction based on position-specific scoring matrices." *J Mol Biol* 292(2): 195-202.

Khan, S. A. **(1997)**. "Rolling-circle replication of bacterial plasmids." *Microbiol Mol Biol Rev* 61(4): 442-55.

Khan, S. A. **(2000)**. "Plasmid rolling-circle replication: recent developments." *Mol Microbiol* 37(3): 477-84.

Khan, S. A. (2005). "Plasmid rolling-circle replication: highlights of two decades of research." *Plasmid* 53(2): 126-36.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." *Nature* 227(5259): 680-5.

Lander, G. C., E. Estrin, M. E. Matyskiela, C. Bashore, E. Nogales and A. Martin (2012). "Complete subunit architecture of the proteasome regulatory particle." *Nature* 482(7384): 186-91.

Lander, G. C., A. Martin and E. Nogales (2013). "The proteasome under the microscope: the regulatory particle in focus." *Curr Opin Struct Biol* 23(2): 243-51.

Leggett, D. S., M. H. Glickman and D. Finley (2005). "Purification of proteasomes, proteasome subcomplexes, and proteasome-associated proteins from budding yeast." *Methods Mol Biol* 301: 57-70.

Leslie, A. and H. Powell (2007). *Processing Diffraction Data with Mosflm Evolving Methods for Macromolecular Crystallography*, Springer Verlag.

Letunic, I., T. Doerks and P. Bork (2012). "SMART 7: recent updates to the protein domain annotation resource." *Nucleic Acids Res* 40(Database issue): D302-5.

Lovell, S. C., I. W. Davis, W. B. Arendall, 3rd, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson and D. C. Richardson (2003). "Structure validation by Calpha geometry: phi,psi and Cbeta deviation." *Proteins* 50(3): 437-50.

Lowe, J., D. Stock, B. Jap, P. Zwickl, W. Baumeister and R. Huber (1995). "Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution." *Science* 268(5210): 533-9.

Lu, X. J. and W. K. Olson (2003). "3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures." *Nucleic Acids Res* 31(17): 5108-21.

Ludtke, S. J. (2010). "3-D structures of macromolecules using single-particle analysis in EMAN." *Methods Mol Biol* 673: 157-73.

Luo, X., D. G. Sanford, P. A. Bullock and W. W. Bachovchin (1996). "Solution structure of the origin DNA-binding domain of SV40 T-antigen." *Nat Struct Biol* 3(12): 1034-9.

Lupas, A., M. Van Dyke and J. Stock (1991). "Predicting coiled coils from protein sequences." *Science* 252(5009): 1162-4.

Luscombe, N. M., R. A. Laskowski and J. M. Thornton (1997). "NUCPLLOT: a program to generate schematic diagrams of protein-nucleic acid interactions." *Nucleic Acids Res* 25(24): 4940-5.

Madigan, M. T. (2009). Brock Biology of Microorganisms, Pearson Addison Wesley.

Matyskiela, M. E., G. C. Lander and A. Martin (2013). "Conformational switching of the 26S proteasome enables substrate degradation." *Nat Struct Mol Biol* 20(7): 781-8.

Matyskiela, M. E. and A. Martin (2013). "Design principles of a universal protein degradation machine." *J Mol Biol* 425(2): 199-213.

Mooij, W. T., E. Mitsiki and A. Perrakis (2009). "ProteinCCD: enabling the design of protein truncation constructs for expression and crystallization experiments." *Nucleic Acids Res* 37(Web Server issue): W402-5.

Moscoso, M., R. Eritja and M. Espinosa (1997). "Initiation of replication of plasmid pMV158: mechanisms of DNA strand-transfer reactions mediated by the initiator RepB protein." *J Mol Biol* 268(5): 840-56.

Murata, S., H. Yashiroda and K. Tanaka (2009). "Molecular mechanisms of proteasome assembly." *Nat Rev Mol Cell Biol* 10(2): 104-15.

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): 240-55.

Nagy, V. and I. Dikic (2010). "Ubiquitin ligase complexes: from substrate selectivity to conjugational specificity." *Biol Chem* 391(2-3): 163-9.

Painter, J. and E. A. Merritt (2006). "Optimal description of a protein structure in terms of multiple groups undergoing TLS motion." *Acta Crystallogr D Biol Crystallogr* 62(Pt 4): 439-50.

Pathare, G. R., I. Nagy, S. Bohn, P. Unverdorben, A. Hubert, R. Korner, S. Nickell, K. Lasker, A. Sali, T. Tamura, T. Nishioka, F. Forster, W. Baumeister and A. Bracher (2012). "The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together." *Proc Natl Acad Sci U S A* 109(1): 149-54.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin (2004). "UCSF Chimera--a visualization system for exploratory research and analysis." *J Comput Chem* 25(13): 1605-12.

Pickart, C. M. and R. E. Cohen (2004). "Proteasomes and their kin: proteases in the machine age." *Nat Rev Mol Cell Biol* 5(3): 177-87.

Pintilie, G. D., J. Zhang, T. D. Goddard, W. Chiu and D. C. Gossard (2010). "Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions." *J Struct Biol* 170(3): 427-38.

Proteau, A., R. Shi and M. Cygler (2010). "Application of dynamic light scattering in protein crystallization." *Curr Protoc Protein Sci* Chapter 17: Unit 17 10.

Quillin, M. L. and B. W. Matthews (2000). "Accurate calculation of the density of proteins." *Acta Crystallogr D Biol Crystallogr* 56(Pt 7): 791-4.

Riedinger, C., J. Boehringer, J. F. Trempe, E. D. Lowe, N. R. Brown, K. Gehring, M. E. Noble, C. Gordon and J. A. Endicott (2010). "Structure of Rpn10 and its interactions with polyubiquitin chains and the proteasome subunit Rpn12." *J Biol Chem* 285(44): 33992-4003.

Rosenzweig, R., P. A. Osmulski, M. Gaczynska and M. H. Glickman (2008). "The central unit within the 19S regulatory particle of the proteasome." *Nat Struct Mol Biol* 15(6): 573-80.

Rosenzweig, R., V. Bronner, D. Zhang, D. Fushman and M. H. Glickman (2012). "Rpn1 and Rpn2 coordinate ubiquitin processing factors at proteasome." *J Biol Chem* 287(18): 14659-71.

Ruiz-Maso, J. A., C. Lopez-Zumel, M. Menendez, M. Espinosa and G. del Solar (2004). "Structural features of the initiator of replication protein RepB encoded by the promiscuous plasmid pMV158." *Biochim Biophys Acta* 1696(1): 113-9.

Ruiz-Maso, J. A., R. Lurz, M. Espinosa and G. del Solar (2007). "Interactions between the RepB initiator protein of plasmid pMV158 and two distant DNA regions within the origin of replication." *Nucleic Acids Res* 35(4): 1230-44.

Saeki, Y. and K. Tanaka (2007). "Unlocking the proteasome door." *Mol Cell* 27(6): 865-7.

Sambrook, J. and D. W. Russell (2006). *The condensed protocols from Molecular cloning : a laboratory manual.* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

Samso, M., M. Radermacher, J. Frank and M. P. Koonce (1998). "Structural characterization of a dynein motor domain." *J Mol Biol* 276(5): 927-37.

Sanches, M., B. S. Alves, N. I. Zanchin and B. G. Guimaraes (2007). "The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer." *J Mol Biol* 370(5): 846-55.

Scott, A. M., J. D. Wolchok and L. J. Old (2012). "Antibody therapy of cancer." *Nat Rev Cancer* 12(4): 278-87.

Schaechter, M., J. L. Ingraham and F. C. Neidhardt (2006). *Microbe.* Washington, D.C., American Society for Microbiology.

Schreiner, P., X. Chen, K. Husnjak, L. Randles, N. Zhang, S. Elsassner, D. Finley, I. Dikic, K. J. Walters and M. Groll (2008). "Ubiquitin docking at the

proteasome through a novel pleckstrin-homology domain interaction." *Nature* 453(7194): 548-52.

Schrödinger, L. The PyMOL Molecular Graphics System.

Schultz, J., F. Milpetz, P. Bork and C. P. Ponting (1998). "SMART, a simple modular architecture research tool: identification of signaling domains." *Proc Natl Acad Sci U S A* 95(11): 5857-64.

Schwartz, A. L. and A. Ciechanover (1999). "The ubiquitin-proteasome pathway and pathogenesis of human diseases." *Annu Rev Med* 50: 57-74.

Sessa, F., M. Mapelli, C. Ciferri, C. Tarricone, L. B. Areces, T. R. Schneider, P. T. Stukenberg and A. Musacchio (2005). "Mechanism of Aurora B activation by INCENP and inhibition by hesperadin." *Mol Cell* 18(3): 379-91.

Sharon, M., T. Taverner, X. I. Ambroggio, R. J. Deshaies and C. V. Robinson (2006). "Structural organization of the 19S proteasome lid: insights from MS of intact complexes." *PLoS Biol* 4(8): e267.

Smith, D. M., S. C. Chang, S. Park, D. Finley, Y. Cheng and A. L. Goldberg (2007). "Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry." *Mol Cell* 27(5): 731-44.

Tanaka, K. (2009). "The proteasome: overview of structure and functions." *Proc Jpn Acad Ser B Phys Biol Sci* 85(1): 12-36.

Tomko, R. J., Jr. and M. Hochstrasser (2011). "Incorporation of the Rpn12 subunit couples completion of proteasome regulatory particle lid assembly to lid-base joining." *Mol Cell* 44(6): 907-17.

Tomko, R. J., Jr. and M. Hochstrasser (2013). "Molecular architecture and assembly of the eukaryotic proteasome." *Annu Rev Biochem* 82: 415-45.

VanLoock, M. S., A. Alexandrov, X. Yu, N. R. Cozzarelli and E. H. Egelman (2002). "SV40 large T antigen hexamer structure: domain organization and DNA-induced conformational changes." *Curr Biol* 12(6): 472-6.

Verma, R., L. Aravind, R. Oania, W. H. McDonald, J. R. Yates, 3rd, E. V. Koonin and R. J. Deshaies (2002). "Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome." *Science* 298(5593): 611-5.

Zhang, F., M. Hu, G. Tian, P. Zhang, D. Finley, P. D. Jeffrey and Y. Shi (2009). "Structural insights into the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*." *Mol Cell* 34(4): 473-84.