

**Estudios estructurales en la
proteína helicasa MCM de
*Methanobacterium
thermoautotrophicum***

**Tesis doctoral de:
*Yacob Gómez Llorente***

Dirigida por:
Carmen San Martín Pastrana
José María Carazo García

Abstract

Methanobacterium thermoautotrophicum minichromosome maintenance complex (mtMCM), a cellular replicative helicase, is a useful model for the more complex eukaryotic MCMs. Biochemical and crystallographic evidence indicates that mtMCM assembles as a double hexamer (dHex), but previous electron microscopy studies reported only the presence of single heptamers or single hexamers (Pape, T., Meka, H., Chen, S., Vicentini, G., Van Heel, M., and Onesti, S. (2003) *EMBO Rep.* 4, 1079–1083; Yu, X., VanLoock, M. S., Poplawski, A., Kelman, Z., Xiang, T., Tye, B. K., and Egelman, E. H. (2002) *EMBO Rep.* 3, 792–797). Here we present the first three-dimensional electron microscopy reconstruction of the full-length mtMCM dHex in which two hexamers contact each other via the structurally well defined N-terminal domains. The dHex has obvious side openings that resemble the side channels of LTag (large T antigen). 6-fold and 7-fold rings were observed in the same mtMCM preparation, but we determined that assembly as a double ring favors 6-fold structures. Additionally, open rings were also detected, which suggests a direct mtMCM loading mechanism onto DNA. Furthermore we demonstrated that changing arginine to alanine at amino acid position 161 or the insertion of a six-aminoacid peptide at the hexamer-hexamer interface in N-terminal-mtMCM dodecamer, disrupted dHex formation and produced stable single hexamers (sHex).

Cláusula de buen uso

Queda prohibida la utilización, investigación y desarrollo, de forma directa o indirecta, de cualquiera de las aportaciones científicas propias del autor que se presentan en esta memoria, por parte de cualquier ejército o grupo armado, para cualquier uso militar. Queda prohibida del mismo modo, para cualquier uso que atente contra los derechos humanos o contra el medio ambiente

Good use right

It is prohibited to use, to investigate or to develop, in a direct or indirect way, any of the scientific contributions of the author contained in this work by any army or armed group, for any military purposes. It is prohibited for any use which is against human rights or the environment.

Índice

Abstract	I
Cláusula de buen uso / Good use right	II
Índice	III
Introducción	1
<hr/>	
I.1 – Replicación del DNA celular	2
I.1.1 – Entendiendo la replicación. Eventos fundamentales	2
I.1.2 – Replicación en eucariotas	4
I.1.3 – Replicación en arqueas	9
I.2 – Helicasas y la familia AAA+ de proteínas	11
I.2.1 – Clasificación filogenética de la familia MCM: La superfamilia AAA+ de proteínas NTPasas	11
I.2.2 – El dominio P-loop de unión e hidrólisis del nucleótido	12
I.2.3 – La familia de helicasas MCM en el grupo PS1BH	15
I.2.4 – Estructura del sitio de unión al nucleótido en helicasas	16
I.2.5 – Modelos de funcionamiento en helicasas	17
I.3 – Proteínas MCM	18
I.3.1 – Ensamblaje del complejo MCM eucariota	19
I.3.2 – MCM helicasa replicativa? Otros candidatos	21
I.4 – MCM de <i>Methanobacterium thermoautotrophicum</i>	22
I.4.1 – Dominio helicasa AAA+ C-terminal	24
I.4.2 – Dominio N-terminal de unión a DNA	26
I.5 – Microscopia electrónica	30
I.5.1 – Funcionamiento básico del microscopio electrónico	31
I.5.2 – Trabajo con las muestras	31
I.5.3 – Características de las imágenes de microscopía	33
I.5.4 – Procesamiento inicial de las imágenes	33
I.5.5 – Reconstrucción tridimensional	34
Objetivos	37
<hr/>	

Materiales y métodos	39
M.1 – Proteínas	39
M.2 – Secuencias de DNA: Fragmentos de restricción y oligonucleótidos	40
M.3 – Muestreo de condiciones físico-químicas	41
M.4 – Fosforilación radiactiva de moléculas de DNA celular contaminante	43
M.5 – Detección de complejos proteína:DNA en condiciones nativas	43
M.5.1 – Retardo en geles de acrilamida	43
M.5.2 – Retardo en geles de agarosa	44
M.5.3 – Transferencia de DNA a membranas: <i>southern-blot</i>	44
M.6 – Marcaje con estreptavidina o estreptavidina-oro	45
M.7 – Preparación de las muestras para microscopía electrónica	47
M.7.1 – Tinción negativa	47
M.7.2 – Criomicroscopía electrónica	47
M.8 – Microscopía electrónica	48
M.9 – Procesamiento digital de imágenes	48
M.10 – Reconstrucción tridimensional	50
M.11 – Modelado de estructuras atómicas	51
Resultados	53
R.1 – Estructura cuaternaria de mtMCM en estado libre	53
R.1.1 – Polimorfismo estructural	53
R.1.2 – Flexibilidad del extremo C-terminal	58
R.1.3 – Las vistas laterales son mayoritariamente proyecciones de dobles hexámeros	59
R.1.4 – Reconstrucción del doble hexámero	62
R.1.5 – Estructura del doble hexámero	64
R.1.6 – Ajuste de la estructura atómica de N-mtMCM	66

R.2	–	Análisis estructural de mutaciones en los residuos de la interfaz entre hexámeros del fragmento N-mtMCM	68
R.2.1	–	Interacciones en el fragmento amino terminal de mtMCM	68
R.2.2	–	Estructura de N-mtMCM en solución	68
R.2.3	–	Mutantes N-mtMCM-R161A y 6Ins	72
R.2.4	–	Estructura de los mutantes R161A y 6Ins	72
R.3	–	Caracterización de las condiciones para la obtención de complejos mtMCM:DNA y su estudio mediante ME	74
R.3.1	–	Descartando contaminación con DNA celular	75
R.3.2	–	Análisis de complejos N-mtMCM:DNA	76
R.3.3	–	Análisis de complejos mtMCM:DNA	85
R.4	–	Criomicroscopía de mtMCM	96
R.4.1	–	CrioME en agujeros (sin carbón)	96
R.4.2	–	CrioME sobre película de carbón	96
Discusión			101
<hr/>			
D.1	–	Polimorfismo estructural	102
D.1.1	–	Hexámeros y heptámeros, dobles y sencillos	102
D.1.2	–	Anillos abiertos	103
D.2	–	Estructura del doble anillo. Similitudes con LTag de SV40	104
D.2.1	–	Estructura de doble hexámero	104
D.2.2	–	Similitudes con LTag. Flexibilidad del dominio C-terminal	105
D.3	–	Formación de complejos con DNA	107
D.3.1	–	Triple hexámero de N-mtMCM	108
D.3.2	–	Unión de múltiples moléculas de MCM a una molécula de DNA	109
D.4	–	Modelos de funcionamiento	110
Conclusiones			115
Glosario			117
Bibliografía			121
<hr/>			

Introducción

La replicación del material genético es un proceso crucial del desarrollo y la perpetuación de todos los seres vivos. Se define como el mecanismo a través del cual una célula genera una copia exacta de su material genético. Es el procedimiento bioquímico por el que los organismos transmiten su herencia individual y de especie a través del tiempo.

Para llevar a cabo la replicación de una manera precisa y en el momento adecuado del ciclo de vida de la célula, es necesaria la intervención ordenada y coordinada de un gran número de elementos. La estructura y el funcionamiento de estos elementos han sido muy conservados en los seres vivos, de manera que a menudo las claves funcionales del proceso son compartidas por muchas especies.

Existen fundamentalmente dos esquemas replicativos: el de los seres procariontes (bacterias) y el de los eucariontes (protistas, hongos, vegetales y animales). El más estudiado y el mejor conocido por su simplicidad es el primero, siendo por su complejidad menos comprendido el eucarionte; debido al gran número de factores implicados, aún se está recopilando el conocimiento estructural y funcional necesario para comprender el proceso en su totalidad.

El conocimiento adquirido sobre el mecanismo bioquímico de la replicación deriva en importantes aplicaciones médicas, ya que es posible interrumpir el proceso bloqueando la actividad de proteínas específicas mediante el uso de drogas y antibióticos (norfloxacina (Goldstein 1987), ciprofloxacina (Hilliard *et al.*, 1995), camptotecina (Hsiang *et al.*, 1985), etc.), siendo éste un método valioso de control de procesos infecciosos o cancerígenos.

I.1 – Replicación del DNA celular

I.1.1 – *Entendiendo la replicación. Eventos fundamentales*

La replicación se divide mecánicamente en tres etapas diferenciadas: iniciación, elongación y terminación.

Durante la iniciación el DNA es preparado para ser duplicado y se ensambla la maquinaria proteica necesaria para realizar la copia. Ésta actúa en la fase de elongación, leyendo base a base la secuencia de cada hebra y sintetizando, base a base, una hebra hija complementaria. Finalmente, una vez se ha completado la copia, se desensambla la maquinaria en el momento de la terminación.

Todos los organismos comparten esencialmente el mismo mecanismo replicativo. En general, en todas las especies el proceso se lleva a cabo mediante una serie de eventos funcionales necesarios:

En primer lugar, la replicación requiere que la doble hélice de DNA se abra, es decir, que las dos cadenas de DNA se encuentren separadas, para que las proteínas DNA polimerasas puedan leer la secuencia de bases y sintetizar las hebras copia. Durante la iniciación esto implica los tres primeros eventos:

- Determinación de los puntos de inicio para el proceso (orígenes de replicación): Las proteínas denominadas iniciadoras se unen a lo largo del DNA en los lugares de inicio de la síntesis.
- Apertura de la doble hélice: Las proteínas iniciadoras también se ocupan de desestabilizar la doble hélice e iniciar su apertura, que luego será continuada a lo largo de toda la longitud del DNA por las enzimas helicasas.
- Carga y activación de enzimas helicasas: Antes de entrar en funcionamiento, las helicasas deben ser colocadas en el origen de replicación y activadas, a menudo mediante otras proteínas, denominadas cargadoras.

En los sistemas procariotas cada una de estas acciones es llevada a cabo por una proteína concreta, incluso en sistemas más sencillos como los virales una única proteína puede reunir las tres actividades necesarias. Sin embargo, en un sistema eucariota típico, puede ser necesaria la

actividad de más de 20 polipéptidos diferentes para completar la iniciación (Bell y Dutta 2002).

A lo largo de todo el proceso se generan hebras de DNA desapareadas, también llamadas regiones de DNA de banda sencilla (*single-stranded* DNA o ssDNA) susceptibles de ser degradadas por nucleasas de la propia célula. Estas regiones son protegidas mediante proteínas de unión a ssDNA.

Una vez se ha abierto la molécula de DNA se generan dos horquillas a través de las cuales pueden avanzar las enzimas que sintetizan las nuevas hebras de DNA, las DNA polimerasas. En cada horquilla se encuentran dos hebras de ssDNA. Para leer y replicar cada una se necesita un núcleo de DNA polimerasa (denominado *core*). Estos dos núcleos permanecen unidos entre sí en cada horquilla gracias a la acción de otras proteínas que hacen de nexo entre ambos.

Durante la elongación, los eventos necesarios se derivan en gran parte de varias características de las DNA polimerasas:

- Estas enzimas sólo puede sintetizar a partir de un extremo de ácido nucleico 3'-OH libre.
- Sintetizan sólo en un sentido (5'→3', en el del esqueleto azúcar-fosfato). Pero el DNA está formado por dos hebras complementarias antiparalelas (de sentidos opuestos).
- Cuando la polimerasa actúa individualmente no tiene procesividad suficiente, es decir, no completaría la síntesis de una molécula de DNA de gran longitud, la finalizaría prematuramente.

La primera circunstancia la resuelven enzimas RNA polimerasas, que sintetizan pequeños fragmentos de RNA que sirven como cebadores (también denominados iniciadores o *primers*) para la DNA polimerasa, ya que proporcionan el extremo 3'-OH libre que ésta necesita para iniciar la síntesis de DNA. En segundo lugar, dado que en una horquilla tenemos dos hebras de ssDNA de sentidos opuestos y la polimerasa sólo sintetiza en un sentido, una de las hebras debe ser reorientada para igualar el sentido de síntesis con el de avance de la replicación. En procariontes esto se resuelve formando bucles en la hebra 'retardada', de modo que el DNA en esta hebra se sintetiza de forma discontinua, iniciando una ronda nueva de replicación en cada bucle. Los fragmentos discontinuos de DNA generados,

denominados fragmentos de Okazaki, se unen mediante una proteína 'DNA ligasa'.

Para aumentar la procesividad de la polimerasa e impedir que ésta se libere del DNA antes de finalizar completamente la copia, la enzima se acopla a una proteína anular denominada pinza deslizante β (β *sliding clamp*), que se enhebra en la hebra desapareada y la recorre longitudinalmente. Esta proteína es cargada sobre el DNA por una proteína cargadora específica (β -*clamp loader*).

Fundamentalmente, las funciones necesarias y los eventos replicativos mencionados son compartidos por todos los seres vivos. En los organismos más evolucionados, como los eucariotas superiores, el aumento del nivel de organización implica necesariamente un control preciso del momento de la replicación celular, y por tanto un aumento de la complejidad del proceso. En el lado opuesto, organismos más sencillos como los virus pueden simplificar el proceso, incluyendo varias de las funciones necesarias en un único polipéptido y reduciendo los mecanismos de control, permitiendo así economizar recursos.

1.1.2 – Replicación en eucariotas

En los organismos eucariontes hay al menos dos circunstancias que requieren un aumento en la complejidad del sistema replicativo: la coordinación necesaria de la replicación con el ciclo celular, ya mencionada, y el gran tamaño, organización y compactación del DNA.

En procariontes existe un único origen con una secuencia bien determinada. El origen de *Escherichia coli* (*E. coli*) por ejemplo, denominado *oriC*, contiene una secuencia consenso de 9 pares de bases (pb) repetida 4 veces, denominada "DnaA box", que es reconocida por la proteína iniciadora DnaA. Sin embargo en eucariotas existen normalmente varios orígenes de replicación, y éstos, poco frecuentemente se definen mediante una secuencia consenso (como ocurre en *Saccharomyces cerevisiae* (Bell 1995)), sino más bien en regiones enriquecidas en pares de bases A/T (en *Schizosaccharomyces pombe* (Clyne y Kelly 1995; Dubey *et al.*, 1996)), motivos estructurales del DNA (en metazoos (Bell 1995)) o simplemente al azar (en los ovocitos de *Xenopus laevis* (Blow 2001)).

El complejo de reconocimiento del origen (ORC, *origin recognition complex*) está compuesto por 6 proteínas conservadas de diferentes

tamaños denominadas Orc1-6 (entre 120 y 50 kDa en *S. cerevisiae* (Bell y Stillman 1992)). Actúan como proteínas iniciadoras, análogamente a DnaA de *E. coli*, reconociendo cada complejo heterohexamérico un origen de replicación y uniéndose a él para comenzar a reclutar la maquinaria replicativa. En la mayoría de las especies el complejo permanece unido a los orígenes durante todo el ciclo celular (Bell y Stillman 1992; Diffley y Cocker 1992).

Orc1, 4 y 6 son enzimas que hidrolizan ATP (ATPasas) de la familia AAA+ (Bell y Dutta 2002). La unión de ORC al origen se produce sólo cuando Orc1 se encuentra unido una molécula de ATP (Bell y Stillman 1992; Klemm *et al.*, 1997).

Aunque ORC permanezca unido al un origen de replicación determinado durante todo el ciclo celular, este no se encuentra necesariamente activo, es decir, en estado competente para el ensamblaje de la maquinaria e iniciar la replicación (Fox *et al.*, 1995; Rusche *et al.*, 2002). Para que el origen se active es necesaria la fosforilación de ORC y otras proteínas mediante enzimas quinasas (CDK, *cyclin-dependent kinases* y DDK, *Dbf4-dependent kinase*). La activación o inactivación de estas enzimas se realiza en respuesta a otras señales de regulación celulares, en el momento adecuado del ciclo celular (Tabancay y Forsburg 2006).

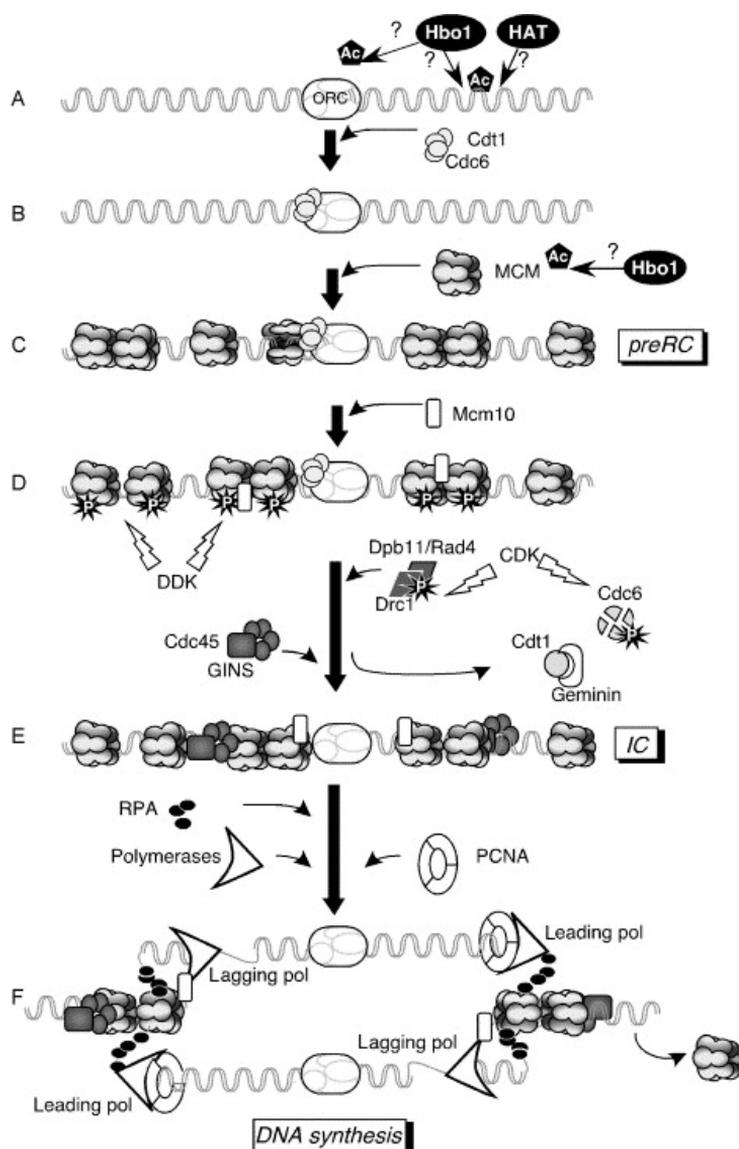


Figura I.1: Modelo de la replicación eucariota. (Extraído de Tabancay y Forsburg 2006).

Al menos dos proteínas más actúan posibilitando la carga de la helicasa en el origen, Cdc6 y Cdt1. Ambas se asocian de forma independiente al ORC y promueven la carga del complejo de proteínas MCM (proteínas de mantenimiento del minicromosoma, *minichromosome maintenance proteins*). Este complejo, formado por las seis proteínas homólogas Mcm2-7, posee actividad helicasa y se ha comprobado que es necesario para las etapas de iniciación y elongación de la replicación del DNA, siendo considerado la helicasa replicativa eucariota.

La proteína Cdc6 también posee actividad ATPasa y pertenece a la familia AAA+. Esta proteína guarda cierta homología con Orc1 y en menor medida también con Orc4 y 5 (Neuwald *et al.*, 1999). También tiene ciertas similitudes estructurales con proteínas cargadoras de la pinza deslizante β (Liu *et al.*, 2000). La hidrólisis del ATP unido en Orc1 y Cdc6 promueve la carga de múltiples complejos Mcm2-7 (Randell *et al.*, 2006). Posiblemente estas cuatro proteínas, Cdc6, Orc1, 4 y 5, actúen a modo de cargador de MCM sobre el DNA, tal y como lo hacen las β -clamp loaders (Mendez y Stillman 2003) (Figura I.2).

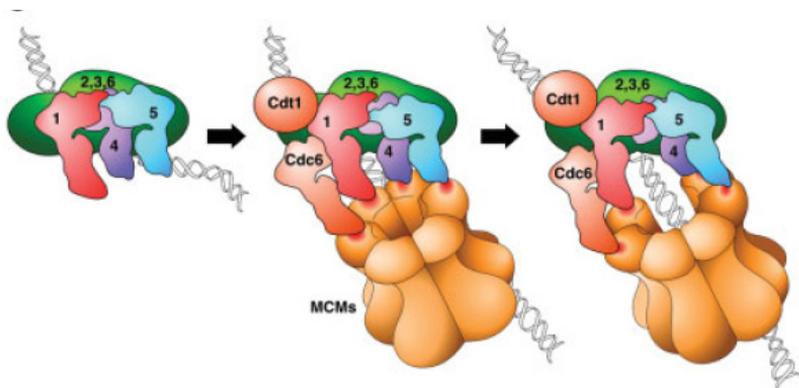


Figura I.2: Modelo propuesto por Mendez y Stillman 2003 para la carga de MCM en el DNA.

Mecanismo propuesto para la carga de MCM en el DNA a través de Cdc6, Orc1, 4 y 5 mediante el mecanismo de apertura del anillo.

A partir de la carga del complejo MCM en el origen éste se encuentra preparado para activarse y comenzar el desapareamiento de las dos hebras y el proceso replicativo. En este momento DDK fosforila a MCM sobre el DNA activando cada origen independientemente (Tabancay y Forsburg 2006), mientras que CDK, por su parte, lleva a cabo una complicada red de fosforilaciones que vinculan el momento de inicio de la replicación con el ciclo celular. Por medio de éstas, CDK activa definitivamente la replicación y posteriormente impide que los orígenes vuelvan a activarse de nuevo una vez han lanzado el proceso.

Las proteínas que regula CDK mediante fosforilación incluyen ORC, MCM y Cdc6, al igual que otros factores que actúan en la activación de la helicasa, como Dpb11 y TopBP1, y también proteínas que intervienen en la etapa de elongación, como la DNA polimerasa α y la proteína RPA de unión a ssDNA (revisado en Tabancay y Forsburg 2006).

La activación definitiva del origen se lleva a cabo mediante la carga de Cdc45 y la actividad del complejo heterotetramérico GINS (Kubota *et al.*, 2003; Takayama *et al.*, 2003). Cdc45, al igual que MCM, es requerida en las etapas de iniciación y elongación de la replicación (Hopwood y Dalton 1996; Tercero *et al.*, 2000) y parece ser el factor que definitivamente activa los complejos de MCM cargados próximos al origen. De este modo, mientras que la proporción celular de MCM:ORC es 40:1, la de Cdc45:ORC es 2:1 (Edwards *et al.*, 2002) sugiriendo que ejerce de “factor limitante” de la replicación, delimitando el número de orígenes en los que se puede iniciar la replicación simultáneamente.

El ensamblaje de Cdc45 actúa abriendo el origen de replicación y permitiendo la carga de RPA (Walter y Newport 2000). Para su carga en el origen es requerida la actividad de un complejo que incluye las proteínas TopBP1 y Dpb11 (Kamimura *et al.*, 1998), así como la presencia de la proteína MCM10 (Christensen y Tye 2003).

Esta proteína, MCM10, parece ser finalmente la responsable de vincular iniciación y elongación. Su nombre proviene de haber sido descubierta en los mismos experimentos en que lo fueron las proteínas del complejo Mcm2-7, pero no guarda ninguna homología con ellas. Además de intervenir en la carga de Cdc45, actúa estabilizando la DNA polimerasa α (Christensen y Tye 2003; Yang *et al.*, 2005).

Una vez el DNA del origen de replicación se ha abierto, la célula ensambla la maquinaria de síntesis de DNA y comienza la elongación.

Como hemos mencionado, es función de CDK impedir el reinicio de la replicación en un origen determinado, en el que ésta ya se haya iniciado. Esta proteína actúa inactivando por fosforilación las subunidades TopBP1 y Dpb11 del complejo cargador de Cdc45 (Tabancay y Forsburg 2006), y fosforilando Cdc6 cuya función quedará inhabilitada, según la especie, por degradación o bien por exportación del núcleo (Jiang *et al.*, 1999; Mimura *et al.*, 2004). Cdt1, por su parte se inactiva mediante la interacción con la proteína geminina, cuya concentración en la célula varía coordinadamente con el ciclo celular (Tada *et al.*, 2001).

A pesar del gran número de interacciones descritas, la función de cada uno de los elementos puede no ser exactamente la misma en las diferentes especies, complicando el establecimiento de cada una de las funciones del proceso y de las actividades requeridas para llevarlo a cabo. Esta circunstancia, combinada con la gran cantidad de interacciones que establecen entre sí todos los elementos, y la complicada regulación a la que se ven sometidos, mantiene aún en incógnita la solución a la fórmula del proceso replicativo.

1.1.3 – Replicación en arqueas

Las arqueas comparten con las bacterias una gran cantidad de características, como el código genético, los procesos de biosíntesis de los sillares de las macromoléculas: aminoácidos y nucleótidos; y los fundamentos de los procesos de transcripción y traducción. Sin embargo, ciertas características de estos dos últimos procesos, como la estructura y propiedades funcionales de la RNA polimerasa, los factores de transcripción y la secuencia de los promotores (Reeve *et al.*, 1997; Thomm 1996), son compartidas con los organismos eucariotas, al igual que algunos de los elementos de los aparatos de plegamiento de proteínas (Hofman-Bang *et al.*, 1999). Todas estas características sugieren la existencia de un antecesor común para ambos, no relacionado con las bacterias (Whitman *et al.*, 1999).

Como hemos visto en el apartado anterior, la replicación en eucariotas aún no se ha comprendido totalmente. Del mismo modo, tampoco se ha llegado a determinar en arqueas. Sin embargo, la mayoría de los componentes de la replicación que han sido identificados en estos microorganismos son compartidos con los eucariontes, como las proteínas

iniciadoras (ORC y Cdc6), el complejo MCM, las proteínas RPA, la pinza deslizante β (denominada PCNA, *proliferating cell nuclear antigen*, antígeno nuclear de proliferación celular) y su cargador (RFC, *replication factor C*, factor C de replicación) (Kelman y White 2005). Sin embargo, el menor grado de organización celular de las arqueas con respecto a los eucariotas sugiere una regulación más sencilla y por tanto una simplificación en el proceso, de hecho muchos de los componentes conservados muestran considerables simplificaciones con respecto a sus homólogos eucariotas:

El número de orígenes de replicación, por ejemplo, es muy pequeño. Así, aunque experimentalmente sólo han sido determinados en unas pocas especies, en las que se ha encontrado sólo un origen (*Pyrococcus abyssi* (Matsunaga *et al.*, 2001; Oyama *et al.*, 2001) y *Halobacterium* sp. (Zhang y Zhang 2003)), o bien 2 ó 3 que inician la replicación simultáneamente (*Sulfolobus solfataricus* (Robinson *et al.*, 2004) y *S. acidocalcaricus* (Lundgren *et al.*, 2004), en muchas otras el número de orígenes ha sido predicho informáticamente encontrándose resultados similares. En *Methanobacterium thermoautotrophicum*, por ejemplo, se ha calculado un único origen (Barry y Bell 2006).

Del mismo modo, en los genomas de arqueas se ha encontrado un número más reducido de homólogos para las proteínas replicativas, que en los de eucariotas. Así, la gran mayoría poseen entre 1 y 3 genes homólogos a Cdc6/ORC eucariota y un único homólogo para MCM (aunque algunas especies presentan hasta 4) (Barry y Bell 2006). No se ha encontrado equivalente para Cdt1, sugiriendo que tanto las funciones desempeñadas por los 7 polipéptidos de Cdc6 y ORC, como por Cdt1, deben ser desempeñados por los apenas 2 ó 3 homólogos de Cdc6/ORC encontrados, del mismo modo que las diferentes funcionalidades de Mcm2-7, habitualmente por un único polipéptido.

I.2 – Helicasas y la familia AAA+ de proteínas

I.2.1 – **Clasificación filogenética de la familia MCM: La superfamilia AAA+ de proteínas NTPasas**

Una gran parte del proteoma de cualquier organismo se encuentra formado por proteínas que unen nucleósidos trifosfato (NTP's) que utilizan normalmente como sustrato de reacciones bioquímicas. Dentro de las enzimas que unen NTP, las *P-loop* NTPasas (NTPasas que contienen un bucle de unión a fosfato (*phosphate-binding loop*)), son las más abundantes, constituyendo entre el 10 y el 18% de los productos génicos de los genomas procariotas y eucariotas (Saraste *et al.*, 1990). Las proteínas P-loop se caracterizan por contener un motivo de unión a NTP altamente conservado y distribuido en todos los reinos biológicos, al que también se denomina motivo Walker A.

Las proteínas NTPasas catalizan la hidrólisis de nucleótidos trifosfato en una reacción denominada desfosforilación. Esta reacción libera una gran cantidad de energía que es utilizada normalmente por la enzima para catalizar una reacción acoplada desfavorable.

La familia AAA de NTPasas (ATPasas asociadas con diferentes actividades celulares, (*ATPases associated with various cellular activities*)) está formada por miembros con diversas funcionalidades: degradación de proteínas (subunidades del proteasoma y metaloproteasas), fusión de vesículas, secreción de neurotransmisores, control del ciclo celular, progresión de la mitosis y de la meiosis, etc. Todas las proteínas estudiadas de esta familia que forman estructuras oligoméricas se ensamblan en anillos hexaméricos (Frickey y Lupas 2004).

Esta familia se expandió posteriormente en AAA+ mediante análisis comparativo de secuencia y estructura, incluyendo otras proteínas relacionadas, más distanciadas filogenéticamente, todas implicadas en el metabolismo de ácidos nucleicos, tales como las proteínas de la familia DnaA de iniciadores de la replicación bacteriana, ciertos factores de transcripción, las helicasas de la superfamilia III (SF3), y la familia de helicasas eucariotas MCM (Neuwald *et al.*, 1999). Un alineamiento posterior de secuencias basado en estructura amplió aún más esta familia, incluyendo además las familias de cargadores de helicasas y pinzas deslizantes, tanto bacterianos como eucariotas (DnaX, RFC, Cdc6), los iniciadores de la replicación eucariotas (ORC), la familia de chaperonas

ClpAB y otras familias de proteínas como RuvB, dineína, quelatasas, etc. (Iyer *et al.*, 2004).

Un análisis filogenético de estas familias en base a sus motivos estructurales, permite comprender el funcionamiento de las proteínas que las componen mediante el establecimiento de relaciones funcionales entre los motivos conservados.

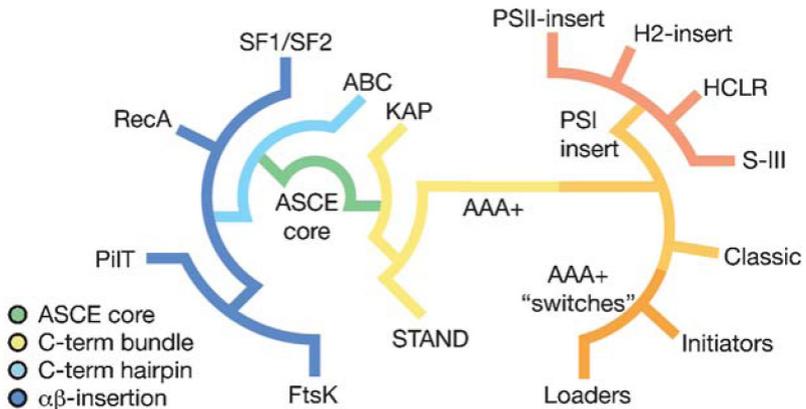


Figura 1.3: Clasificación filogenética de la familia de ATPasas AAA+ (Extraído de Iyer *et al.*, 2004).

1.2.2 – El dominio P-loop de unión e hidrólisis del nucleótido

Estructuralmente, un dominio *P-loop* típico contiene un conjunto de 5 láminas β paralelas dispuestas en un plano, rodeado a su vez por hélices α dispuestas por encima y por debajo de éste (Milner-White *et al.*, 1991). Dentro del dominio *P-loop* encontramos dos secuencias altamente conservadas, los motivos Walker A y B. Numerando los motivos de estructura secundaria desde el extremo amino al carboxilo, el motivo Walker A (el motivo *P-loop*, propiamente dicho, que da nombre a todo el dominio) se encuentra en el bucle entre la lámina β_1 y la hélice α posterior. El motivo Walker B se encuentra en la lámina β_3 (Guenther *et al.*, 1997; Iyer *et al.*, 2004; Neuwald *et al.*, 1999).

La secuencia del Walker A, en la familia AAA+, es típicamente GXXGXGK[S/T] con ligeras variaciones, donde 'X' representa cualquier aminoácido, aunque habitualmente al menos uno de los dos residuos situados entre las dos primeras glicinas, es prolina. La secuencia Walker B, en la familia AAA+, es típicamente hhhhDE, donde 'h' representa un residuo hidrofóbico. El glutamato conservado (E) cataliza la unión de una molécula de agua que realiza el ataque nucleófilo al fosfato y del ATP. La lisina (K) del motivo Walker A une por su parte el fosfato β y ambos retienen un catión Mg^{2+} también necesario para la catálisis. La lámina β_4 contiene otro motivo conservado, el motivo "sensor 1", que interviene en la hidrólisis del ATP interaccionando con el fosfato γ (Iyer *et al.*, 2004; Milner-White *et al.*, 1991; Saraste *et al.*, 1990; Story y Steitz 1992; Vetter y Wittinghofer 1999; Walker *et al.*, 1982).

La clase AAA+ incluye además otras características particulares en la estructura del dominio P-loop ATPasa, comparado con otras proteínas *P-loop* (Iyer *et al.*, 2004; Ogura y Wilkinson 2001): En primer lugar, la secuencia N-terminal que precede a la lámina β_1 forma un bucle de residuos que recorre perpendicularmente el plano de láminas β y termina en una hélice α (denominada α_0) en la que podemos encontrar algunos residuos conservados. También es característico de esta clase un haz de 4 hélices α con el que finaliza el dominio, situado posteriormente, es decir, C-terminal en secuencia, a la lámina β_5 . Esta estructura contiene el motivo conservado "sensor 2", que parece tener un importante papel en mediar cambios conformacionales asociados al estado de fosforilación del nucleótido unido, habiéndose detectado un desplazamiento relativo del haz, según el nucleótido unido a la proteína se trate de ATP o ADP. Por último, la hélice α previa a la lámina β_5 contiene un motivo dedo de arginina que, como veremos más adelante, en las helicasas oligoméricas se desplaza hasta el *P-loop* del monómero adyacente favoreciendo la hidrólisis de ATP.

classical AAA+ ATPases

CipAB-N
(1KSF,1JBK)
E. coli

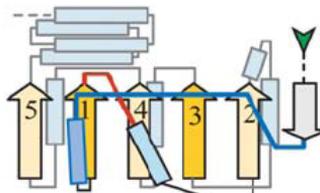


Figura I.4: Esquema del dominio P-loop clásico en las AAA+ ATPasas.

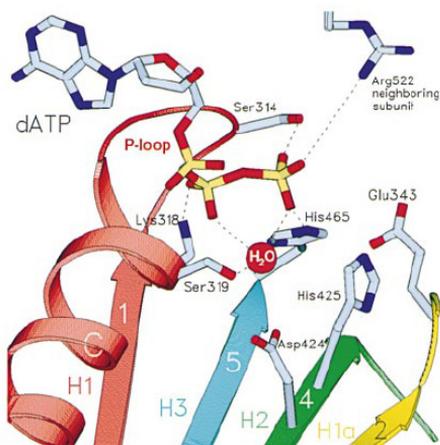


Figura I.5: Detalle estructural de un NTP-binding domain en gp4 de T7.

1.2.3 – La familia de helicasas MCM en el grupo PS1BH

La familia de helicasas MCM pertenece, dentro de la clase AAA+, al grupo PS1BH (*pre-sensor 1 β -hairpin*, bucle β previo al sensor 1), y a su vez, al subgrupo h2i (*helix 2 insert*). El grupo PS1BH incluye una gran cantidad de familias de proteínas, y se caracteriza por la inserción de 2 láminas β antiparalelas antes del motivo sensor 1, es decir, antes de la lámina $\beta 4$ (Iyer *et al.*, 2004). Se ha observado que este motivo se desplaza casi 2 nm entre las etapas de unión de ATP y la posterior hidrólisis (Gai *et al.*, 2004).

El grupo h2i se caracteriza por un inserto β - α - β que divide en dos la hélice $\alpha 2$, posterior a la lámina $\beta 2$. Estas dos láminas forman puentes de hidrógeno entre ellas (Fodje *et al.*, 2001; Iyer *et al.*, 2004) y ha sido propuesto que pueden actuar separando físicamente las dos hebras del DNA, actuando como una cuña (Jenkinson y Chong 2006; Takahashi *et al.*, 2005). También es característico del grupo h2i la inserción de una larga hélice α tras la lámina $\beta 5$, situada antes del haz sensor 2, que cambia la orientación de éste con respecto al núcleo β de la región ATPasa.

Se ha estudiado la función de algunas de estas regiones características en proteínas de la familia MCM en arqueas. Se ha determinado que la eliminación del inserto h2i, en el caso de MCM de *M. thermoautotrophicum*, supone un aumento en la afinidad de la proteína por ssDNA ó dsDNA, y un aumento también en su actividad ATPasa estimulada por dsDNA. Sin embargo, esta mutación causa la pérdida de la actividad helicasa (Jenkinson y Chong 2006).

Las mutaciones en los residuos conservados del PS1BH en el caso de la arquea *S. solfataricus*, produce una disminución de la capacidad de unión al DNA y la eliminación completa de la actividad helicasa (McGeoch *et al.*, 2005).

Además de los motivos mencionados, característicos de los grupos PS1BH y h2i, la familia de proteínas MCM se caracteriza estructuralmente por compartir un motivo de unión a zinc N-terminal (Tye 1999) y un motivo HTH (*helix-turn-helix*, hélice-giro-hélice) C-terminal (Aravind y Koonin 1999). Ambos funcionan usualmente como motivos de unión a DNA (Aravind y Koonin 1999; Brennan y Matthews 1989), aunque el dedo de Zn también puede intervenir en interacciones proteína-proteína (Laity *et al.*, 2001). En secuencia, la familia comparte sus propios motivos Walker específicos, como veremos más adelante.

I.2.4 – Estructura del sitio de unión al nucleótido en helicasas

Las enzimas helicasas se clasifican tradicionalmente en familias atendiendo a similitudes en sus secuencias conservadas (Gorbalenya y Koonin 1993). Las superfamilias 1 y 2 (SF1 y SF2) están constituidas por helicasas a menudo monoméricas o dimericas, muy relacionadas entre sí y formadas por varios dominios (PcrA (Subramanya *et al.*, 1996), Rep de *E. coli* (Korolev *et al.*, 1997), NS3 de HCV (Yao *et al.*, 1997)). A menudo uno o varios de estos dominios son estructuralmente similares a la proteína RecA (una ATPasa que funciona en reparación y recombinación del DNA (Story *et al.*, 1992)). Las proteínas de las SF1 y 2 contienen un único sitio de unión a NTP tipo *P-loop* situado en la interfaz entre dos dominios, frecuentemente ambos similares a RecA.

Las superfamilias SF3 y SF4 están constituidas por helicasas oligoméricas en forma de anillo, muy a menudo hexaméricas. Los miembros de estas familias forman los bolsillos de unión a NTP en la interfaz entre dos monómeros contiguos, frecuentemente utilizando cada bolsillo diferentes motivos de un monómero u otro (Patel y Donmez 2006). Algunas miembros de estas familias, como gp4 de T7 (Sawaya *et al.*, 1999) o DnaB de *Thermus aquaticus* (Bailey *et al.*, 2007), guardan también homología con RecA en sus dominios helicasa C-terminales.

1.2.5 – Modelos de funcionamiento en helicasas

Actualmente, el conocimiento estructural sobre los cambios conformacionales implicados en la relación entre las diferentes funciones de las proteínas helicasas es limitado. En los últimos años se está comenzando a recabar esta información, que permitirá determinar cómo la unión e hidrólisis del nucleótido impulsa la translocación de la enzima sobre el ácido nucleico y el desapareamiento de la doble hebra o la eliminación de otras estructuras secundarias. Por el momento, la resolución de determinadas estructuras ha permitido plantear mecanismos para determinadas proteínas:

Así, se han propuesto varios mecanismos de carga de las helicasas hexaméricas en el DNA, estos son: enhebrado, en el que la helicasa se carga por un extremo libre del DNA, introduciendo una de las dos hebras por el canal central del anillo; reorganización, por el cual el oligómero se formaría directamente sobre el DNA a partir de sus monómeros; y apertura del anillo, en el que la helicasa se forma previamente y posteriormente se carga sobre el DNA abriendo la interfaz entre dos monómeros y cerrándose el anillo posteriormente (Patel y Picha 2000). Actualmente el modelo de apertura del anillo se ha postulado para la carga de gp4 de T7 (Ahnert *et al.*, 2000; Crampton *et al.*, 2006) y la de DnaB de *E. coli* (Bujalowski y Jezewska 2000), y se ha determinado estructuralmente para la carga sobre el RNA del factor de transcripción ρ (Skordalakes y Berger 2003).

Del mismo modo, se han propuesto diferentes mecanismos en base a resultados bioquímicos, cinéticos y estructurales, para la translocación de la helicasa sobre el ácido nucleico, para la separación de las hebras en la molécula de dsDNA y para la transducción de la energía de la hidrólisis del nucleótido en translocación y desapareamiento de las hebras (Patel y Picha 2000; Patel y Donmez 2006). La publicación recientemente de la estructura de la helicasa E1 del virus del papiloma bovino (BPV), en complejo con DNA y en diferentes estados de unión a ATP, ha proporcionado valiosa información para entender, en esta proteína, cómo se acopla la hidrólisis secuencial del nucleótido en las diferentes subunidades del hexámero a la translocación sobre la hebra de ssDNA naciente (Enemark y Joshua-Tor 2006).

I.3 – Proteínas MCM

En todas las especies eucariontes y arqueas estudiadas han sido encontrados miembros de la familia de helicasas MCM (Forsburg 2004, Myllykallio y Forterre 2000).

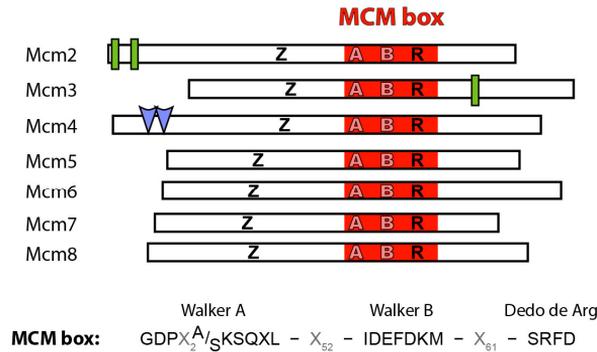
En esta familia, la región AAA+ se encuentra definida mediante sus propias secuencias específicas (Kearsey y Labib 1998; Koonin 1993), delimitando una región MCM consenso (*MCM box*) que es la versión de esta familia de la región AAA+ de 200 aminoácidos. La secuencia MCM consenso incluye de este modo versiones específicas de los motivos Walker y dedo de arginina (ver Figura I.6).

Además de las secuencias conservadas de la región consenso, todas las proteínas MCM contienen en su secuencia un dedo de zinc N-terminal, que en algunas especies es necesario para la viabilidad celular, ensamblado de la proteína, o la actividad ATPasa (Fletcher *et al.*, 2003; Forsburg *et al.*, 1997; Poplawski *et al.*, 2001; Sherman *et al.*, 1998; You *et al.*, 2002).

En todos los eucariotas estudiados se han encontrado al menos, 6 proteínas homólogas de la familia MCM, denominadas Mcm2-7 (Tye 1999). A parte de la región MCM consenso y el dedo de zinc, las proteínas Mcm2-7 no comparten una homología particular entre ellas. Sin embargo comparando sus secuencias entre diferentes especies se ha comprobado que, por ejemplo, la proteína Mcm2 humana guarda mayor homología con Mcm2 de *S. cerevisiae* (un ortólogo, es decir, una proteína homóloga perteneciente a una especie diferente) que con la proteína Mcm4 humana (un parólogo, es decir, un homólogo dentro de la misma especie). Esto permite considerar a las 6 proteínas Mcm2-7 como pertenecientes a 6 clases de proteínas MCM diferentes, y sugiere una función específica para cada una de las clases.

Adicionalmente, algunas de las clases muestran secuencias específicas, no encontradas en las demás: Por ejemplo, Mcm2 y Mcm3 contienen secuencias de localización nuclear (Ishimi *et al.*, 2001; Pasion y Forsburg 1999) y Mcm4 contiene una secuencia para CDK (Forsburg 2004) (Figura I.6).

A



▽ : Regulación por CDK █ : Señal de localización nuclear

B

	<i>S. cerevisiae</i>	<i>X. laevis</i>	<i>H. sapiens</i>
Mcm2	868	886	904
Mcm3	971	807	808
Mcm4	933	863	863
Mcm5	776	735	734
Mcm6	1017	796/825(Z)	821
Mcm7	845	720	719
Mcm8			840

Figura 1.6: Secuencias conservadas en la familia MCM.

A: Secuencias conservadas en los miembros de la familia MCM. Se detalla la secuencia “MCM box” C-terminal con los motivos conservados Walker A y B (A y B en la figura) y el dedo de arginina (R). Z indica la posición del dedo de zinc N-terminal. **B:** Número de aminoácidos de cada proteína para las especies indicadas. Z indica la forma embrionaria (zygotic) de Mcm6. (Basado en Forsburg 2004 y Tye 1999)

1.3.1 – Ensamblaje del complejo MCM eucariota

Numerosos estudios indican que Mcm2-7 se asocia *in vivo* formando un heterohexámero de estequiometría 1:1:1:1:1:1 (Adachi *et al.*, 1997). Se ha comprobado que esta oligomerización es necesaria para el transporte al núcleo y para la función, ya que no hay mutantes en ensamblaje que sean viables (Lei *et al.*, 2002; Sherman *et al.*, 1998).

In vitro, sin embargo, las subunidades de MCM muestran diferentes afinidades de oligomerización (Coue *et al.*, 1998; Davey *et al.*, 2003; Lee y Hurwitz 2000):

Mcm4, 6 y 7 tienen una fuerte afinidad y forman un complejo trimérico, denominado “MCM core” (Ishimi 1997; Lee y Hurwitz 2001). Por su parte, Mcm2 muestra una reducida afinidad de ensamblaje con el core, mientras que Mcm3 y 5 pueden formar dímeros y unirse a otras subunidades con una débil afinidad.

El *core* puede dimerizar en (Mcm4,6,7)₂, aunque sólo en ausencia de las otras subunidades y se deshace nuevamente en trímeros añadiendo Mcm2. Este complejo doble-trimérico posee una débil actividad helicasa 3'→5', es decir, avanza desapareando la doble hélice del DNA sobre una de las hebras desde el extremo 3'-OH hasta el 5'-P. Esta actividad desaparece en presencia de las otras subunidades. La influencia que ejercen Mcm2, 3 y 5 sobre la actividad ATPasa y la capacidad de dimerización del *core* sugieren una posible actividad reguladora para estas proteínas.

A pesar de las diferentes funciones de cada una de las clases, las evidencias indican que el complejo Mcm2-7 comparte una función común, ya que la eliminación de cualquiera de las proteínas que lo constituyen produce similares fenotipos erróneos en replicación (Madine *et al.*, 1995; Romanowski *et al.*, 1996).

Estudios de actividad ATPasa realizados en *S. cerevisiae* sugieren un modelo de organización del posible heterohexamero (Davey *et al.*, 2003): Los monómeros de MCM aislados no poseen actividad ATPasa, pero se ha observado que ciertas parejas de proteínas MCM (Mcm3,7, Mcm4,7 y Mcm2,6) pueden asociarse y generar esta actividad. Estos trabajos junto al estudio de fenotipos de mutantes en los motivos Walker (Crevel *et al.*, 2001) sugieren un modelo en que las unidades reguladoras se alternan con las proteínas del *core* (Davey *et al.*, 2003).

A pesar de que la única forma con actividad helicasa *in vitro* es el *core*, la forma heterohexamérica es la predominante *in vivo*. Esto podría inducir a pensar que la forma activa durante la replicación es Mcm4,6,7, y que el heterohexamero es una forma inactiva de la helicasa predominante cuando la célula no se encuentra en replicación. Sin embargo, Mcm2 y 3 son necesarias no sólo durante la iniciación de la replicación, sino durante toda la fase S, y sus mutantes son indistinguibles de aquellos en los que se inactiva Mcm4, 6 ó 7 (Labib *et al.*, 2000).

Aunque la forma activa del complejo y la función de cada subunidad MCM, son todavía un misterio, es posible que su correcto funcionamiento dependa de la actuación de otras proteínas. De acuerdo a esta suposición, se ha observado que el complejo MCM de *X. laevis* aumenta su actividad helicasa si Cdc45 se encuentra presente (Masuda *et al.*, 2003)

I.3.2 – MCM helicasa replicativa? Otros candidatos

A pesar de ser el complejo Mcm2-7 el mejor candidato a helicasa replicativa en eucariotas y de encontrarse su implicación en la etapa de iniciación de la replicación firmemente establecida, aún quedan cuestiones por resolver para poder afirmar que ésta es la función de estas proteínas. Una de ellas es la ya explicada dificultad para determinar la forma activa de este complejo, la cual debe poseer una actividad helicasa dotada de la procesividad característica de las helicasas replicativas.

En segundo lugar, la denominada “paradoja de MCM” (Dimitrova *et al.*, 1999; Hyrien *et al.*, 2003), mediante la cual se hace referencia a la extensa distribución que adopta *in vivo* el complejo MCM sobre toda la cromatina, no permaneciendo concentrada en los lugares de síntesis (los *foci* de replicación), a diferencia de otras helicasas replicativas (Edwards *et al.*, 2002; Romanowski *et al.*, 1996). Este puede ser el motivo por el que aún no se han encontrado interacciones físicas claras entre Mcm2-7 y el resto de componentes de la maquinaria replicativa en la horquilla de replicación, como las DNA polimerasas o las proteínas RPA de unión a DNA de banda simple.

Algunas proteínas con actividad helicasa implicadas en la replicación, tales como Dna2, o las helicasas homólogas a RecQ de *E. coli*, Sgs1 de *S. cerevisiae*, BLM ó WRN (helicasas de los síndromes de Bloom y Werner)), y también otras helicasas como Pif1 y Rrm3, fueron candidatos a helicasa replicativa, pero, o bien se ha demostrado que desempeñan otros papeles, como la maduración de los fragmentos de Okazaki (Dna2, Labib y Diffley 2001), o bien fueron descartadas porque sus mutantes de delección no producen el fenotipo letal esperado (Pif1 y Rrm3, Foury y Lahaye 1987; Ivessa *et al.*, 2000), o bien sus mutantes no impiden la replicación (Sgs1, Gangloff *et al.*, 2000).

Recientemente ha sido identificado un nuevo miembro de la familia MCM en células de vertebrados, denominado Mcm8 (Gozuacik *et al.*, 2003), que comparte la secuencia consenso *MCM box*. Esta proteína posee actividades ATPasa y helicasa *in vitro*, colocaliza con los *foci* de replicación e interacciona con la proteína RPA, uniéndose al replisoma después de la carga de Mcm2-7. Sin embargo, su presencia no es esencial ya que otras helicasas, como Mcm2-7, pueden compensar el efecto de su ausencia. Además sólo existe en vertebrados, de modo que el papel de helicasa replicativa en el resto de eucariotas debería ser por tanto desempeñado por

otra proteína (Maiorano *et al.*, 2006). Mcm9, otra proteína de la familia, está muy relacionada con Mcm8 considerándose que fue originada por medio de una duplicación génica. Esta proteína se encuentra del mismo modo sólo en vertebrados, y su función no ha sido aún estudiada (Maiorano *et al.*, 2006).

Los diferentes comportamientos *in vivo* del complejo MCM en las diferentes especies dificultan la elucidación de su función en la replicación. Así, mientras que el mejor argumento a favor del complejo Mcm2-7 como helicasa replicativa es que sus mutantes son incapaces de llevar a cabo la replicación del DNA, por ejemplo, en *X. laevis*; sin embargo, esto no ocurre así en *S. pombe*, donde los mutantes son capaces de sintetizar abundantes cantidades de DNA (Forsburg 2004). Del mismo modo, mientras que el papel del complejo en la iniciación de la replicación se encuentra establecido y aceptado por la comunidad científica, su funcionamiento como helicasa replicativa durante la elongación es una cuestión en debate, existiendo argumentos a favor (Aparicio *et al.*, 1997; Labib *et al.*, 2000) y en contra (Hennessy *et al.*, 1990; Nasmyth y Nurse 1981).

I.4 – MCM de *Methanobacterium thermoautotrophicum*

Hay una única proteína homóloga de MCM en la arquea *Methanobacterium thermoautotrophicum* (mtMCM). Esta proteína consta de 666 aminoácidos y un peso molecular de 75,5 kDa (Smith *et al.*, 1997). Mediante cromatografía de filtración en gel eluye como un pico de elevado peso molecular (~950 kDa), cuyo tamaño concuerda con el de un oligómero dodecamérico, que en ocasiones aparece acompañado de un pico minoritario de proteína en forma monomérica (Chong *et al.*, 2000; Kelman *et al.*, 1999; Shechter *et al.*, 2000). En determinadas condiciones (concentración de NaCl > 750 mM) puede aparecer un pico de un oligómero de peso molecular intermedio, probablemente debido a la disociación del dodecámero en hexámeros sencillos (Shechter *et al.*, 2000).

En otras arqueas también se ha determinado la presencia de helicasas de la familia MCM, existiendo sólo un homólogo en el genoma de *Archaeoglobus fulgidus* (afuMCM), *Sulfolobus solfataricus* (ssoMCM) y en la mayoría de arqueas estudiadas (*Pyrococcus* sp., *Aeropyrum pernix*,

Thermoplasma acidophilum, etc.), salvo en *Methanococcus jannaschii* donde existen de tres a cuatro secuencias relacionadas, según genoma (Myllykallio y Forterre 2000). Tanto afuMCM como ssoMCM forman hexámeros en solución (Carpentieri *et al.*, 2002; Grainge *et al.*, 2003; McGeoch *et al.*, 2005), aunque se ha informado que en altas concentraciones ssoMCM puede formar oligómeros de mayor tamaño, compatibles con una estequiometría dodecamérica (Barry *et al.*, 2007). También se ha informado de la presencia de la proteína ssoMCM unida a ssDNA en forma hexamérica (McGeoch *et al.*, 2005).

mtMCM es la proteína de la familia MCM que muestra el polimorfismo más amplio. Además de haberse caracterizado mediante cromatografía de exclusión las formas dodecamérica y hexamérica, se han reconstruido a partir de datos de microscopía electrónica estructuras hexaméricas (Pape *et al.*, 2003), heptaméricas (Yu *et al.*, 2002) y helicoidales (Chen *et al.*, 2005) de esta tesis se determinó por primera vez la estructura del doble hexámero (Gómez-Llorente *et al.*, 2005). Y más recientemente, se publicó la reconstrucción de una forma doble heptamérica (Costa *et al.*, 2006a, 2006b).

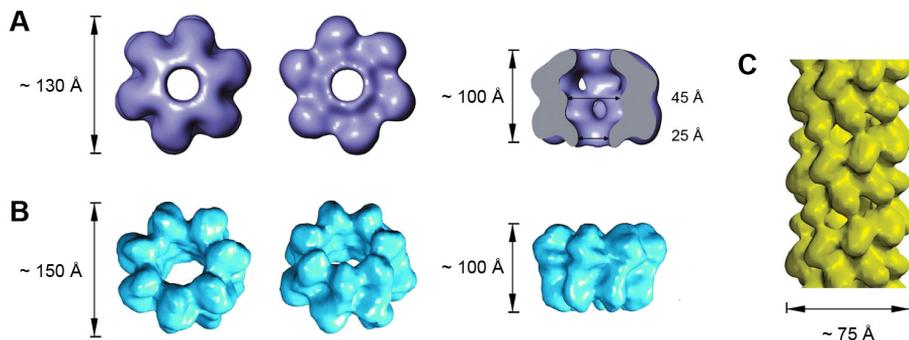


Figura 1.7: Estructuras de mtMCM resueltas mediante microscopía electrónica.

A: Hexámero sencillo en vistas frontales (superior e inferior) y en sección lateral. **B:** Heptámero sencillo en dos vistas frontales (superior e inferior) ligeramente inclinadas y en vista lateral. **C:** Estructura helicoidal. Todas las estructuras se encuentran atravesadas por canales centrales de diferentes dimensiones.

Cualquiera de las dos formas observadas en cromatografía, tanto la de alto peso molecular, como la monomérica, pueden unir ssDNA, hidrolizan ATP en presencia de DNA y poseen actividad helicasa 3'→5' DNA dependiente (Kelman *et al.*, 1999). Su actividad ATPasa se estimula tanto con ssDNA como con dsDNA (Chong *et al.*, 2000; Shechter *et al.*, 2000), pero una concentración alta de ssDNA termina inhibiendo esta actividad, mientras que una alta concentración de dsDNA no produce ese efecto (Chong *et al.*, 2000). Para el funcionamiento de la actividad helicasa es necesaria la hidrólisis de ATP ó dATP (Kelman *et al.*, 1999), y no es necesaria la presencia de una cola 5' ó 3' protuberante (Chong *et al.*, 2000) como para muchas helicasas. La procesividad de mtMCM es alta, siendo capaz de desaparecer 500pb de dsDNA *in vitro* (Chong *et al.*, 2000; Shechter *et al.*, 2000).

1.4.1 – Dominio helicasa AAA+ C-terminal

Un alineamiento de secuencia en la familia MCM permite distinguir en mtMCM dos regiones diferenciadas (Tye 1999): una región N-terminal no conservada (que incluye los primeros 280 aminoácidos); y una región C-terminal (aproximadamente los 380 aminoácidos restantes hasta el extremo C-terminal) conservada en secuencia en la familia y que contiene el dominio ATPasa AAA+, responsable de la actividad helicasa (Chong *et al.*, 2000) (Figura 1.8). Por tanto, la región C-terminal contiene los motivos característicos de la familia AAA+, Walker A, Walker B y dedo de arginina, y también los motivos PS1BH y h2i (Iyer *et al.*, 2004). Los 100 últimos aminoácidos de la región C-terminal forman el motivo HTH de unión a dsDNA presente en todos los miembros de la familia MCM (ver Introducción, 1.2.3).

Un análisis de la capacidad de unión a DNA de cada dominio en MCM de *S. solfataricus* muestra que el dominio N-terminal (aa's 1-266) une ssDNA, pero en una cantidad considerablemente menor que la proteína completa. Por su parte, el dominio C-terminal por sí sólo une aún menos que el fragmento N-terminal, mientras que el dominio HTH no une DNA de una forma detectable (Barry *et al.*, 2007). Por otro lado, la eliminación del dominio HTH no tiene apenas efecto en la capacidad de unión (Barry *et al.*, 2007). El dominio HTH en ssoMCM no parece por tanto, ocuparse de unir el DNA a la proteína.

Sin embargo, algunos estudios apuntan que su función podría ser la de modular el comportamiento de la proteína en función de su estado de unión al sustrato. Así, se ha informado de que la eliminación de este motivo influye en la estimulación de la actividad ATPasa mediante DNA en mtMCM (Jenkinson y Chong 2006), y en la estimulación de las actividades ATPasa y helicasa en ssoMCM (Grainge *et al.*, 2003), inhibiendo esta estimulación cuando no hay DNA presente.

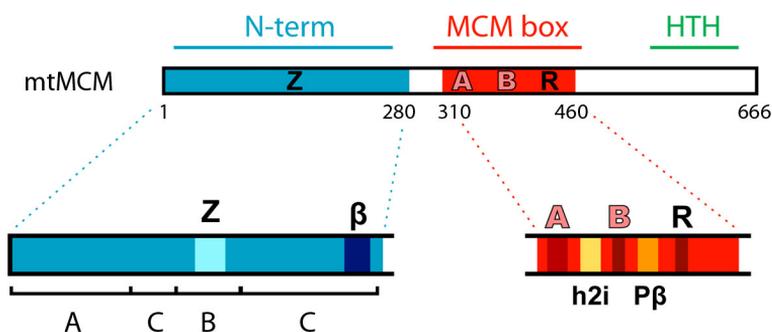


Figura 1.8: Mapa de la secuencia de mtMCM.

Distribución de los motivos conservados de la familia MCM y h2i a lo largo de la secuencia de mtMCM. Se indica la localización aproximada de las regiones *MCM box* (rojo) HTH y la región N-terminal (azul), de estructura atómica conocida. Para esta última se especifica la localización en la secuencia de sus tres dominios estructurales y las posiciones de los motivos de unión a zinc (Z) y β -hairpin (β) (consultar Introducción, 1.4.2). En la región *MCM box* se especifican los motivos Walker A y B (A y B) y el dedo de arginina (R) que forman el bolsillo de unión a ATP, y los motivos h2i y PS1BH (P β).

I.4.2 – Dominio N-terminal de unión a DNA

El fragmento amino terminal de mtMCM (N-mtMCM) (aminoácidos 2-286, 32,3 kDa) ha sido cristalizado y su estructura ha sido resuelta a 3,1 Å de resolución (Fletcher et al., 2003) (Figura I.9, A). Es la única estructura perteneciente a una proteína de la familia MCM que ha sido resuelta a nivel atómico. Su perfil cromatográfico indica que eluye, sin necesidad de cofactores, como un oligómero muy estable de alto peso molecular (~380 kDa), probablemente un doble hexámero, que sólo se separa en dos hexámeros sencillos a altas concentraciones de urea (4 M en Tris·Cl 100 mM pH 8,0, NaCl 250 mM, DTT 1 mM) (Fletcher et al., 2003; 2005).

En un tampón Tris·Cl 50 mM pH 8,0, citrato sódico 0,8 M y glicerol 10% (v/v), cristaliza en un doble hexámero (dHex) formado por dos hexámeros individuales anulares enfrentados cabeza con cabeza por sus extremos N-terminales, quedando en el interior de la estructura un largo canal central que la atraviesa de un extremo a otro (Figura I.10 y Figura R.9, A)

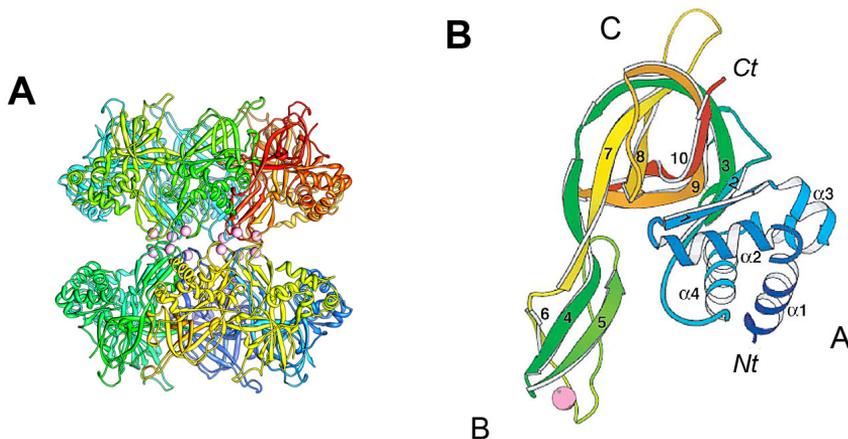


Figura I.9: Estructura atómica de N-mtMCM. Representación de estructuras secundarias.

A: Estructura global del doble hexámero en vista lateral. Cada monómero se ha representado coloreado independientemente. **B:** Estructura de un monómero mostrando los tres dominios (A, B y C), los extremos N- y C-terminal (Nt y Ct) y

el átomo de zinc, en color rosado. (La figura en B ha sido tomada de Fletcher *et al.*, 2003.)

Un alineamiento de secuencia dirigido por estructura del fragmento N-mtMCM frente a otros miembros de la familia MCM muestra que, a pesar de encontrarse esta región muy poco conservada en secuencia, si hay una gran conservación estructural (Fletcher *et al.*, 2003).

Cada monómero individual se pliega en 3 dominios distintos: A, B y C, siendo el dominio A el más N-terminal y el C el más C-terminal (Figura I.8 e Figura I.9, B). El dominio A se encuentra formado por las primeras 4 hélices α ($\alpha 1 - 4$) que se empaquetan formando en el dHex una protuberancia independiente que se proyecta hacia el exterior de la estructura. Aún no se ha determinado el papel de este dominio aunque se baraja un posible papel regulador (Kasiviswanathan *et al.*, 2004). Su eliminación en ssoMCM no tiene influencia en la capacidad de unión a DNA (Barry *et al.*, 2007).

Mientras que el fragmento N-mtMCM posee capacidad de unión a DNA de cadena doble o sencilla (Fletcher *et al.*, 2003), en *A. fulgidus*, la delección de todo el dominio N-terminal produce una proteína que conserva tanto la capacidad de hexamerización, como la capacidad de unión a DNA y la actividad helicasa (Grainge *et al.*, 2003). En *M. thermoautotrophicum*, sin embargo, el dominio N-terminal parece ser el principal responsable de la capacidad de oligomerización y de unión a DNA, siendo los dominios B y C los que parecen tener mayor influencia en estas funciones:

El dominio B está formado completamente por láminas β ($\beta 4 - 6$). Contiene un motivo de unión a Zn (Figura I.9, B), que en la proteína N-mtMCM es el único que se encuentra en posición adecuada para producir interacciones hexámero-hexámero. La estructura sugiere que este motivo mantiene la interacción entre ambos hexámeros mediante puentes de hidrógeno e interacciones hidrofóbicas cruzadas entre los residuos de ambos. El átomo de Zn proporcionaría la estabilidad estructural necesaria a los bucles para mantener las interacciones (Fletcher *et al.*, 2003).

Sin embargo, la información sobre la función del motivo de unión a Zn parece de nuevo contradictoria: Por un lado, la eliminación completa del dominio B ha sido estudiada y se ha informado que los mutantes generados sobre la proteína mtMCM completa son capaces de generar estructuras dodecaméricas, pero incapaces de unir DNA (Kasiviswanathan *et al.*, 2004). Adicionalmente, una mutación sustitutiva en la cisteína 158 del motivo provoca una reducción tanto de la capacidad de unión a Zn, como de la capacidad de unión a DNA, pero no afecta a la estabilidad del doble

hexámero (Poplawski *et al.*, 2001). Sin embargo, mutantes de las proteínas MCM4 y 7 eucariotas en el mismo motivo se ven afectados en la capacidad de oligomerización del complejo hexamérico MCM4/6/7 y disminuidos en su capacidad de unión a DNA (You *et al.*, 2002).

Los estudios publicados en esta tesis demuestran que mutaciones que alteran la funcionalidad de la arginina 161, perteneciente al motivo de unión a Zn, eliminan la capacidad de formación del dHex para el fragmento N-mtMCM. Paralelamente a este trabajo se confirmó el efecto de esta misma mutación en la proteína mtMCM completa, que se muestra incapacitada para formar dHex y disminuida en su capacidad de unión a DNA (Fletcher *et al.*, 2005).

El dominio C se encuentra situado entre el A y B y está constituido por dos regiones separadas en secuencia, incluyendo una región intermedia (lámina β 3) y la región más C-terminal de la estructura (β 7 – 10). El dominio C es el principal responsable de la oligomerización entre monómeros, siendo incapaz de oligomerizar un mutante de delección de este dominio. El mismo estudio afirma también que el dominio C aislado es capaz de formar por sí solo estructuras dodecaméricas (Kasiswathan *et al.*, 2004). Otro estudio afirma que un mutante de delección de N-mtMCM en los primeros 111 aminoácidos (que incluyen los primeros 18 aminoácidos del dominio C), forma sólo monómeros aislados (Chong *et al.*, 2000).

Los dominios B y C forman la pared del canal central, constituida exclusivamente por láminas β . Este canal se encuentra extraordinariamente cargado positivamente, mientras que el exterior de la molécula se encuentra muy cargado negativamente (Figura I.10). Sus dimensiones (23 Å de diámetro mínimo) permitirían que una molécula de dsDNA atravesara la estructura de un extremo a otro, del mismo modo que se ha demostrado que ocurre en otras helicasas hexaméricas (E1 de BPV (Enemark *et al.*, 2000), DnaB (Jezewska *et al.*, 1998; Kaplan y O'Donnell 2002) o gp4 de T7 (Egelman *et al.*, 1995)).

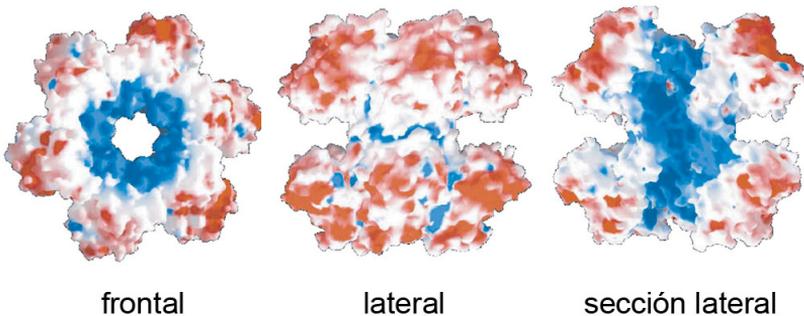


Figura 1.10: Cargas de los átomos en la estructura de N-mtMCM.

Tres tipos de vistas de la estructura atómica de N-mtMCM mostrando una superficie externa constituida mayoritariamente por átomos con carga negativa (en color rojo) mientras que el canal central se encuentra extraordinariamente cargado positivamente (azul). (Tomado de Fletcher *et al.*, 2003).

Las dos últimas laminas β del dominio C ($\beta_9 - 10$) forman un bucle β (β -hairpin) que penetra en el canal, produciendo los bucles de los 6 monómeros en conjunto, el punto de mayor estrechamiento de éste. La mutación de los residuos conservados del bucle β produce, en *M. thermoautotrophicum*, la eliminación de la capacidad de unión a DNA (Fletcher *et al.*, 2003). Sin embargo, la misma mutación en *S. solfataricus*, produce tan sólo una disminución de esta capacidad, y una reducción en la actividad helicasa (McGeoch *et al.*, 2005). La disposición de este bucle en *M. thermoautotrophicum*, próxima a la región C-terminal, sugiere una función en la unión a DNA durante la apertura de la doble hélice, quizás de remodelado del dsDNA (Fletcher *et al.*, 2003) o bien de translocación a lo largo del ssDNA, de modo similar al mecanismo sugerido para la helicasa E1 de BPV (Enemark y Joshua-Tor 2006).

Como hemos podido comprobar, la gran mayoría de las cuestiones que conciernen a la familia de proteínas MCM se encuentran actualmente en debate, desde la propia determinación de su función como helicasa replicativa eucariota, su ubicación celular, su momento de actuación (iniciación o también elongación) y la función de cada uno de sus monómeros, hasta la función de cada uno de sus motivos conservados y de sus dominios resueltos.

I.5 – Microscopía electrónica

La microscopía electrónica (ME) es una de las principales técnicas para obtener la estructura tridimensional de macromoléculas biológicas. Frente a las extensamente usadas resonancia magnética nuclear (RMN) y difracción de rayos X (o cristalografía de rayos X), la ME permite tratar especímenes grandes y flexibles, como proteínas multiméricas o grandes complejos macromoleculares. Este tipo de estructuras resultan intratables por RMN, ya que esta técnica impone restricciones de tamaño (Pervushin *et al.*, 1997; Wuthrich 1998, actualizar), y en muchas ocasiones también quedan fuera del alcance de la cristalografía, bien porque se resisten a cristalizar (salvo que sean eliminadas sus partes flexibles), bien porque su gran tamaño dificulta la determinación de las fases, necesaria para resolver la estructura (Frank 2002).

Incluso en los casos en los que se consigue obtener un cristal, la formación de éste puede interferir en el estudio de las interacciones de las macromoléculas que lo componen, debido a que el empaquetamiento favorezca determinados modos de interacción e impida o directamente prohíba otros (Frank 2002). Una excepción notable a esta regla es la cristalización del ribosoma (Nissen *et al.*, 2000; Schluenzen *et al.*, 2000; Wimberly *et al.*, 2000; Yusupov *et al.*, 2001), que por otro lado la confirma, ya que ha llevado décadas conseguirla.

La ME permite en ocasiones resolver estructuras a un nivel de resolución atómico, cuando éstas se ordenan en cristales bidimensionales o filamentos helicoidales (Henderson 2004) en los que el elevado grado de periodicidad permite filtrar el ruido y recuperar la información de la estructura. En la mayoría de los casos sin embargo, las proteínas o los complejos no interaccionan entre sí para dar lugar a estructuras periódicas o cristalinas, si no que se comportan en una preparación como múltiples partículas individuales.

Podemos también resolver estructuralmente este tipo de muestras mediante microscopía alcanzando rangos de media resolución (10 – 30 Å) obteniendo los mejores resultados con partículas que gozan de una elevada simetría, como por ejemplo los virus icosaédricos [

1.5.1 – Funcionamiento básico del microscopio electrónico

El concepto de funcionamiento de un microscopio electrónico es similar en al de uno óptico: La radiación, en este caso electrones, es emitida desde un filamento, y posteriormente conducida y focalizada sobre la muestra mediante bobinas electromagnéticas, similarmente a como lo hacen las lentes en el microscopio óptico. La muestra, depositada sobre una delgada capa de carbón o colodión que soporta una rejilla metálica, es atravesada por la radiación en una dirección perpendicular, formando una imagen de proyección que es amplificada sobre una pantalla fluorescente o bien registrada en una película fotográfica (micrografía) o en un dispositivo digital (CCD, *charge couple device*).

Los electrones son acelerados desde el filamento donde se emiten (cátodo) mediante una diferencia de potencial con un ánodo próximo. La magnitud de este voltaje define la longitud de onda del electrón y por tanto la capacidad de resolución de la observación realizada. Así, aunque en teoría podrían alcanzarse resoluciones subatómicas, en la práctica, las aberraciones de las lentes, difíciles de corregir, junto con los problemas inherentes a la preparación y observación de las muestras, que veremos a continuación, disminuyen la resolución práctica hasta aproximadamente 10 Å.

1.5.2 – Trabajo con las muestras

En el interior del microscopio la muestra se ve sometida a dos elementos que la degradan rápidamente. El primero de ellos es la propia radiación electrónica, que rompe los enlaces químicos de las moléculas destruyendo la muestra. En segundo lugar, el uso de electrones como radiación impone el trabajo en condiciones de alto vacío, ya que éstos sufrirían dispersiones al colisionar con los átomos de cualquier gas. La muestra queda expuesta a este vacío en el interior de la columna del microscopio, sufriendo una rápida deshidratación que altera la estructura de las moléculas observadas, alejándolas de su estructura nativa.

Varias técnicas han sido desarrolladas para evitar los efectos de la radiación y de la deshidratación. La técnica de dosis mínima (Williams y Fisher 1970) se basa en irradiar el objeto de observación con la cantidad mínima necesaria de radiación (menos de $10 \text{ e}^-/\text{Å}^2$), de forma que el ajuste

del enfoque, previo a la toma de cada micrografía y durante el cual la muestra recibe la mayor cantidad de radiación, se realiza en una posición adyacente a la zona a fotografiar, que se conserva intacta hasta el momento de tomar la micrografía. Esta técnica se aplica rutinariamente en el trabajo con muestras susceptibles de ser procesadas y reconstruidas.

Para proteger a la muestra tanto de la radiación electrónica como de la deshidratación, se han desarrollado la tinción negativa y la criomicroscopía (Dubochet *et al.*, 1988; Lepault *et al.*, 1983) (para una revisión de ambas técnicas, consultar Frank 2006b). La tinción negativa es la técnica habitual para realizar la caracterización básica de una muestra. De rápida implementación, consiste en embeber las partículas de la proteína a observar en una fina capa amorfa de un metal pesado (uranio, wolframio, etc.), que rodea las partículas a modo de molde. Esta capa preserva parcialmente a la estructura de los efectos de la radiación y de la deshidratación. Sin embargo, el metal pesado supone la principal fuente de contraste en la imagen formada, de modo que la información estructural que obtenemos se refiere fundamentalmente a la estructura externa del espécimen y no tanto a su estructura interna. Sin embargo, la tinción negativa puede producir distorsiones en la estructura por efectos de aplastamiento. La resolución se encuentra limitada al tamaño del grano del agente de tinción, pudiéndose alcanzar los 20 Å.

La criomicroscopía electrónica (crioME) supone mejoras importantes con respecto a la preservación de la muestra y permite alcanzar mayor resolución (aproximadamente hasta 10 Å), aunque es de más difícil desarrollo: La muestra, depositada sobre una rejilla bien análoga a las utilizadas para tinción negativa, bien con agujeros perforados en el soporte de carbón, es sumergida en una solución de etano líquido a temperatura del nitrógeno líquido (aproximadamente -170 °C) mediante una guillotina que permite su inmersión rápida y controlada. El pequeño tamaño de la muestra, combinado con la velocidad de la inmersión y la capacidad calorífica del etano, permiten una velocidad de enfriamiento lo suficientemente rápida para que las moléculas de agua no se organicen en hielo cristalino al congelarse, manteniéndose en un estado amorfo similar al líquido denominado hielo vítreo. La muestra se mantiene en todo momento, incluso durante su observación a una temperatura no superior a -160 °C. La baja temperatura de observación reduce los efectos del daño por radiación, y permite conservar a la muestra suspendida en el estado acuoso en el que se encontraba en el momento de su preparación,

conservando por tanto su estructura nativa. Por otro lado, la ausencia de agente de tinción permite alcanzar una mayor resolución, a costa de disminuir el contraste de las imágenes resultantes.

1.5.3 – Características de las imágenes de microscopía

La interacción de los electrones con los átomos de la muestra provoca diferentes tipos de dispersiones permitiendo la aparición de contraste en la imagen. Como consecuencia de estas interacciones y debido a imperfecciones instrumentales (aberraciones de las lentes, astigmatismo, etc.) el proceso de formación de la imagen se ve afectado dando lugar a una versión alterada del objeto observado. Es posible simular los efectos que tiene el microscopio en la formación de la imagen aplicando la denominada “función de transferencia de contraste” (CTF, *contrast transfer function*) a la transformada de Fourier de la imagen del objeto.

Los dos principales efectos de estas distorsiones son la inversión del contraste de la imagen en determinadas frecuencias de su transformada de Fourier, y la atenuación de la señal en las altas frecuencias, que representan los detalles del objeto representado. Cuando se está llevando a cabo una reconstrucción con objeto de resolver una estructura a alta resolución es necesario estimar la CTF para cada imagen, e intentar corregir estos efectos usando la estimación como modelo en la imagen tomada por el microscopio.

1.5.4 – Procesamiento inicial de las imágenes

Habitualmente las imágenes tomadas en condiciones de mínima dosis presentan una relación señal ruido (SNR, *signal to noise ratio*) muy baja, siendo necesario combinar la señal proveniente de multitud de partículas idénticas para conseguir incrementar esta relación. Sin embargo, las partículas de una misma muestra proteica a menudo no son idénticas sino que más bien gozan de una extensa variabilidad, que debe ser tratada convenientemente mediante algoritmos de procesamiento de imágenes para lograr este objetivo.

Las dos fuentes de variabilidad que originan mayores dificultades en el trabajo con las imágenes son la variabilidad estructural y la de orientación. La primera de ellas se refiere a cada una de las diferentes conformaciones,

organizaciones cuaternarias o estados de agregación que puede adoptar una proteína o complejo determinado en una misma preparación. Cada una de ellas genera básicamente un objeto distinto, siendo necesario diferenciar las imágenes provenientes de cada uno de los objetos diferentes. La variabilidad de orientación, por su parte, se refiere a las diferentes orientaciones que puede adoptar un objeto determinado al colocarse sobre el soporte de la rejilla.

Una muestra determinada puede presentar diferentes grados de heterogeneidad dentro de los que hemos mencionados. Así, como veremos, una misma muestra de mtMCM puede presentar partículas de diferente constitución oligomérica aunque dispuestas únicamente en dos orientaciones preferentes.

La estrategia concreta a seguir para tratar la heterogeneidad presente depende de cada tipo de muestra, pero en general parte de combinar iteraciones de alineamiento traslacional y rotacional con iteraciones de algoritmos de reconocimiento de patrones y clasificación (PCA (*principal components analysis*, análisis de componentes principales, van Heel 1984), KerDenSOM (Pascual-Montano *et al.*, 2001), Maximum Likelihood (Scheres *et al.*, 2005a)). Mediante las sucesivas iteraciones, es posible separar un conjunto inicial de imágenes en clases homogéneas, que representen cada una de las orientaciones de cada uno de los objetos presentes.

En el caso concreto de las helicasas que oligomerizan en forma de anillo, el análisis de la simetría de las imágenes mediante descomposición en sus armónicos de Fourier, componiendo el denominado “espectro rotacional” (Crowther y Amos 1971), puede ser una metodología útil para analizar la estequiometría del anillo o las diferentes conformaciones que éste presente (Barcena *et al.*, 1998; Yu *et al.*, 1996).

1.5.5 – Reconstrucción tridimensional

Una vez determinada la estructura origen de cada una de las clases encontradas, es posible combinar las diferentes imágenes de las diferentes orientaciones para reconstruir su estructura tridimensional.

Si el número de orientaciones que adopta la proteína sobre el plano de la rejilla es elevado y cubre todo el espacio angular, es posible utilizar esa información directamente para llevar a cabo la reconstrucción. Se han desarrollado varios métodos que trabajan en diversos espacios (real (Penczek *et al.*, 1994), de Radon (Radermacher 1994), de Fourier (Jonic *et*

al., 2005), etc.) basados en un volumen de referencia inicial, que sirve de base para buscar las orientaciones (*projection matching*) de las imágenes experimentales. De este modo, el modelo inicial se proyecta y se comparan las diferentes proyecciones con las imágenes experimentales, asignando a cada una el ángulo correspondiente a la dirección de proyección de la imagen del modelo con la que mayor correlación encuentra. Estas técnicas tienen el inconveniente de que a menudo el resultado final es altamente dependiente de la referencia inicial elegida (Frank 2006a; Shaikh *et al.*, 2003) por lo que ésta debe ser elegida cuidadosamente.

El método de “líneas comunes” (Crowther *et al.*, 1970; Penczek *et al.*, 1996) no requiere referencia inicial. Se basa en el teorema de la sección central, que determina la existencia de una línea común entre dos proyecciones de un mismo objeto en el espacio de Fourier. Este método requiere imágenes con muy poco ruido, de modo que se aplica con éxito una vez se han obtenido imágenes medias de clases representativas de las diferentes orientaciones.

En muchos casos la proteína adopta tan sólo unas pocas orientaciones, en cuyo caso es necesario obtener la cantidad suficiente de imágenes en otras orientaciones para cubrir el espacio angular, mediante la inclinación de la muestra en el interior del microscopio (Radermacher *et al.*, 1987). Este método tampoco requiere una referencia inicial, puesto que las orientaciones se asignan a partir de los ángulos de las partículas de la imagen inclinada y de las rotaciones de las mismas partículas en una imagen tomada sin inclinación. A menudo las reconstrucciones realizadas mediante este último método adolecen de falta de resolución en determinadas direcciones (anisotropía), debido a la limitación técnica del ángulo de giro de la muestra en el microscopio, que no puede rebasar los $\pm 60^\circ$.

Finalmente, una vez obtenidas las orientaciones, el volumen se reconstruye por diferentes métodos (Frank 2006c), bien mediante retroproyección en el espacio de Fourier, bien mediante algoritmos algebraicos (Marabini *et al.*, 1998).

Objetivos

El objetivo general de esta tesis es profundizar en el conocimiento estructural de la helicasa MCM de *Methanobacterium thermoautotrophicum*, relacionando la información obtenida con su función durante la replicación del DNA.

- Investigar la relación entre los diferentes estados oligoméricos de mtMCM y las condiciones del medio, y definir si existen condiciones que proporcionan una población homogénea para su estudio mediante microscopía electrónica tridimensional (ME3D) de alta resolución.
- Obtener un mapa tridimensional de mtMCM en su probable forma activa, el doble hexámero.
- Analizar el estado oligomérico de mutantes diseñados del fragmento amino terminal, en la medida que aporten evidencias para determinar la forma activa de la proteína completa.
- Definir las condiciones óptimas de formación y marcaje de complejos MCM:DNA para su estudio mediante ME3D.

Materiales y métodos

M.1 – Proteínas

La proteína silvestre (WT) MCM de la arquea *Methanobacterium thermoautotrophicum* (mtMCM) (75,5 kDa) fue expresada en *Escherichia coli* (*E. coli*) como proteína de fusión GST-mtMCM y posteriormente purificada mediante cromatografía de afinidad por glutatión, proteólisis con trombina y cromatografía de filtración en gel en una columna Superose 6. El proceso total es el descrito en Fletcher *et al.*, 2005, y fue realizado por Ryan J. Fletcher bajo la dirección del Dr. Xiaojiang S. Chen del Departamento de Biología Molecular y Computacional de la Universidad de California, en EE.UU.

Igualmente el Dr. X. S. Chen proporcionó el fragmento amino terminal de mtMCM (N-mtMCM, 32,3 kDa), y sus mutantes N-mtMCM-R161A y N-mtMCM-6Ins. Las tres proteínas fueron expresadas y purificadas del mismo modo que la proteína mtMCM WT según se describe en Fletcher *et al.*, 2003 y en Fletcher *et al.*, 2005, respectivamente.

Diferentes muestras de proteínas WT fueron utilizadas, purificadas cada una independientemente. En todas ellas la proteína se encontraba en un tampón Tris·Cl pH 8,0, DTT 10 mM, conteniendo bien NaCl 250 mM o bien 500 mM, según el stock. La proteína N-mtMCM y sus mutantes se encontraban en tampón Tris·Cl pH 8,0, NaCl 250 mM y DTT 10mM.

El mutante mtMCM Δ h2i, conjuntamente a una muestra WT control, fueron expresadas en *E. coli* como proteínas de fusión con cola de

histidinas, tal y como se describe en Chong *et al.*, 2000. Ambas fueron purificadas mediante cromatografía de afinidad por metal en una columna TALON, seguida de cromatografía de intercambio iónico en una columna Resource-Q, eliminación de sales y concentración mediante membrana en microcentrífuga, tal y como se describe en Jenkinson y Chong 2006. Las muestras fueron provistas en un tampón final constituido únicamente por Tris-Cl pH 7,5 a baja concentración (5 ó 10 mM según el stock), ambas proporcionadas por el Dr. James P. J. Chong del Departamento de Biología de la Universidad de York del Reino Unido.

Todas las proteínas fueron caracterizadas sistemáticamente tras ser purificadas para asegurar que reunían las condiciones necesarias de trabajo en microscopía electrónica. Esta validación consideró tres cuestiones: 1 – Verificación de la pureza e integridad de la proteína mediante electroforesis desnaturalizante, descartando proteólisis y posibles contaminaciones con otras proteínas. 2 – Análisis del grado de contaminación con DNA celular. 3 – Observación preliminar mediante microscopía electrónica de la muestra teñida negativamente. En la mayoría de los casos los resultados de estas comprobaciones no se muestran en este trabajo, reuniendo todas las proteínas las condiciones necesarias para los experimentos realizados.

M.2 – Secuencias de DNA: Fragmentos de restricción y oligonucleótidos

El fragmento de dsDNA de 300 pb utilizado en los experimentos de retardo en gel fue obtenido a partir de la digestión del plásmido comercial pGEM-T Easy de Promega (www.promega.com), que fue clonado en *E.coli* y purificado. Su digestión con XmnI y EcoRI produce un fragmento de dsDNA de 1 kb y otro de 300 pb que fueron separados mediante electroforesis y purificados del gel utilizando el kit de purificación “DNA gel extraction kit” de Qiagen (www.qiagen.com). La cantidad final de DNA obtenido se calculó cuantificando la absorción de luz ultravioleta a 260 nm de una pequeña cantidad de muestra en un espectrofotómetro Nanodrop (www.nanodrop.com).

Las moléculas de DNA de pequeño tamaño (<100 pb) utilizadas en los experimentos de formación de complejos MCM:DNA, fueron generadas

hibridando dos moléculas complementarias de DNA monocatenario. Los oligonucleótidos complementarios fueron sintetizados y purificados por la empresa Invitrogen (www.invitrogen.com), mediante cromatografía HPLC o bien mediante geles de acrilamida.

Los oligonucleótidos, de secuencia inespecífica, fueron diseñados para minimizar la formación de estructuras secundarias no deseadas. En la Tabla M.1, se indican las secuencias utilizadas. Las construcciones de horquilla (fork) y burbuja (bble) se consiguieron introduciendo en la secuencia regiones no complementarias de poly-T.

Los oligonucleótidos liofilizados se resuspendieron en tampón TE (10 mM Tris-Cl pH 8,0, 1 mM EDTA) conteniendo 50 mM NaCl. Para generar las moléculas bicatenarias se mezclaron cantidades equimolares de los oligos a hibridar, a continuación se calentaron a 90 – 95 °C durante 2 minutos y se dejaron enfriar lentamente hasta temperatura ambiente. La correcta hibridación se verificó mediante electroforesis.

De esta forma, por ejemplo, hibridando los oligos dsNMCM-1 con dsNMCM-2 se obtiene una molécula de dsDNA de 40 pb; hibridando los oligos bbleFLMCM-70-1bio y bbleFLMCM-70-2 se obtiene una molécula de DNA de 70 pb marcada con biotina en un extremo, y que contiene una burbuja de banda sencilla de 30 pb flanqueada por dos regiones de dsDNA de 20 pb cada una.

M.3 – Muestreo de condiciones físico-químicas

Las proteínas, diluidas en su tampón final de purificación, se dializaron frente a un volumen 1000 veces superior de solución Tris-Cl 50 mM pH 8,0, DTT 1 mM, con la concentración adecuada de NaCl (0, 50, 100, 250, 500 ó 1000 mM) y de Mg^{2+} , ATP ó AMPNP (Tabla R.1). La diálisis se realizó a la temperatura indicada, por defecto 25 °C, durante 1 h (30 min en el caso de la muestra incubada a 60 °C), por medio de una membrana de nitrocelulosa de Millipore (www.millipore.com) con un tamaño de poro de 0,025 μ m.

Oligonucleótido Secuencia (5' → 3')	Longitud (bases)
dsNMCM-1 TCTACCTGGACTCGACCGGG GGGCCAGCTCAGGTCCATCA	(40)
dsNMCM-2 TGATGGACCTGAGCTGGCCC CCCGGTCGAGTCCAGGTAGA	(40)
bbleNMCM-1 CTCGACCGGGTTTTTTTTTT TTTTTTTTTTGGGCCAGCTC	(40)
bbleNMCM-2 GAGCTGGCCCTTTTTTTTTT TTTTTTTTTTCCCGGTCGAG	(40)
dsFLMCM-1 TCTGCTCGGTAGGACCTGGA CTCGACCGGGGGGCCAGCTC AGGTCCAGGATGGCTCGTCA	(60)
dsFLMCM-2 TGACGAGCCATCCTGGACCT GAGCTGGCCCCCGGTCGAG TCCAGGTCTACCGAGCAGA	(60)
bbleFLMCM-1 TCTACCTGGACTCGACCGGG TTTTTTTTTTTTTTTTTTTT GGGCCAGCTCAGGTCCATCA	(60)
bbleFLMCM-2 TGATGGACCTGAGCTGGCCC TTTTTTTTTTTTTTTTTTTT CCCGGTCGAGTCCAGGTAGA	(60)
dsFLMCM-70-1bio ●-TCTACTAGGCTCGGTAGGAC CTGACTCGACCGGGGGGCC AGCTCAGGTCCAGGATGGCT CGGATCATCA	(70)
dsFLMCM-70-2 TGATGATCCGAGCCATCCTG GACCTGAGCTGGCCCCCGG TCGAGTCCAGGTCTACCGA GCCTAGTAGA	(70)
forkFLMCM-70-1bio ●-TCTACTACGCTCAGCGTACC GGGACCTGGACTCGACCGGG TTTTTTTTTTTTTTTTTTTT TTTTTTTTTT	(70)
forkFLMCM-70-2 TTTTTTTTTTTTTTTTTTTT TTTTTTTTTTCCCGGTCGAG TCCAGGTCCCGGTACGCTGA GCCTAGTAGA	(70)
bbleFLMCM-70-1bio ●-TCTACCTGGACTCGACCGGG TTTTTTTTTTTTTTTTTTTT TTTTTTTTTTGGGCCAGCTC AGGTCCATCA	(70)
bbleFLMCM-70-2 TGATGGACCTGAGCTGGCCC TTTTTTTTTTTTTTTTTTTT TTTTTTTTTTCCCGGTCGAG TCCAGGTAGA	(70)

Tabla M.1: Oligonucleótidos utilizados en los experimentos de formación de complejos DNA:MCM.

Lista completa de oligonucleótidos utilizados. El símbolo “●-” indica una molécula de biotina unida covalentemente en el extremo 5'. Las moléculas dsDNA se obtienen hibridando oligonucleótidos complementarios, por ejemplo, la hibridación de dsFLMCM-70-1bio y dsFLMCM-70-2 produce una molécula de dsDNA de 70 pb marcada con biotina en un extremo.

M.4 – Fosforilación radiactiva de moléculas de DNA celular contaminante

El método seguido para detectar posibles contaminantes de DNA celular en las muestras de proteínas utilizadas se basa en marcar con fósforo radiactivo el extremo 5'-fosfato de todas las moléculas de DNA existentes en la muestra. Posteriormente el DNA marcado radiactivamente se detecta mediante autorradiografía. El protocolo empleado es el que aparece en Sambrook y Russell 2001, con algunas modificaciones:

Fueron utilizadas tan sólo 5 pmol de γ - ^{32}P -ATP en una mezcla de reacción que contenía 1 μl de muestra problema de proteína. Las reacciones fueron detenidas incubándolas 20 min a 70 °C y el ATP marcado no incorporado fue eliminado utilizando cromatografía de filtración en gel. El eluido fue cargado en electroforesis de acrilamida y revelado mediante autorradiografía.

M.5 – Detección de complejos proteína:DNA en condiciones nativas

M.5.1 – Retardo en geles de acrilamida

Las reacciones de formación de complejos proteína:DNA se llevaron a cabo incubando durante 30 min a la temperatura apropiada (por defecto 25°C) y en agitación suave, una mezcla conteniendo las cantidades adecuadas de proteína y DNA, diluidas en el tampón inicial de la proteína y conteniendo o no, según el experimento, diferentes concentraciones de NaCl, MgCl_2 5 mM y/o AMPPNP 3 mM (consultar Resultados R.3.2 – R.3.2.3).

En los experimentos en que todas las muestras fueron incubadas en presencia de MgCl_2 , los tampones de electroforesis fueron preparados con MgCl_2 a la misma concentración de las muestras. En aquellos en los que se utilizó AMPPNP 3 mM, se tuvo la precaución de añadir una pequeña cantidad de AMPPNP hasta una concentración 20 μM .

El tampón de carga utilizado estaba formado exclusivamente por una disolución acuosa de glicerol y azul de bromofenol. Los geles nativos se prepararon sin SDS en Tris-Cl pH 8,8, al 5% de acrilamida, salvo que se

indique lo contrario. Algunos de los geles fueron preparados con concentrador (*stacking*) buscando optimizar la resolución de las bandas. En estos casos, el gel concentrador era preparado en un tampón Tris-Cl pH 6,8 al porcentaje estándar de acrilamida (~4%), también sin SDS.

La presencia de DNA fue detectada mediante 30 min de tinción con SYBR Safe (Invitrogen) o bromuro de etidio (BrEt) y revelada mediante luz ultravioleta. Posteriormente, para visualizar las bandas de proteína el gel fue teñido con azul Coomassie. La identidad de las bandas de proteína y DNA fue comprobada superponiendo computacionalmente ambos revelados.

M.5.2 – Retardo en geles de agarosa

Las muestras se prepararon tal y como se indica en el apartado anterior. Se tuvo la precaución de diluirlas hasta un volumen final no inferior a 10 μ l para formar una banda continua en el interior de cada pocillo, claramente identificable.

Se emplearon geles de agarosa al 1% preparados en tampón TAE (Tris-acetato 40 mM pH ~8,5, EDTA 2 mM), suplementados en cada caso, de ser requerido, con $MgCl_2$ 5 mM y AMPPNP 20 μ M. Se verificó la presencia de DNA tiñendo el gel con SYBR Safe y revelándolo con luz ultravioleta.

M.5.3 – Transferencia de DNA a membranas: *southern-blot*

La detección del retardo en los experimentos de formación de complejos de mtMCM con DNAs de pequeño tamaño marcados con biotina, se realizó además mediante *southern-blot*, técnica de mayor sensibilidad, que asegura la detección de pequeñas cantidades de DNA retardado. Mediante este método, se realiza la transferencia del DNA del gel a una membrana de nitrocelulosa y se fija térmicamente, siguiendo un protocolo ligeramente modificado para DNAs de pequeño tamaño:

Los geles se lavan para desnaturalizar el DNA en tampón de desnaturalización (NaCl 1,5 M, NaOH 0,5 M) durante 30 min; posteriormente su pH se neutraliza lavando con Tris-Cl 0,5 M pH 7,0, NaCl 1,5 M durante 45 min. Una membrana de nitrocelulosa se coloca sobre el gel en un montaje estándar para una transferencia semi-húmeda que se

lleva a cabo durante al menos 12 horas con SSC 20x (NaCl 3 M, citrato sódico 0,3 M, pH 7,0). Una vez finalizada la transferencia, la membrana se retira, se equilibra en SSC 2x y el DNA se fija a ella mediante tratamiento térmico en estufa durante 3 h a 80 °C.

La membrana se bloquea incubándola durante 1 h en una solución de leche en polvo desnatada (leche en polvo desnatada al 0,6% en TTBS 1x (Tris·Cl 0,1 M pH 7,5, NaCl 150 mM, Tween 20 0,1% (v/v))). La detección del DNA se realiza incubando la membrana 1 h con estreptavidina-peroxidasa disuelta en solución de bloqueo (1 a 10 µl de solución comercial de estreptavidina-peroxidasa (General Electric Healthcare, GE, www.gelifesciences.com) en 10 ml de solución de bloqueo), seguida de tres lavados de 30 min cada uno (en TTBS 1x) para eliminar la estreptavidina no unida.

De esta forma, el conjugado estreptavidina-peroxidasa se une específicamente a la biotina del extremo del DNA. La peroxidasa se utiliza para catalizar una reacción quimioluminiscente acoplada que delimita aquellas zonas de la membrana donde la enzima ha quedado retenida mediante la unión de la estreptavidina a la biotina del DNA. En nuestros experimentos se utilizó ECL (GE) como reactivo quimioluminiscente y el revelado se llevó a cabo mediante autorradiografía.

M.6 – Marcaje con estreptavidina o estreptavidina-oro

La detección de moléculas de DNA en las micrografías se realizó marcando con conjugados de estreptavidina-oro (strep-Au) las moléculas de biotina del extremo de nuestras moléculas de DNA. En nuestros experimentos se utilizaron conjugados de oro coloidal (British Biocell Internacional, www.bbigold.com) de pequeño tamaño, 5 nm, adecuados para marcar partículas pequeñas como proteínas individuales, pero de un tamaño suficiente para ser distinguidos del grano del agente de tinción.

Para evitar los falsos positivos, se eliminaron las marcas libres utilizando una resina de níquel, que une específicamente la cola de histidinas de las proteínas mtMCM que fueron utilizadas en estos experimentos. De este modo se eliminaron las partículas de strep-Au y

strep-Au:DNA, uniéndose a la resina sólo las partículas de mtMCM, tanto aquellas que se encontraban formando complejos como las que no.

La formación de los complejos mtMCM:DNA se produjo en primer lugar, de la manera descrita en cada experimento (ver Resultados R.3.3.3). Posteriormente la mezcla fue añadida a la resina junto con las moléculas de strep-Au, incubándola durante 10 min para que se produzca la reacción de unión estreptavidina:biotina. A continuación la resina fue lavada y la proteína eluida por medio de un aumento de la concentración de imidazol hasta 300 mM. El eluido se utilizó entonces directamente para preparar rejillas de microscopía.

En estas condiciones, los marcajes con estreptavidina (sin oro), no producen falsos positivos como strep-Au. Las moléculas de estreptavidina libres sobre la rejilla tienen un tamaño pequeño que las hace fácilmente distinguibles de los oligómeros de mtMCM pero difíciles de apreciar en la selección individual de las partículas. El conjunto de partículas marcadas se revela posteriormente durante el procesamiento y clasificación de una gran población de imágenes, momento en el que las partículas marcadas se distinguen de las no marcadas por la estreptavidina unida a la biotina, cuya posición es relativamente poco variable con respecto al ruido de fondo.

En nuestros experimentos fue utilizada estreptavidina tetramérica (~52 kDa) (Roche, www.roche-applied-science.com) disuelta en agua y preparada fresca, añadida en exceso a la mezcla de complejos de mtMCM con DNA, una vez formados éstos. La mezcla resultante se utilizó directamente para su observación al microscopio.

M.7 – Preparación de las muestras para microscopía electrónica

M.7.1 – Tinción negativa

Las muestras fueron diluidas, cuando fue necesario, en el tampón original de la proteína o en el tampón de reacción (en el caso de complejos) hasta la concentración adecuada. Usualmente para cada tipo de proteína y para cada muestra se realizó un chequeo de concentraciones hasta encontrar la idónea para microscopía. La proteína fue adsorbida sobre rejillas de cobre recubiertas de carbón-colodión, que previamente han sido sometidas a una descarga iónica para aumentar su carácter hidrófilo. Para cada experimento, 3 μ l de la muestra son incubados sobre una rejilla durante 3 minutos, retirando posteriormente el tampón excedente junto con las partículas que no han quedado adheridas a la rejilla. Se realizan varios lavados con agua destilada y se tiñe con acetato de uranilo al 2% durante 45 seg.

M.7.2 – Criomicroscopía electrónica

La muestra (habitualmente un volumen de 3 μ l) fue incubada sobre una rejilla de cobre/rodio Quantifoil (www.quantifoil.com) previamente sometida a una descarga iónica. Estas rejillas se encuentran recubiertas por una película de carbón sobre la que se han practicado regularmente agujeros de 2 μ m de diámetro. Tras una incubación de 3 min se elimina el exceso de muestra, quedando en cada agujero un delgado menisco de solución expuesto al aire por ambas caras, es decir, una delgada película de tampón en el que se encuentran diluidas las moléculas de proteína.

Inmediatamente la muestra es vitrificada, sumergiendo la rejilla en etano líquido a aproximadamente -180 °C, utilizando guillotinas especialmente diseñadas para este procedimiento. La muestra vitrificada se conserva en nitrógeno líquido a -196 °C.

M.8 – Microscopía electrónica

La observación de las rejillas de tinción negativa fue realizada en un microscopio de transmisión Jeol 1200 EX-II (www.jeol.com) con un voltaje de aceleración de 80 kV. Las imágenes fueron tomadas a una magnificación de 60.000X.

Las muestras vitrificadas se examinaron utilizando un microscopio Tecnai F20 (www.fei.com) equipado con un cañón de emisión por campo (*field emission gun*) y un anticontaminador Gatan, y utilizando un criobrazo y una unidad de criotransferencia fabricados también por Gatan (www.gatan.com). Las muestras fueron examinadas a 200 kV y fotografiadas a 50.000X a un subenfoco variable en un rango entre -3 y -7 μm . En todo momento la muestra se mantuvo a una temperatura no superior a -170 °C.

En ambos casos las micrografías fueron tomadas sobre placas Kodak SO-163 (www.kodak.com) en condiciones de mínima dosis ($\sim 10 \text{ e}^-/\text{Å}^2$).

M.9 – Procesamiento digital de imágenes

Toas las micrografías fueron digitalizadas en un microdensitómetro Zeiss/Intergraph SCAI (www.zeiss.com, www.intergraph.com). Aquellas que fueron tomadas de muestras teñidas negativamente se digitalizaron utilizando una frecuencia de muestreo de 21 μm o menor y posteriormente fueron submuestreadas hasta un tamaño final de píxel de 3,5 Å en la muestra. Las micrografías de muestras vitrificadas se digitalizaron a 7 μm y fueron submuestreadas hasta 4,2 Å/pix.

La calidad de cada micrografía fue evaluada en cuanto a astigmatismo y deriva, mediante el análisis de la transformada de Fourier de la imagen. Para las micrografías de criomicroscopía se calculó su grado de desenfoque real utilizando el programa ctfiffind3 (Mindell y Grigorieff 2003) y se llevó a cabo la corrección de la fase de la función de transferencia de contraste utilizando el programa Bsoft (Heymann 2001).

Pre-procesamiento, alineamiento bidimensional y clasificación de imágenes fueron llevados a cabo utilizando el paquete de software Xmipp (xmipp.cnb.csic.es) (Marabini *et al.*, 1996; Sorzano *et al.*, 2004). Las partículas fueron seleccionadas manualmente utilizando tanto Xmipp como

EMAN (blake.bcm.tmc.edu/eman) (Ludtke *et al.*, 1999). En el caso de mtMCM, las partículas fueron agrupadas durante su selección en dos familias diferentes atendiendo a sus diferencias morfológicas claramente visibles: rectangulares o anulares. Para cada familia las partículas fueron recortadas a un tamaño adecuado (128x128 y 64x64 pix respectivamente), realizando el procesamiento de cada familia independientemente. Las imágenes individuales fueron normalizadas, centradas y alineadas bidimensionalmente usando un método libre de referencia (Marco *et al.*, 1996; Penczek *et al.*, 1992).

La clasificación de imágenes atendiendo a su heterogeneidad se llevó a cabo en diferentes momentos a lo largo del alineamiento, mediante mapas autoorganizativos KerDenSOM (Kernel Probability Density Estimator Self-Organizing Maps, Pascual-Montano *et al.*, 2001) o bien mediante el método de Maximum Likelihood (Scheres *et al.*, 2005a; Scheres *et al.*, 2005b; Scheres *et al.*, 2007a; Scheres *et al.*, 2007b). Los grupos resultantes fueron realineados y vueltos a clasificar iterativamente hasta que se estimaron suficientemente homogéneos.

La resolución de las imágenes medias fue calculada utilizando el método de la relación señal/ruido espectral (*spectral signal-to-noise ratio*), con un umbral de 4 (Unser *et al.*, 1987).

M.10 – Reconstrucción tridimensional

Para obtener la reconstrucción tridimensional del doble anillo de mtMCM se partió de varios modelos iniciales, generados a partir de las coordenadas atómicas de la estructura de N-mtMCM, obtenidas de la base de datos de estructuras de proteínas Protein Data Bank (PDB, www.pdb.org) (código PDB de N-mtMCM: 1LTL). Las coordenadas atómicas de un único monómero fueron utilizadas para generar los dobles pentámero, heptámero y octámero, mediante las translaciones y rotaciones adecuadas, que fueron calculadas y posteriormente aplicadas utilizando Amira (www.amiravis.com). Los dobles oligómeros se crearon concatenando los monómeros individuales correctamente colocados mediante MOLEMAN2 (Kleywegt *et al.*, 2001). Los mapas tridimensionales de densidad correspondientes a los modelos sintéticos fueron creados mediante SPIDER (http://www.wadsworth.org/spider_doc/spider/docs/spider.html) (Frank *et al.*, 1996), asumiendo una frecuencia de muestreo de 3,5 Å/pix. Posteriormente se filtraron a 22 Å y fueron promediados rotacionalmente utilizando Xmipp, para minimizar la influencia de la simetría del modelo inicial en los algoritmos de reconstrucción.

Las reconstrucciones se llevaron a cabo en dos fases: En primer lugar la búsqueda de orientaciones se realizó utilizando un algoritmo de refinamiento integrado en el paquete SPIDER. El algoritmo comienza comparando las proyecciones de un volumen inicial de referencia con las imágenes experimentales, asignando a cada imagen experimental los ángulos de la proyección con la que mejor correlación encuentra. Un nuevo modelo inicial se genera entonces retroproyectando las imágenes experimentales de acuerdo a los ángulos asignados. El algoritmo itera de este modo, refinando los ángulos y las translaciones encontradas para cada imagen hasta que la resolución del volumen reconstruido no aumenta significativamente, momento en el cual se obtienen las orientaciones definitivas. La segunda fase de la reconstrucción consistió en utilizar las orientaciones calculadas para reconstruir el volumen final utilizando ART (Marabini *et al.*, 1998). Por último, el análisis plano a plano de la simetría de los volúmenes se llevó a cabo utilizando Xmipp.

La resolución de los volúmenes resultantes fue estimada en base al criterio que asigna un valor de 1 a la relación señal/ruido espectral tridimensional (*three-dimensional spectral signal-to-noise ratio*) (Unser *et al.*, 2005).

Todos los mapas generados se visualizaron utilizando Amira, aplicando un umbral tal, que el volumen corresponde a la masa indicada (usualmente 100% de la masa de la proteína), considerando una densidad de $1,33 \text{ g/cm}^3$ (Squire y Himmel 1979).

M.11 – Modelado de estructuras atómicas

La modelización de la estructura de mtMCM se realizó ajustando las coordenadas atómicas del fragmento N-mtMCM y de la estructura atómica del dominio helicasa de LTag de SV40 (código PDB: 1N25), sobre el volumen final reconstruido del doble hexámero. Se utilizó la herramienta CoLoRes (Contour low resolution) para modelado (*fitting*) rígido automático dentro del paquete Situs (<http://situs.biomachina.org/fguide.html>) (Chacon y Wriggers 2002; Wriggers *et al.*, 1999). El modelado manual y el ajuste de la posición relativa de los hexámeros de la estructura de N-mtMCM se realizó utilizando Amira.

La evaluación del ajuste óptimo de las moléculas sobre el volumen del doble hexámero se realizó mediante SPIDER, calculando el coeficiente de correlación entre las posibles estructuras en las diferentes posiciones y el volumen del doble hexámero enmascarado con una máscara cilíndrica.

Resultados

R.1 – Estructura cuaternaria de mtMCM en estado libre¹

R.1.1 – Polimorfismo estructural

mtMCM ha mostrado un polimorfismo considerable en anteriores investigaciones. En el momento de la realización de los experimentos de esta tesis, se había descrito la existencia de hexámeros sencillos (Pape *et al.*, 2003), heptámeros sencillos (Yu *et al.*, 2002) y estructuras helicoidales (Chen *et al.*, 2005), mientras que el fragmento amino terminal (N-mtMCM) había sido cristalizado como doble hexámero (Fletcher *et al.*, 2003).

La muestra de mtMCM utilizada en este trabajo ha sido expresada en *E.coli* y purificada, y mediante cromatografía de filtración en gel presenta un pico de absorbancia mayoritario que concuerda con el tamaño de ~900 kDa, probablemente un doble hexámero (Figura R.1)., tal y como se ha observado en otras ocasiones (Chong *et al.*, 2000; Kelman *et al.*, 1999; Shechter *et al.*, 2000).

¹ Gómez-Llorente, Y., Fletcher, R.J., Chen, X.S., Carazo, J.M. and San Martín, C. (2005). *J Biol Chem* **280**(49): 40909-40915.

Observando la proteína mediante ME y tinción negativa (TN) encontramos principalmente partículas con dos morfologías diferentes: bien redondeadas y con forma de anillo, o bien alargadas, rectangulares y que presentan cuatro zonas principales de densidad, apreciables como cuatro barras paralelas (Figura R.2). Algunas de las partículas con forma de anillo presentan simetría 6 mientras que otras muestran simetría 7.

Identificamos las partículas con forma de anillo con proyecciones en vista frontal de oligómeros de mtMCM, pero no es posible determinar con certeza si corresponden a anillos dobles o sencillos. Del mismo modo las partículas rectangulares se corresponden probablemente con vistas laterales de dobles anillos, pero dada la existencia de vistas frontales con simetrías 6 ó 7, no es posible determinar si las vistas laterales provienen de dobles hexámeros o de dobles heptámeros, debido a la gran similitud que existiría entre dos proyecciones laterales de estas conformaciones.

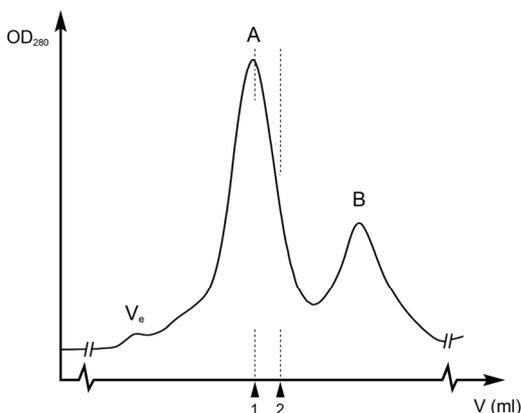


Figura R.1: Perfil de elución de mtMCM en cromatografía de filtración.

Perfil de elución de mtMCM en una columna de filtración Superose 6. 1 y 2 indican los volúmenes de elución para proteínas globulares de tamaños aproximados de 900 y 450 kDa, respectivamente. V_e es el volumen de exclusión. El pico **A** corresponde a un oligómero de mtMCM de elevado peso molecular (~900 kDa). La masa molecular de mtMCM es 75,5 kDa, de modo que el pico **A** posiblemente corresponde a un dodecámero. El pico **B** está formado por proteína degradada de bajo peso molecular.

Dada la heterogeneidad presente en las primeras observaciones de mtMCM, se hizo un barrido de distintas condiciones fisico-químicas para determinar si éstas afectaban al estado oligomérico de la proteína, y si era

posible obtener una población más homogénea para un posterior análisis tridimensional. De este modo se barajaron diferentes concentraciones de NaCl, temperaturas y tiempo de incubación, y la presencia de NTP y Mg^{2+} .

En cada condición, se tomaron imágenes de ME y fueron seleccionadas manualmente las partículas separándolas en dos familias independientes: vistas frontales y laterales. Cada familia es procesada independientemente y clasificada en grupos homogéneos usando mapas autoorganizativos (ver Materiales y métodos, M.9). Los resultados muestran la presencia de oligómeros con simetría 6 y 7 en cualquiera de las condiciones ensayadas, así como una proporción variable de vistas laterales (Tabla R.1).

En ocasiones encontramos formas oligoméricas de simetría 8 e incluso vistas frontales de anillos abiertos (Tabla R.1, última fila y tres últimas filas respectivamente).

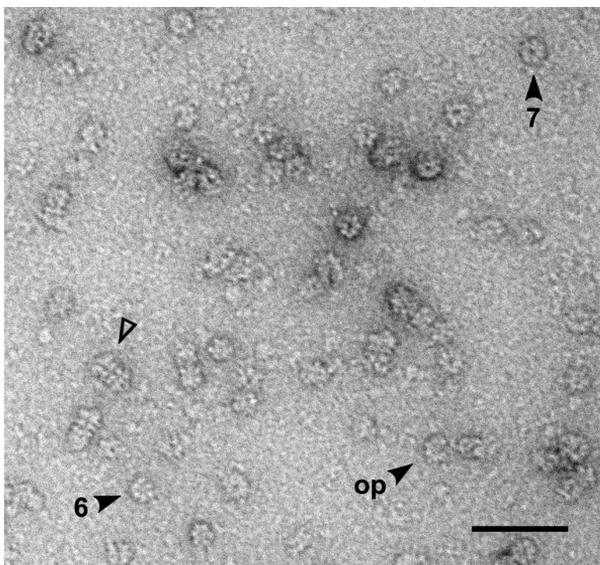


Figura R.2: Heterogeneidad de vistas en una micrografía de mtMCM.

Preparación de ME teñida negativamente. Es posible observar vistas frontales de partículas con forma de anillo de simetrías 6 ó 7 (marcadas con “▶”), anillos abiertos (op ▶), y vistas rectangulares que probablemente corresponden a vistas laterales de dobles oligómeros (▷). La barra representa 50 nm.

Condiciones experimentales: [NaCl], temp. (tiempo inc.), [Mg ²⁺], [NTP]					
<i>Part.</i> (% <i>d</i>)	Vistas frontales		Vistas laterales	Otras vistas	
	Sim. 6	Sim. 7		Sim. 2	Anillos ab.
50 mM, 25 °C (1h), 0 mM, 0 mM					
5064 (13%)					
	298 6%	172 3%	2854 54%	1085 21%	
100 mM, 25 °C (1h), 0 mM, 0 mM					
4846 (26%)					
	837 16%	336 6%	2715 52%		
250 mM, 25 °C (1h), 0 mM, 0 mM					
7869 (18%)					
	3430 30%	1538 19%	1565 20%		
500 mM, 25 °C (1h), 0 mM, 0 mM					
3202 (9%)					
	1306 41%	474 15%	1130 35%		

Tabla R.1: Polimorfismo estructural en las muestras de mtMCM.

Muestreo de condiciones físico-químicas para mtMCM. Imágenes medias de las diferentes poblaciones encontradas en preparaciones de mtMCM según cada condición experimental. Debajo de cada imagen se indica el número de partículas incluidas en la media y el porcentaje que representa con respecto al total de partículas seleccionadas. A la izquierda de cada fila y en cursiva se indica el número total de partículas seleccionadas (***Part***) y el porcentaje de partículas descartadas o no clasificables (%***d***), entre paréntesis. Las resoluciones de las imágenes varía entre 19 y 26 Å (salvo la vista lateral con AMPNP, que tiene 40 Å), aunque todas se muestran filtradas a 26 Å. La barra representa 100 Å.

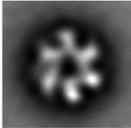
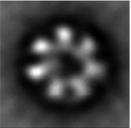
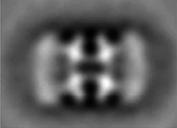
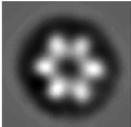
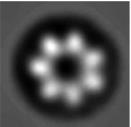
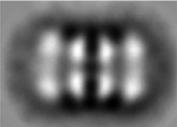
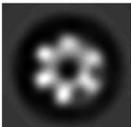
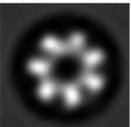
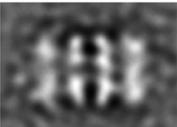
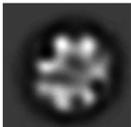
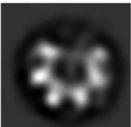
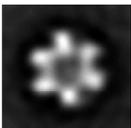
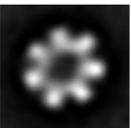
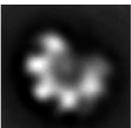
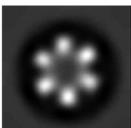
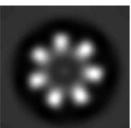
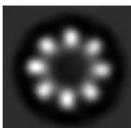
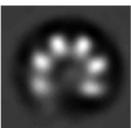
Condiciones experimentales: [NaCl], temp. (tiempo inc.), [Mg ²⁺], [NTP]					
Part. (%d)	Vistas frontales		Vistas laterales	Otras vistas	
	Sim. 6	Sim. 7		Sim. 2 ó 8	Anillos ab.
1 mM, 25 °C (1h), 0 mM, 0 mM					
1406 (12%)					
	439 31%	184 13%	619 44%		
250 mM, 37 °C (1h), 0 mM, 0 mM					
2830 (33%)					
	528 19%	915 32%	453 16%		
250 mM, 37 °C (1h), AMPPNP 2 mM, 5 mM					
2777 (9%)					
	1001 36%	517 19%	20 1%	454 16%	535 19%
250 mM, 60 °C (30 min), 0 mM, 0 mM					
2726 (20%)					
	237 9%	379 14%			1553 57%
500 mM, 60 °C (30 min), 0 mM, 0 mM					
4760 (9%)					
	647 14%	1867 39%		375 8%	1592 33%

Tabla R.1 – (Continuación).

R.1.2 – Flexibilidad del extremo C-terminal

Es posible diferenciar cuatro zonas principales de densidad en las vistas laterales de los dobles oligómeros, a modo de barras paralelas perpendiculares al eje longitudinal. La observación detallada de la imagen media de todas las poblaciones permite apreciar diferencias de definición entre estas zonas: Mientras que las regiones interiores se encuentran bien definidas y delimitadas, las exteriores parecen menos densas y también menos definidas, siendo incluso “borrosas”, lo cual sugiere una mayor flexibilidad del doble anillo en estas zonas.

El análisis de imágenes de tinción negativa de una muestra del fragmento N-mtMCM resulta en una serie de vistas de la molécula claramente identificables con diferentes proyecciones de la estructura cristalina del doble hexámero de N-mtMCM (Fletcher *et al.*, 2003) (ver Introducción, I.4.2; Resultados, R.2.1, R.2.2). La proyección lateral de la molécula tiene un aspecto muy similar a las dos barras interiores de las vistas laterales de mtMCM (Figura R.3). Podemos concluir, por tanto, que mientras las dos barras interiores corresponden a las regiones N-terminales de la molécula (residuos 2-286), las barras exteriores corresponden a las C-terminales (287-666), siendo estas regiones más flexibles que la N-terminal.

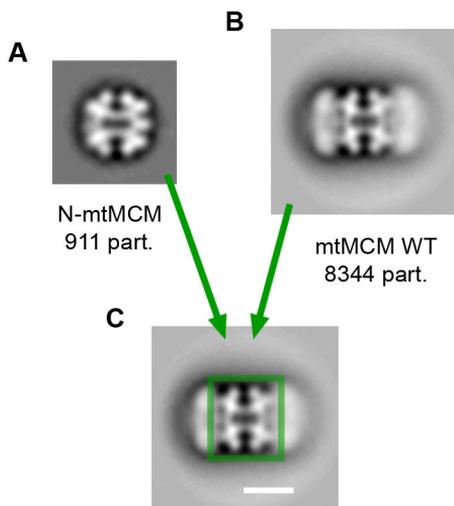


Figura R.3: Localización del dominio N-terminal de mtMCM sobre las imágenes de vistas laterales.

A – B: Imágenes medias de proyección de N-mtMCM y mtMCM WT, respectivamente, en vista lateral. El número de partículas incluidas en el promediado se indica debajo de cada imagen. La imagen **B** incluye las imágenes de todas las poblaciones de vistas laterales del muestreo de condiciones físico-químicas. **C:** Ubicación del dominio N-terminal en mtMCM por comparación entre ambas imágenes. Todas las imágenes se encuentran filtradas a 21 Å. La barra representa 100 Å.

R.1.3 – Las vistas laterales son mayoritariamente proyecciones de dobles hexámeros

Las vistas laterales son por tanto proyecciones de un doble anillo de mtMCM a lo largo de un eje transversal. Un conjunto de proyecciones transversales que se encuentren distribuidas cubriendo el espacio angular, es decir, que correspondan a rotaciones del doble anillo en torno a su eje longitudinal, contiene la información necesaria para reconstruir la estructura tridimensional de la molécula. Utilizamos la técnica de asignación de direcciones de proyección (*projection matching*) (Materiales y métodos, M.10) para reconstruir el mapa 3D del doble anillo. Este algoritmo necesita una estructura inicial de referencia para la primera iteración de asignación de direcciones. La elevada similitud de la región N-terminal de las vistas laterales con las proyecciones laterales del fragmento N-mtMCM, sugiere inmediatamente a este fragmento como candidato ideal para la referencia inicial. Sin embargo, dado que los resultados obtenidos con esta técnica de reconstrucción son muy sensibles a la elección correcta de la referencia inicial (Frank 2006a; Shaikh *et al.*, 2003), y puesto que no existe certeza de la constitución oligomérica de las vistas laterales, que podrían ser dobles hexámeros, dobles heptámeros, o una mezcla de proyecciones de ambos objetos, decidimos realizar varios experimentos de asignación angular sobre el mismo conjunto de imágenes cambiando únicamente las referencias iniciales.

Como modelos iniciales utilizamos la estructura cristalográfica de N-mtMCM (un doble hexámero) y tres modelos sintéticos: un doble pentámero, un doble heptámero y un doble octámero, generados a partir de esta estructura realizando los desplazamientos y rotaciones apropiados (ver Materiales y métodos, M.10). Todos los modelos fueron generados sobre la estructura atómica y posteriormente filtrados. El conjunto inicial de proyecciones experimentales lo constituyen el conjunto de las vistas laterales de los experimentos en la Tabla R.1, en ausencia de NTP y a 25 °C. En total se utilizaron 8883 imágenes.

La técnica de asignación de direcciones de proyección parte de proyectar el volumen inicial de referencia y comparar cada imagen experimental con cada una de las proyecciones, asignando a la imagen los ángulos de la proyección con la que mejor correlación ha encontrado. Una vez asignados ángulos a todas las imágenes éstas se retro-proyectan en una estructura tridimensional, que se correspondería con una mejor

aproximación al objeto experimental que las generó, y que es vuelta a utilizar como modelo inicial para una siguiente iteración de búsqueda de direcciones de proyección. El método itera de esta manera hasta converger, es decir, hasta que el resultado no varía a lo largo de las iteraciones.

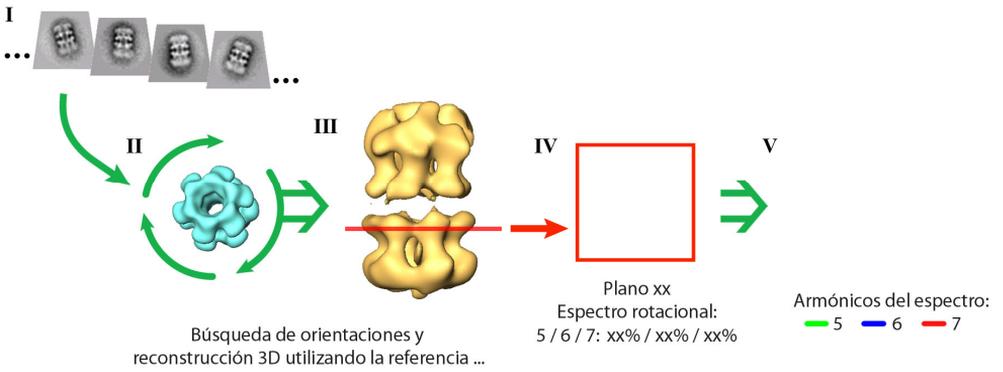
Durante el proceso, realizamos la búsqueda de orientaciones sólo en una región central de las imágenes de 77 Å de radio, que incluye la región N-terminal con amplitud. En el momento de la retro-proyección utilizamos sin embargo las imágenes completas. No fueron impuestas condiciones de simetría a lo largo del proceso, sin embargo los volúmenes fueron simetrizados, cuando se indica, una vez reconstruidos.

El resultado de las reconstrucciones estuvo claramente condicionado por la simetría de los modelos iniciales, de modo que utilizando el doble pentámero, el doble hexámero y el doble heptámero de N-mtMCM como referencias, el algoritmo converge a una estructura de dos anillos con simetrías 5, 6 ó 7 respectivamente (Figura R.4, B). En el caso del doble octámero el resultado de la reconstrucción no corresponde a ninguna estructura reconocible.

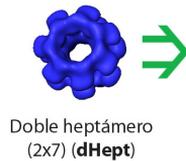
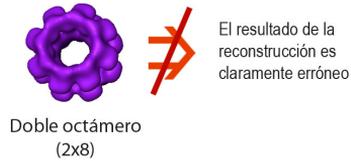
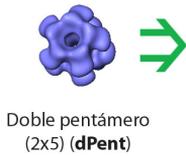
Figura R.4: Proceso seguido para determinar la estructura del doble oligómero de mfMCM.

- A:** Para cada imagen del conjunto de vistas laterales (I) se realiza una búsqueda iterativa de orientaciones basada en una referencia inicial (II). Una vez el método converge las imágenes se retroproyectan en un mapa tridimensional (III) cuya simetría puede ser analizada plano a plano (IV). El análisis global de la simetría en todos los planos del volumen puede utilizarse para determinar la constitución oligomérica de la reconstrucción (V).
- B:** Los volúmenes reconstruidos utilizando como referencias modelos generados a partir de N-mtMCM tienen la misma simetría que éstos (salvo en el caso del doble octámero). **C:** Utilizando referencias toroidales sólo se obtienen reconstrucciones con simetría 6 (toroide del d5 y dHex) o sin simetría rotacional mayoritaria (toroide del d7).

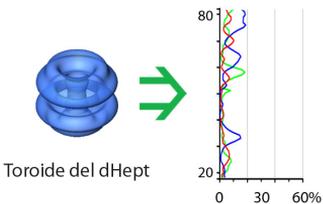
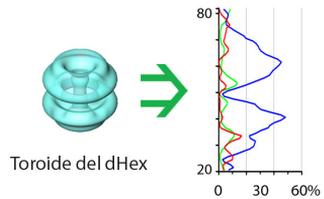
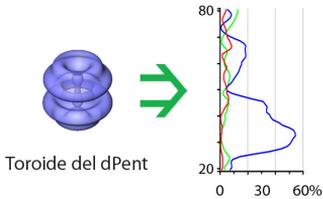
A



B



C



Para minimizar la influencia de la simetría de la referencia inicial, se generaron modelos toroidales, promediando rotacionalmente los tres modelos sintéticos de 5, 6 y 7 monómeros. Estos toroides fueron empleados como referencias iniciales para el cálculo independiente de tres nuevos mapas 3D, tal y como se describió anteriormente. La simetría de los volúmenes resultantes fue analizada posteriormente analizando plano a plano (Figura R.4, B). Este análisis reveló una clara predominancia del armónico 6 sobre el 5 y el 7 en dos de las reconstrucciones, mientras que en la tercera de ellas el mapa final, aún habiendo sido generado a partir de las orientaciones calculadas para una referencia inicial toroidal generada de una estructura heptamérica, no presenta simetría 7 ni ninguna simetría clara en ninguno de sus planos. De este resultado podemos concluir que la mayor parte de las partículas que han generado las vistas laterales, si no todas, son dobles hexámeros (dHex's), y que por tanto la forma en doble anillo mayoritaria de mtMCM es la doble hexamérica.

R.1.4 – Reconstrucción del doble hexámero

Habiendo determinado que las vistas laterales corresponden mayoritariamente a proyecciones de dHex's de mtMCM realizamos una reconstrucción tridimensional del doble hexámero de mtMCM utilizando como referencia inicial el dHex de N-mtMCM, filtrado.

Partiendo de las orientaciones determinadas en la asignación angular que ha utilizado como modelo inicial el toroide construido a partir del dHex de N-mtMCM, descartamos aquellas imágenes cuyo ángulo de inclinación se desvía más de 15° respecto del eje transversal, es decir, aquellas vistas que no son estrictamente laterales.

Analizamos las partículas restantes (5427 imágenes, el 61% del conjunto inicial) y descartamos aquellas que presentan uno o ambos extremos C-terminales mal definidos, clasificando mediante KerDenSOM (Figura R.5, B). El resultado es un nuevo conjunto de imágenes (1668 en total) que representa la variabilidad debida exclusivamente a las diferentes orientaciones que puede tomar el dHex al ser proyectado rotando sobre su eje longitudinal. El conjunto final de datos está por tanto constituido por una población homogénea de vistas estrictamente laterales, proyecciones del

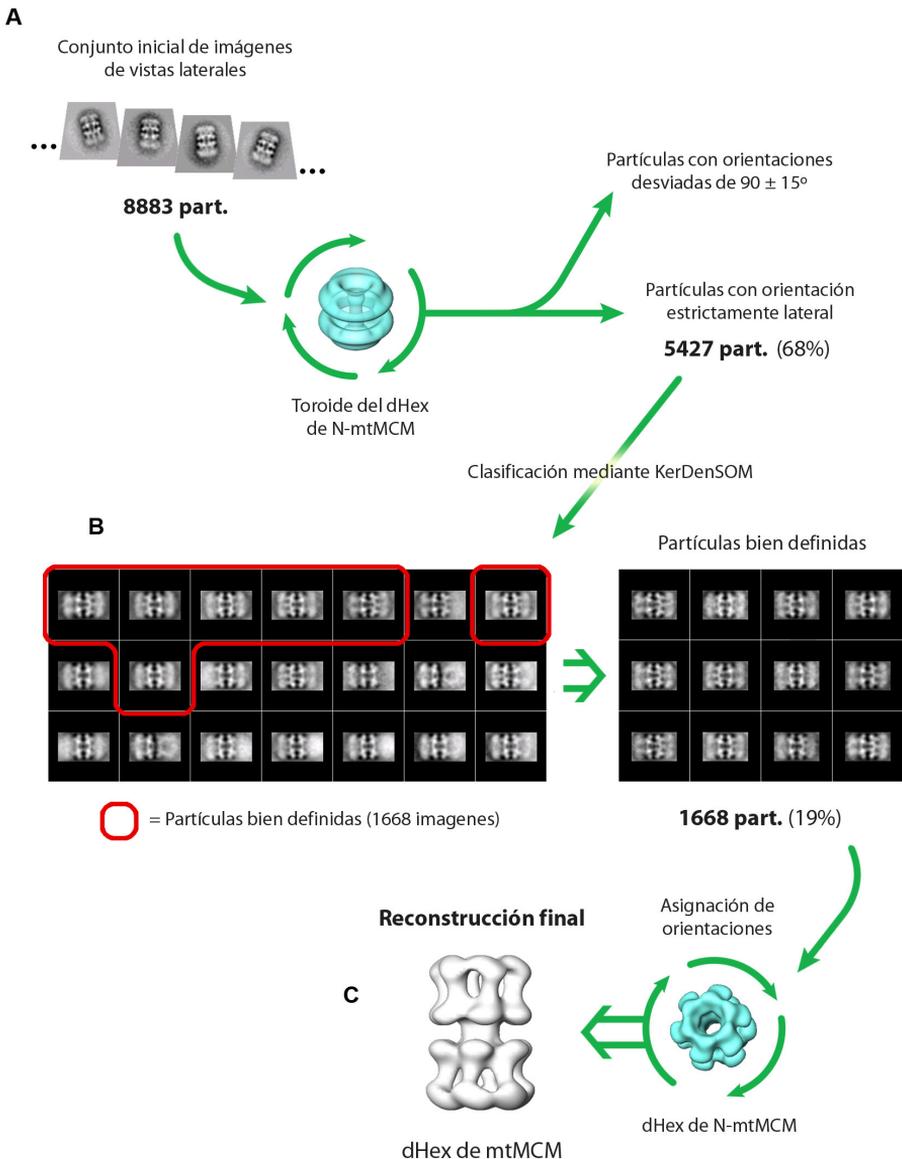


Figura R.5: Reconstrucción final del doble hexámero de mtMCM

A: Del conjunto inicial de imágenes con las asignaciones de ángulos y translaciones deducidos a partir del toroide del dHex de N-mtMCM, se eliminan aquellas que no corresponden a vistas estrictamente laterales. **B:** Posteriormente se clasifican y eliminan las partículas peor definidas. **C:** Con el resultado se realiza una nueva búsqueda de orientaciones y la reconstrucción final, tomando como referencia inicial el dHex de N-mtMCM.

dHex que cubren todo el espacio angular de rotación sobre el eje longitudinal de éste, y cuyos extremos C- y N-terminales se encuentran bien definidos.

Con este conjunto de datos se calcula una reconstrucción tridimensional que alcanza 25 Å de resolución, utilizando la estructura cristalina filtrada de N-mtMCM como modelo inicial, realizando la búsqueda de orientaciones en un radio estricto corresponde a la región N-terminal. Durante el proceso de reconstrucción no imponemos condiciones de simetría. El volumen final resultante es filtrado a la resolución calculada y finalmente se le aplica simetría C_6 . El análisis de las vistas laterales mostró ausencia de simetría 2 con respecto al plano ecuatorial entre anillos, de modo que la simetría D_6 no sería aplicable.

R.1.5 – Estructura del doble hexámero

El volumen final del dHex es un cilindro formado por dos hexámeros enfrentados cabeza con cabeza por sus extremos N-terminales, con regiones paralelas diferenciadas dispuestas en sectores transversales al eje longitudinal, de tal manera que las dos regiones C-terminales se sitúan en los dos sectores exteriores y las N-terminales en los dos interiores (Figura R.6, A). Cada uno de estos sectores se encuentra ligeramente rotado con respecto a sus vecinos, adquiriendo la estructura una torsión total de $\sim 24^\circ$ entre ambos extremos (Figura R.6, B).

Un canal central recorre el dHex a lo largo de su eje longitudinal. Sus dimensiones mínimas (25 Å de ancho y 182 Å de largo) podrían acomodar una molécula de dsDNA de 54 pb de longitud (Figura R.6, A, sección lateral), tal y como ha sido sugerido para N-mtMCM, considerando la alta concentración de carga positiva en el interior del canal (Fletcher *et al.*, 2003), del mismo modo que para otras helicasas recorridas por canales centrales (Li *et al.*, 2003).

En la interfase entre monómeros, a la altura de la transición entre C- y N-terminal, aparecen ventanas laterales ovaladas, con radios mínimos de 11 y 25 Å.

Una débil densidad está presente en algunos puntos a lo largo del canal central, apreciable también en las imágenes medias de las vistas laterales (Figura R.7; Figura R.6, A).

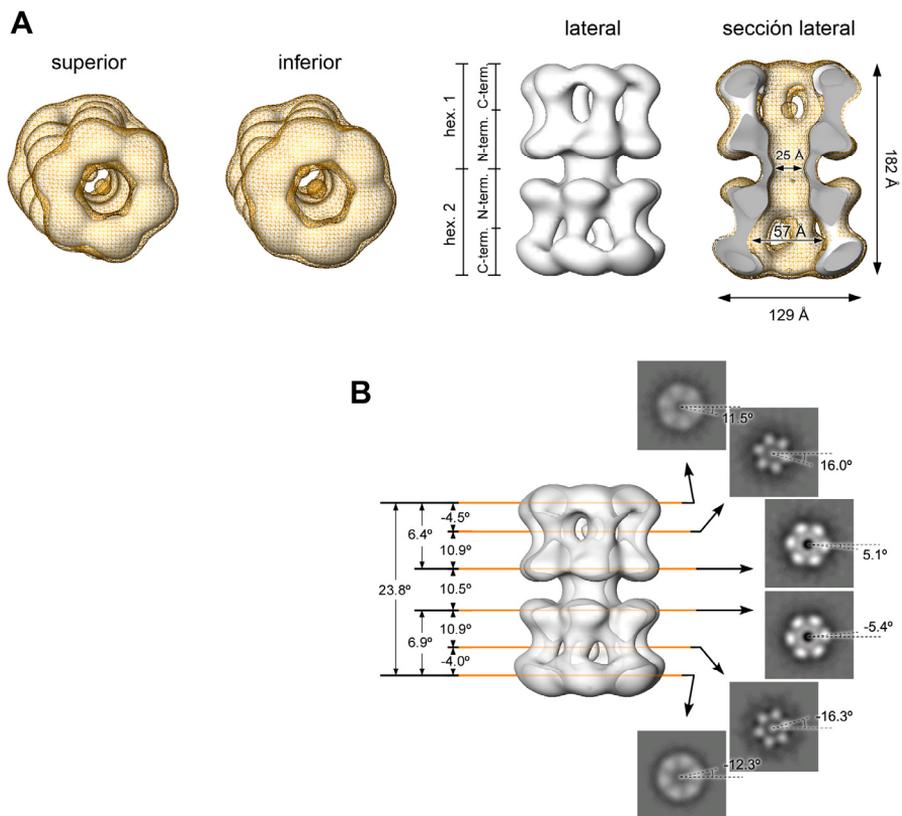
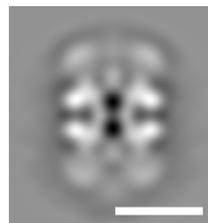


Figura R.6: Estructura del doble hexámero de mtMCM.

A: Reconstrucción tridimensional final mostrada en diferentes orientaciones. La superficie amarilla de malla muestra el volumen generado a un umbral correspondiente al 100% de la masa del dodecámero. La superficie sólida de color blanco está generada a un umbral mayor, correspondiente al 100% de la masa de N-mtMCM para las dos regiones N-terminales. Las dimensiones que se indican han sido calculadas al umbral del 100% de la masa de mtMCM. **B:** Rotación entre regiones: Se indican las posiciones y rotaciones relativas de diferentes planos representativos de cada región del dHex. Los valores de rotación han sido calculados mediante correlación cruzada de cada plano con respecto una proyección del volumen a lo largo del eje longitudinal de la estructura completa.

Figura R.7: Sección central longitudinal de la reconstrucción del dHex de mtMCM.
Una densidad débil se observa entre las regiones N-terminales y C-terminales a lo largo del canal central. La barra representa 100 Å.



R.1.6 – Ajuste de la estructura atómica de N-mtMCM

Es posible encajar la estructura atómica de N-mtMCM en la región N-terminal de nuestra reconstrucción (Figura R.8). Alineamientos de secuencia basados en estructura han demostrado que la región N-terminal de mtMCM, que contiene los dominios de unión a DNA y se encuentra poco conservada en secuencia en la familia MCM, se encuentra sin embargo bastante conservada estructuralmente (Fletcher *et al.*, 2003). Por su parte, la región C-terminal incluye los motivos Walker A y B siendo, por tanto, la responsable de la actividad helicasa y, aunque se encuentra más conservada en secuencia que la región N-terminal, no hay ningún homólogo cuya estructura cristalina se haya resuelto para este dominio (Forsburg 2004; Jenkinson y Chong 2006; Tye 1999; Tye y Sawyer 2000) (ver Introducción, I.4.2).

Sin embargo es posible encajar una región equivalente de alguna helicasa resuelta no homóloga, que incluya los dominios Walker A y B; por ejemplo, es posible encajar los dominios D2 y D3 (residuos 355 a 627) responsables de la actividad helicasa de la proteína hexamérica LTag de SV40, que han sido resueltos en complejo hexamérico (Li *et al.*, 2003).

El fragmento N-terminal y el dominio helicasa de LTag suponen un total de 510 residuos. Su encaje sobre la estructura del dHex ocupa totalmente el volumen asignado a un umbral correspondiente al 100% de la masa (datos no mostrados). Dado que la proteína mtMCM completa tiene 666 aminoácidos de longitud, falta el espacio para encajar ~150 residuos en la estructura reconstruida.

Se ha propuesto un modelo donde estos residuos C-terminales forman una corona sobre la región C-terminal de la estructura final (Costa *et al.*, 2006a, 2006b). En ocasiones, con poca frecuencia, hemos observado

densidad en esta región en algunas reconstrucciones de mtMCM. Sin embargo, encontramos una débil densidad deslocalizada frecuentemente en el canal central de nuestra reconstrucción (Figura R.7; Figura R.6, A). Considerando este hecho junto con la extrema flexibilidad que hemos observado en la región C-terminal, planteamos la posibilidad de que los residuos no modelados estén dotados de una flexibilidad que les permita adoptar posiciones variables alrededor del dominio helicasa C-terminal e incluso ocasionalmente en el canal central. Esta flexibilidad impediría determinar su posición mediante técnicas de promediado de imágenes, motivo por el cual no aparecen en la reconstrucción.

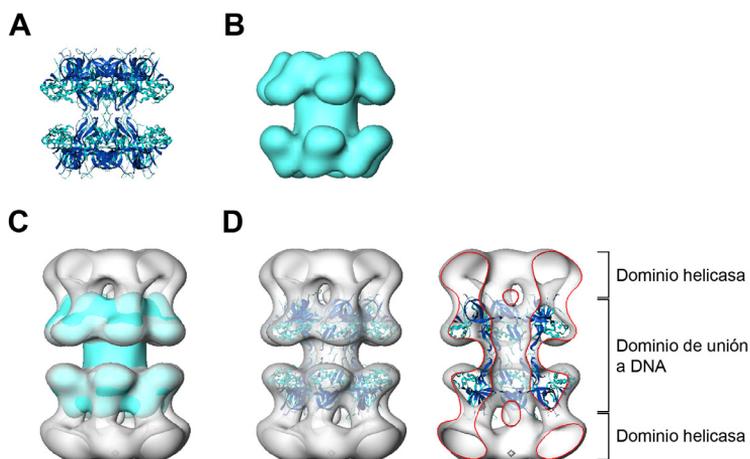


Figura R.8: Ajuste de la estructura atómica de N-mtMCM en el volumen reconstruido del dHex.

A – B: Estructura atómica y mapa de densidad filtrado a 25 Å del dHex de N-mtMCM. **C – D:** Ajuste sobre la estructura del dHex de mtMCM de las estructuras de A y B con los dos hexámeros ligeramente rotados. En la sección lateral se indica el perfil del dHex en rojo.

R.2 – Análisis estructural de mutaciones en los residuos de la interfaz entre hexámeros del fragmento N-mtMCM¹

R.2.1 – Interacciones en el fragmento amino terminal de mtMCM

La resolución de la estructura atómica del fragmento amino terminal de mtMCM (N-mtMCM, aminoácidos 2-286. Ver Introducción, I.4.2) (Fletcher *et al.*, 2003) ha mostrado que mientras que un gran número de interacciones conservan los monómeros unidos en los hexámeros individuales sólo unos pocos aminoácidos de cada monómero interaccionan para mantener los dos hexámeros unidos entre sí.

Uno de estos residuos, la arginina 161 (Arg-161 ó R161), forma puentes de hidrógeno con cadenas laterales de monómeros adyacentes, tanto en del mismo anillo como del anillo opuesto (Fletcher *et al.*, 2003; Fletcher *et al.*, 2005) (Figura R.9). Puesto que hay 12 R161 en un doble hexámero, es posible que éste residuo tenga un papel importante en mantener las interacciones hexámero–hexámero.

Se han documentado otras mutaciones en el motivo de unión a Zn que tienen influencia en las actividades de la helicasa, sin efectos, sin embargo, en la estabilidad del doble hexámero. Por ejemplo, una sustitución de la cisteína 158 por una serina reduce las actividades de unión a Zn, de unión a DNA y la actividad helicasa, pero no produce cambios en la estructura oligomérica (Poplawski *et al.*, 2001).

R.2.2 – Estructura de N-mtMCM en solución

Una muestra de N-mtMCM Wild Type (WT), expresada y purificada de la misma forma que la proteína que fue cristalizada (Fletcher *et al.*, 2003) (ver Materiales y métodos, M.1) y en un tampón prácticamente idéntico al utilizado para estimar su estado oligomérico mediante cromatografía de filtración molecular, fue teñida negativamente, observada al microscopio y los resultados obtenidos sometidas a procesamiento bidimensional.

¹ Fletcher, R.J., Shen, J., Gomez-Llorente, Y., Martin, C.S., Carazo, J.M. and Chen, X.S. (2005). *J Biol Chem* **280**(51): 42405-42410.

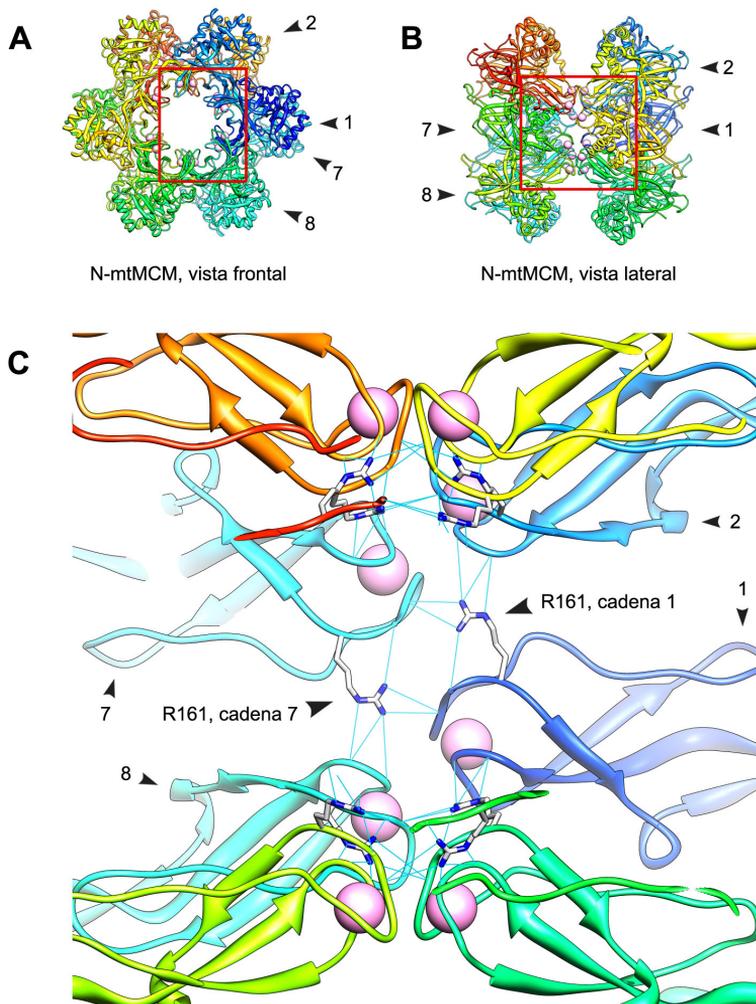


Figura R.9: Puentes de hidrógeno formados por la arginina 161 en el doble hexámero de N-mtMCM.

A – B: Vistas frontal y lateral de la estructura cristalina del doble hexámero de N-mtMCM. Los monómeros se indican en diferentes colores. Dos monómeros opuestos de hexámeros diferentes se indican como 1 y 7, y dos de sus monómeros laterales adyacentes como 2 y 8 (los 4 en diferentes tonalidades de azul). **C:** Posibles puentes de hidrógeno entre la arginina 161 (R161) de los monómeros 1 y 7 con el esqueleto polipeptídico de su misma cadena y de cadenas cercanas.

Las partículas se alinearon, clasificaron y reagruparon en dos clases diferentes (Tabla R.2, A-C) utilizando el programa Xmipp para el alineamiento y mapas autoorganizativos KerDenSOM para la clasificación (ver Materiales y métodos, M.9). Las imágenes medias finales se filtraron a la resolución indicada.

Podemos calcular un mapa de densidad a partir de la estructura cristalográfica de N-mtMCM y filtrarla a la resolución obtenida para las imágenes de microscopía. De esta forma podemos intentar identificar cada uno de nuestros dos grupos de partículas con vistas diferentes del doble hexámero. En la Tabla R.2 (F-H) se muestran proyecciones filtradas del volumen correspondiente a la estructura cristalina de N-mtMCM. Las proyecciones se realizan en dos orientaciones distintas del volumen: la primera se corresponde al volumen proyectado a lo largo de un eje longitudinal que atraviesa la estructura a través del canal central, proyección que denominamos “vista frontal”. La segunda es una proyección perpendicular a la primera, que denominamos “vista lateral”.

Claramente nuestros grupos de imágenes experimentales corresponden a vistas frontales y laterales del doble hexámero (Tabla R.2, A-B). La similitud se aprecia más claramente si aplicamos simetría 6 en el caso de las vistas con forma de anillo, que equivaldrían a las vistas frontales clásicas de algunas helicasas hexaméricas (Tabla R.2, C).

R.2.3 – Mutantes N-mtMCM-R161A y 6Ins

Con el fin de determinar qué aminoácidos están implicados en la capacidad de doble oligomerizar de la proteína mtMCM, se diseñaron mutantes del fragmento N-mtMCM, más estable y mejor caracterizado, alterados en la arginina 161 (R161), el residuo presumiblemente más implicado en las interacciones hexámero–hexámero (Figura R.9). En el laboratorio de nuestro colaborador X. S. Chen, se purificaron dos mutantes en los que la función de este aminoácido pudiera quedar alterada, bien por sustitución (en el mutante R161A, en que la arginina 161 fue sustituida por una Alanina), bien por inserción (en el mutante 6Ins, en que la secuencia GGS GSG se insertó entre la Cys158 y la Gly159).

Tanto mediante cromatografía de exclusión en gel, como mediante geles nativos se comprobó que, mientras el fragmento N-mtMCM WT (32,3 kDa) se comporta como un oligómero de un peso molecular elevado (~380 kDa), consistente con el tamaño de un dHex (387 kDa); los mutantes R161A y 6Ins forman oligómeros de menor tamaño (~190 kDa) consistentes con la estequiometría de hexámeros sencillos (193 kDa). Adicionalmente, se realizaron ensayos de capacidad de unión a DNA, actividad ATPasa y actividad helicasa en proteínas mtMCM completas, mutadas en R161 de la misma manera. Estos mutantes mostraron una disminución en la actividad helicasa con respecto a la proteína WT, sin embargo, su actividad ATPasa y la capacidad de unión a DNA no mostró variaciones con respecto a la proteína WT (Fletcher *et al.*, 2005).

R.2.4 – Estructura de los mutantes R161A y 6Ins

Las muestras de R161A y 6Ins fueron analizadas paralelamente a la microscopía y procesamiento del fragmento N-mtMCM. Las partículas se alinearon y clasificaron utilizando Xmipp y KerDenSOM siguiendo el mismo procedimiento que para N-mtMCM WT (ver Materiales y métodos, M.9).

En ambas muestras mutantes se encontraron solamente partículas correspondientes a vistas frontales de anillos hexaméricos, con un valor muy alto del armónico 6 (Tabla R.2, D y E). En el caso de R161A, fueron descartadas un elevado número de imágenes de vistas frontales, bien debido a que se encontraban ligeramente inclinadas, o bien debido a ser excesivamente ruidosas.

Una característica llamativa de los resultados obtenidos es la quiralidad presente en las imágenes experimentales de R161A y 6Ins. Quiralidad es el término aplicable a una imagen (o a una proyección) que no es posible superponer sobre su imagen especular. La quiralidad de una proyección desaparece si sumamos a ésta su imagen especular. De este modo, la proyección del doble hexámero de N-mtMCM carece de quiralidad, es decir, es aquiral, porque está constituida por dos hexámeros sencillos con orientación opuesta, cuyas proyecciones individuales serían imágenes especulares.

Por el contrario, la marcada quiralidad de las vistas frontales hexaméricas de las muestras de R161A y 6Ins sugiere que la estequiometría de los oligómeros formados es la de anillos sencillos.

Otra característica interesante de los resultados, que sugiere también que los mutantes forman sólo hexámeros sencillos, es la ausencia de vistas laterales similares a las de N-mtMCM WT (Tabla R.2, A). Esta ausencia puede deberse a la inestabilidad que supondría una orientación lateral para un anillo sencillo sobre la superficie cargada de la rejilla, la cual probablemente induciría al anillo a tumbarse, adoptando las partículas de las proteínas mutantes siempre, de este modo, una orientación frontal.

Para confirmarlo, fue tomado uno de los dos hexámeros de la estructura atómica de N-mtMCM, y posteriormente fue filtrado y proyectado a lo largo de un eje que recorre longitudinalmente su canal central (Tabla R.2, I). Al igual que para N-mtMCM WT, la similitud entre las proyecciones de la estructura cristalina y las imágenes experimentales es evidente, indicando que efectivamente R161A y 6Ins forman hexámeros sencillos, de acuerdo a los resultados bioquímicos previos (Fletcher *et al.*, 2005).

Podemos confirmar por tanto que las mutaciones permiten a la proteína mantener una organización anular hexamérica pero impiden la estructuración en doble hexámero. El análisis bioquímico de los mutantes indica que éstos se encuentran disminuidos en su actividad helicasa con respecto a la proteína WT, sugiriendo que probablemente la forma doble-hexamérica es imprescindible en la función.

R.3 – Caracterización de las condiciones para la obtención de complejos mtMCM:DNA y su estudio mediante ME

Abordar el estudio estructural de la función de una proteína requiere a menudo determinar su estructura en varios estados funcionales. La determinación de las diferencias conformacionales entre cada estado funcional puede permitir profundizar en la comprensión de su mecanismo de funcionamiento. En el caso de una helicasa, diferentes estados funcionales susceptibles de provocar cambios conformacionales con respecto a la forma 'libre' son, por ejemplo, el estado de unión al nucleótido y el de unión al sustrato ácido nucleico (denominado de aquí en adelante estado '*acomplejado*') (Dillingham 2006; Enemark y Joshua-Tor 2006; Li *et al.*, 2003; Skordalakes y Berger 2003).

Existe poca información estructural sobre proteínas MCM en estado de unión al DNA (Costa *et al.*, 2006a, 2006b). Esta carencia se debe principalmente al carácter de la interacción helicasa:ácido-nucleico, intrínsecamente inestable, que permite a la proteína recorrer el DNA en lugar de quedarse firmemente unido a él. Otra circunstancia que dificulta el estudio de los complejos MCM:DNA es, en el caso de mtMCM, su elevado polimorfismo, el cual dificulta aún más la detección e interpretación de los cambios conformacionales asociados a la interacción con el ácido nucleico.

Desde un punto de vista técnico, la ME acarrea una dificultad adicional: Utilizando metodologías que permiten la reconstrucción tridimensional de proteínas, tales como la tinción negativa o la criomicroscopía electrónica (crioME), las moléculas de ácidos nucleicos son difíciles de visualizar. De esta forma, en la observación de una muestra mtMCM:DNA no sería posible diferenciar a priori una partícula de mtMCM libre de otra que está formando complejo, simplemente por la visualización del DNA presente en la partícula.

El abordaje ideal para sortear esta dificultad técnica, implicaría la purificación de la forma *acomplejada* para realizar así el trabajo sobre la forma aislada. Sin embargo la inestabilidad de la interacción probablemente impida el aislamiento de la forma *acomplejada*.

Una posibilidad para conseguir estabilizar el complejo podría consistir en retener la proteína sobre el DNA en complejos similares a los formados en la iniciación de la replicación, mediante la interacción con proteínas cargadoras o con las proteínas iniciadoras. Sin embargo, el

desconocimiento existente sobre el funcionamiento *in vivo* de estas proteínas, que abarca tanto la incertidumbre existente acerca de su forma activa, sus cofactores proteicos y su regulación, como el desconocimiento sobre la secuencia o motivo estructural que determina el punto de origen de replicación, dificulta enormemente el planteamiento de los posibles experimentos.

Una alternativa razonable, en el caso de que aislar la forma unida a DNA no fuera posible, consistiría en marcar una de las dos formas (libre o unida a DNA) para así seleccionar las partículas de cada tipo y trabajar con ellas independientemente. En estas circunstancias, es necesario estimar la proporción de proteína unida a DNA en cada condición experimental, con el fin de trabajar en las condiciones en que se maximice la relación proteína acomplejada/libre. Esta estimación la podemos realizar mediante electroforesis de retardo en gel o bien mediante el marcaje selectivo de partículas de complejos en los que el DNA se encuentra biotililado, con conjugados estreptavidina-oro, utilizando partículas de oro de pequeño tamaño. El marcaje con estreptavidina-oro permitiría una estimación de la proporción proteína acomplejada/libre, pero las partículas de oro imposibilitarían la reconstrucción.

Una vez estimado que en las condiciones de experimentación la cantidad de partículas acomplejadas es razonable, la utilización de marcaje con estreptavidina sin oro en el DNA biotililado sí permitiría una reconstrucción posterior. Aunque no permite una preselección a priori, la distinción de ambos tipos de partículas podría llevarse a cabo mediante el uso de algoritmos de clasificación (KerDenSOM, Maximum likelihood).

R.3.1 – Descartando contaminación con DNA celular

Como es lógico, la determinación de diferencias conformacionales entre las formas libre y acomplejada, requiere trabajar con muestras totalmente libres de contaminación con DNA celular, en las que se tenga control certeza de la ausencia de complejos a la hora de reconstruir la forma libre.

Todas las muestras de proteínas purificadas que se han utilizado en este trabajo han sido sometidas a un protocolo estándar de caracterización antes de su primer uso, que incluye la verificación de la ausencia de contaminación con DNA celular (ver Materiales y métodos, M.1).

Los ácidos nucleicos de presentes en una célula presentan su extremo 5' fosforilado. Este fosfato se puede intercambiar por un fosfato radiactivo

de una molécula γ -³²P-ATP mediante la acción de la polinucleótido quinasa del bacteriófago T4 (ver Materiales y métodos, M.4). De esta manera todas las moléculas de ácido nucleico contaminante presentes en la muestra de proteína quedan marcadas radiactivamente en su extremo 5'-P, ya sean moléculas de ssDNA, dsDNA o de otros ácidos nucleicos. Las moléculas radiactivas se pueden detectar en cantidades muy pequeñas mediante métodos de elevada sensibilidad, siendo posible por tanto comprobar la presencia de pequeñas cantidades de DNA contaminante.

Los experimentos demostraron que todas las muestras de mtMCM WT se encontraban ausentes de contaminación (Figura R.10).]

R.3.2 – Análisis de complejos N-mtMCM:DNA

R.3.2.1 – Complejos con DNA de 300 pb

La detección de la cantidad de proteína en estado libre es por tanto crucial en nuestros experimentos de formación de complejos MCM:DNA.

Un método rápido y asequible de detección de ambas formas es la electroforesis en condiciones nativas. Para ensayar la formación de complejos realizamos experimentos con el mutante N-mtMCM, más homogéneo y que forma bandas más definidas y discretas en geles nativos que la proteína mtMCM WT (datos no mostrados). Hemos utilizado electroforesis nativa en acrilamida al 6% (ver Materiales y métodos M.5) en la que detectamos la posición del DNA mediante tinción con SYBR Safe o BrEt, y la posición de la proteína mediante tinción con Coomassie.

Del mismo modo utilizamos una molécula de dsDNA de 300 pb, cuyo retardo sea fácilmente apreciable y que proporcione una intensidad de señal suficiente para su detección con SYBR Safe o BrEt.

Realizamos un conjunto de experimentos de saturación de la unión de N-mtMCM a DNA, preparando varias reacciones en las que mezclamos una cantidad constante de N-mtMCM frente a concentraciones crecientes de dsDNA, hasta comprobar que la cantidad de proteína acomplejada, detectada como bandas de proteína retardadas, deja de aumentar.

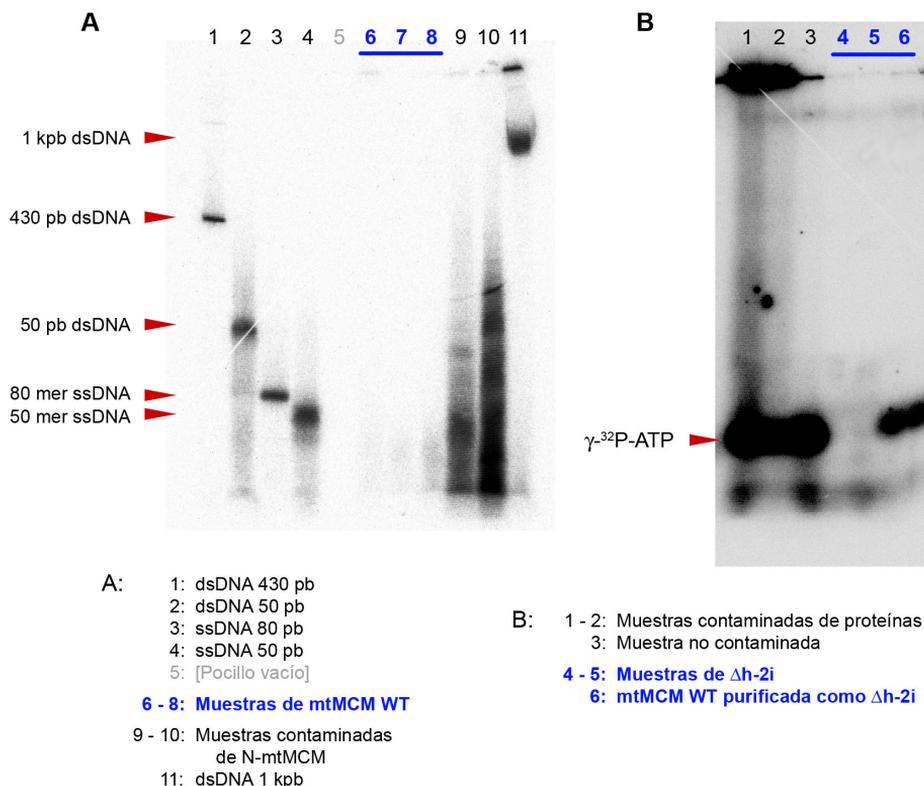


Figura R.10: Detección del DNA celular contaminante mediante fosforilación radiactiva, electroforesis y autorradiografía.

A: Resultados de diversas muestras marcadas. Pocillos 1 – 4 y 11: DNAs control marcados, de longitud conocida. 6 – 8: Muestras de mtMCM WT utilizadas por defecto en este trabajo, libres de contaminación. 9 – 10: Muestras de N-mtMCM con diversos grados de contaminación. En esta imagen no se aprecia el frente del gel.

B: Resultados de muestras marcadas del mutante Δh2i y WT, purificadas de la misma manera. Pocillos 1 – 3: Otras proteínas mtMCM no utilizadas en los experimentos, mostradas con propósito comparativo, de las cuales 1 y 2 tienen diversos grados de contaminación y 3 se encuentra no contaminada. 4 – 6: mtMCM-Δh2i y WT, no contaminadas. La banda de γ -³²P-ATP libre residual marca el frente de la electroforesis.

Puesto que una molécula de N-mtMCM en forma doble hexamérica (N-mtMCM₁₂) tiene una longitud de ~120 Å a lo largo de su canal central, y una molécula de DNA de 300 pb tiene ~1000 Å de longitud, es posible que cada molécula de DNA pueda unir simultáneamente varias moléculas de dHex de N-mtMCM. Por tanto, posiblemente la saturación, en la que todas las moléculas de proteína se encuentran unidas a DNA, se produzca por debajo de la proporción equimolar DNA:N-mtMCM₁₂.

La reacción de formación del complejo se llevó a cabo en ausencia de NaCl y Mg²⁺ o en presencia de NaCl 50 mM y/o Mg²⁺ 5 mM, sin que se detectaran diferencias apreciables entre las distintas condiciones (Figura R.11, A). En todas ellas se encontró una cantidad de proteína libre, aún quedando un excedente de DNA susceptible de formar complejo (Figura R.11, A, pocillo 6). Podemos concluir que existe una cantidad determinada de dsDNA que satura la capacidad de unión de N-mtMCM en estas condiciones, es decir, existe una proporción de N-mtMCM₁₂ libre:acomplejada a partir de la cual no aumenta la cantidad de complejo formado, permaneciendo siempre una cantidad de proteína en estado libre.

Las bandas de DNA retardado coinciden con bandas de proteína, indicando que en estas posiciones se encuentra un complejo proteína:DNA. En los primeros pasos del gradiente no se detecta la presencia de DNA en la banda retardada, probablemente por la pequeña cantidad de DNA presente, que produce una débil señal.

Recorriendo el gradiente de menor a mayor concentración de DNA, se observa un “patrón en escalera”: para las mayores concentraciones utilizadas aparecen varias bandas de proteína poco retardadas (pocillos 5 y 6); mientras que a menores concentraciones aparece una única banda de retardo muy retrasada (pocillos 2 y 3). En las condiciones de baja concentración de DNA de los primeros pocillos, es difícil la detección de éste mediante SYBR o BrEt, de modo que no apreciamos su presencia en la banda de retardo.

Es posible explicar este patrón considerando que en condiciones de baja proporción dsDNA/N-mtMCM₁₂, la posibilidad de que varias moléculas de proteína se unan a una sola molécula de DNA aumenta, sufriendo estos complejos, por tanto, un fuerte retardo. Por otro lado, en proporciones cercanas a la equimolecularidad se formarían complejos en los que tan sólo un doble hexámero o dos de N-mtMCM se unen por molécula de DNA, produciendo bandas cuyo retardo es considerablemente menor (Figura R.11, B).

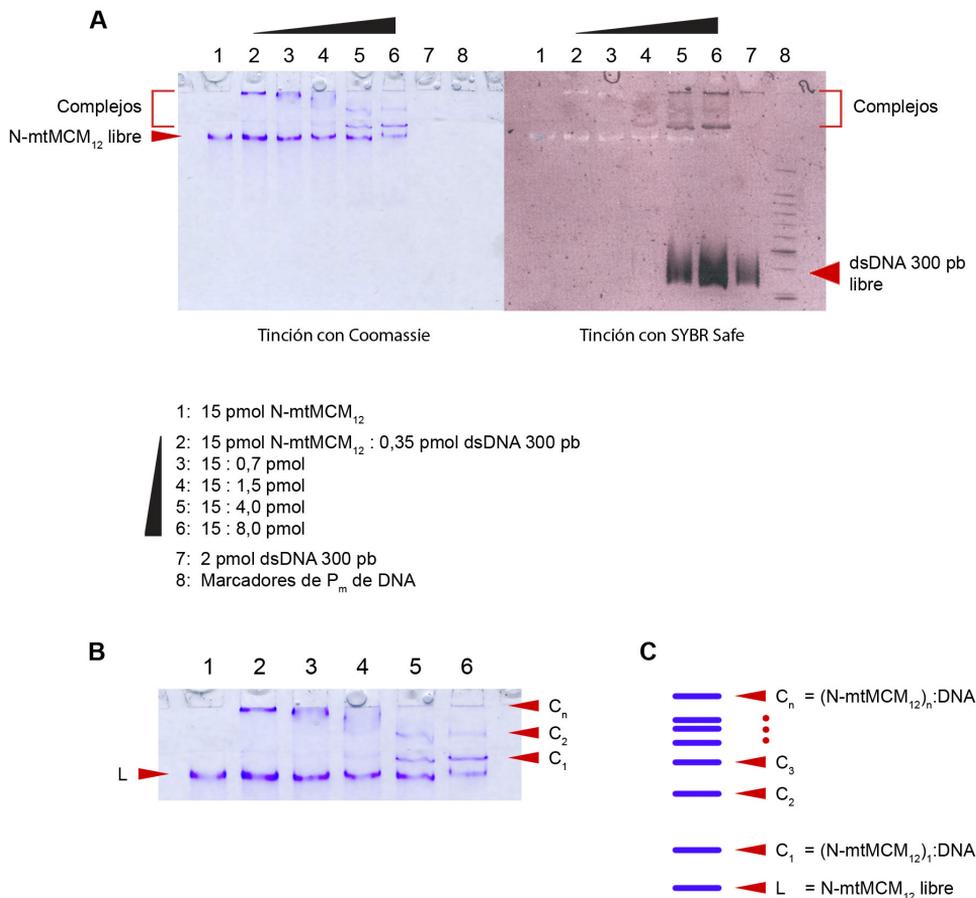


Figura R.11: Retardo en gel para los complejos N-mtMCM:DNA en un gradiente de dsDNA de 300 pb. Detección de la cantidad de proteína libre.

Las mezclas de reacción fueron incubadas 30 min a temperatura ambiente en el tampón original de la proteína, en ausencia de NaCl ó Mg^{2+} . **A:** Detección de las bandas de proteína en el gel mediante Coomassie y de las bandas de DNA mediante SYBR Safe. Los pocillos 1 y 7 indican la posición de las bandas de proteína libre y DNA libre, respectivamente. Los pocillos 2 a 6 forman un gradiente de DNA para una concentración constante de proteína. En estos pocillos se puede apreciar la aparición de diferentes bandas de retardo (que contienen complejos DNA:proteína), diferentes en cada paso del gradiente, que forman un "patrón en escalera". **B y C:** Hipótesis que explica el patrón en escalera. A una gran concentración de DNA se forman complejos con un único dodecámero de N-mtMCM por molécula de dsDNA (C_1). A medida que disminuye la cantidad de DNA, cada molécula de DNA une más moléculas de N-mtMCM12 (C_2 , C_3 , etc.), y disminuye la cantidad de complejos con uno o pocos dodecámeros unidos.

R.3.2.2 – Complejos N-mtMCM:DNA de 40/60 pb

Experimentos similares fueron realizados con moléculas más pequeñas de DNA, con la intención de formar complejos de 1 dodecámero de N-mtMCM por cada molécula de DNA y ajustar el tamaño de ambas moléculas, de cara al trabajo con partículas individuales en ME.

Empleamos construcciones de ácidos nucleicos diseñadas para minimizar la formación de estructuras secundarias no deseadas. Utilizamos dos estructuras diferentes, una de ellas es simplemente DNA de banda doble de 40 pb de longitud, diseñado para acomodar tan sólo una molécula de dHex (40 pb = 136 Å de longitud) (Figura R.12). La segunda de ellas es una molécula de doble banda de 40 pb que contiene una burbuja de 20 pb (= 68 Å) (ver Materiales y métodos M.2). Esta construcción pretende mimetizar la estructura del origen de replicación en el momento de la carga de la helicasa.

Los experimentos también son realizados con ssDNA de 40 pb, y DNA's (ss y ds) de un tamaño ligeramente superior, 60 pb (~200 Å), en presencia de diferentes concentraciones de Mg^{2+} (0; 0,5 ó 5,0 mM), AMPPNP 5 mM o/y NaCl 50 mM y en ausencia de cofactores.

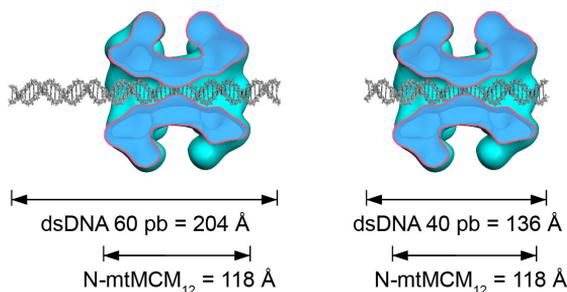
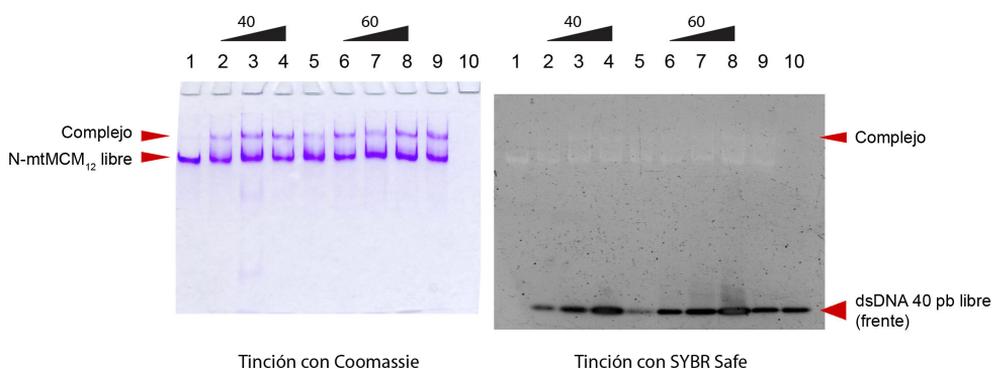


Figura R.12: Comparación de tamaños de los DNA's de 40 y 60 pb utilizados frente al doble hexámero de N-mtMCM.

Representación de los DNA's de 40 y 60 pb utilizados en los experimentos, anillados en el canal central de la estructura atómica del doble hexámero de N-mtMCM, filtrado y seccionado lateralmente.

En todos los casos los resultados fueron similares (Figura R.13, se muestran solamente los resultados para dsDNA). Podemos observar como en este experimento no aparece un patron “en escalera”, formándose en su lugar una única banda de retardo en cualquier punto del gradiente, indicando la formación del mismo tipo de complejo para todas las concentraciones de DNA utilizadas. En la mayoría de los casos no conseguimos detectar la posición de las bandas de DNA, probablemente debido a la débil señal que emite la banda retardada, proveniente de una pequeña concentración de un DNA muy pequeño.



- 1: 15 pmol N-mtMCM₁₂
 2: 15 pmol N-mtMCM₁₂ : 20 pmol dsDNA 40 pb
 3: 15 : 50 pmol
 4: 15 : 150 pmol
 5: 15 : 20 pmol dsDNA 40 pb + 5 mM AMPPNP
 6: 15 pmol N-mtMCM₁₂ : 20 pmol dsDNA 60 pb
 7: 15 : 50 pmol
 8: 15 : 150 pmol
 9: 15 : 20 pmol dsDNA 60 pb + 5 mM AMPPNP
 10: 20 pmol dsDNA 60 pb

Figura R.13: Retardo en gel para los complejos N-mtMCM:DNA en un gradiente de dsDNA de 40 pb.

Las mezclas de reacción fueron incubadas 30 min a temperatura ambiente en el tampón original de la proteína, 50 mM NaCl y suplementado con 5mM Mg²⁺. El retardo en gel se ha realizado para dsDNA de 40 y 60 pb. En cada punto del gradiente se aprecia una única banda de retardo, que aparece en la misma posición independientemente de la concentración de DNA utilizada. Esta banda probablemente corresponde al complejo DNA:proteína, aunque debido a que la detección de DNA ha fallado en esa banda, no es posible asegurarlo.

R.3.2.3 – Triple hexámero de N-mtMCM

Los complejos N-mtMCM:DNA formados como se indica en el apartado anterior se analizaron mediante ME, realizando procesamientos de las partículas individuales de varias muestras: una muestra control que contiene sólo la proteína N-mtMCM, y dos muestras de formación de complejos N-mtMCM:DNA, una de ellas con un dsDNA de 40 pb y la otra con un dsDNA de 60 pb (en el gel de la Figura R.13, los pocillos 1, 4 y 8 respectivamente).

En los resultados (Tabla R.3) podemos observar por primera vez un triple hexámero (3Hex) de N-mtMCM (Tabla R.3, D). El 3Hex está constituido por un doble hexámero de N-mtMCM, en su disposición habitual cabeza con cabeza, sobre el que se encuentra superpuesto un hexámero sencillo de N-mtMCM, cabeza con cola.

	A	B	C	D
N-mtMCM libre	6343 (52%)	4339 (35%)	1342 (11%)	286 (2%)
N-mtMCM: dsDNA (40 pb)	9828 (59%)	0 (0%)	6861 (24%)	3307 (17%)
N-mtMCM: dsDNA (60 pb)	2823 (35%)	0 (0%)	4676 (59%)	469 (6%)

Tabla R.3: Resultados del procesamiento de las muestras de N-mtMCM:DNA con dsDNA de 40 y 60 pb. Presencia del 3Hex de N-mtMCM.

Se indican los 4 diferentes tipos de partículas observados en cada caso y su proporción con respecto al total de partículas. **A:** Partículas en vista frontal (con el canal libre): Probablemente en este grupo están incluidas vistas frontales de dHex y 3Hex, de ahí su quiralidad parcial. **B:** Partículas en vista frontal con el canal ocluido: En contra de lo que cabría esperar, este grupo sólo está presente en el control sin DNA (muestra “N-mtMCM libre”). Posiblemente la densidad en el canal se deba a un artefacto de tinción, observado en otras ocasiones (ver Resultados, R.3.3.4). **C:** Vista lateral de un dHex. **D:** Vista lateral de un 3Hex. La barra indica 100 Å.

Es posible por tanto completar la interpretación de los geles de retardo con DNAs de pequeño tamaño (Figura R.14, A). De esta manera, la banda retardada probablemente se encuentre constituida por la superposición de dos bandas, la correspondiente al 3Hex y una segunda banda formada por el complejo 3Hex:DNA. De la misma manera, la banda de proteína no retardada posiblemente corresponda también a otras dos bandas superpuestas, el complejo N-mtMCM₁₂:DNA y N-mtMCM₁₂ en estado libre.

Suponemos que las bandas de proteína libre y proteína acomplejada se encuentran superpuestas en ambos casos dada la pequeña diferencia de peso molecular entre las dos formas, y suponiendo también que la diferencia de carga entre la proteína libre y el complejo no es determinante a la hora de su avance a lo largo del gel (Figura R.14, C).

Interpretando la electroforesis podemos deducir que el 3Hex se forma mayoritariamente en presencia de DNA, pero no podemos confirmar este resultado mediante ME, ya que la gran similitud entre la vista frontal de un dHex y de un 3Hex dificultaría la clasificación y por tanto la interpretación del procesamiento.

Una última electroforesis parece confirmar que la forma triple hexamérica también puede existir aunque en menor proporción, en ausencia de DNA (Figura R.14, B), aunque no podemos confirmarlo, ya que una ligera contaminación de DNA exógeno en las muestras de N-mtMCM fue detectada, posteriormente a la realización de estos experimentos.

Adicionalmente, los experimentos realizados con ssDNA o con construcciones con una alta proporción de ssDNA (por ejemplo, construcciones con burbujas u horquillas), mostraron ausencia de la banda de retardo, correspondiente al 3Hex (datos no mostrados), lo cual sugiere que esta forma aparece predominante en presencia de dsDNA.

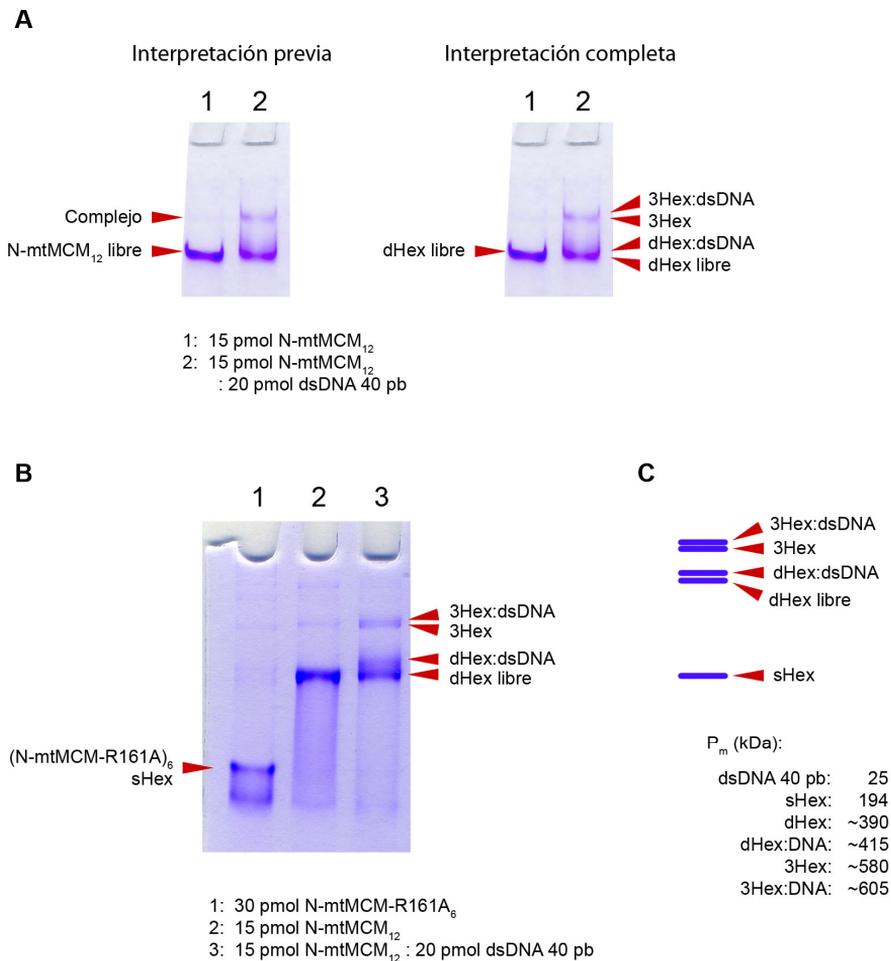


Figura R.14: Interpretación de las bandas de retardo en los geles nativos de complejos N-mtMCM:DNA.

A: Podemos reinterpretar los resultados de los geles de retardo para DNAs de pequeño tamaño, tras haber observado la presencia de 3Hex en nuestras preparaciones. **C:** De esta forma, dada la pequeña diferencia de peso molecular entre las formas libres de dHex y 3Hex y las formas unidas a DNA, la banda retardada probablemente corresponde a la banda de 3Hex superpuesta a la del complejo con DNA e igualmente la banda no retardada al dHex de N-mtMCM y su complejo con DNA. **B:** En una nueva electroforesis parecen apreciarse todas las bandas, incluso la forma 3Hex en ausencia de DNA. En el gel que se muestra se ha incluido también una muestra del mutante R161A, que forma únicamente hexámeros sencillos (sHex) (pocillo 1).

R.3.3 – Análisis de complejos mtMCM:DNA

R.3.3.1 – Análisis de condiciones de formación de complejos con DNAs de 70 pb

Es posible deducir dos conclusiones importantes de los experimentos del apartado anterior: En primer lugar, hemos comprobado que siempre existe una proporción de proteína que permanece libre en presencia de DNA, en multitud de condiciones, independientemente del exceso de DNA utilizado.

La segunda conclusión es que no podemos distinguir mediante electroforesis la proteína que se encuentra formando complejo de la que se encuentra en estado libre cuando el DNA que utilizamos es muy pequeño. En estas condiciones el retardo que el DNA produce sobre la banda de proteína, crucial para distinguir que ésta se encuentra acomplejada, es demasiado pequeño, y ambas bandas (proteína libre y proteína en complejo) solapan.

El propósito de estos experimentos fue encontrar condiciones en que la gran mayoría de la proteína o toda, si fuera posible, se encontrara acomplejada. Una vez fue comprobada la imposibilidad de conseguir una alta eficiencia de unión en nuestras condiciones de trabajo, el siguiente paso fue encontrar condiciones en que la eficiencia de formación de complejo fuera lo más alta posible. Dado que no ha sido posible diferenciar los dos estados de la proteína (libre y acomplejada), es posible detectar la mejora en la eficiencia de acomplejación en dos condiciones diferentes teniendo en cuenta tan sólo la cantidad de DNA retardado.

De este modo, en los siguientes experimentos fue utilizada tinción de DNA en geles de agarosa para detectar exclusivamente bandas de DNA (ver Materiales y métodos, M.5.2). Los electroforesis nativa en acrilamida fue descartada para la proteína mtMCM completa además, debido a que parte de la proteína permanece retenida en el pocillo (datos no mostrados), probablemente debido a su gran tamaño en forma nativa.

Para detectar con mayor sensibilidad las pequeñas cantidades de DNA retardado, el mismo gel que fue teñido con SYBR fue posteriormente transferido a una membrana mediante *southern-blot*, detectando el DNA previamente marcado con biotina mediante estreptavidina-peroxidasa (Materiales y métodos, M.5.3).

Para intentar maximizar la formación de complejos fue utilizada la proteína mtMCM mutante de delección de la región h2i (mtMCM- Δ h2i), cuya afinidad por DNA se ha determinado mayor que la de mtMCM WT

(Jenkinson y Chong 2006) (ver Introducción, I.2.3), y una muestra de mtMCM WT purificada de la misma manera (ver Materiales y métodos, M.1). Utilizamos diferentes estructuras de DNA, todas con una longitud total de 70 pb (~240 Å de longitud) que, de acuerdo al tamaño calculado en la estructura del doble hexámero de mtMCM (~180 Å de longitud), debe impedir impedimentos estéricos en las interacciones de la biotina del extremo con la estreptavidina utilizada en los experimentos de marcaje.

Las construcciones de DNA utilizadas fueron (ver Materiales y métodos, M.2): un oligonucleótido de ssDNA (forkFLMCM-70-1bio), dsDNA (dsFLMCM-70), una horquilla de 30 pb de banda sencilla (forkFLMCM-70) y una burbuja de 30 pb (bbleFLMCM-70).

Fueron realizados diferentes experimentos de saturación con DNA con las cuatro construcciones diferentes y las dos proteínas (Figura R.15), no encontrándose variación apreciable en los resultados entre proteínas, ni dependiendo de las condiciones (fueron comprobadas condiciones de presencia o ausencia de ATP, AMPPNP, Mg^{2+} y diferentes concentraciones de NaCl). Sin embargo es posible apreciar una ligera mejora en la unión con ssDNA con respecto a las demás construcciones (Figura R.15, comparar DNA libre con banda retardada en pocillos num. 3 en todas las construcciones, en la tinción con SYBR). Este resultado está de acuerdo con el comportamiento de la mayoría de helicasas (revisado en Delagoutte y von Hippel 2002; Patel y Picha 2000; Tuteja y Tuteja 2004), pero sin embargo no ha sido apreciado para mtMCM en otras investigaciones (Chong *et al.*, 2000).

La utilización de detección mediante transferencia y estreptavidina-peroxidasa es un método sensible pero no fácil de cuantificar. Mientras que utilizando tinción de DNA (SYBR o BrEt) la intensidad de la señal es proporcional a la masa del ácido nucleico; utilizando este otro método, la señal depende de la proporción de las moléculas que se encuentren correctamente biotiniladas, de la cantidad de DNA transferido (que varía según la longitud del DNA o sus estructuras secundarias) y de la calidad del reacción de detección a lo largo de la membrana (la exposición real de una zona de membrana al complejo estreptavidina-peroxidasa), que a menudo es heterogénea en membranas de gran tamaño.

De esta forma la cuantificación de la unión a DNA se puede deducir de la tinción de éste, pero la detección más sensible se realiza mediante *southern-blot*, siendo posible, en todo caso cuantificar el resultado para cada tipo de DNA independientemente, no conjuntamente.

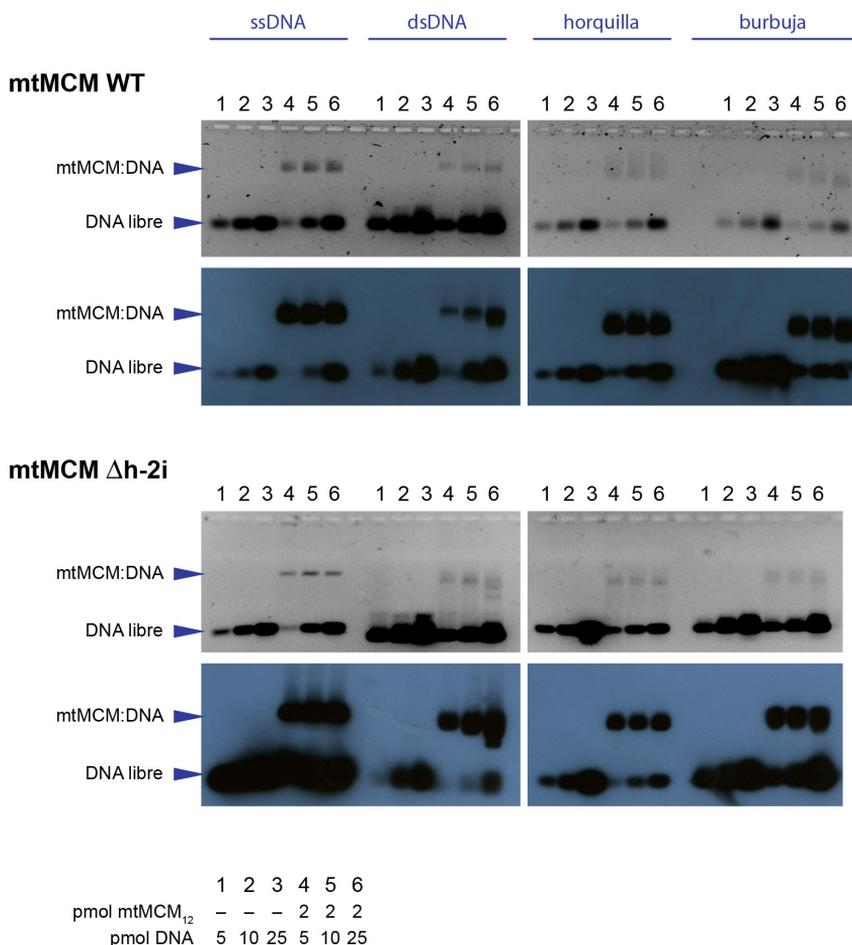


Figura R.15: Gels de retardo con diferentes construcciones de DNA para mtMCM WT y Δ h2i.

Electroforesis de retardo en gel para DNAs de 70 pb y las construcciones ssDNA, dsDNA, horquilla y burbuja, en gradientes de concentración de DNA, para las proteínas mtMCM WT (en 4 mM ATP, 7 mM Mg²⁺ y 50 mM NaCl) y mtMCM Δ h2i (en 0 mM ATP, Mg²⁺ y NaCl). Se muestran los resultados del gel al 1% de agarosa teñido con SYBR (imágenes en blanco y negro) y del mismo gel transferido a una membrana y revelado mediante estreptavidina peroxidasa (imágenes en azul y negro). Es posible apreciar un aumento en la cantidad de complejo formado para ssDNA, con ambas proteínas.

R.3.3.2 – Titulación de la unión mtMCM:ssDNA

Con el fin de estimar la cantidad de DNA necesaria para saturar la capacidad de unión de la proteína fue realizado un experimento de saturación detallado partiendo de una cantidad de DNA muy inferior a la cantidad de proteína utilizada.

Los resultados mostraron que para ssDNA de 70 pb, 1,5 pmol de mtMCM- $\Delta h2i_{12}$ es saturado con 1 pmol de ssDNA. Es decir, aproximadamente un tercio de la proteína permanece en estado libre, independientemente de la cantidad de DNA añadido (Figura R.16).

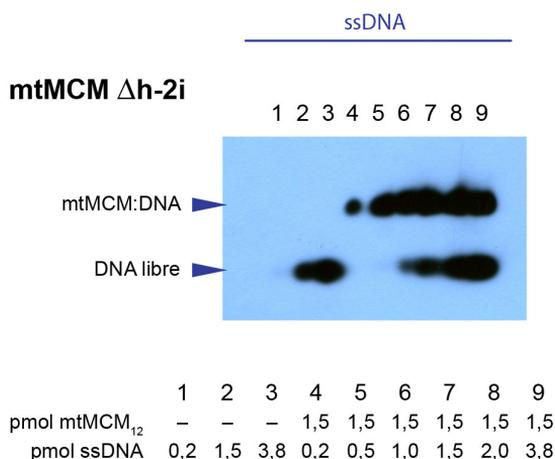


Figura R.16: Titulación de la unión de mtMCM- $\Delta h2i$ a ssDNA.

Southern-blot de un gel de saturación de la capacidad de unión de mtMCM- $\Delta h2i$ a DNA de banda sencilla. Se aprecia como la banda retardada no aumenta de a partir del pocillo 6 en que se la proporción MCM:DNA es 1,5:1.

R.3.3.3 – Marcaje con estreptavidina-oro

Es posible obtener una estimación cuantitativa de la proporción de complejos formados frente a la cantidad de proteína libre y de la viabilidad de las muestras de complejo para el trabajo con ME, utilizando marcaje individual de los complejos con estreptavidina-oro (strep-Au) y observando las muestras así marcadas al microscopio.

Este tipo de marcaje puede generar falsos positivos. Para evitarlos, fue utilizada la cola de histidinas de mtMCM para, mediante una columna de níquel y tras la formación del complejo, separar las partículas de MCM en el eluido quedando cualquier otra proteína o partícula libre eliminada en los lavados de la columna (ver Materiales y métodos, M.6).

El análisis de los lavados y eluidos de la columna reveló que, mientras mtMCM queda retenido en la columna interaccionando con ella correctamente y eluye posteriormente, el DNA aparece principalmente en el primer lavado (Figura R.17 y Figura R.18).

Estos resultados pueden ser la consecuencia de la esperada inestabilidad del complejo mtMCM:DNA, causante de que la interacción no se conserve en las condiciones de los sucesivos lavados, siendo el DNA eliminado del complejo y quedando la proteína retenida en la columna.

Por otro lado, es posible que el complejo no interaccione con la columna pero sí lo haga la proteína libre. En esta circunstancia debería observarse un incremento de la proteína eluida en los lavados frente a una muestra control de acomplejación, en la que no hayamos añadido DNA.

De cualquier forma, en cualquiera de estas circunstancias no fue logrado el propósito del experimento, separar el complejo de las partículas libres de strep-Au.

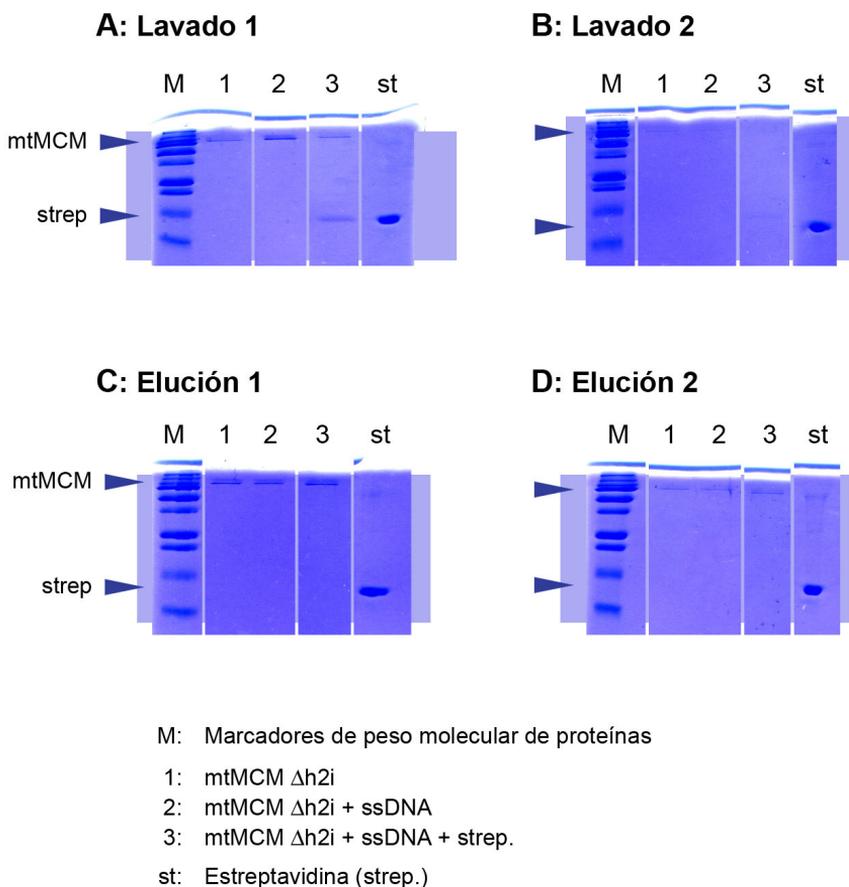
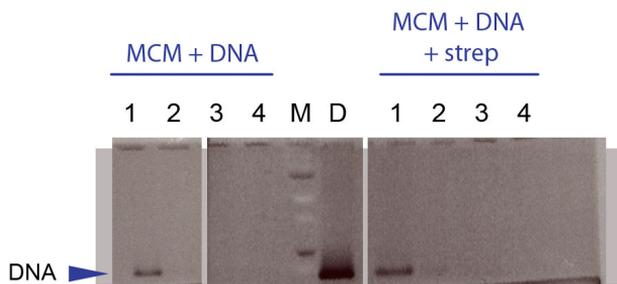


Figura R.17: Análisis de las proteínas del complejo ternario mtMCM- Δ h2i:ssDNA:strep retenidas en columnas de níquel.

Electroforesis desnaturante (en acrilamida al 15%) de una muestra de complejo mtMCM Δ h2i:ssDNA, bien marcada con estreptavidina (sin oro) (pocillo 3), bien no marcada (pocillo 2), junto con un control de proteína libre (pocillo 1). El DNA utilizado tiene 70 nucleótidos. **A y B:** Lavados de la columna tras la unión del complejo. En el primer lavado eluye parte de la proteína en todos los casos y la totalidad de la estreptavidina. **C y D:** Eluciones: En la primera elución se libera gran parte de la proteína unida, pero no aparece señal de estreptavidina, sugiriendo que esta no ha permanecido retenida en la columna. El tampón utilizado fue Tris-Cl 10 mM pH 7,5; 50 mM NaCl.



M: Marcadores de peso molecular de DNA

D: DNA

1: Lavado 1

2: Lavado 2

3: Elución 1

4: Elución 2

Figura R.18: Análisis del DNA del complejo ternario mtMCM- Δ h2i:ssDNA: strep retenido en columnas de níquel.

Electroforesis desnaturizante en acrilamida al 15% para la mismas muestras de complejo mtMCM Δ h2i:ssDNA de la Figura R.17 (en las que el DNA se encuentra marcado con estreptavidina (sin oro) (rótulo “MCM+DNA+strep”), o bien no se encuentra marcado (rótulo “MCM+DNA”). En ambas muestras, todo el DNA eluye en el primer lavado, no quedando retenido por la proteína mtMCM unida a la columna.

En estas condiciones, fueron preparadas rejillas de ME con el primer eluido de muestras de formación de complejos con las cuatro construcciones de DNA (ssDNA, dsDNA, forkDNA y bbleDNA). Las muestras fueron observadas al microscopio electrónico (Figura R.19) y se estimó la proporción de partículas marcadas sobre el total de partículas en las micrografías tomadas (Figura R.20). En todos los casos la proporción de partículas marcadas es muy pequeña, como cabe esperar puesto que el DNA ha eluido principalmente en los lavados. Es posible apreciar una pequeña diferencia en el recuento entre las diferentes construcciones, que indicaría una mejora en la interacción con DNA de banda simple, de acuerdo a los resultados bioquímicos.

Aunque los resultados permiten cuantificar preliminarmente la interacción, las partículas así marcadas no son susceptibles de alineamiento o clasificación, ya que la partícula de strep-Au solapa con la molécula de mtMCM, imposibilitando cualquier tipo de alineamiento o reconstrucción.

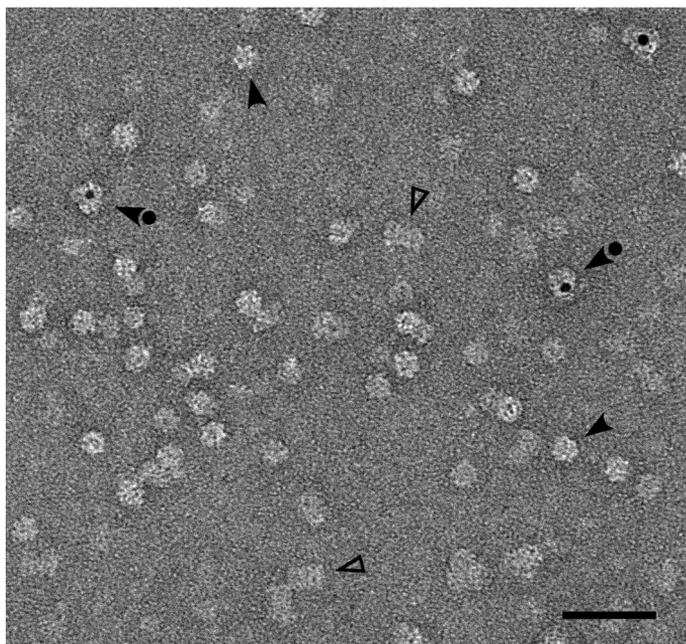


Figura R.19: Campo de microscopía de una muestra de mtMCM:DNA marcada con estreptavidina-oro.

Preparación de una muestra de complejo mtMCM- Δ h2i:bbleDNA:strep-Au teñida negativamente. Es posible observar el mismo tipo de vistas que en una preparación de mtMCM libre (ver Figura R.2), apreciándose tanto vistas frontales (marcadas con "▶") y laterales (▷), como partículas marcadas (●▶). Se ha utilizado la construcción bbleFLMCM-70 y oro coloidal de 5 nm. La barra representa 50 nm.

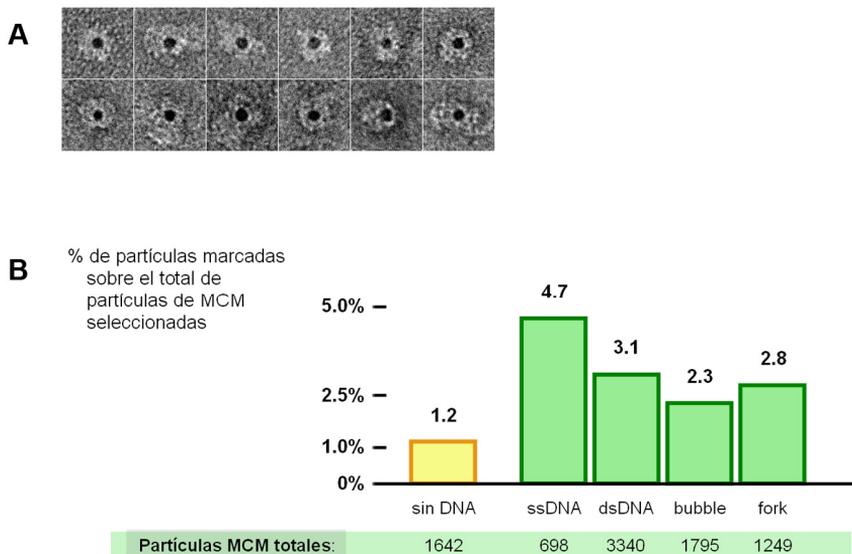


Figura R.20: Recuento de partículas marcadas con oro en los complejos formados con diferentes DNA's sustrato.

Recuento de los complejos mtMCM- Δ h2i:DNA:strep-Au eluidos de una columna de níquel, formados con diferentes DNA's sustrato. **A:** Galería de partículas ejemplo marcadas con oro. En todos los casos la partícula de strep-Au solapa con la partícula de MCM. Es posible apreciar tanto partículas rectangulares (probablemente vistas laterales) como redondeadas (vistas frontales). **B:** Recuento para las diferentes construcciones de DNA utilizadas. En todos los casos la proporción de partículas marcadas es muy baja con respecto al total de partículas, probablemente debido a que la mayoría de los complejos se elimina al perderse el DNA en los lavados. Sin embargo, hay un ligero aumento del número de partículas marcadas para ssDNA.

R.3.3.4 – Marcaje con estreptavidina

El marcaje con estreptavidina (sin oro) permite el alineamiento y clasificación de las partículas, tras el cual es posible distinguir las moléculas marcadas acomplejadas de las partículas de proteína en estado libre (ver Materiales y métodos, M.6).

La estreptavidina se une a la biotina del DNA, situada en el extremo de éste, que posiblemente en los complejos con DNAs de 70 pb se

encuentre próximo al canal central, de producirse la interacción DNA:mtMCM en el interior del canal de ésta.

Utilizamos estreptavidina tetramérica, de ~52 kDa, el tamaño de aproximadamente 2/3 de un monómero de mtMCM (75,5 kDa). En el procesamiento de las partículas así marcadas debería por tanto aparecer en forma de una masa de posición variable próxima al canal central de mtMCM.

Comprobamos la formación del complejo ternario mtMCM:DNA:estreptavidina, mediante retardo en gel (Figura R.21). En los resultados se observa como una fracción de la proteína acomplejada une a su vez estreptavidina. Esta misma muestra de reacción (Figura R.21, ssDNA y dsDNA, pocillo 4) la utilizamos para preparar rejillas de ME junto a dos controles negativos adicionales, uno constituido simplemente por mtMCM libre y otro por mtMCM con estreptavidina pero sin DNA. El experimento fue realizado por duplicado, observando al microscopio y procesando en cada caso las partículas de cada preparación.

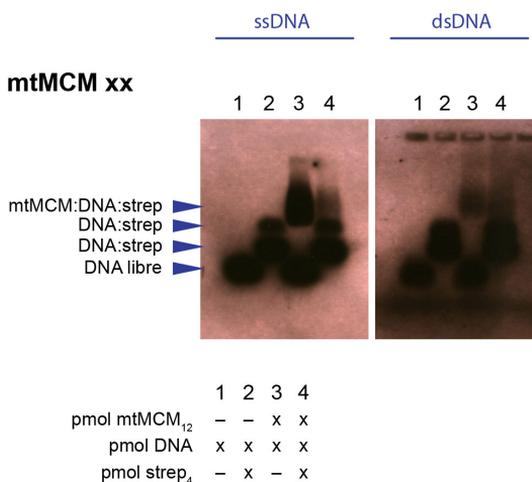


Figura R.21: Formación del complejo ternario estreptavidina:DNA:mtMCM. Retardo en gel.

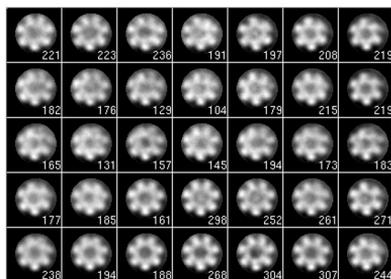
Retardo en gel de agarosa al 1% del DNA previamente marcado con estreptavidina, acomplejado posteriormente a mtMCM.

En ambos experimentos se observa señal positiva en el interior del canal central, en una fracción de las vistas frontales. Esta señal podría corresponder a la estreptavidina. Sin embargo, aparece no sólo en las muestras de complejo sino también idénticamente en el control de proteína libre de uno de los experimentos (Figura R.22).

La observación de las vistas laterales no apoya la idea de que esta señal corresponda a estreptavidina, pues no hay una masa adicional, ni diferencia clara entre las imágenes de las tres muestras, como tampoco la hay entre estas imágenes y las vistas laterales de mtMCM de otros experimentos anteriores (datos no mostrados).

Es posible que la señal corresponda a una tinción parcial o incorrecta del canal, que ya hemos observado en otras ocasiones (ver Tabla R.3, B), o bien a la presencia de zonas desordenadas de la proteína, observadas en la reconstrucción del doble hexámero (ver Figura R.7).

mtMCM- Δ h-2i:ssDNA:strep



mtMCM- Δ h-2i

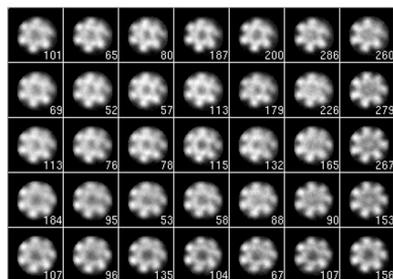


Figura R.22. Mapas autoorganizativos KerDenSOM de las imágenes de vistas frontales de muestras de mtMCM- Δ h2i:DNA:strep y de mtMCM- Δ h2i en estado libre.

Para ambas muestras se aprecia señal en el interior del canal, no pudiendo por tanto relacionar esta señal con la presencia de estreptavidina unida al DNA. El DNA utilizado para formar el complejo es ssDNA de 70 pb. El resultado es similar con dsDNA de 70 pb.

R.4 – Criomicroscopía de mtMCM

Utilizando tinción negativa de muestras de mtMCM, los artefactos de tinción del canal central pueden conducir a una interpretación errónea de los resultados, siendo posible confundir esta tinción con la presencia de una masa colocada en esa posición. De este modo, la tinción negativa no es la técnica adecuada para trabajar con muestras de complejos mtMCM:DNA:estrepavidina. Sin embargo, la utilización de vitrificación en ausencia de agente de tinción como técnica de preparación puede eliminar este problema.

R.4.1 – CriOME en agujeros (sin carbón)

Fueron realizadas observaciones de preparaciones de mtMCM en ausencia de DNA vitrificadas en ausencia de carbón. La muestra se vitrifica en forma de delgados meniscos en el interior de agujeros perforados en el soporte de carbón que cubre la rejilla. Las observaciones en estas condiciones muestran mayoritariamente vistas frontales de partículas con simetría 6 (Figura R.23). Una clasificación posterior mediante maximum likelihood revela un pequeño grupo de partículas con simetría 7 (Figura R.23, B, 4), sin embargo no se consiguió encontrar vistas laterales o inclinadas con respecto a la vista frontal, haciendo inviable la reconstrucción de un volumen en estas condiciones.

Una causa posible de este fenómeno podría consistir en la interacción de la región frontal de las partículas con la interfase hielo-aire, induciéndolas a permanecer en vista frontal.

R.4.2 – CriOME sobre película de carbón

Como alternativa para inducir a las partículas a disponerse en otras orientaciones, fueron utilizadas rejillas sobre las que se depositó una delgada capa de carbón hidrófilo en los agujeros, de modo que se posibilite una interacción lateral de las partículas con este soporte.

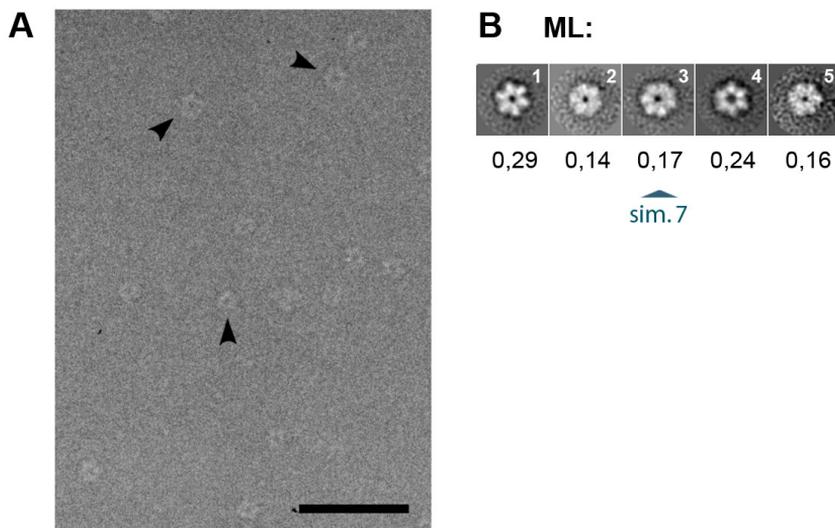


Figura R.23: Campo de crioME de una muestra de mtMCM vitrificada en una rejilla de agujeros. Análisis mediante maximum likelihood.

A: Campo de crioME en el que se pueden apreciar claramente vistas frontales con forma de anillo (marcadas como “▶”). No ocurre lo mismo sin embargo con las vistas laterales. La barra representa 50 nm. **B:** Clases generadas mediante maximum likelihood en espacio real. La mayoría representan partículas con simetría 6, sólo una de ellas partículas con simetría 7 (clase 3). Probablemente el resto de clases se hayan generado en base a diferencias de foco. En la parte inferior se indica el peso de cada clase, que se puede interpretar como la fracción de partículas en tanto por 1.

En estas condiciones, si fue posible encontrar vistas laterales (Figura R.24), aunque en baja proporción, probablemente debido a la dificultad de su visualización con respecto a las vistas frontales, y por tanto de su selección como partículas individuales. La presencia de estas vistas laterales se consiguió determinar principalmente mediante el uso una nueva implementación de Maximum Likelihood que trabaja en el espacio de Fourier y tiene en cuenta los diferentes grados de desenfoque de las imágenes (Scheres *et al.*, 2007a)¹. Estos resultados no se apreciaron en implementaciones anteriores de Maximum Likelihood (ver Figura R.25).

¹ Scheres, S.H., Núñez-Ramírez, R., Gómez-Llorente, Y., San Martín, C., Eggermont, P.P. and Carazo, J.M. (2007). *Structure* (in press).

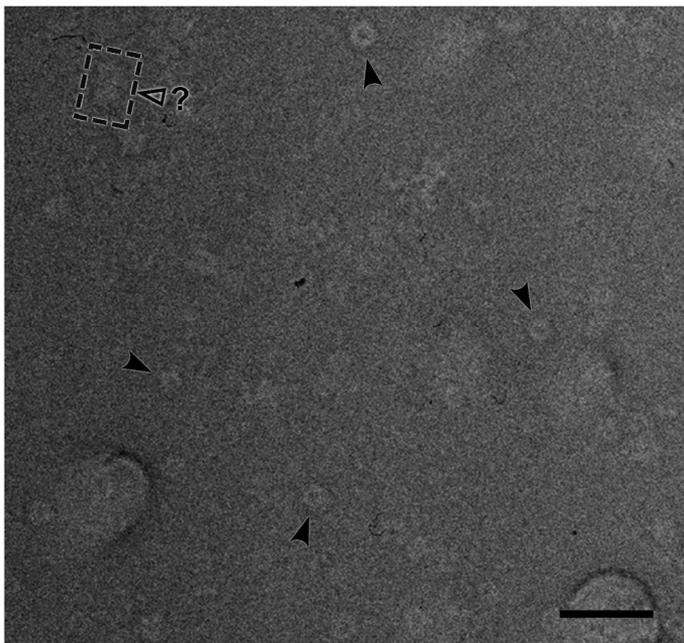


Figura R.24: Campo de crioME de una muestra de mtMCM vitrificada sobre una película de carbón.

Preparación de mtMCM- $\Delta h2i$ vitrificada sobre una rejilla cubierta de una delgada película de carbón. Se distinguen fácilmente vistas frontales con forma de anillo (marcadas con "▶"). Es posible apreciar vistas laterales, aunque con mayor dificultad (? ▶) a pesar de que la imagen ha sido tomada con un alto grado de desenfoque. La barra representa 50 nm

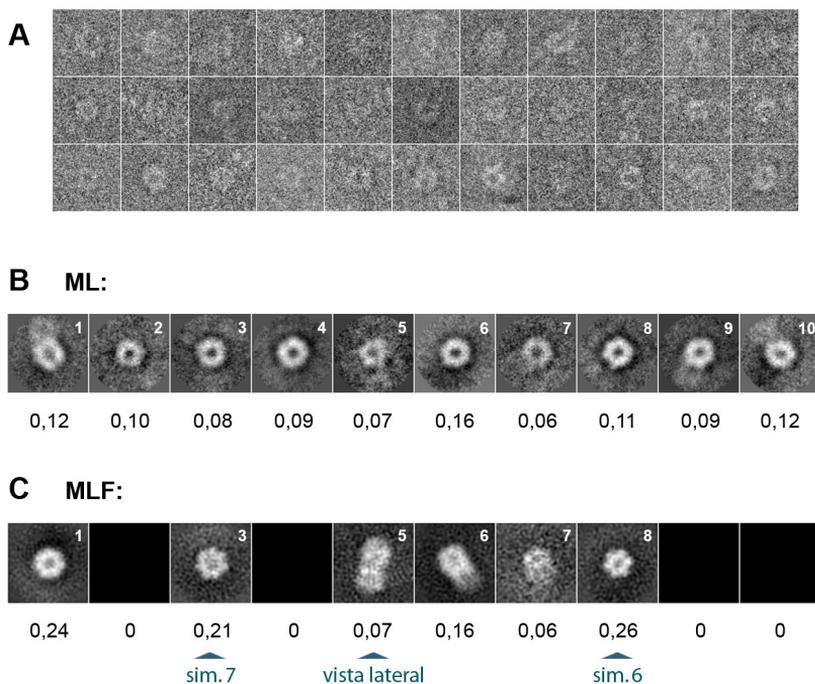


Figura R.25. Procesamiento de la criome con carbón de una muestra de mtMCM- Δ h2i. Análisis de heterogeneidad mediante ML y MLF.

A: Galería de partículas. **B y C:** Clases generadas mediante maximum likelihood en espacio real (ML) (**B**) y maximum likelihood en espacio de Fourier (MLF) (**C**). Bajo las clases se indica la proporción de partículas en tanto por uno. Sólo mediante MLF se consiguen alinear y clasificar las imágenes, mientras que ML en el espacio real converge rápidamente en clases no significativas. Utilizando MLF aparecen tanto vistas frontales de simetrías 6 y 7 (clases 8 y 3 respectivamente), como imágenes probablemente correspondientes a vistas laterales (clase 5). Cuatro de las cinco clases formadas aparecen "vacías", circunstancia que ocurre a menudo con MLF sobre imágenes de mtMCM (Scheres *et al.*, 2007a).

Discusión

Las helicasas son enzimas que se encuentran presentes en todos los procesos fisiológicos de metabolismo de los ácidos nucleicos (replicación, recombinación, transcripción, reparación), catalizando la eliminación de estructuras secundarias, formando parte integral de la compleja maquinaria macromolecular que lleva a cabo cada uno de estos procesos.

Las reacciones que catalizan (como por ejemplo, el desapareamiento de la doble hélice de DNA durante la replicación) son termodinámicamente desfavorables, y son llevadas a cabo mediante el aporte energético que supone la hidrólisis de nucleótidos trifosfato, que es acoplada a la reacción catalizada. Esta reacción debe producirse con la suficiente procesividad, coordinada con el resto de actores proteicos y con el ciclo celular, de modo que las helicasas deben ser capaces de interactuar con elementos reguladores. La integridad de la información genética depende de que estas reacciones se lleven a cabo de la manera adecuada, produciendo los defectos de estas enzimas diversas enfermedades en humanos (síndromes de Bloom (Ellis *et al.*, 1995; Kaneko y Kondo 2004), Cockayne (Ellis 1997), Rothmund-Thomson (Kitao *et al.*, 1999; Macris *et al.*, 2006) y Werner (Gray *et al.*, 1997; Mohaghegh *et al.*, 2001), anemia de Fanconi (Niedernhofer *et al.*, 2005), xeroderma pigmentoso (Berneburg y Lehmann 2001; Vermeulen *et al.*, 2000), tricotodistrofia (Vermeulen *et al.*, 2000), retraso mental con α -talasemia ligado al cromosoma X (Villard *et al.*, 1997)).

D.1 – Polimorfismo estructural

D.1.1 – Hexámeros y heptámeros, dobles y sencillos

Estudios estructurales previos realizados sobre mtMCM han resuelto su estructura tridimensional mediante microscopía electrónica (ME) en diferentes organizaciones cuaternarias, tales como hexámeros sencillos (Pape *et al.*, 2003), heptámeros sencillos (Yu *et al.*, 2002) y estructuras helicoidales (Chen *et al.*, 2005). Por otro lado, la estructura cristalográfica de N-mtMCM se ha resuelto, determinando que este fragmento se organiza en dobles hexámeros (Fletcher *et al.*, 2003).

Mediante cromatografía de filtración en gel, se ha observado en varias ocasiones que la proteína mtMCM eluye en un pico de elevado peso molecular, que concuerda bien con el tamaño de un doble hexámero (Chong *et al.*, 2000; Kelman *et al.*, 1999; Shechter *et al.*, 2000).

En este trabajo se ha observado por primera vez preparaciones en las que mtMCM oligomeriza en estructuras cuaternarias coexistentes de diferentes simetrías, principalmente 6 ó 7. Igualmente, se han encontrado anillos abiertos en las mismas preparaciones, en las que también es posible observar dobles anillos, cuyo estado oligomérico ha sido posible determinar que es mayoritariamente doble-hexamérico.

Sin embargo, todos los experimentos que han sido realizados sobre el fragmento N-terminal de mtMCM, tanto sobre la proteína WT como sobre los mutantes 6Ins y R161A, han encontrado principalmente estructuras con simetría 6, de acuerdo también a otras investigaciones (Chen *et al.*, 2005; Fletcher *et al.*, 2003; Fletcher *et al.*, 2005)

La presencia mayoritaria de dobles anillos exclusivamente doble-hexaméricos, tanto en proteínas mtMCM como N-mtMCM WT parece sugerir que la doble-oligomerización es un proceso favorecido en estructuras hexaméricas de esta proteína o bien que es un proceso que impulsa a la hexamerización de la estructura en la que se genera.

Es posible que la región C-terminal de la proteína mtMCM tenga algún papel regulador en la organización en doble anillo. Así, aunque no tenemos certeza del carácter doble o sencillo de los anillos de vistas frontales que observamos con mtMCM, puesto que las vistas laterales son mayoritariamente dHex posiblemente las vistas frontales con simetría 7 son fundamentalmente heptámeros sencillos. Las estructuras cuaternarias de un solo anillo y las no-hexaméricas, no se observan o aparecen

minoritariamente en la proteína N-mtMCM WT, en la que la región C-terminal se encuentra ausente.

Sin embargo, las interacciones que mantienen la estructura del doble oligómero, establecidas mediante los aminoácidos R161 de cada monómero con monómeros adyacentes, como se ha demostrado en esta tesis (y que sido publicado en Fletcher *et al.*, 2005), no parecen ser incompatibles con estequiometrías de mayor orden, tales como dobles heptámeros, puesto que los átomos que las establecen pueden mantener sus posiciones relativas en ambas estequiometrías.

De cualquier forma, se debe proceder cautelosamente al establecer consecuencias funcionales *in vivo* a partir del estado oligomérico de proteínas que han sido manipuladas genéticamente y sobreexpresadas, y cuyos experimentos y ensayos de actividad se realizan *in vitro*. Sin embargo, el polimorfismo que muestra mtMCM es un hecho corroborado por numerosas investigaciones, que posiblemente indica la existencia de un equilibrio entre varias formas polimórficas.

Un equilibrio entre las diferentes formas oligoméricas explicaría bien la existencia de diferentes proporciones hexaméricas y heptaméricas en preparaciones distintas de una misma muestra de proteína, en la que los oligómeros se encuentran ya formados.

D.1.2 – Anillos abiertos

Los anillos abiertos encontrados se han observado en coexistencia con otras formas en anillo cerrado. En estas preparaciones, las partículas se encontraban bien estructuradas, no existiendo indicios de degradación por la aparición de una cantidad fuera de lo normal de proteína agregada o amorfa.

Los mecanismos de funcionamiento de las helicasas son en general poco conocidos. Se han propuesto tres modelos para explicar la carga de la familia de helicasas hexaméricas en el origen de replicación: enhebrado, reorganización y apertura del anillo (Patel y Picha 2000). El modelo de enhebrado es difícil de aplicar en un DNA circular como el de una arquea, ya que es necesario la presencia de un extremo libre en el DNA que pueda ser introducido a través del canal del anillo. Por otro lado, la presencia de anillos abiertos es fuerte indicativo de que posiblemente el método de carga de mtMCM es el de apertura y cierre del anillo.

Aun cuando debemos ser cautelosos al interpretar estos resultados e inferir conclusiones sobre el funcionamiento de la proteína *in vivo*, como se ha explicado en la sección anterior, nuestros resultados concordarían con una situación en la cual la apertura y cierre del anillo fuera un mecanismo espontáneo de la proteína (Delagoutte y von Hippel 2003), quizás regulado *in vivo* mediante cofactores químicos o la presencia de otras proteínas, quizás proteínas cargadoras. Por otro lado, esta forma puede servir de estado de transición entre las diferentes estequiometrías observadas.

D.2 – Estructura del doble anillo. Similitudes con LTag de SV40

D.2.1 – Estructura de doble hexámero

En nuestras preparaciones hemos observado tanto vistas frontales de diferentes simetrías como vistas laterales de dobles hexámeros. Aunque la relevancia funcional de cada estructura no es aún clara, un doble hexámero (dHex), estructura con simetría 6, correlaciona mejor con la presencia de seis componentes en el complejo eucariótico Mcm2-7 que otras organizaciones (revisado en Dutta y Bell 1997; Kearsey y Labib 1998; Tye 1999).

Una estructura de doble anillo es necesaria en el caso de LTag de SV40, para que la replicación del DNA viral se produzca. Aunque los hexámeros sencillos tienen actividad helicasa por sí solos, sin embargo, su actividad es menor que la del hexámero completo (revisado en Sclafani *et al.*, 2004; Simmons 2000). Del mismo modo, mutaciones similares a las de los mutantes R161A y 6Ins analizados en este trabajo, generadas en la proteína mtMCM completa, producen hexámeros sencillos de mtMCM cuya actividad es menor que la de la proteína WT (Fletcher *et al.*, 2005).

D.2.2 – Similitudes con LTag. Flexibilidad del dominio C-terminal

La estructura tridimensional del doble anillo de mtMCM guarda muchas similitudes con el doble anillo de LTag, además de ser ambos dobles hexámeros. En primer lugar, en las dos proteínas los hexámeros se encuentran orientados enfrentados cabeza con cabeza, apareciendo ventanas laterales, entre cada dos monómeros de cada hexámero. Los fragmentos de ambas estructuras que han sido resueltos se encuentran atravesados por un canal central cargado positivamente con anchura suficiente para acomodar una doble hebra de DNA (Fletcher *et al.*, 2003; Gomez-Lorenzo *et al.*, 2003; Li *et al.*, 2003; Pape *et al.*, 2003; Valle *et al.*, 2000; Valle *et al.*, 2006). Estas semejanzas sugieren que ambas proteínas pueden compartir un mismo modelo de actividad helicasa durante la replicación, quizás similar al modelo de “iris” sugerido para LTag en Li *et al.*, 2003.

Sin embargo la flexibilidad relativa de los dominios se encuentra invertida en ambos sistemas: Tal y como hemos observado a lo largo del alineamiento y la clasificación de las vistas laterales, y también en la reconstrucción, el dominio de unión a DNA en mtMCM es muy rígido mientras que la región C-terminal, responsable de la actividad helicasa, es altamente flexible. En LTag ocurre al contrario, siendo la región helicasa C-terminal bastante rígida y la región N-terminal flexible (Arthur *et al.*, 1988; Li *et al.*, 2003; Valle *et al.*, 2000; Valle *et al.*, 2006).

La helicasa LTag se ensambla específicamente sobre el origen de replicación viral en presencia de ATP y sin la ayuda de factores proteicos accesorios. Los monómeros se ensamblan alrededor del origen para formar dos anillos hexaméricos (Dean *et al.*, 1987; Dean *et al.*, 1992; Fanning y Knippers 1992; Mastrangelo *et al.*, 1989; Simmons 2000). mtMCM, por el contrario, forma dobles hexámeros en solución sin la necesidad de ningún cofactor ni proteína adicional. Algunos estudios consideran, sin embargo, que la organización doble hexamérica de mtMCM se encuentra condicionada a su interacción con dsDNA y sólo se forma en su presencia (Costa *et al.*, 2006a, 2006b). Sin embargo, nosotros hemos comprobado que el dHex se forma en gran cantidad en muestras que libre de DNA, prestando especial atención a trabajar con muestras libres de contaminación por ácidos nucleicos (ver Resultados, R.3.1).

Mientras que LTag es capaz de reconocer por sí solo la secuencia específica del origen, valiéndose del dominio OBD (Arthur *et al.*, 1988; Bradshaw *et al.*, 2004; Kim *et al.*, 1999), mtMCM no es capaz de un reconocimiento específico, no existiendo en su secuencia ningún dominio homólogo a OBD o de función similar. Probablemente mtMCM se sirva de alguna proteína adicional para realizar esta función, tal como mtCdc6-Orc, que actúe como cargador (Bell y Dutta 2002; Fujita *et al.*, 1999; Perkins y Diffley 1998; Weinreich *et al.*, 1999).

La región N-terminal de LTag es una estructura flexible que sólo se organiza correctamente tras la interacción entre ambos hexámeros sobre el origen viral (Valle *et al.*, 2006). Probablemente su flexibilidad es necesaria para el reconocimiento específico del origen de replicación del genoma viral. mtMCM no realiza por sí misma la función de reconocimiento del origen de replicación. Este hecho quizás tenga relación con la mucho mejor estructuración de la región N-terminal, confirmada también en otras ocasiones (Chen *et al.*, 2005; Pape *et al.*, 2003).

Podemos especular con que la flexibilidad del dominio N-terminal de LTag se debe a su necesidad de adaptarse específicamente a un sustrato, mientras que la región de unión al DNA de mtMCM no realiza una unión específica. Del mismo modo, probablemente la flexibilidad del dominio helicasa de mtMCM se encuentra relacionada con una fuerte regulación celular, tal y como ocurre en su homólogo eucariota Mcm2-7 que se encuentra múltiplemente regulado por tres de sus seis subunidades y por otros cofactores proteicos. (Forsburg 2004; Fujita *et al.*, 1998; Ishimi *et al.*, 1998; Kimura *et al.*, 1996).

Parte de la secuencia C-terminal de LTag, los últimos 80 aminoácidos, presenta una gran flexibilidad, pero esta región no tiene actividad en la replicación y es prescindible para llevarla a cabo, siendo su función la de producir el reconocimiento de la célula hospedadora (Pipas 1985; Simmons 2000)

Concluyendo, la región C-terminal mucho más flexible que la N-terminal, parece regular la aparición de anillos abiertos e influir en el equilibrio de formación de dobles hexámeros, permitiendo otras estequiometrías. Posiblemente constituye un punto fundamental de regulación e interacción con otros componentes del replisoma. Las similitudes estructurales de mtMCM con LTag indican que posiblemente comparten muchas características de su modelo de funcionamiento.

D.3 – Formación de complejos con DNA

No existe información estructural a alta resolución de proteínas MCM unidas a DNA. El fragmento N-mtMCM ha sido intentado cristalizar unido a una secuencia de DNA no-específica en varias ocasiones sin éxito por el momento. La elevada flexibilidad de la región C-terminal es probablemente uno de los motivos que dificulta la obtención de un cristal estable.

Por otro lado, el desconocimiento sobre el motivo origen de replicación del *M. thermoautotrophicum* así como la incertidumbre sobre las proteínas y cofactores necesarios para la unión de la helicasa al origen, son circunstancias que limitan la obtención de complejos a aquellos obtenidos mediante interacción libre inespecífica, y por ello poco eficiente, entre la helicasa y el DNA, formando por tanto complejos de reducida estabilidad.

La baja eficiencia en la acomplejación, combinada con la imposibilidad de visualizar el DNA en la selección de partículas individuales, limita las posibilidades de trabajo a condiciones en que la formación de complejos se encuentre altamente favorecida y la cantidad de proteína libre sea mínima. En estas condiciones, la detección de las formas libre y acomplejada de la proteína es crucial. Una posibilidad para facilitar la distinción de las partículas acomplejadas en la selección de partículas es marcar diferencialmente una de las dos formas.

En nuestros experimentos hemos comprobado que efectivamente, en las condiciones ensayadas, no es posible alcanzar alta eficiencia de formación de complejos, obteniéndose los mejores resultados con DNA de banda simple. Muchas otras proteínas helicicas muestran una mayor eficiencia de unión a DNA de banda sencilla que a DNA de banda doble (revisado en Delagoutte y von Hippel 2002; Patel y Picha 2000; Tuteja y Tuteja 2004), aunque este resultado no ha sido comprobado con proteínas MCM y en el caso de mtMCM tan sólo se han observado resultados similares para ssDNA y dsDNA (Chong *et al.*, 2000).

La detección diferencial de proteína libre / acomplejada no es posible para DNAs de pequeño tamaño por métodos rápidos y asequibles como retardo en gel. Por otro lado, el marcaje mediante estreptavidina en muestras teñidas negativamente es inviable como método de detección de densidad en el canal central para mtMCM, siendo la alternativa más útil el marcaje con estreptavidina y vitrificación, que en estos momentos estamos poniendo en práctica.

A pesar de estas limitaciones, los experimentos de unión a DNA han proporcionado algunos resultados interesantes: el patrón de bandas en escalera que se forma en experimentos de formación de N-mtMCM con DNAs de 300 pb (ver Resultados R.3.2) y la estructura de 3Hex (ver Resultados R.3.2.3) sugieren ideas sobre el funcionamiento de mtMCM en su unión al DNA:

D.3.1 – Triple hexámero de N-mtMCM

Las interacciones que mantienen el 3Hex de N-mtMCM no pueden ocurrir en la proteína mtMCM completa, de modo que el 3Hex no tiene relevancia fisiológica. La interfase de unión de los hexámeros que, dentro del 3Hex, se encuentran enfrentados cabeza con cola, se produce entre residuos que sólo se encuentran expuestos y son susceptibles de interactuar en N-mtMCM, mientras que en la proteína mtMCM completa se encuentran colocados en el interior de la estructura de cada monómero, no siendo posible el mismo tipo de interacción.

La existencia del 3Hex de N-mtMCM en presencia de DNAs de 40 ó 60 pb, cuando esta proteína ha mostrado en otros experimentos, en ausencia de DNA, una forma mayoritariamente doble hexamérica, sugiere que probablemente la formación de dHex es un proceso dinámico, existiendo un equilibrio entre la forma de hexámero sencillo (sHex) y de hexámero doble. Así, una cantidad presente de sHex debe entrar en juego para que el 3Hex pueda ser formado a partir de una mayoría de hexámeros dobles.

El 3Hex aparece especialmente cuando N-mtMCM interactúa con DNA (Figura R.14), sugiriendo que esta estructura probablemente se forma sobre el ácido nucleico y que posiblemente, al menos uno de los tres anillos se cargue como anillo sencillo.

D.3.2 – Unión de múltiples moléculas de MCM a una molécula de DNA

Los resultados bioquímicos indican que la gran mayoría de los triples hexámeros de los ensayos de formación de complejos, contienen DNA. La revisión de las dimensiones de los DNA's utilizados sugieren que éstos se deben encontrar mayoritariamente cubiertos por N-mtMCM, si no totalmente (Figura D.1).

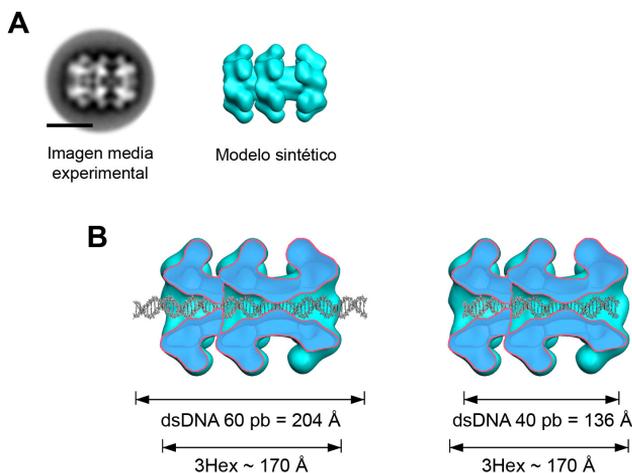


Figura D.1: Dimensiones del 3Hex de N-mtMCM sobre los DNA's utilizados.

A: Imagen media del 3Hex obtenida a partir de las imágenes experimentales y modelo sintético del 3Hex generado a partir de la estructura cristalina filtrada de N-mtMCM. La barra indica 100 Å. **B:** Esquema de un 3Hex de N-mtMCM sobre DNA's de 40 y 60 pb.

La “escalera de bandas” que se aprecia en los experimentos con DNA de 300 pb, sugiere que varios hexámeros de N-mtMCM (1, 2, 3, 4, ..., n) pueden cargarse en una única molécula de DNA, formándose en la electroforesis diferentes bandas sucesivas, cada una para un complejo de una determinada estequiometría ((N-mtMCM₆)_n:DNA) (ver Figura R.11, C). Esta situación recuerda a la “paradoja de MCM” (Dimitrova *et al.*, 1999; Hyrien *et al.*, 2003), observada *in vivo* en numerosos tipos celulares en la que, según el organismo, de 10 a 100 complejos de MCM son cargados por

origen de replicación (Edwards *et al.*, 2002; Lei *et al.*, 1996; Mahbubani *et al.*, 1997). Los cálculos en eucariotas indican que aproximadamente un complejo MCM es cargado por cada 1 kb de dsDNA.

Del mismo modo, se ha determinado que las proteínas MCM (Kaplan *et al.*, 2003; Shin *et al.*, 2003) y también otras proteínas (como DnaB de *Thermus aquaticus* o gp4 de T7 (Kaplan 2000; Kaplan y O'Donnell 2002)) se unen a dsDNA y avanzan sobre él sin desaparecerlo, en ausencia de estructuras en horquilla o con una región no apareada (Forsburg 2004; Patel y Donmez 2006).

Sin embargo, aún no se ha determinado con claridad el papel en la replicación de MCM ni su ubicación con respecto al replisoma durante la misma, aunque algunos estudios indican que MCM se carga en el DNA en el origen de replicación y posteriormente se desplaza sobre el DNA de doble banda, lejos de la horquilla de replicación ejerciendo su función lejos del replisoma. Se ha sugerido un modelo de acción conocido como "modelo de bombeo" (*pumping model*) que intenta explicar estos hechos (Egelman 2001; Kaplan *et al.*, 2003; Laskey y Madine 2003).

De esta forma, es posible que la unión observada de varios dodecámeros de N-mtMCM a moléculas de dsDNA no sea una situación circunstancial, si no que refleje la realidad fisiológica del interior celular.

D.4 – Modelos de funcionamiento

Toda la información obtenida en estos experimentos se puede intentar entender integrada en una red de equilibrios de oligomerización, cuyos estados principales son las diferentes formas que hemos observado mediante microscopía electrónica. Esta red podría ser el núcleo central de un mecanismo muy flexible de oligomerización y carga en el DNA cuyas características hemos esbozado en los modelos de la Figura D.2.

A partir de los monómeros libres de mtMCM recién sintetizados se produce la nucleación, probablemente por concatenación de monómeros en una estructura de anillo abierto, hasta que se alcanza una cantidad mínima, de 6 ó 7 subunidades, que posibilita el cierre del anillo sin tensiones que dificulten las interacciones entre monómeros (Figura D.2, I).

La reacción de nucleación es espontánea, no siendo requerida ninguna proteína específica de *M. thermoautotrophicum* para llevarla a

cabo, y se encuentra muy desplazada hacia la formación del oligómero. Podemos sustentar estas afirmaciones en el hecho de que mtMCM se expresa en *E. coli* formando complejos de elevado peso molecular sin necesidad de otras proteínas de arquea, eluyendo en una columna de filtración en gel limpiamente en especies oligoméricas de alto peso molecular, permaneciendo tan sólo una pequeña proporción de la proteína en forma monomérica (Chong *et al.*, 2000; Fletcher *et al.*, 2003; Fletcher *et al.*, 2005; Gómez-Llorente *et al.*, 2005; Kelman *et al.*, 1999; Shechter *et al.*, 2000).

La forma en anillo abierto, observada por primera vez en nuestros experimentos, puede representar tanto un estado de transición entre diferentes estequiometrías, como la forma de carga de la molécula en el DNA (Figura D.2, II). Otras helicasas como RuvB (Miyata *et al.*, 2000) y la helicasa/primasa de T7 (Crampton *et al.*, 2006) también pueden alternar entre diferentes estados oligoméricos antes de cargarse en el DNA.

Probablemente una única organización cuaternaria de todas las observadas es la forma funcional, constituyendo las demás intermedios de formación de ésta, interconvertibles. La pequeña proporción de anillos abiertos frente a oligómeros cerrados sugiere que posiblemente la forma abierta sea inestable y rápidamente se cierra. La existencia de oligómeros sencillos y dobles en una misma preparación, también observada mediante cromatografía, sugiere que la forma doble probablemente se forma a partir de anillos sencillos (Figura D.2, III y IV).

Los anillos dobles son mayoritariamente hexaméricos, a pesar de la existencia de formas con simetría 7. Este hallazgo apunta que, o bien los dobles heptámeros son inestables y su formación se encuentra impedida, o bien la forma doble hexámerica es muy estable desplazando a los dobles heptámeros formados hacia su estequiometría. Sin embargo, ha sido reportada la existencia de anillos dobles heptaméricos en algunas ocasiones (Costa *et al.*, 2006a, 2006b). Es posible que su inestabilidad disminuya su número con respecto al de dobles hexámeros por debajo de lo que hemos podido detectar mediante las técnicas que hemos utilizado en nuestros experimentos. Probablemente la utilización de criome, que permite un aumento de calidad en las imágenes obtenidas, y el uso de maximum-likelihood 3D como método de alineamiento y clasificación, posibiliten la determinación de la existencia o no de la forma doble heptamérica.

El doble hexámero es una especie que se forma en solución fácilmente, tanto en ausencia como en presencia de DNA. Por otro lado, no hemos detectado anillos doble-oligoméricos abiertos en nuestro análisis de vistas laterales. De este modo, no podemos afirmar, en experimentos *in vitro*, en ausencia de proteínas cargadoras, si la apertura y carga del anillo se produce sobre el dHex (Figura D.2, Modelo 2, V) o bien sobre hexámeros sencillos (Modelo 1, V) formándose el dHex posteriormente por unión de dos hexámeros previamente cargados, al modo de LTag.

La presencia de una proporción de proteína no acomplejada en nuestros experimentos de unión a DNA en cualquier concentración de ácido nucleico, sugiere nuevamente un equilibrio entre formas libre y acomplejada, que en la célula probablemente se encuentre regulado exquisitamente mediante proteínas cargadoras u otros cofactores. Numerosas publicaciones describen dos poblaciones coexistentes de proteínas MCM, una libre y otra unida al DNA, en el interior celular (Forsburg 2004).

Podemos resumir estas teorías en dos hipótesis: probablemente todas las formas observadas de mtMCM son estados interconvertibles, intermedios de formación del dHex y de carga en el DNA, que se transforman mediante apertura del anillo y entrada o salida de monómeros, encontrándose probablemente el estado de transición, el anillo abierto, fuertemente regulado en la célula. Todos los equilibrios se encuentran desplazados hacia la forma más estable, el dHex, probablemente la forma funcional.

Modelo 1

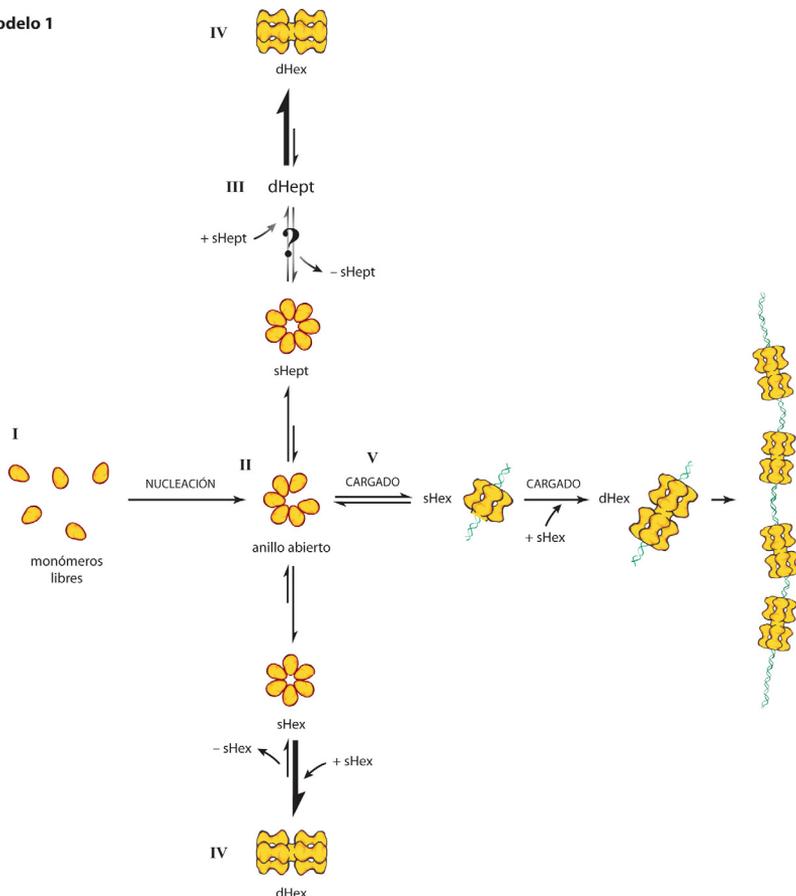


Figura D.2: Dos modelos diferentes de equilibrios de oligomerización y carga en el DNA de mtMCM.

A partir de los monómeros libres se produce la nucleación espontánea de la proteína (I) en oligómeros anulares de 6 ó 7 subunidades, por medio de un estado de transición en anillo abierto (II). Los anillos sencillos formados acomplejan en anillos dobles mayoritariamente hexaméricos (IV). En nuestra experimentación no hemos encontrado anillos dobles heptaméricos (III), de modo que no podemos asegurar su existencia. El anillo abierto es la forma de carga en el DNA (V) que, aunque no tenemos certeza de en qué forma ocurre, deducimos que se puede producir como dHex (modelo 2) o bien como dos anillos sencillos que luego se unen en uno doble (modelo 1), tal como ocurre con LTag.

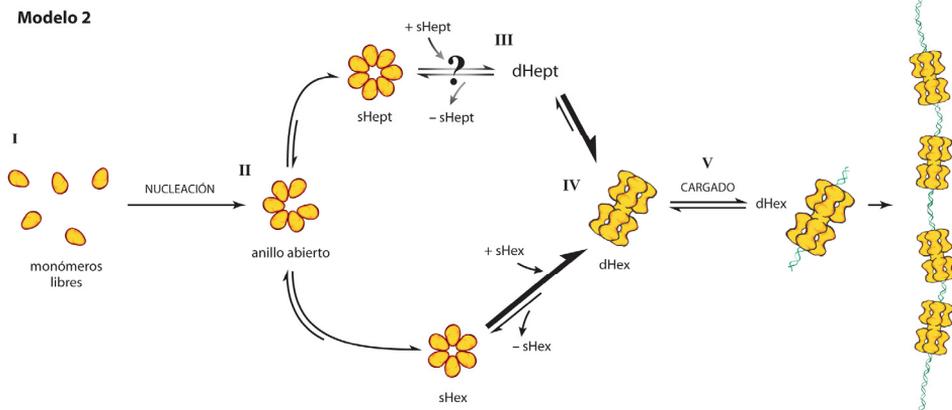


Figura D.2 – Continuación.

Conclusiones

- 1.** FL-mtMCM oligomeriza en varios estados oligoméricos coexistentes en multitud de condiciones bioquímicas. Anillos con simetría 6, con simetría 7, dobles anillos y anillos abiertos coexisten en ausencia o presencia de ATP, AMPPNP, Mg^{2+} , diferentes concentraciones de NaCl y diferentes temperaturas.
- 2.** La presencia de anillos abiertos sugiere que la forma de carga de mtMCM en el DNA es mediante el mecanismo de “apertura del anillo”.
- 3.** Las formas doble-oligoméricas son mayoritariamente dobles hexámeros:
 - a.** Cada doble hexámero de mtMCM está constituido por dos hexámeros enfrentados por sus extremos N-terminales, del mismo modo que en la estructura resuelta para el fragmento N-terminal.
 - b.** El doble hexámero es atravesado longitudinalmente por un canal de más de 2,5 nm de diámetro, capaz de acomodar una molécula de dsDNA.
 - c.** La región C-terminal se encuentra peor definida que la N-terminal, posiblemente debida a una mayor flexibilidad.
 - d.** La molécula muestra ventanas laterales intermonoméricas que sugieren un mecanismo de translocación del DNA similar al propuesto para LTag de SV40, en el que la

molécula de dsDNA entra por el canal central y sale como ssDNA por las ventanas laterales.

- e.** El doble hexámero presenta un eje de giro helicoidal de modo que sus cuatro regiones cuaternarias (C-term-1, N-term-1, N-term-2 y C-term-2) rotan ligeramente a favor de las agujas del reloj a lo largo del eje longitudinal del oligómero.
- 4.** Mutaciones en la arginina 161 del fragmento N-mtMCM impiden la formación del doble hexámero. Este hecho correlaciona con una menor actividad helicasa, reforzando la hipótesis de que la forma activa es la doble hexamérica.
- 5.** Se ha comprobado que en un rango muy amplio de condiciones de unión a DNA existe una proporción elevada de mtMCM libre y por tanto es necesario recurrir a técnicas de marcaje para distinguir las moléculas de MCM unidas a DNA.
- 6.** La tinción negativa no es la técnica adecuada para identificar la presencia de DNA en el canal central de mtMCM debido a la tendencia de esta proteína a sufrir artefactos de tinción en dicha región.

Glosario

- **3D**: Tridimensional
- **Aquiral**: Objeto o imagen no-quiral (ver Quiralidad).
- **crioME**: Crio-microscopía electrónica.
- **C-terminal**: COOH-terminal o carboxilo terminal, los aminoácidos más próximos de la secuencia polipeptídica de una proteína a su extremo C-terminal, en su último aminoácido, éste incluido, por oposición a N-terminal.
- **h2i**: Helix-2 insert, inserción en la hélice 2. Familia de proteínas AAA+ caracterizada por una inserción β - α - β en el interior de la hélice α 2 del dominio P-loop.
- **HPLC**: High performance liquid chromatography, cromatografía líquida de alta resolución.
- **HTH**: Helix-turn-helix, hélice-giro-hélice, motivo estructural de unión a DNA muy común en proteínas reguladoras de la expresión génica.
- **kb**: kilo-pares de bases, equivalente a 1000 pb.
- **kDa**: kilo-Dalton, equivalente a 1000 unidades de masa atómica.
- **KerDenSOM**: Kernel probability density estimator self-organizing maps, mapa auto-organizativo estimador de la probabilidad de densidad del núcleo.
- **ME**: Microscopía electrónica.
- **N-terminal**: NH₃-terminal o amino terminal, los aminoácidos más próximos de la secuencia polipeptídica de una proteína a su extremo N-terminal, en su primer aminoácido, éste incluido, por oposición a C-terminal.
- **pb**: par de bases, unidad base de un ácido nucleico de doble cadena.
- **PDB**: Protein data bank, base de datos de proteínas (www.pdb.org).
- **pix**: Píxel, unidad base de una imagen digital.
- **Quiralidad**: Propiedad de un objeto de no ser superponible con su imagen especular.
- **TN**: Tinción negativa.
- **WT**: Wild type, fenotipo “silvestre”, es decir, tal y como se encuentra en la naturaleza, sin mutaciones.

Especies:

- ***E. coli***: *Escherichia coli*.
- ***M. thermoautotrophicum***: *Methanobacterium thermoautotrophicum*.
- ***S. cerevisiae***: *Saccharomyces cerevisiae*.
- ***S. pombe***: *Schizosaccharomyces pombe*.
- **SV40**: *Simian-virus 40*, virus número 40 de mono.
- ***X. laevis***: *Xenopus laevis*.

Proteínas:

- **6Ins**: Mutante de inserción de N-mtMCM: 6 aminoácidos se insertan entre la cisteína 158 y la glicina 159.
- **AAA**: *ATPases associated with various cellular activities*, ATPasas asociadas con diferentes actividades celulares.
- **CDK**: *Cyclin-dependent kinase*, proteínas quinasas dependientes de ciclinas.
- **DDK**: *Dbf4-dependent kinase*, proteína quinasa dependiente de la proteína Dbf4.
- **EcoRI**: Enzima de restricción de *E. coli*, cepa RY13.
- **LTag**: *Large T-antigen*, antígeno T grande.
- **MCM**: *Mini-chromosome maintenance proteins*, proteínas de mantenimiento del minicromosoma.
- **mtMCM**: MCM de *M. thermoautotrophicum*.
- **N-mtMCM**: Fragmento N-terminal de mtMCM (aminoácidos 2-286).
- **ORC**: *Origin binding complex*, complejo de reconocimiento del origen.
- **R161A**: Mutante de sustitución de N-mtMCM: Arg161 se sustituye por alanina.
- **XmnI**: Enzima de restricción de *Xanthomonas manihotis*.

Compuestos químicos / Tampones:

- **AMPPNP**: *Adenylylimidodiphosphate*, análogo no hidrolizable de ATP.
- **ATP**: *Adenosine triphosphate*, trifosfato de adenosina o adenosín trifosfato
- **γ -³²P-ATP**: ATP marcado radiactivamente con fósforo³² en el tercer fosfato (γ).
- **bio**: Biotina.
- **BrEt**: Bromuro de etidio.
- **DNA**: *Deoxyribonucleic acid*, ácido desoxirribonucleico (ADN).
- **dsDNA**: *Double-stranded DNA*, DNA de doble cadena.

- **DTT:** Ditiotreitol.
- **EDTA:** Ethylenediamine tetraacetic acid, ácido etilendiaminotetraacético (AEDT)
- **Glutation:** También GSH, es un tripéptido, tres aminoácidos unidos covalentemente mediante enlaces peptídicos: Glu-Gly-Cys.
- **NTP:** Nucleoside triphosphate, trifosfato de un nucleósido, es decir cualquiera de los siguientes: ATP, GTP, TTP o/y CTP.
- **poly-T:** Oligonucleótido de secuencia (T)_n.
- **RNA:** Ribonucleic acid, ácido ribonucleico (ARN).
- **Ser** (ó S): Serina, un aminoácido.
- **ssDNA:** Single-stranded DNA, DNA de cadena sencilla.
- **strep:** Estreptavidina.
- **strep-Au:** Estreptavidina-oro.
- **Tampon SSC 20x:** (Saline sodium-citrate) NaCl 3 M, citrato sódico 0,3 M, pH 7,0
- **Tampon TAE:** Tris-acetato 40 mM pH ~8,5, EDTA 2 mM.
- **Tampon TE:** Tris-Cl 10 mM pH 8,0, EDTA 1 mM.
- **Tampon TTBS:** (Tris-Tween buffer saline) Tris-Cl 0,1 M pH 7,5, NaCl 150 mM, Tween 20 0,1% (v/v)
- **Tris:** Tris(hidroximetil)aminometano.

Direcciones web de proveedores de productos, software y bases de datos:

- Amira: www.amiravis.com
- British Biocell International: www.bbigold.com
- EMAN: <http://blake.bcm.tmc.edu/eman/>
- Fei Company: www.fei.com
- Gatan, Inc.: www.gatan.com
- General Electric Healthcare; GE: www.gelifesciences.com
- Intergraph Corporation: www.intergraph.com
- Invitrogen Corporation: www.invitrogen.com
- Jeol Ltd.: www.jeol.com
- Kodak; Eastman Kodak Company: www.kodak.com
- Millipore Corporation: www.millipore.com
- Nanodrop Technologies: www.nanodrop.com
- PDB, RCSB Protein Data Bank: www.pdb.org
- Promega Corporation: www.promega.com
- Qiagen: www.qiagen.com
- Quantifoil Micro Tools GmbH: www.quantifoil.com
- Roche Diagnostics Corporation, Applied Science Dept.: www.roche-applied-science.com
- Situs: <http://situs.biomachina.org/fguide.html>

- SPIDER, *S*ystem for *P*rocessing *I*mage *D*ata from *E*lectron *m*icroscopy and *R*elated fields:
http://www.wadsworth.org/spider_doc/spider/docs/spider.html
- Xmipp, *X*-*W*indow-based *M*icroscopy *I*mage *P*rocessing *P*ackage:
<http://xmipp.cnb.csic.es/>
- Zeiss; Carl Zeiss, Inc.: www.zeiss.com

Bibliografía

- Adachi, Y., Usukura, J. and Yanagida, M. (1997). A globular complex formation by Nda1 and the other five members of the MCM protein family in fission yeast. *Genes Cells* **2**(7): 467-479.
- Ahnert, P., Picha, K.M. and Patel, S.S. (2000). A ring-opening mechanism for DNA binding in the central channel of the T7 helicase-primase protein. *Embo J* **19**(13): 3418-3427.
- Aparicio, O.M., Weinstein, D.M. and Bell, S.P. (1997). Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell* **91**(1): 59-69.
- Aravind, L. and Koonin, E.V. (1999). DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Res* **27**(23): 4658-4670.
- Arthur, A.K., Hoss, A. and Fanning, E. (1988). Expression of simian virus 40 T antigen in *Escherichia coli*: localization of T-antigen origin DNA-binding domain to within 129 amino acids. *J Virol* **62**(6): 1999-2006.
- Bailey, S., Eliason, W.K. and Steitz, T.A. (2007). The crystal structure of the *Thermus aquaticus* DnaB helicase monomer. *Nucleic Acids Res* **35**(14): 4728-4736.
- Barcena, M., Martin, C.S., Weise, F., Ayora, S., Alonso, J.C. and Carazo, J.M. (1998). Polymorphic quaternary organization of the *Bacillus subtilis* bacteriophage SPP1 replicative helicase (G40 P). *J Mol Biol* **283**(4): 809-819.
- Barry, E.R. and Bell, S.D. (2006). DNA replication in the archaea. *Microbiol Mol Biol Rev* **70**(4): 876-887.
- Barry, E.R., McGeoch, A.T., Kelman, Z. and Bell, S.D. (2007). Archaeal MCM has separable processivity, substrate choice and helicase domains. *Nucleic Acids Res* **35**(3): 988-998.
- Bell, S.P. and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**(6374): 128-134.
- Bell, S.P. (1995). Eukaryotic replicators and associated protein complexes. *Curr Opin Genet Dev* **5**(2): 162-167.
- Bell, S.P. and Dutta, A. (2002). DNA replication in eukaryotic cells. *Annu Rev Biochem* **71**: 333-374.
- Berneburg, M. and Lehmann, A.R. (2001). Xeroderma pigmentosum and related disorders: defects in DNA repair and transcription. *Adv Genet* **43**: 71-102.

- Blow, J.J. (2001). Control of chromosomal DNA replication in the early *Xenopus* embryo. *Embo J* **20**(13): 3293-3297.
- Bradshaw, E.M., Sanford, D.G., Luo, X., Sudmeier, J.L., Gurard-Levin, Z.A., Bullock, P.A. and Bachovchin, W.W. (2004). T antigen origin-binding domain of simian virus 40: determinants of specific DNA binding. *Biochemistry* **43**(22): 6928-6936.
- Brennan, R.G. and Matthews, B.W. (1989). The helix-turn-helix DNA binding motif. *J Biol Chem* **264**(4): 1903-1906.
- Bujalowski, W. and Jezewska, M.J. (2000). Kinetic mechanism of the single-stranded DNA recognition by *Escherichia coli* replicative helicase DnaB protein. Application of the matrix projection operator technique to analyze stopped-flow kinetics. *J Mol Biol* **295**(4): 831-852.
- Carpentieri, F., De Felice, M., De Falco, M., Rossi, M. and Pisani, F.M. (2002). Physical and functional interaction between the mini-chromosome maintenance-like DNA helicase and the single-stranded DNA binding protein from the crenarchaeon *Sulfolobus solfataricus*. *J Biol Chem* **277**(14): 12118-12127.
- Chacon, P. and Wriggers, W. (2002). Multi-resolution contour-based fitting of macromolecular structures. *J Mol Biol* **317**(3): 375-384.
- Chen, Y.J., Yu, X., Kasiviswanathan, R., Shin, J.H., Kelman, Z. and Egelman, E.H. (2005). Structural polymorphism of *Methanothermobacter thermautotrophicus* MCM. *J Mol Biol* **346**(2): 389-394.
- Chong, J.P., Hayashi, M.K., Simon, M.N., Xu, R.M. and Stillman, B. (2000). A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc Natl Acad Sci U S A* **97**(4): 1530-1535.
- Christensen, T.W. and Tye, B.K. (2003). *Drosophila* MCM10 interacts with members of the prereplication complex and is required for proper chromosome condensation. *Mol Biol Cell* **14**(6): 2206-2215.
- Clyne, R.K. and Kelly, T.J. (1995). Genetic analysis of an ARS element from the fission yeast *Schizosaccharomyces pombe*. *Embo J* **14**(24): 6348-6357.
- Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. and Onesti, S. (2006a). Structural basis of the *Methanothermobacter thermautotrophicus* MCM helicase activity. *Nucleic Acids Res* **34**(20): 5829-5838.
- Costa, A., Pape, T., van Heel, M., Brick, P., Patwardhan, A. and Onesti, S. (2006b). Structural studies of the archaeal MCM complex in different functional states. *J Struct Biol* **156**(1): 210-219.
- Coue, M., Amariglio, F., Maiorano, D., Bocquet, S. and Mechali, M. (1998). Evidence for different MCM subcomplexes with differential binding to chromatin in *Xenopus*. *Exp Cell Res* **245**(2): 282-289.

- Crampton, D.J., Ohi, M., Qimron, U., Walz, T. and Richardson, C.C. (2006). Oligomeric states of bacteriophage T7 gene 4 primase/helicase. *J Mol Biol* **360**(3): 667-677.
- Crevel, G., Ivetic, A., Ohno, K., Yamaguchi, M. and Cotterill, S. (2001). Nearest neighbour analysis of MCM protein complexes in *Drosophila melanogaster*. *Nucleic Acids Res* **29**(23): 4834-4842.
- Crowther, R.A., Amos, L.A., Finch, J.T., De Rosier, D.J. and Klug, A. (1970). Three dimensional reconstructions of spherical viruses by fourier synthesis from electron micrographs. *Nature* **226**(5244): 421-425.
- Crowther, R.A. and Amos, L.A. (1971). Harmonic analysis of electron microscope images with rotational symmetry. *J Mol Biol* **60**(1): 123-130.
- Davey, M.J., Indiani, C. and O'Donnell, M. (2003). Reconstitution of the Mcm2-7p heterohexamer, subunit arrangement, and ATP site architecture. *J Biol Chem* **278**(7): 4491-4499.
- Dean, F.B., Dodson, M., Echols, H. and Hurwitz, J. (1987). ATP-dependent formation of a specialized nucleoprotein structure by simian virus 40 (SV40) large tumor antigen at the SV40 replication origin. *Proc Natl Acad Sci U S A* **84**(24): 8981-8985.
- Dean, F.B., Borowiec, J.A., Eki, T. and Hurwitz, J. (1992). The simian virus 40 T antigen double hexamer assembles around the DNA at the replication origin. *J Biol Chem* **267**(20): 14129-14137.
- Delagoutte, E. and von Hippel, P.H. (2002). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases. *Q Rev Biophys* **35**(4): 431-478.
- Delagoutte, E. and von Hippel, P.H. (2003). Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: Integration of helicases into cellular processes. *Q Rev Biophys* **36**(1): 1-69.
- Diffley, J.F. and Cocker, J.H. (1992). Protein-DNA interactions at a yeast replication origin. *Nature* **357**(6374): 169-172.
- Dillingham, M.S. (2006). Replicative helicases: a staircase with a twist. *Curr Biol* **16**(19): R844-847.
- Dimitrova, D.S., Todorov, I.T., Melendy, T. and Gilbert, D.M. (1999). Mcm2, but not RPA, is a component of the mammalian early G1-phase prereplication complex. *J Cell Biol* **146**(4): 709-722.
- Dubey, D.D., Kim, S.M., Todorov, I.T. and Huberman, J.A. (1996). Large, complex modular structure of a fission yeast DNA replication origin. *Curr Biol* **6**(4): 467-473.
- Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowell, A.W. and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* **21**(2): 129-228.

- Dutta, A. and Bell, S.P. (1997). Initiation of DNA replication in eukaryotic cells. *Annu Rev Cell Dev Biol* **13**: 293-332.
- Edwards, M.C., Tutter, A.V., Cvetic, C., Gilbert, C.H., Prokhorova, T.A. and Walter, J.C. (2002). MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in *Xenopus* egg extracts. *J Biol Chem* **277**(36): 33049-33057.
- Egelman, E.H., Yu, X., Wild, R., Hingorani, M.M. and Patel, S.S. (1995). Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases. *Proc Natl Acad Sci U S A* **92**(9): 3869-3873.
- Egelman, E.H. (2001). Structural biology. Pumping DNA. *Nature* **409**(6820): 573, 575.
- Ellis, N.A., Groden, J., Ye, T.Z., Straughen, J., Lennon, D.J., Ciocchi, S., Proytcheva, M. and German, J. (1995). The Bloom's syndrome gene product is homologous to RecQ helicases. *Cell* **83**(4): 655-666.
- Ellis, N.A. (1997). DNA helicases in inherited human disorders. *Curr Opin Genet Dev* **7**(3): 354-363.
- Enemark, E.J., Chen, G., Vaughn, D.E., Stenlund, A. and Joshua-Tor, L. (2000). Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. *Mol Cell* **6**(1): 149-158.
- Enemark, E.J. and Joshua-Tor, L. (2006). Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442**(7100): 270-275.
- Fanning, E. and Knippers, R. (1992). Structure and function of simian virus 40 large tumor antigen. *Annu Rev Biochem* **61**: 55-85.
- Fletcher, R.J., Bishop, B.E., Leon, R.P., Sclafani, R.A., Ogata, C.M. and Chen, X.S. (2003). The structure and function of MCM from archaeal *M. Thermoautotrophicum*. *Nat Struct Biol* **10**(3): 160-167.
- Fletcher, R.J., Shen, J., Gomez-Llorente, Y., Martin, C.S., Carazo, J.M. and Chen, X.S. (2005). Double hexamer disruption and biochemical activities of *Methanobacterium thermoautotrophicum* MCM. *J Biol Chem* **280**(51): 42405-42410.
- Fodje, M.N., Hansson, A., Hansson, M., Olsen, J.G., Gough, S., Willows, R.D. and Al-Karadaghi, S. (2001). Interplay between an AAA module and an integrin I domain may regulate the function of magnesium chelatase. *J Mol Biol* **311**(1): 111-122.
- Forsburg, S.L., Sherman, D.A., Otilie, S., Yasuda, J.R. and Hodson, J.A. (1997). Mutational analysis of Cdc19p, a *Schizosaccharomyces pombe* MCM protein. *Genetics* **147**(3): 1025-1041.
- Forsburg, S.L. (2004). Eukaryotic MCM proteins: beyond replication initiation. *Microbiol Mol Biol Rev* **68**(1): 109-131.

- Foury, F. and Lahaye, A. (1987). Cloning and sequencing of the PIF gene involved in repair and recombination of yeast mitochondrial DNA. *Embo J* **6**(5): 1441-1449.
- Fox, C.A., Loo, S., Dillin, A. and Rine, J. (1995). The origin recognition complex has essential functions in transcriptional silencing and chromosomal replication. *Genes Dev* **9**(8): 911-924.
- Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**(1): 190-199.
- Frank, J. (2002). Single-particle imaging of macromolecules by cryo-electron microscopy. *Annu Rev Biophys Biomol Struct* **31**: 303-319.
- Frank, J. (2006a). Chapter 5: Three-Dimensional Reconstruction. 7.6: Model Dependence of the Reconstruction. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Oxford, N.Y., Oxford University Press: 246-247.
- Frank, J. (2006b). Chapter 2: Specimen Preparation Methods. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Oxford, N.Y., Oxford University Press: 246-247.
- Frank, J. (2006c). Chapter 5: Three-Dimensional Reconstruction. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Oxford, N.Y., Oxford University Press: 246-247.
- Frickey, T. and Lupas, A.N. (2004). Phylogenetic analysis of AAA proteins. *J Struct Biol* **146**(1-2): 2-10.
- Fujita, M., Yamada, C., Tsurumi, T., Hanaoka, F., Matsuzawa, K. and Inagaki, M. (1998). Cell cycle- and chromatin binding state-dependent phosphorylation of human MCM heterohexameric complexes. A role for cdc2 kinase. *J Biol Chem* **273**(27): 17095-17101.
- Fujita, M., Yamada, C., Goto, H., Yokoyama, N., Kuzushima, K., Inagaki, M. and Tsurumi, T. (1999). Cell cycle regulation of human CDC6 protein. Intracellular localization, interaction with the human mcm complex, and CDC2 kinase-mediated hyperphosphorylation. *J Biol Chem* **274**(36): 25927-25932.
- Gai, D., Zhao, R., Li, D., Finkielstein, C.V. and Chen, X.S. (2004). Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* **119**(1): 47-60.
- Gangloff, S., Soustelle, C. and Fabre, F. (2000). Homologous recombination is responsible for cell death in the absence of the Sgs1 and Srs2 helicases. *Nat Genet* **25**(2): 192-194.
- Goldstein, E.J. (1987). Norfloxacin, a fluoroquinolone antibacterial agent. Classification, mechanism of action, and in vitro activity. *Am J Med* **82**(6B): 3-17.

- Gómez-Llorente, Y., Fletcher, R.J., Chen, X.S., Carazo, J.M. and San Martín, C. (2005). Polymorphism and double hexamer structure in the archaeal minichromosome maintenance (MCM) helicase from *Methanobacterium thermoautotrophicum*. *J Biol Chem* **280**(49): 40909-40915.
- Gomez-Lorenzo, M.G., Valle, M., Frank, J., Gruss, C., Sorzano, C.O., Chen, X.S., Donate, L.E. and Carazo, J.M. (2003). Large T antigen on the simian virus 40 origin of replication: a 3D snapshot prior to DNA replication. *Embo J* **22**(23): 6205-6213.
- Gorbalenya, A.E. and Koonin, E.V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. *Curr Opin Struct Biol* **3**(3): 419-429.
- Gozuacik, D., Chami, M., Lagorce, D., Faivre, J., Murakami, Y., Poch, O., Biermann, E., Knippers, R., Brechot, C. and Paterlini-Brechot, P. (2003). Identification and functional characterization of a new member of the human Mcm protein family: hMcm8. *Nucleic Acids Res* **31**(2): 570-579.
- Grainge, I., Scaife, S. and Wigley, D.B. (2003). Biochemical analysis of components of the pre-replication complex of *Archaeoglobus fulgidus*. *Nucleic Acids Res* **31**(16): 4888-4898.
- Gray, M.D., Shen, J.C., Kamath-Loeb, A.S., Blank, A., Sopher, B.L., Martin, G.M., Oshima, J. and Loeb, L.A. (1997). The Werner syndrome protein is a DNA helicase. *Nat Genet* **17**(1): 100-103.
- Guenther, B., Onrust, R., Sali, A., O'Donnell, M. and Kuriyan, J. (1997). Crystal structure of the delta' subunit of the clamp-loader complex of *E. coli* DNA polymerase III. *Cell* **91**(3): 335-345.
- Henderson, R. (2004). Realizing the potential of electron cryo-microscopy. *Q Rev Biophys* **37**(1): 3-13.
- Hennessy, K.M., Clark, C.D. and Botstein, D. (1990). Subcellular localization of yeast CDC46 varies with the cell cycle. *Genes Dev* **4**(12B): 2252-2263.
- Heymann, J.B. (2001). Bsoft: image and molecular processing in electron microscopy. *J Struct Biol* **133**(2-3): 156-169.
- Hilliard, J.J., Krause, H.M., Bernstein, J.I., Fernandez, J.A., Nguyen, V., Ohemeng, K.A. and Barrett, J.F. (1995). A comparison of active site binding of 4-quinolones and novel flavone gyrase inhibitors to DNA gyrase. *Adv Exp Med Biol* **390**: 59-69.
- Hofman-Bang, J., Lange, M., Conway de Macario, E., Macario, A.J. and Ahring, B.K. (1999). The genes coding for the hsp70 (dnaK) molecular chaperone machine occur in the moderate thermophilic archaeon *Methanosarcina thermophila* TM-1. *Gene* **238**(2): 387-395.
- Hopwood, B. and Dalton, S. (1996). Cdc45p assembles into a complex with Cdc46p/Mcm5p, is required for minichromosome maintenance,

- and is essential for chromosomal DNA replication. *Proc Natl Acad Sci U S A* **93**(22): 12309-12314.
- Hsiang, Y.H., Hertzberg, R., Hecht, S. and Liu, L.F. (1985). Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I. *J Biol Chem* **260**(27): 14873-14878.
- Hyrien, O., Marheineke, K. and Goldar, A. (2003). Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. *Bioessays* **25**(2): 116-125.
- Ishimi, Y. (1997). A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J Biol Chem* **272**(39): 24508-24513.
- Ishimi, Y., Komamura, Y., You, Z. and Kimura, H. (1998). Biochemical function of mouse minichromosome maintenance 2 protein. *J Biol Chem* **273**(14): 8369-8375.
- Ishimi, Y., Komamura-Kohno, Y., Arai, K. and Masai, H. (2001). Biochemical activities associated with mouse Mcm2 protein. *J Biol Chem* **276**(46): 42744-42752.
- Ivessa, A.S., Zhou, J.Q. and Zakian, V.A. (2000). The *Saccharomyces* Pif1p DNA helicase and the highly related Rrm3p have opposite effects on replication fork progression in ribosomal DNA. *Cell* **100**(4): 479-489.
- Iyer, L.M., Leipe, D.D., Koonin, E.V. and Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* **146**(1-2): 11-31.
- Jenkinson, E.R. and Chong, J.P. (2006). Minichromosome maintenance helicase activity is controlled by N- and C-terminal motifs and requires the ATPase domain helix-2 insert. *Proc Natl Acad Sci U S A* **103**(20): 7613-7618.
- Jezewska, M.J., Rajendran, S., Bujalowska, D. and Bujalowski, W. (1998). Does single-stranded DNA pass through the inner channel of the protein hexamer in the complex with the *Escherichia coli* DnaB Helicase? Fluorescence energy transfer studies. *J Biol Chem* **273**(17): 10515-10529.
- Jiang, W., Wells, N.J. and Hunter, T. (1999). Multistep regulation of DNA replication by Cdk phosphorylation of HsCdc6. *Proc Natl Acad Sci U S A* **96**(11): 6193-6198.
- Jonic, S., Sorzano, C.O., Thevenaz, P., El-Bez, C., De Carlo, S. and Unser, M. (2005). Spline-based image-to-volume registration for three-dimensional electron microscopy. *Ultramicroscopy* **103**(4): 303-317.
- Kamimura, Y., Masumoto, H., Sugino, A. and Araki, H. (1998). Sld2, which interacts with Dpb11 in *Saccharomyces cerevisiae*, is required for chromosomal DNA replication. *Mol Cell Biol* **18**(10): 6102-6109.

- Kaneko, H. and Kondo, N. (2004). Clinical features of Bloom syndrome and function of the causative gene, BLM helicase. *Expert Rev Mol Diagn* **4**(3): 393-401.
- Kaplan, D.L. (2000). The 3'-tail of a forked-duplex sterically determines whether one or two DNA strands pass through the central channel of a replication-fork helicase. *J Mol Biol* **301**(2): 285-299.
- Kaplan, D.L. and O'Donnell, M. (2002). DnaB drives DNA branch migration and dislodges proteins while encircling two DNA strands. *Mol Cell* **10**(3): 647-657.
- Kaplan, D.L., Davey, M.J. and O'Donnell, M. (2003). Mcm4,6,7 uses a "pump in ring" mechanism to unwind DNA by steric exclusion and actively translocate along a duplex. *J Biol Chem* **278**(49): 49171-49182.
- Kasiviswanathan, R., Shin, J.H., Melamud, E. and Kelman, Z. (2004). Biochemical characterization of the Methanothermobacter thermoautotrophicus minichromosome maintenance (MCM) helicase N-terminal domains. *J Biol Chem* **279**(27): 28358-28366.
- Kearsey, S.E. and Labib, K. (1998). MCM proteins: evolution, properties, and role in DNA replication. *Biochim Biophys Acta* **1398**(2): 113-136.
- Kelman, Z., Lee, J.K. and Hurwitz, J. (1999). The single minichromosome maintenance protein of Methanobacterium thermoautotrophicum DeltaH contains DNA helicase activity. *Proc Natl Acad Sci U S A* **96**(26): 14783-14788.
- Kelman, Z. and White, M.F. (2005). Archaeal DNA replication and repair. *Curr Opin Microbiol* **8**(6): 669-676.
- Kim, H.Y., Barbaro, B.A., Joo, W.S., Prack, A.E., Sreekumar, K.R. and Bullock, P.A. (1999). Sequence requirements for the assembly of simian virus 40 T antigen and the T-antigen origin binding domain on the viral core origin of replication. *J Virol* **73**(9): 7543-7555.
- Kimura, H., Ohtomo, T., Yamaguchi, M., Ishii, A. and Sugimoto, K. (1996). Mouse MCM proteins: complex formation and transportation to the nucleus. *Genes Cells* **1**(11): 977-993.
- Kitao, S., Shimamoto, A., Goto, M., Miller, R.W., Smithson, W.A., Lindor, N.M. and Furuichi, Y. (1999). Mutations in RECQL4 cause a subset of cases of Rothmund-Thomson syndrome. *Nat Genet* **22**(1): 82-84.
- Klemm, R.D., Austin, R.J. and Bell, S.P. (1997). Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell* **88**(4): 493-502.
- Kleywegt, G.J., Zou, J.Y., Kjeldgaard, M. and Jones, T.A. (2001). International Tables for Crystallography: Crystallography of Biological Macromolecules. Rossmann, M.G. and Arnold, E.

- Dordrecht, The Netherlands, Kluwer Academic Publishers. **Vol. F:** 353-356.
- Koonin, E.V. (1993). A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res* **21**(11): 2541-2547.
- Korolev, S., Hsieh, J., Gauss, G.H., Lohman, T.M. and Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**(4): 635-647.
- Kubota, Y., Takase, Y., Komori, Y., Hashimoto, Y., Arata, T., Kamimura, Y., Araki, H. and Takisawa, H. (2003). A novel ring-like complex of Xenopus proteins essential for the initiation of DNA replication. *Genes Dev* **17**(9): 1141-1152.
- Labib, K., Tercero, J.A. and Diffley, J.F. (2000). Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science* **288**(5471): 1643-1647.
- Labib, K. and Diffley, J.F. (2001). Is the MCM2-7 complex the eukaryotic DNA replication fork helicase? *Curr Opin Genet Dev* **11**(1): 64-70.
- Laity, J.H., Lee, B.M. and Wright, P.E. (2001). Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol* **11**(1): 39-46.
- Laskey, R.A. and Madine, M.A. (2003). A rotary pumping model for helicase function of MCM proteins at a distance from replication forks. *EMBO Rep* **4**(1): 26-30.
- Lee, J.K. and Hurwitz, J. (2000). Isolation and characterization of various complexes of the minichromosome maintenance proteins of *Schizosaccharomyces pombe*. *J Biol Chem* **275**(25): 18871-18878.
- Lee, J.K. and Hurwitz, J. (2001). Processive DNA helicase activity of the minichromosome maintenance proteins 4, 6, and 7 complex requires forked DNA structures. *Proc Natl Acad Sci U S A* **98**(1): 54-59.
- Lei, M., Kawasaki, Y. and Tye, B.K. (1996). Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in *Saccharomyces cerevisiae*. *Mol Cell Biol* **16**(9): 5081-5090.
- Lei, M., Cheng, I.H., Roberts, L.A., McAlear, M.A. and Tye, B.K. (2002). Two mcm3 mutations affect different steps in the initiation of DNA replication. *J Biol Chem* **277**(34): 30824-30831.
- Lepault, J., Booy, F.P. and Dubochet, J. (1983). Electron microscopy of frozen biological suspensions. *J Microsc* **129**(Pt 1): 89-102.
- Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E., Jochimiak, A., Szakonyi, G. and Chen, X.S. (2003). Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature* **423**(6939): 512-518.

- Liu, J., Smith, C.L., DeRyckere, D., DeAngelis, K., Martin, G.S. and Berger, J.M. (2000). Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol Cell* **6**(3): 637-648.
- Ludtke, S.J., Baldwin, P.R. and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**(1): 82-97.
- Lundgren, M., Andersson, A., Chen, L., Nilsson, P. and Bernander, R. (2004). Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci U S A* **101**(18): 7046-7051.
- Macris, M.A., Krejci, L., Bussen, W., Shimamoto, A. and Sung, P. (2006). Biochemical characterization of the RECQ4 protein, mutated in Rothmund-Thomson syndrome. *DNA Repair (Amst)* **5**(2): 172-180.
- Madine, M.A., Khoo, C.Y., Mills, A.D. and Laskey, R.A. (1995). MCM3 complex required for cell cycle regulation of DNA replication in vertebrate cells. *Nature* **375**(6530): 421-424.
- Mahbubani, H.M., Chong, J.P., Chevalier, S., Thommes, P. and Blow, J.J. (1997). Cell cycle regulation of the replication licensing system: involvement of a Cdk-dependent inhibitor. *J Cell Biol* **136**(1): 125-135.
- Maiorano, D., Lutzmann, M. and Mechali, M. (2006). MCM proteins and DNA replication. *Curr Opin Cell Biol* **18**(2): 130-136.
- Marabini, R., Masegosa, I.M., San Martin, M.C., Marco, S., Fernandez, J.J., de la Fraga, L.G., Vaquerizo, C. and Carazo, J.M. (1996). Xmipp: An Image Processing Package for Electron Microscopy. *J Struct Biol* **116**(1): 237-240.
- Marabini, R., Herman, G.T. and Carazo, J.M. (1998). 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). *Ultramicroscopy* **72**(1-2): 53-65.
- Marco, S., Chagoyen, M., de la Fraga, L.G., Carazo, J.M. and Carrascosa, J.L. (1996). A variant to the "random approximation" of the reference-free alignment algorithm. *Ultramicroscopy* **66**(1-2): 5-10.
- Mastrangelo, I.A., Hough, P.V., Wall, J.S., Dodson, M., Dean, F.B. and Hurwitz, J. (1989). ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature* **338**(6217): 658-662.
- Masuda, T., Mimura, S. and Takisawa, H. (2003). CDK- and Cdc45-dependent priming of the MCM complex on chromatin during S-phase in *Xenopus* egg extracts: possible activation of MCM helicase by association with Cdc45. *Genes Cells* **8**(2): 145-161.
- Matsunaga, F., Forterre, P., Ishino, Y. and Mylykallio, H. (2001). In vivo interactions of archaeal Cdc6/Orc1 and minichromosome

- maintenance proteins with the replication origin. *Proc Natl Acad Sci U S A* **98**(20): 11152-11157.
- McGeoch, A.T., Trakselis, M.A., Laskey, R.A. and Bell, S.D. (2005). Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat Struct Mol Biol* **12**(9): 756-762.
- Mendez, J. and Stillman, B. (2003). Perpetuating the double helix: molecular machines at eukaryotic DNA replication origins. *Bioessays* **25**(12): 1158-1167.
- Milner-White, E.J., Coggins, J.R. and Anton, I.A. (1991). Evidence for an ancestral core structure in nucleotide-binding proteins with the type A motif. *J Mol Biol* **221**(3): 751-754.
- Mimura, S., Seki, T., Tanaka, S. and Diffley, J.F. (2004). Phosphorylation-dependent binding of mitotic cyclins to Cdc6 contributes to DNA replication control. *Nature* **431**(7012): 1118-1123.
- Mindell, J.A. and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**(3): 334-347.
- Miyata, T., Yamada, K., Iwasaki, H., Shinagawa, H., Morikawa, K. and Mayanagi, K. (2000). Two different oligomeric states of the RuvB branch migration motor protein as revealed by electron microscopy. *J Struct Biol* **131**(2): 83-89.
- Mohaghegh, P., Karow, J.K., Brosh Jr, R.M., Jr., Bohr, V.A. and Hickson, I.D. (2001). The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res* **29**(13): 2843-2849.
- Myllykallio, H. and Forterre, P. (2000). Mapping of a chromosome replication origin in an archaeon: response. *Trends Microbiol* **8**(12): 537-539.
- Nasmyth, K. and Nurse, P. (1981). Cell division cycle mutants altered in DNA replication and mitosis in the fission yeast *Schizosaccharomyces pombe*. *Mol Gen Genet* **182**(1): 119-124.
- Neuwald, A.F., Aravind, L., Spouge, J.L. and Koonin, E.V. (1999). AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res* **9**(1): 27-43.
- Niedernhofer, L.J., Lalai, A.S. and Hoeijmakers, J.H. (2005). Fanconi anemia (cross)linked to DNA repair. *Cell* **123**(7): 1191-1198.
- Nissen, P., Hansen, J., Ban, N., Moore, P.B. and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**(5481): 920-930.
- Ogura, T. and Wilkinson, A.J. (2001). AAA+ superfamily ATPases: common structure--diverse function. *Genes Cells* **6**(7): 575-597.

- Oyama, T., Ishino, Y., Cann, I.K., Ishino, S. and Morikawa, K. (2001). Atomic structure of the clamp loader small subunit from *Pyrococcus furiosus*. *Mol Cell* **8**(2): 455-463.
- Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M. and Onesti, S. (2003). Hexameric ring structure of the full-length archaeal MCM protein complex. *EMBO Rep* **4**(11): 1079-1083.
- Pascual-Montano, A., Donate, L.E., Valle, M., Barcena, M., Pascual-Marqui, R.D. and Carazo, J.M. (2001). A novel neural network technique for analysis and classification of EM single-particle images. *J Struct Biol* **133**(2-3): 233-245.
- Pasion, S.G. and Forsburg, S.L. (1999). Nuclear localization of *Schizosaccharomyces pombe* Mcm2/Cdc19p requires MCM complex assembly. *Mol Biol Cell* **10**(12): 4043-4057.
- Patel, S.S. and Picha, K.M. (2000). Structure and function of hexameric helicases. *Annu Rev Biochem* **69**: 651-697.
- Patel, S.S. and Donmez, I. (2006). Mechanisms of helicases. *J Biol Chem* **281**(27): 18265-18268.
- Penczek, P., Radermacher, M. and Frank, J. (1992). Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy* **40**(1): 33-53.
- Penczek, P.A., Grassucci, R.A. and Frank, J. (1994). The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy* **53**(3): 251-270.
- Penczek, P.A., Zhu, J. and Frank, J. (1996). A common-lines based method for determining orientations for $N > 3$ particle projections simultaneously. *Ultramicroscopy* **63**(3-4): 205-218.
- Perkins, G. and Diffley, J.F. (1998). Nucleotide-dependent prereplicative complex assembly by Cdc6p, a homolog of eukaryotic and prokaryotic clamp-loaders. *Mol Cell* **2**(1): 23-32.
- Pervushin, K., Riek, R., Wider, G. and Wuthrich, K. (1997). Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* **94**(23): 12366-12371.
- Pipas, J.M. (1985). Mutations near the carboxyl terminus of the simian virus 40 large tumor antigen alter viral host range. *J Virol* **54**(2): 569-575.
- Poplawski, A., Grabowski, B., Long, S.E. and Kelman, Z. (2001). The zinc finger domain of the archaeal minichromosome maintenance protein is required for helicase activity. *J Biol Chem* **276**(52): 49371-49377.
- Radermacher, M., Wagenknecht, T., Verschoor, A. and Frank, J. (1987). Three-dimensional reconstruction from a single-exposure, random

- conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J Microsc* **146**(Pt 2): 113-136.
- Radermacher, M. (1994). Three-dimensional reconstruction from random projections: orientational alignment via Radon transforms. *Ultramicroscopy* **53**(2): 121-136.
- Randell, J.C., Bowers, J.L., Rodriguez, H.K. and Bell, S.P. (2006). Sequential ATP hydrolysis by Cdc6 and ORC directs loading of the Mcm2-7 helicase. *Mol Cell* **21**(1): 29-39.
- Reeve, J.N., Sandman, K. and Daniels, C.J. (1997). Archaeal histones, nucleosomes, and transcription initiation. *Cell* **89**(7): 999-1002.
- Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernander, R. and Bell, S.D. (2004). Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell* **116**(1): 25-38.
- Romanowski, P., Madine, M.A. and Laskey, R.A. (1996). XMCM7, a novel member of the *Xenopus* MCM family, interacts with XMCM3 and colocalizes with it throughout replication. *Proc Natl Acad Sci U S A* **93**(19): 10189-10194.
- Rusche, L.N., Kirchmaier, A.L. and Rine, J. (2002). Ordered nucleation and spreading of silenced chromatin in *Saccharomyces cerevisiae*. *Mol Biol Cell* **13**(7): 2207-2222.
- Sambrook, J. and Russell, D.W. (2001). 9.16: Phosphorylation of DNA Molecules with Protruding 5' Termini by the Exchange Reaction. *Molecular cloning: a laboratory manual*. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press. **Vol. 2**: 9.73-75.
- Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990). The P-loop—a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**(11): 430-434.
- Sawaya, M.R., Guo, S., Tabor, S., Richardson, C.C. and Ellenberger, T. (1999). Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* **99**(2): 167-177.
- Scheres, S.H., Valle, M., Nunez, R., Sorzano, C.O., Marabini, R., Herman, G.T. and Carazo, J.M. (2005a). Maximum-likelihood multi-reference refinement for electron microscopy images. *J Mol Biol* **348**(1): 139-149.
- Scheres, S.H., Valle, M. and Carazo, J.M. (2005b). Fast maximum-likelihood refinement of electron microscopy images. *Bioinformatics* **21 Suppl 2**: ii243-ii244.
- Scheres, S.H., Núñez-Ramírez, R., Gómez-Llorente, Y., San Martín, C., Eggermont, P.P. and Carazo, J.M. (2007a). Modelling experimental image formation for maximum-likelihood classification of electron microscopy data. *Structure*: (in press).

- Scheres, S.H., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P., Frank, J. and Carazo, J.M. (2007b). Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods* **4**(1): 27-29.
- Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. and Yonath, A. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* **102**(5): 615-623.
- Sclafani, R.A., Fletcher, R.J. and Chen, X.S. (2004). Two heads are better than one: regulation of DNA replication by hexameric helicases. *Genes Dev* **18**(17): 2039-2045.
- Shaikh, T.R., Hegerl, R. and Frank, J. (2003). An approach to examining model dependence in EM reconstructions using cross-validation. *J Struct Biol* **142**(2): 301-310.
- Shechter, D.F., Ying, C.Y. and Gautier, J. (2000). The intrinsic DNA helicase activity of Methanobacterium thermoautotrophicum delta H minichromosome maintenance protein. *J Biol Chem* **275**(20): 15049-15059.
- Sherman, D.A., Pasion, S.G. and Forsburg, S.L. (1998). Multiple domains of fission yeast Cdc19p (MCM2) are required for its association with the core MCM complex. *Mol Biol Cell* **9**(7): 1833-1845.
- Shin, J.H., Jiang, Y., Grabowski, B., Hurwitz, J. and Kelman, Z. (2003). Substrate requirements for duplex DNA translocation by the eukaryal and archaeal minichromosome maintenance helicases. *J Biol Chem* **278**(49): 49053-49062.
- Simmons, D.T. (2000). SV40 large T antigen functions in DNA replication and transformation. *Adv Virus Res* **55**: 75-134.
- Skordalakes, E. and Berger, J.M. (2003). Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* **114**(1): 135-146.
- Smith, D.R., Doucette-Stamm, L.A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Reeve, J.N. and et al. (1997). Complete genome sequence of Methanobacterium thermoautotrophicum deltaH: functional analysis and comparative genomics. *J Bacteriol* **179**(22): 7135-7155.
- Sorzano, C.O., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J.R., Scheres, S.H., Carazo, J.M. and Pascual-Montano, A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol* **148**(2): 194-204.
- Squire, P.G. and Himmel, M.E. (1979). Hydrodynamics and protein hydration. *Arch Biochem Biophys* **196**(1): 165-177.

- Story, R.M., Weber, I.T. and Steitz, T.A. (1992). The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355**(6358): 318-325.
- Story, R.M. and Steitz, T.A. (1992). Structure of the recA protein-ADP complex. *Nature* **355**(6358): 374-376.
- Subramanya, H.S., Bird, L.E., Brannigan, J.A. and Wigley, D.B. (1996). Crystal structure of a DExx box DNA helicase. *Nature* **384**(6607): 379-383.
- Tabancay, A.P., Jr. and Forsburg, S.L. (2006). Eukaryotic DNA replication in a chromatin context. *Curr Top Dev Biol* **76**: 129-184.
- Tada, S., Li, A., Maiorano, D., Mechali, M. and Blow, J.J. (2001). Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nat Cell Biol* **3**(2): 107-113.
- Takahashi, T.S., Wigley, D.B. and Walter, J.C. (2005). Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. *Trends Biochem Sci* **30**(8): 437-444.
- Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A. and Araki, H. (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev* **17**(9): 1153-1165.
- Tercero, J.A., Labib, K. and Diffley, J.F. (2000). DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. *Embo J* **19**(9): 2082-2093.
- Thomm, M. (1996). Archaeal transcription factors and their role in transcription initiation. *FEMS Microbiol Rev* **18**(2-3): 159-171.
- Tuteja, N. and Tuteja, R. (2004). Prokaryotic and eukaryotic DNA helicases. Essential molecular motor proteins for cellular machinery. *Eur J Biochem* **271**(10): 1835-1848.
- Tye, B.K. (1999). MCM proteins in DNA replication. *Annu Rev Biochem* **68**: 649-686.
- Tye, B.K. and Sawyer, S. (2000). The hexameric eukaryotic MCM helicase: building symmetry from nonidentical parts. *J Biol Chem* **275**(45): 34833-34836.
- Unser, M., Trus, B.L. and Steven, A.C. (1987). A new resolution criterion based on spectral signal-to-noise ratios. *Ultramicroscopy* **23**(1): 39-51.
- Unser, M., Sorzano, C.O., Thevenaz, P., Jonic, S., El-Bez, C., De Carlo, S., Conway, J.F. and Trus, B.L. (2005). Spectral signal-to-noise ratio and resolution assessment of 3D reconstructions. *J Struct Biol* **149**(3): 243-255.
- Valle, M., Gruss, C., Halmer, L., Carazo, J.M. and Donate, L.E. (2000). Large T-antigen double hexamers imaged at the simian virus 40 origin of replication. *Mol Cell Biol* **20**(1): 34-41.

- Valle, M., Chen, X.S., Donate, L.E., Fanning, E. and Carazo, J.M. (2006). Structural basis for the cooperative assembly of large T antigen on the origin of replication. *J Mol Biol* **357**(4): 1295-1305.
- van Heel, M. (1984). Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). *Ultramicroscopy* **13**(1-2): 165-183.
- Vermeulen, W., Bergmann, E., Auriol, J., Rademakers, S., Frit, P., Appeldoorn, E., Hoeijmakers, J.H. and Egly, J.M. (2000). Sublimiting concentration of TFIIH transcription/DNA repair factor causes TTD-A trichothiodystrophy disorder. *Nat Genet* **26**(3): 307-313.
- Vetter, I.R. and Wittinghofer, A. (1999). Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Q Rev Biophys* **32**(1): 1-56.
- Villard, L., Lossi, A.M., Cardoso, C., Proud, V., Chiaroni, P., Colleaux, L., Schwartz, C. and Fontes, M. (1997). Determination of the genomic structure of the XNP/ATR_X gene encoding a potential zinc finger helicase. *Genomics* **43**(2): 149-155.
- Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo J* **1**(8): 945-951.
- Walter, J. and Newport, J. (2000). Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. *Mol Cell* **5**(4): 617-627.
- Weinreich, M., Liang, C. and Stillman, B. (1999). The Cdc6p nucleotide-binding motif is required for loading mcm proteins onto chromatin. *Proc Natl Acad Sci U S A* **96**(2): 441-446.
- Whitman, W.B., Pfeifer, F., Blum, P. and Klein, A. (1999). What archaea have to tell biologists. *Genetics* **152**(4): 1245-1248.
- Williams, R.C. and Fisher, H.W. (1970). Electron microscopy of tobacco mosaic virus under conditions of minimal beam exposure. *J Mol Biol* **52**(1): 121-123.
- Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vornrhein, C., Hartsch, T. and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature* **407**(6802): 327-339.
- Wriggers, W., Milligan, R.A. and McCammon, J.A. (1999). Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. *J Struct Biol* **125**(2-3): 185-195.
- Wuthrich, K. (1998). The second decade--into the third millenium. *Nat Struct Biol* **5** **Suppl**: 492-495.
- Yang, X., Gregan, J., Lindner, K., Young, H. and Kearsy, S.E. (2005). Nuclear distribution and chromatin association of DNA polymerase

- alpha-primase is affected by TEV protease cleavage of Cdc23 (Mcm10) in fission yeast. *BMC Mol Biol* **6**: 13.
- Yao, N., Hesson, T., Cable, M., Hong, Z., Kwong, A.D., Le, H.V. and Weber, P.C. (1997). Structure of the hepatitis C virus RNA helicase domain. *Nat Struct Biol* **4**(6): 463-467.
- You, Z., Ishimi, Y., Masai, H. and Hanaoka, F. (2002). Roles of Mcm7 and Mcm4 subunits in the DNA helicase activity of the mouse Mcm4/6/7 complex. *J Biol Chem* **277**(45): 42471-42479.
- Yu, X., Jezewska, M.J., Bujalowski, W. and Egelman, E.H. (1996). The hexameric E. coli DnaB helicase can exist in different Quaternary states. *J Mol Biol* **259**(1): 7-14.
- Yu, X., VanLoock, M.S., Poplawski, A., Kelman, Z., Xiang, T., Tye, B.K. and Egelman, E.H. (2002). The Methanobacterium thermoautotrophicum MCM protein can form heptameric rings. *EMBO Rep* **3**(8): 792-797.
- Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H. and Noller, H.F. (2001). Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**(5518): 883-896.
- Zhang, R. and Zhang, C.T. (2003). Multiple replication origins of the archaeon Halobacterium species NRC-1. *Biochem Biophys Res Commun* **302**(4): 728-734.