



Universitat Autònoma  
de Barcelona

*Estudios estructurales de relaxasas  
involucradas en  
el proceso de conjugación bacteriana*

Silvia Russi

Barcelona, 2009





**Universitat Autònoma  
de Barcelona**

*Institut de Biologia Molecular de Barcelona (IBMB-CSIC)*

*Institute for Research in Biomedicine (IRB Barcelona)*

Memoria presentada por Silvia Russi, inscrita en el programa de doctorado “Estructura y Función de Proteínas” para obtener al grado de Doctora.

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

Silvia Russi

Miquel Coll

Barcelona, 2009



*... a Daniel, mi compañero en este viaje.*



¿Por qué la tierra es mi casa?

¿Por qué la noche es oscura?

¿Por qué la luna es blancura  
que engorda como adelgaza?

¿Por qué una estrella se enlaza  
con otra, como un dibujo?

Y ¿por qué el escaramujo  
es de la rosa y del mar?

Yo vivo de preguntar,  
saber no puede ser lujo.

*Silvio Rodríguez*



## AGRADECIMIENTOS

Este trabajo ha sido posible gracias a la colaboración de un extenso grupo de personas que me han brindado, de diversas maneras, el apoyo necesario para la realización de este proyecto de tesis.

He aquí mi más sincero agradecimiento a:

- mi director de tesis, Miquel Coll
- mis colaboradores en el proyecto de estudio de la proteína TrwC: Fernando de la Cruz y María Lucas, de la Universidad de Cantabria
- mis colaboradores en el proyecto de estudio de la proteína MobM: Fabián Lorenzo y Manuel Espinosa, del Centro de Investigaciones Biotecnológicas del CSIC
- Rosa Pérez-Luque, Roeland Boer y Alicia Guasch, quienes trabajaron estrechamente conmigo en los proyectos de la TrwC y MobM
- Judith Linacero y Joan Pous de la Plataforma Automatizada de Cristalografía del Parc Científic de Barcelona
- Jenny Colom del Servicio de Purificación de Proteínas del Instituto de Biología Molecular de Barcelona, CSIC
- Isabel Usón, por sus comentarios y consejos en lo relativo a la colección y procesamiento de datos
- todos y cada uno de mis compañeros de laboratorio, sincrotronadas y anexos: Lionel Costenaro, Leonor Alloza, Sol Cima, Esther Peña, Fernando Garçès, Tomislav Kamenski, Maïlys Boutin, Nereida Jiménez, Anna Rubio, Robert

Janowski, Sèbastien Violot, Esther Ferrando, Albert Canals, Dani Badia, Marta Nadal, Carme Aran, Diana Martínez, Fabio Sessa, Carlo Carolis, Nayibe Guarín, Raquel Arribas, Bárbara Machado, Queralt García, María Adell, David Aparicio, Xavi Carpena, Luca Martinelli, Jordi Querol, Arnau Casañas, Damià Garriga y Cristina Ferrer

- el European Synchrotron Radiation Facility (ESRF) y en particular al personal de las líneas de macromoléculas BM16, ID14-4 e ID23-1, donde se hicieron las recolecciones de datos
- la Generalitat de Catalunya, a través de su beca de doctorado FI
- Xavier Daura, coordinador del programa de doctorado de Estructura y Función de Proteínas y tutor de esta tesis

Y finalmente, quiero agradecer en particular a Daniel, por quererme tal como soy y comprender que por seguir mi vocación soy capaz de mover montañas ... o mudarme a ellas.

## Abreviaturas

ADN	Deoxyribonucleic acid
ATP	Adenosine triphosphate
BSA	Bovine Serum Albumin
C-terminal	Carboxy terminal
CCD	Charge Coupled Device
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic Mobility Shift Assay
ESRF	European Synchrotron Radiation Facility
IHF	Integration Host Factor
MAD	Multiple wavelength Anomalous Diffraction
MPD	2-methyl-2,4-pentenediol
N-terminal	Amino terminal
SAD	Single wavelength Anomalous Diffraction
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
PDB	Protein Data Bank
PEG	Polyethylene glycol
oriT	Origen of transfer
T4SS	Type 4 secretion system

# ÍNDICE

1. INTRODUCCIÓN	1
1.1. <i>La transferencia horizontal de genes en bacteria</i>	3
1.2. <i>El proceso de conjugación bacteriana</i>	4
1.3. <i>La maquinaria molecular de la conjugación</i>	6
1.3.1. El relaxosoma	7
1.3.2. La proteína de acoplamiento	8
1.3.3. El sistema de secreción (T4SS)	8
1.4. <i>El sistema conjugativo R388 de Escherichia coli</i>	9
1.4.1. El ensamblaje del relaxosoma	9
1.4.2. La relaxasa/helicasa TrwC	10
1.4.2.1. <i>Estructura cristalina del dominio N-terminal de TrwC en complejo ADN</i>	12
1.4.2.2. <i>La arquitectura del sitio activo</i>	14
1.4.2.3. <i>El mecanismo catalítico</i>	16
1.4.3. Las proteínas auxiliares TrwA e IHF	17
1.5. <i>El sistema movilizable pMV158 de Streptococcus agalactiae</i>	20
1.5.1. El origen de transferencia ( <i>oriT</i> )	21
1.5.2. La relaxasa MobM	23
2. OBJETIVOS	29
3. MATERIALES Y MÉTODOS	33
3.1. <i>Preparación de la muestra</i>	33
3.1.1. Preparación de la muestra de proteína	33
3.1.2. Preparación de los oligonucleótidos	33
3.1.3. Preparación y purificación de los complejos proteína-ADN	33
3.1.4. Complejos TrwC-ADN	34
3.1.4.1. <i>Preparación y purificación del complejo TrwC-ADN25</i>	34
3.1.4.2. <i>Preparación y purificación del complejo TrwC N293Y18F-ADN27</i>	34
3.1.5. Complejos MobM-ADN	35
3.1.5.1. <i>Preparación y purificación del complejo MobM-ADN24 (IR-2)</i>	35
3.1.5.2. <i>Preparación y purificación del complejo MobM-ADN32 (IR-3)</i>	37
3.1.5.3. <i>Preparación y purificación del complejo MobM N198-ADN32</i>	39

3.1.5.4. Preparación y purificación del complejo MobM N198-ADN26 (IR-3(-6))	41
3.1.5.5. Preparación y purificación del complejo SeMet-MobM N198-ADN26	43
3.2. Obtención de los cristales	45
3.2.1. Cristalización	45
3.2.1.1. Búsqueda de las condiciones de cristalización	46
3.2.1.2. Optimización de los cristales	47
3.2.2. Derivatización de los cristales con metales divalentes	47
3.2.3. Crio-protección	48
3.2.4. Preparación final de los cristales	49
3.2.5. Complejos TrwC-ADN	50
3.2.5.1. Cristalización del complejo TrwC-ADN25	50
3.2.5.2. Derivatización del complejo TrwC-ADN25 con metales divalentes	50
3.2.5.3. Cristalización del complejo TrwC N293 <sup>Y18F</sup> -ADN27	51
3.2.6. Complejos MobM N198-ADN	52
3.2.6.1. Cristalización del complejo MobM N198-ADN32	52
3.2.6.2. Cristalización del complejo MobM N198-ADN26	52
3.2.6.3. Cristalización del complejo SeMetMobM N198-ADN26	54
3.3. Recogida y procesamiento de los datos de difracción	54
3.3.1. Recolección de datos	54
3.3.2. Procesamiento de los datos	57
3.3.3. Complejos TrwC-ADN	59
3.3.3.1. Recolección de datos de los complejos TrwC-ADN25-M <sup>2+</sup>	59
3.3.3.2. Recolección de datos del complejo TrwC <sup>Y18F</sup> -ADN27	62
3.3.4. Complejos MobM N198-ADN	62
3.3.4.1. Recolección de datos del complejo MobM N198-ADN26	62
3.3.4.2. Recolección de datos del complejo SeMet-MobM N198-ADN26	64
3.4. Resolución y refinamiento estructural	66
3.4.1. Resolución estructural	66
3.4.1.1. Métodos de resolución estructural utilizados	67
3.4.1.1.1. Reemplazo molecular	67
3.4.1.1.2. Difracción Anómala	68
3.4.2. Construcción del modelo	69
3.4.3. Refinamiento	70
3.4.4. Complejos TrwC-ADN	71

3.4.4.1. Resolución estructural y refinamiento del complejo TrwC-ADN25	71
3.4.4.2. Resolución estructural y refinamiento del complejo TrwC <sup>Y18F</sup> -ADN27	73
3.4.4.3. Resolución estructural y refinamiento de los complejos MobM N198-ADN26	74
3.4.4.4. Resolución estructural y refinamiento de los complejos SeMetMobM N198-ADN26	75
3.5. Validación del modelo	78
4. RESULTADOS Y DISCUSIÓN	83
4.1. La relaxasa TrwC	83
4.1.1. Detalles estructurales de los complejos TrwC-ADN25-M <sup>+2</sup>	83
4.1.2. Detalles estructurales del complejo TrwCN293 <sup>Y18F</sup> -ADN27	87
4.1.3. La naturaleza del ión metálico en la relaxasa TrwC	89
4.1.4. Rol del ión metálico en el mecanismo catalítico	90
4.1.5. La tirosina catalítica Tyr26	92
4.1.6. Las rutas de salida de la cadena de ADN	92
4.2. La relaxasa MobM	94
4.2.1. Características estructurales generales de la relaxasa MobM	94
4.2.2. Las interacciones MobM-ADN	100
4.2.3. La arquitectura del sitio activo	111
4.2.4. El residuo catalítico	112
4.2.5. El rol y la naturaleza del metal en la actividad catalítica	115
4.2.6. La conformación del extremo 3' de la cadena de ADN	115
4.2.7. El mecanismo catalítico	117
5. CONCLUSIONES	121
5.1. La relaxasa TrwC del plásmido conjugativo R388	121
5.2. La relaxasa MobM del plásmido movilizable pMV158	122
6. BIBLIOGRAFIA	127
7. PUBLICACIONES	139

# ***1. INTRODUCCIÓN***



# 1. Introducción

## 1.1. La transferencia horizontal de genes en bacteria

La transferencia horizontal de genes, también conocida como transferencia lateral, es el proceso mediante el cual un organismo es capaz de transferir material genético a una célula no descendiente. Es un proceso muy común entre los organismos procariontes, en particular las bacterias, y gracias a él, éstas son capaces de adquirir rasgos adaptativos nuevos, como es la resistencia a antibióticos.

Existen tres mecanismos principales a través de los cuales dicha transferencia puede realizarse: la transformación, la transducción y la conjugación. La figura 1.1 representa esquemáticamente estos tres mecanismos.

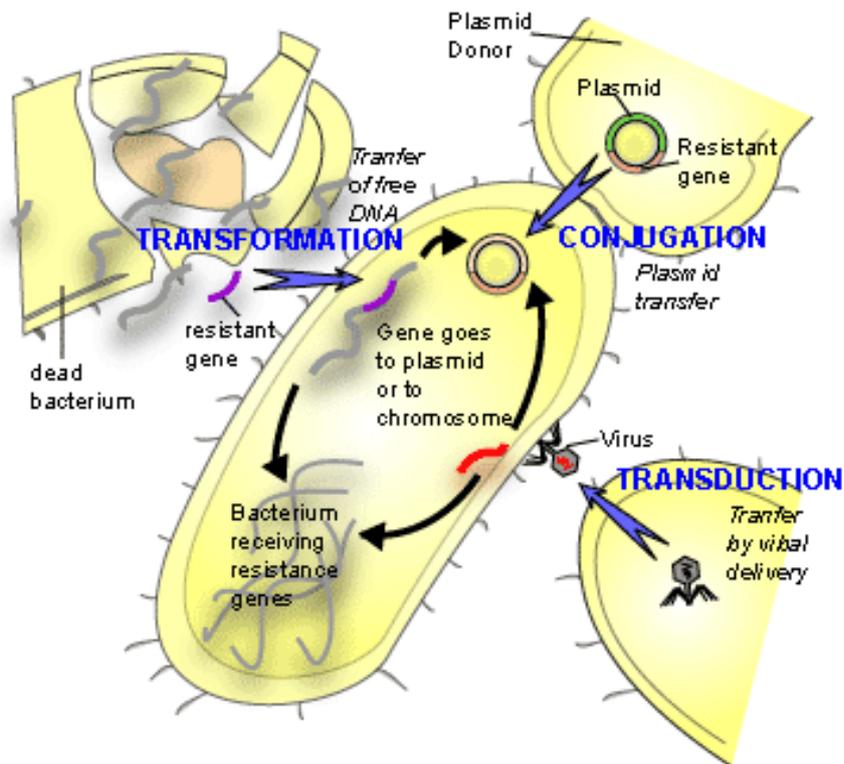


Figura 1.1. Representación esquemática de los tres procesos fundamentales de transmisión de ADN en las bacterias: (a) transformación mediante la adquisición de ADN desnudo del medio, (b) transducción a través de un bacteriófago que actúa como vector de transmisión y (c) conjugación mediada por un plásmido. (<http://www.scq.ubc.ca/attack-of-the-superbugs-antibiotic-resistance/>)

En el proceso de transformación, la célula es capaz de adquirir fragmentos de ADN del medio que la rodea y posteriormente integrarlo a su propio genoma a través de recombinación homologa o por transposición de fragmentos específicos del ADN adquirido.

En el proceso de transducción, el ADN es introducido a la célula receptora con la ayuda de un bacteriófago, que actúa de vector para la transmisión. La infección mediante el fago comienza con la unión de éste a receptores específicos presentes en la superficie de la célula huésped y la posterior inyección de su genoma dentro de la célula.

El tercer, y último mecanismo, es el proceso de conjugación y en él la transferencia se efectúa mediante contacto celular entre dos bacterias produciéndose la transferencia de ADN de una de ellas (dadora) a la otra (receptora). Es en este proceso que se centra el presente trabajo, estudiándolo en detalle a nivel molecular.

## ***1.2. El proceso de conjugación bacteriana***

La habilidad de las bacterias Gram-negativas para transferir su material genético, mediante el proceso de conjugación, fue descubierto en el año 1946 en cultivos de *Escherichia coli* por Lederberg y Tatum (Lederberg and Tatum, 1946). Asimismo, años más tarde, se descubrió que las bacterias Gram-positivas también poseían esa capacidad (Grohmann *et al.*, 2003; Scott and Churchward, 1995; Ton-That and Schneewind, 2003).

La rápida adquisición de rasgos adaptativos nuevos convierte a este proceso en uno de los mecanismos más relevantes en la evolución bacteriana (de la Cruz and Davies, 2000; Ochman *et al.*, 2000) y su estudio puede ser clave para el diseño racional de nuevos fármacos que sean capaces de inhibir este mecanismo de adquisición de resistencia a antibióticos.

La transmisión de material genético, en el proceso de conjugación, es mediada por la transferencia de una molécula circular de ADN extra-cromosómico, llamada

plásmido, y que posee la capacidad de replicarse a sí misma.

Existen dos tipos diferentes de plásmidos que pueden ser transmitidos durante la conjugación: los plásmidos conjugativos y los plásmidos movilizables (Francia *et al.*, 2004).

Los plásmidos conjugativos tienen la característica de ser auto-transmisibles ya que contienen la información genética necesaria para codificar las proteínas involucradas en la formación del sistema de secreción. Estos plásmidos son, no sólo capaces de transferirse entre cepas bacterianas de la misma especie, sino que también pueden hacerlo entre grupos filogénicamente alejados, hongos, plantas e incluso células de mamíferos (Christie, 2001; Christie and Cascales, 2005; Christie and Vogel, 2000; Ding *et al.*, 2003; Seubert *et al.*, 2003; Waters, 2001), y es por este motivo que también se los suele llamar plásmidos promiscuos. Este es el caso de las familias de plásmidos F (IncFI), R388 (IncW), pKM101 (IncN) o RP4 (IncP $\alpha$ ) (Schroder and Lanka, 2005).

Los plásmidos movilizables, por el contrario, no son capaces de auto-transmitirse y deben ser asistidos por los plásmidos conjugativos, coexistentes en la bacteria, quienes les aportan las funciones conjugativas necesarias para su transmisión. Ejemplo de este tipo de plásmidos son los plásmidos de las familias RSF1010, ColE1, CloDF13 o pMV158 (Francia *et al.*, 2004).

Los genes de los plásmidos conjugativos codifican, en dos regiones diferentes, las proteínas involucradas en el proceso de conjugación. La primera, es la región de replicación y transferencia de ADN o Dtf (DNA transfer and replication) y la segunda, es la región de formación del par de unión o Mpf (Mating pair formation) (Backert and Meyer, 2006; Christie *et al.*, 2005; Silverman, 1997).

Las proteínas codificadas por la región Dtf son las responsables de preparar la cadena de ADN a ser transferida y están involucradas en la formación del relaxosoma y su acoplamiento con el sistema de secreción (Ziegelin *et al.*, 1989). Las proteínas del Mpf son, en cambio, las piezas constituyentes del sistema de secreción y las encargadas de formar el conducto trans-membranal que conectará ambas células (incluyendo el apéndice proteico llamado pilus en las bacterias Gram-negativas) y permite el pasaje del

ADN hacia la célula receptora.

### 1.3. La maquinaria molecular de la conjugación

Para lograr ser transferida, la molécula de ADN debe ser capaz de sortear una serie de barreras que le dificultan el pasaje. A modo de ejemplo, en bacterias Gram-negativas podemos citar, la presencia de la membrana interna hidrofóbica, las nucleasas que se alojan en el periplasma, o la carga negativa de la membrana externa (Dreiseikermann, 1994). Las bacterias han desarrollado, para superar estos obstáculos, una compleja maquinaria molecular constituida por diferentes componentes multi-proteicos.

Estos componentes multi-proteicos pueden ser agrupados en tres módulos funcionales diferentes: i) el relaxosoma (Lanka and Wilkins, 1995; Wilkins and Lanka, 1993), ii) el conducto trans-membranal, y iii) la proteína de acoplamiento (Gomis-Ruth *et al.*, 2004; Llosa and de la Cruz, 2005) (figura 1.2).

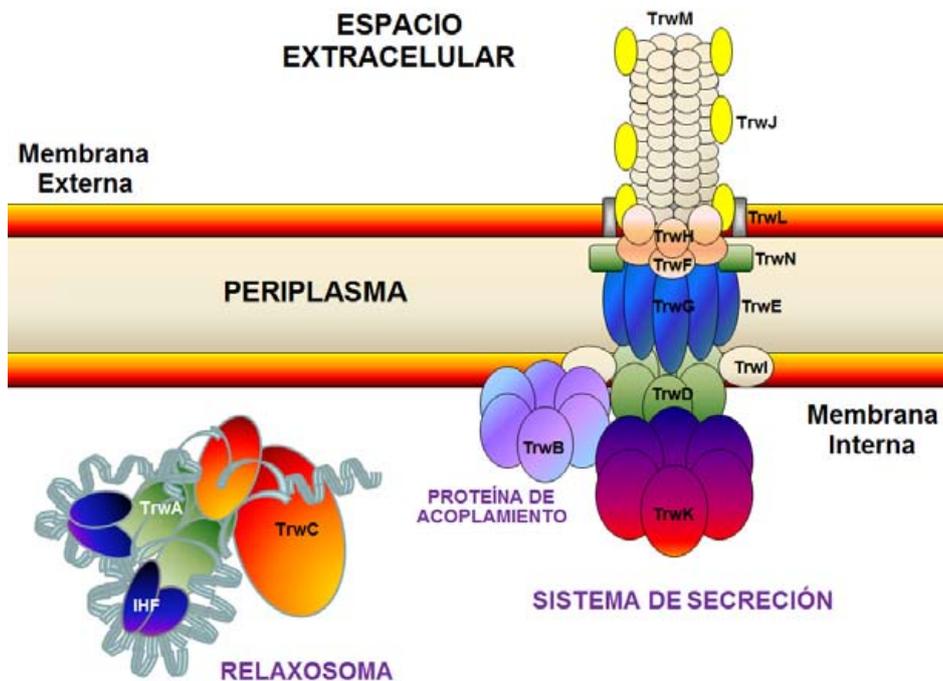


Figura 1.2. Maquinaria conjugativa del plásmido R388 de *Escherichia coli* en el que se indican el relaxosoma, la proteína de acoplamiento y el conducto trans-membranal del sistema de secreción.

El relaxosoma es el complejo núcleo-proteico responsable de iniciar el proceso de transferencia mediante el corte específico de un enlace fosfo-diéster de la cadena de ADN en el origen de transferencia de la misma (*oriT*).

El conducto trans-membranal constituye propiamente el sistema de secreción a través del cual pasará la cadena de ADN y se conecta con el relaxosoma mediante la proteína de acoplamiento.

Este conducto pertenece a la familia de los sistemas de secreción de tipo IV (T4SS). Estos sistemas no sólo son capaces de translocar el ADN a través de la pared celular, sino que, además, pueden transferir macromoléculas, toxinas multiméricas o complejos proteína-ADN a través de ella (Backert and Meyer, 2006; Burns, 2003; Cascales and Christie, 2003; Chen *et al.*, 2005; Christie, 2004; Llosa and O'Callaghan, 2004; Sexton and Vogel, 2002).

### **1.3.1. El relaxosoma**

Está constituido por una proteína relaxasa/helicasa, responsable del corte de la cadena de ADN y del desenrollamiento del ADN plasmídico, así como de un conjunto de proteínas auxiliares. La relaxasa también actúa como proteína piloto, guiando al ADN a través del canal de secreción. El proceso es impulsado energéticamente por las ATPasas que forman parte del mismo.

Las proteínas del relaxosoma se unen de manera específica a una corta secuencia de ADN denominada origen de transferencia (*oriT*), donde el proceso de transferencia se inicia y termina. La activación del relaxosoma se produce una vez que la célula dadora entra en contacto con la célula receptora (Durrenberger *et al.*, 1991). La asociación entre ambas es mediada por el conducto trans-membranal, que está formado por una docena de proteínas. El apéndice extracelular de dicho conducto, el pilus, es el encargado de unirse a la superficie de la célula receptora permitiendo que se establezca un estrecho contacto entre ambas.

### **1.3.2. La proteína de acoplamiento**

La proteína de acoplamiento es una pieza clave en los sistemas conjugativos. Sin ella la translocación de ADN no sería posible debido a su papel como nexo entre el relaxosoma y el sistema de secreción (Cabezón *et al.*, 1997; Firth *et al.*, 1996; Gomis-Ruth *et al.*, 2004; Moncalian *et al.*, 1999a; Santini and Stanisich, 1998).

Es una proteína integral de la membrana interna y presenta un dominio N-terminal trans-membranal y un dominio C-terminal citoplasmático. En su secuencia se reconocen los motivos Walker A y Walker B, característicos de las proteínas de unión a NTP (Walker *et al.*, 1982), y se especula que su actividad NTPasa contribuye al aporte energético de la maquinaria de secreción (Cabezón and de la Cruz, 2006).

Algunos ejemplos de proteínas integrantes de esta familia son: las proteínas TrwB del plásmido R388 de *Escherichia coli*, TraD y TraG de varios plásmidos Gram-negativos y VirD4 de *Agrobacterium tumefaciens*.

### **1.3.3. El sistema de secreción (T4SS)**

De acuerdo a su función, en la mayoría de los casos putativa, y su ubicación celular, las proteínas que componen el sistema de secreción pueden clasificarse en tres grupos: i) las ATPasas citoplasmáticas o asociadas a la membrana interna; ii) las que forman el núcleo central del canal de secreción; y iii) las que componen el pilus.

El ensamblado del sistema es un proceso que se supone que ocurre en cuatro etapas (Christie *et al.*, 2005). El primer paso implica la formación del complejo que constituye el núcleo central del canal de secreción y que atraviesa la membrana celular. Una vez formado, a éste se acoplan las proteínas asociadas a la formación del pilus. El siguiente paso es la formación de la estructura basal de la membrana interna mediante la anexión de una de las ATPasas citoplasmáticas y el proceso finaliza con el ensamblado del pilus.

Aunque han sido propuestos varios modelos que intentan explicar cómo la cadena de ADN, unido covalentemente a la relaxasa, es translocado a través de este conducto, aún se desconocen los detalles precisos de este proceso a nivel molecular. Actualmente, se postulan tres modelos: el modelo de canal (channel model), el modelo ping-pong (ping-pong model) (Atmakuri *et al.*, 2004) y el modelo disparo-bombeo (shoot-and-pump model) (Llosa *et al.*, 2002).

En los tres casos se parte de la base que, previo al pasaje, la cadena polipeptídica de la relaxasa unida al ADN debe ser desplegada por una de las ATPasas citoplasmáticas. En el modelo de canal, se postula que la cadena polipeptídica simplemente se traslada junto con el ADN a través del conducto en un proceso en el que las ATPasas suministran la energía necesaria. En el modelo ping-pong la cadena polipeptídica y el ADN, pasan primero a través de la proteína de acoplamiento, que se dispone en forma de hexámero en torno a la base del sistema de secreción acoplándose al mismo, y luego continúan su camino atravesando la membrana interna por el conducto de secreción (Atmakuri *et al.*, 2004). En el modelo disparo-bombeo, luego de que la cadena polipeptídica ha atravesado la membrana interna, la proteína de acoplamiento se dispone de manera tal que comienza a bombear a la cadena de ADN para ayudarlo a atravesar el canal (Llosa *et al.*, 2002).

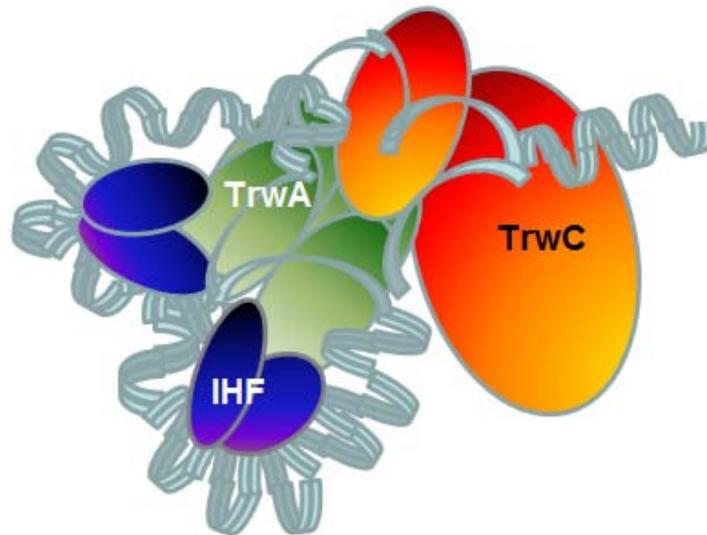
#### ***1.4. El sistema conjugativo R388 de Escherichia coli***

El plásmido conjugativo R388 de *E. coli* pertenece al grupo de incompatibilidad IncW y es uno de sus representantes mejor estudiado de este grupo de plásmidos. En el plásmido R388, las proteínas necesarias para montar la maquinaria que hace posible el proceso de conjugación están codificadas en los genes *trw*.

##### ***1.4.1. El ensamblaje del relaxosoma***

La arquitectura del relaxosoma ha sido estudiada en detalle por Moncalian y colaboradores (Moncalian *et al.*, 1997; Moncalian *et al.*, 1999b). El mismo se ensambla con tres proteínas TrwC (la relaxasa/helicasa), TrwA (reguladora de la transcripción) y

IHF (Integration Host Factor) unidas a la cadena de ADN en el origen de transferencia (*oriT*) (figura 1.3).



*Figura 1.3. Organización del relaxosoma del plásmido R388 de Escherichia coli. Las proteínas IHF y TrwA auxilian a la relaxasa/helicasa TrwC en el procesamiento del ADN*

La actividad helicasa de TrwC separará la cadena de ADN que ha sido cortada de su complementaria, permitiendo la regeneración de la doble cadena mediante la síntesis replicativa de esta última por medio de polimerasas (Kingsman and Willetts, 1978). Este proceso es una variante del mecanismo de replicación conocido como “rolling-circle” (RCR) (Furuya and Komano, 2000; Kramer *et al.*, 1997).

#### **1.4.2. La relaxasa/helicasa TrwC**

TrwC es la relaxasa encargada de iniciar el proceso de conjugación en el sistema conjugativo del plásmido R388 de *E. coli*. Cuenta con 996 residuos y tiene un peso molecular de 108 kDa, distinguiéndose en ella dos dominios funcionales diferentes. El reconocimiento del enlace fosfo-diéster a cortar, las actividades endonucleasa, relaxasa y de trans-esterificación se sitúan en el dominio N-terminal, mientras que su actividad helicasa se concentra en el dominio C-terminal (Llosa *et al.*, 1996).

Los pasos de iniciación y finalización de la replicación conjugativa del ADN son catalizados por dos tirosinas catalíticas (Tyr18 y Tyr26) junto con una tríada de histidinas (His150, His161, His163) que coordinan un metal divalente y que es esencial para la actividad relaxasa de la proteína (Grandoso *et al.*, 2000).

La proteína es capaz de reconocer el sitio de corte en el ADN (*nic*) gracias a una secuencia repetida invertida de 6 pares de bases en la cadena de ADN que adopta un plegamiento característico en forma de horquilla (Guasch *et al.*, 2003). Para ello el ADN es localmente deformado y estabilizado por las proteínas TrwA y el IHF que auxilian a TrwC en el proceso.

Luego del corte inicial de la cadena de ADN, el extremo 5' del *nic* permanece covalentemente unido a la tirosina catalítica Tyr18 mediante un enlace 5'-fosfotirosil, mientras que el nuevo extremo 3' generado es liberado (Llosa *et al.*, 2002; Wilkins and Lanka, 1993). TrwC se mueve entonces de manera procesiva a lo largo de la doble cadena de ADN en la dirección 5' → 3' desenrollándola. La ADN polimerasa III acompaña en su movimiento a la TrwC, encargándose de la síntesis de la cadena complementaria. Cuando TrwC alcanza nuevamente el sitio de corte (*nic*) reconoce la horquilla característica y se produce entonces un segundo corte seguido de la finalización de la reacción y la formación de un nuevo ADN circular de simple cadena.

Sin embargo, estudios recientes han demostrado que la relaxasa es transportada junto con el ADN a la célula receptora, atravesando el sistema de secreción, y por lo tanto es en ella donde la reacción de terminación debe tener lugar (Draper *et al.*, 2005; Garcillan-Barcia *et al.*, 2007).

Para intentar explicar cómo TrwC es capaz de realizar ese segundo corte, se ha propuesto un modelo en el que la proteína modifica su estado de oligomerización. De acuerdo con este modelo, el corte inicial es realizado por la proteína como dímero. Al disociarse posteriormente, de manera tal que el monómero covalentemente unido al ADN se encargaría de conducirlo en su pasaje hacia la célula receptora, mientras que el monómero libre permanecería en la célula dadora desenrollando la cadena de ADN.

#### 1.4.2.1. Estructura cristalina del dominio N-terminal de TrwC en complejo con ADN

En 2003 fue reportada la estructura cristalina del dominio N-terminal de la relaxasa TrwC (Guasch *et al.*, 2003) (figura 1.4), así como la de su proteína homóloga TraI del plásmido F (Datta *et al.*, 2003), ambas en complejo con ADN.

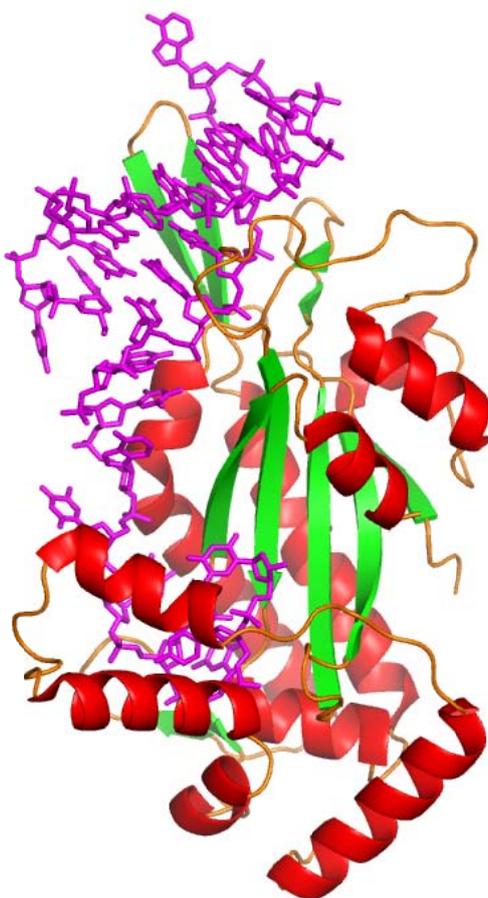
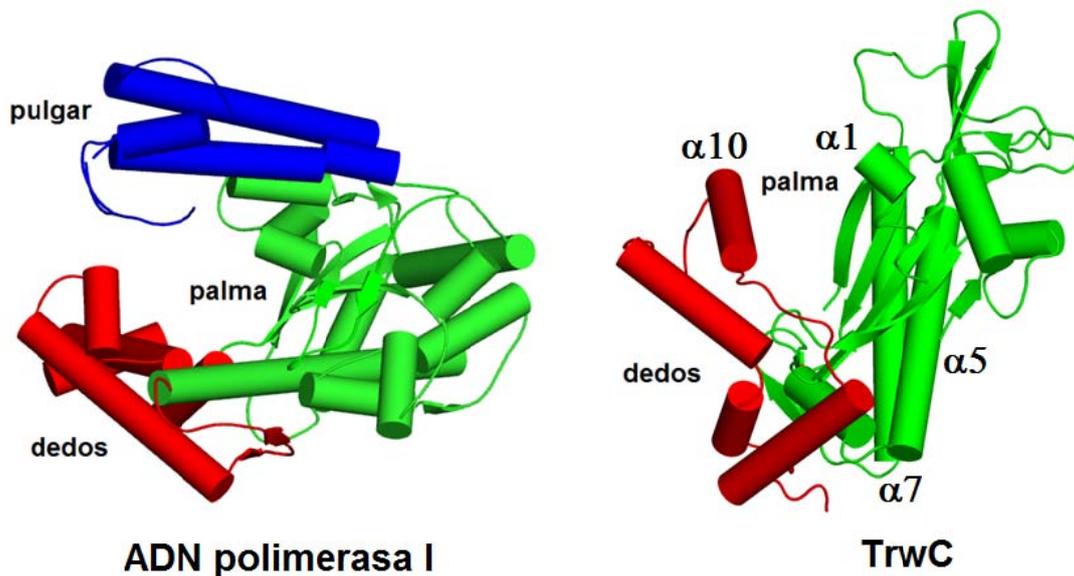


Figura 1.4. Estructura del dominio N-terminal de TrwC en complejo con ADN. El oligonucleótido de 25 bases comprende el origen de transferencia reconocido por TrwC justo antes del sitio de corte.

El dominio N-terminal de TrwC, como el de TraI, presenta un plegamiento global análogo al dominio nucleasa de la proteína iniciadora de la replicación plasmídica RepB (Boer *et al.*, 2009), al de las proteínas virales Rep (Campos-Olivas *et al.*, 2002; Hickman *et al.*, 2002), al del dominio de unión a ADN de la proteína E1 encargada de la iniciación de la replicación del virus del papiloma (Enemark *et al.*, 2002) y al del dominio de unión a ADN en el origen de replicación del antígeno-T SV40 (Meinke *et al.*, 2006).

Por su topología, todas estas proteínas forman parte de la misma familia (Murzin *et al.*, 1995), y exhiben el plegamiento característico del dominio de unión al origen de replicación, el cual comprende un núcleo central formado por cinco hebras beta y dos hélices alfa que se cruzan en diagonal en una cara de la hoja beta (figura 1.4).

Por otra parte, el núcleo estructural central de TrwC también se asemeja al observado en las polimerasas de ADN y recuerda, por su morfología, a una mano derecha. La “palma” está formada por la hoja beta central y dos hélices alfa paralelas ( $\alpha 5$  and  $\alpha 7$ ) que la flanquean por una de las caras, mientras que otras dos hélices alfa lo hacen por la otra cara ( $\alpha 1$  and  $\alpha 10$ ) (figura 1.5).



*Figura 1.5. Plegamiento del dominio N-terminal de TrwC que recuerda la morfología en forma de mano derecha observada en las ADN polimerasas I. Los elementos estructurales que constituyen el pulgar en estas últimas, no obstante, no están presentes en el caso de TrwC.*

La hoja beta central junto con los “dedos” de la proteína delimita, una extensa hendidura donde se sitúa una cadena sencilla de ADN, luego de la horquilla. Esta cadena sencilla de ADN continúa en dirección al sitio activo y en las cercanías de las bases G22, T23 y C24 se produce un marcado giro en forma de U de su extremo 3’.

La base G22 se apila paralela a la base C24, y la T23 se dispone prácticamente

perpendicular a la G22 estableciendo enlaces de hidrógeno con ella y con el grupo amino del residuo SeMet1 de la cadena polipeptídica (figura 1.6). Esta red de interacciones posibilita y estabiliza el pronunciado giro de la cadena de ADN.

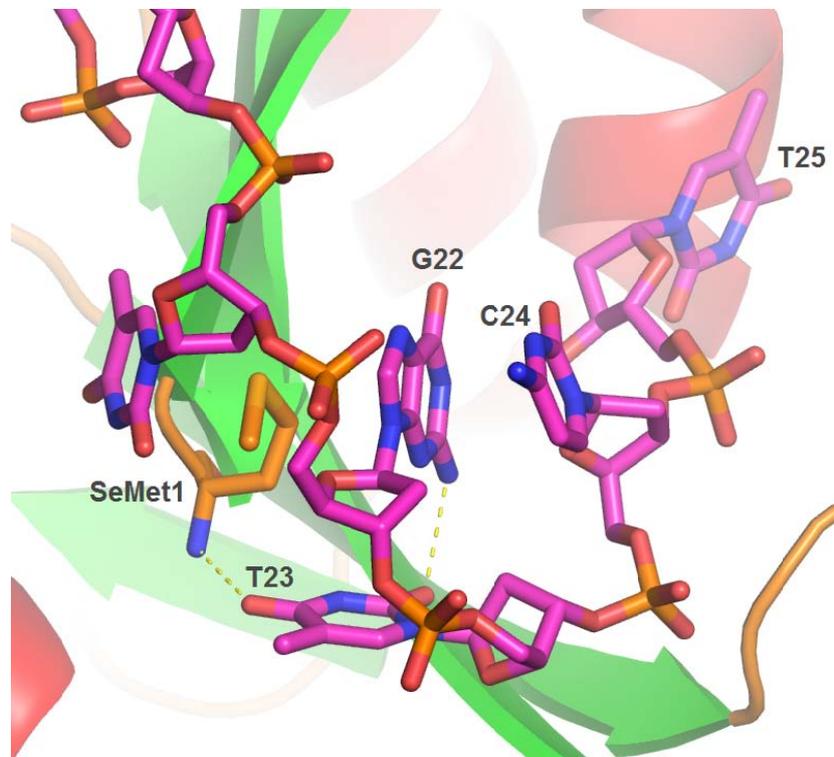


Figura 1.6. Detalle del giro en U del extremo 3' de la cadena de ADN. Las bases G22 y C24 se apilan paralelas, y la base T23 establece enlaces de hidrógeno con la base G22 y la SeMet1.

#### 1.4.2.2. La arquitectura del sitio activo

Las proteínas Rep, así como TrwC, son enzimas involucradas en la transesterificación y sus centros activos están definidos por una tirosina catalítica y un motivo de unión a metal HXH. Sin embargo, una permutación circular entre ellas hace que la posición del residuo catalítico (Tyr) en las relaxasas se sitúe cercano al extremo N-terminal y al C-terminal en las Reps.

En el caso particular de TrwC, el sitio activo se constituye por una tirosina catalítica (Tyr18) y una tríada de histidinas (His150, His 161 y His163) que aloja un metal (figura 1.7). Se observa también la presencia en el sitio activo de un aspártico,

Asp85, y una arginina, Arg154, que establecen contactos con los residuos anteriormente mencionados.

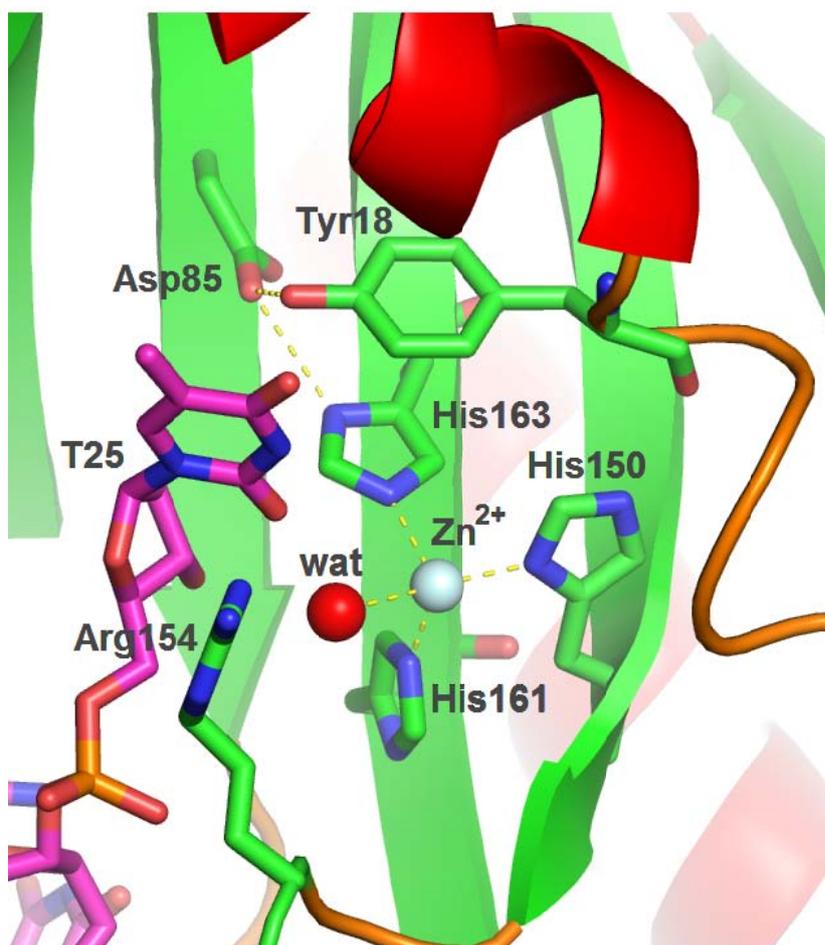


Figura 1.7. Detalle del centro activo del complejo TrwC-ADN25 en la que se observa la tirosina catalítica (Tyr18), la tríada de histidinas (His150, His161 y His 163) y el metal coordinado por ellas.

Un oxígeno del grupo carboxilo del Asp85 participa de un enlace de hidrógeno de tres centros con el oxígeno del grupo hidroxilo de la Tyr18 y el nitrógeno  $\delta 1$  de la His163, quien a su vez se apila paralela a la Tyr18.

El ión metálico ( $Zn^{2+}$ ) se encuentra coordinado por la tríada de histidinas (His150, His161 y His163) y completa su esfera de coordinación tetraédrica con una molécula de agua.

Si bien este trabajo preliminar permitió inferir importantes detalles estructurales de la interacción proteína-ADN, muchos otros interrogantes quedaron pendientes, entre

los que se incluyen la naturaleza química del metal biológicamente activo, el rol preciso de este metal, y la actividad del residuo tirosina en el mecanismo catalítico.

#### ***1.4.2.3. El mecanismo catalítico***

Como en el caso de las topoisomerasas y recombinasas, las relaxasas, también utilizan el ataque nucleofílico por medio de tirosinas en las reacciones de transferencia de ADN. Esto sugiere que sus mecanismos de acción son similares, sin embargo, la presencia del metal en el centro activo en el caso de la relaxasa TrwC y su vinculación con el mecanismo catalítico, indican que los mecanismos tienen que ser diferentes.

La estructura de complejo de TrwC con el oligonucleótido de 25 bases aportó información relevante, sin embargo, no fue posible establecer un mecanismo catalítico preliminar. Asimismo, el papel preciso de la segunda tirosina catalítica Tyr26 no pudo establecerse debido a que el desorden estructural en la zona de localización de este residuo impidió su modelado en la estructura.

Por otra parte, se observó que un par de iones sulfato localizados cerca del sitio de corte podrían estar mimetizando la posición de dos grupos fosfato del ADN que definirían las direcciones de dos rutas de salida que las cadenas de ADN podrían utilizar después del corte.

La separación de los caminos se localiza próxima al sitio de unión del metal y sus direcciones siguen el eje longitudinal horizontal de la TrwC y el perpendicular a este (figura 1.8). Si bien la hipótesis es razonable en función de la evidencia experimental disponible, no pudo ser confirmada debido a la imposibilidad de obtener cristales con oligonucleótidos más largos, que contengan el enlace fosfo-diéster a cortar.

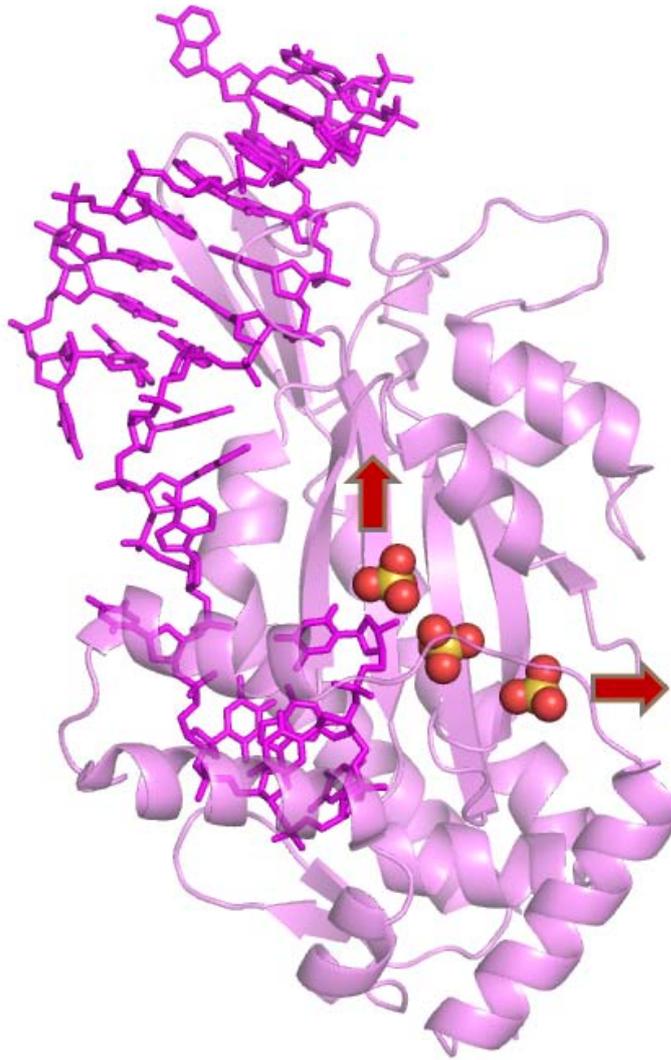


Figura 1.8. Posibles rutas de salida de la cadena de ADN luego del corte indicadas por la posición de los grupos sulfatos que mimetizan a los fosfatos de la cadena de ADN.

#### 1.4.3. Las proteínas auxiliares *TrwA* e *IHF*

*TrwA* se une bajo la forma de tetrámero al origen de transferencia (*oriT*) en dos puntos de unión que se encuentran alejados del sitio de corte (*nic*) (Moncalian *et al.*, 1997) y se ha demostrado que cumple dos roles fundamentales e independientes en el proceso de conjugación.

La unión de *TrwA* a una secuencia específica de reconocimiento en el origen de transferencia (*oriT*) provoca la represión de la transcripción del operón *trwABC*. Este

operón es el encargado de codificar las proteínas TrwA, TrwB y TrwC que son indispensables para el proceso de conjugación (Moncalian and de la Cruz, 2004) y por lo tanto esto sugiere que TrwA tiene un papel clave en la regulación de dicho proceso.

Asimismo, TrwA participa en la preparación del *oriT* ayudando a la relajación de la cadena de ADN superenrollado, asistiendo en este proceso a la proteína TrwC, que es la encargada de catalizar el mismo (Moncalian and de la Cruz, 2004)

La estructura cristalina de la proteína homóloga CopG del plásmido pMV158 en complejo con ADN, demostró como el reconocimiento del ADN se realiza mediante un dímero de dímeros (Gomis-Ruth *et al.*, 1998). Se ha demostrado, por otra parte, que la doble cadena de ADN puede ser curvada más de 120° por la acción de dicho tetrámero (figura 1.9).

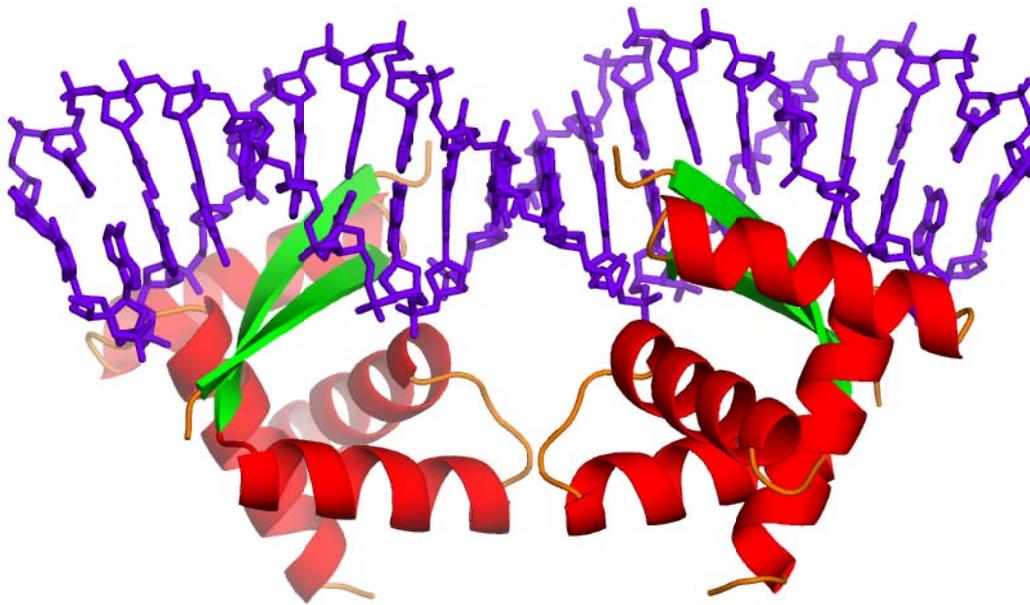


Figura 1.9. Estructura cristalina de la proteína CopG del plásmido pMV158 de *Streptococcus agalactiae*, homóloga a TrwA del plásmido R388 de *Escherichia coli*.

IHF es una proteína heterodimérica que, al igual que TrwA, se une al *oriT* en dos sitios diferentes (Moncalian *et al.*, 1999b). En el relaxosoma, IHF actúa modificando la topología del ADN en el sitio donde la relaxasa tiene que actuar y facilita de esta

manera la reacción de corte de la TrwC (Howard *et al.*, 1995).

La estructura tridimensional de IHF en complejo con ADN (Rice *et al.*, 1996) demostró que las dos subunidades componentes de IHF son estructuralmente equivalentes pese a solo poseer un 30% de identidad de secuencia.

El extremo N-terminal, formado por dos alfa hélices consecutivas y tres hebras beta dispuestas en una hoja beta anti-paralela, son responsables de la dimerización de la proteína. El dominio de unión a ADN penetra en el surco menor de la doble hélice de ADN y lo obliga a adoptar una conformación que forma una vuelta en U de aproximadamente 160° (Rice *et al.*, 1996) (figura 1.10).

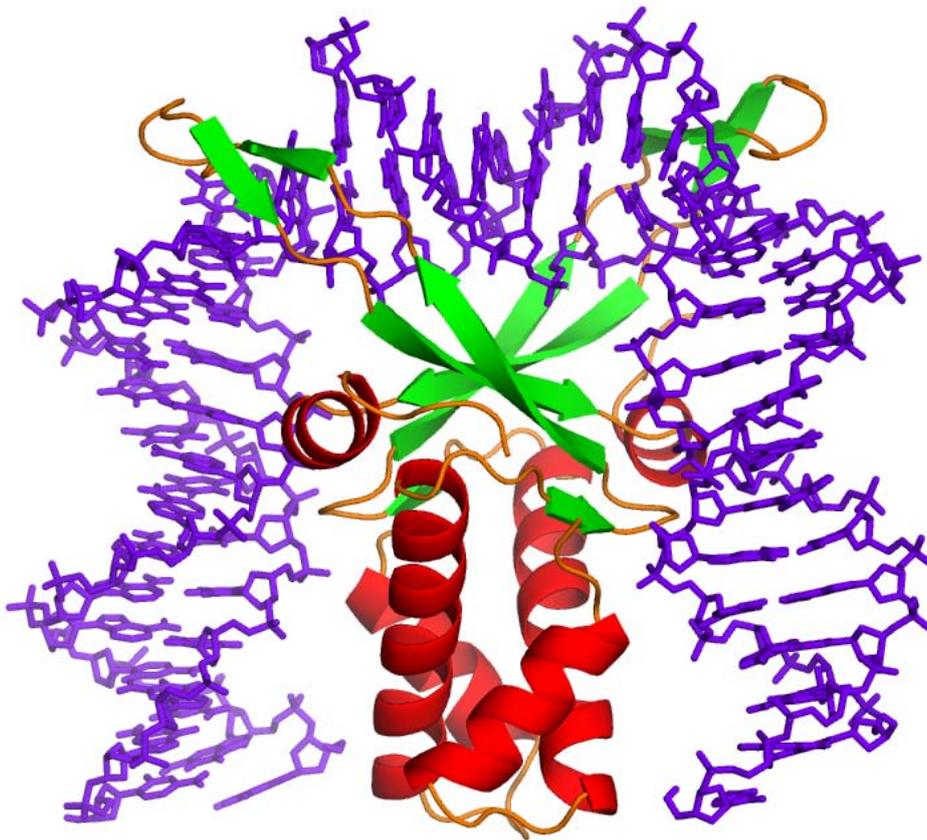


Figura 1.10. Estructura cristalina de IHF de *Escherichia coli* en complejo con ADN.

## 1.5. El sistema movilizable pMV158 de *Streptococcus agalactiae*

El plásmido pMV158, aislado de la bacteria Gram-positiva *Streptococcus agalactiae*, es el representante prototípico de la super-familia de plásmidos movilizables pMV158 (Burdett, 1980).

Este plásmido puede ser movilizado por un amplio rango de bacterias huésped, asociándose a sus respectivos plásmidos conjugativos. Este el caso, por ejemplo, de los plásmidos pIP501/pAM $\beta$ 1 (Priebe and Lacks, 1989) de bacterias Gram-positivas, ó el RP4 y el R388 (Farias *et al.*, 1999) de bacterias Gram-negativas.

Desde el punto de vista de su organización génica el plásmido pMV158 posee tres regiones claramente diferenciadas, que están relacionadas con el control de la replicación, la resistencia de tetraciclina y la movilización del plásmido (figura 1.11). La región de movilización, en particular, está compuesta de un único gen que codifica la relaxasa necesaria para la iniciación del proceso y por la secuencia de iniciación de la transferencia (*oriT*).



Figura 1.11. Esquema de la organización génica del plásmido pMV158 donde se diferencian tres regiones. La región de control de la replicación comprende el *dso* (double strand origin) y codifica las proteínas *CopG* y *RepB*. La resistencia a tetraciclina está dada por la proteína *TetL* codificada en la región de resistencia. Y por último, la región de movilización que codifica la relaxasa *MobM*.

Como en el caso de los plásmidos conjugativos, el inicio de la transferencia de la cadena de ADN comienza con el ataque nucleofílico por parte de una relaxasa a un enlace fosfo-diéster específico (sitio *nic*) que se sitúa en el *oriT* del plásmido (Byrd and Matson, 1997).

Esta reacción es análoga a la que se produce al comienzo del proceso de replicación del plásmido, siendo en ese caso la proteína RepB, codificada en la región de control de la replicación, la encargada de reconocer y cortar el enlace específico en la región *dso* (double strand origin).

Sin embargo, a diferencia de lo que ocurre con las relaxasas involucradas en los procesos de movilización o conjugación, luego del corte la proteína RepB no permanece unida al ADN (Moscoso *et al.*, 1997). Desde el punto de vista bioquímico puede considerarse que estas dos proteínas que codifica el plásmido pMV158 y que son las encargadas de la relajación del ADN en los procesos de replicación y movilización, son funcionalmente diferentes.

### 1.5.1. El origen de transferencia (*oriT*)

Las secuencias específicas de ADN, que son capaces de reconocer y cortar las relaxasas codificadas por la super-familia de plásmidos pMV158, se encuentran muy conservadas (figura 1.12), originando un motivo común que es característico en la familia. Este motivo se define por un conjunto de entre 7 a 10 nucleótidos de secuencias repetidas e invertidas (IR) que forman una horquilla que termina en un lazo que está formado generalmente por 6 nucleótidos.

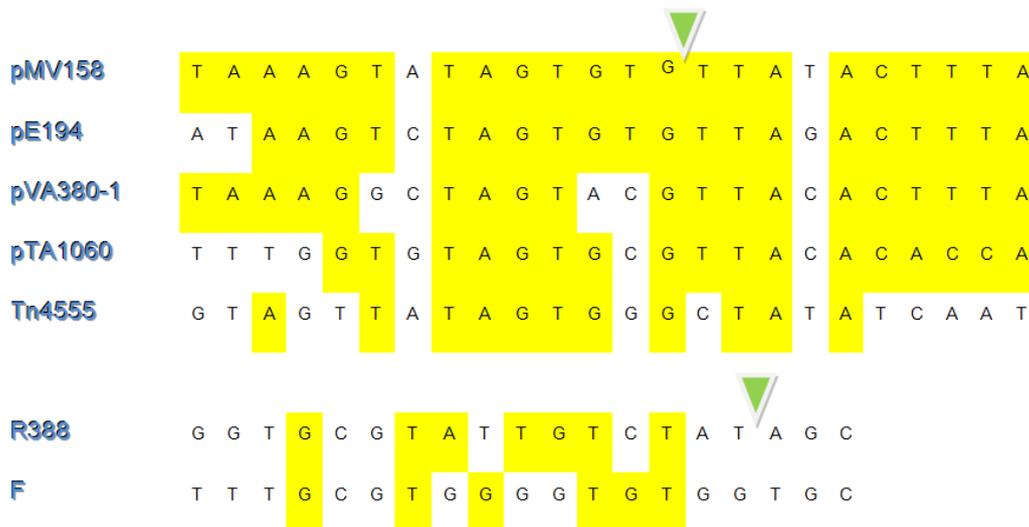


Figura 1.12. Alineamiento de los orígenes de transferencia (*oriT*) en plásmidos de la familia pMV158 y comparación con la de los plásmidos R388 y F. En la figura se han resaltado las bases conservadas en amarillo y se indican los sitios de corte (*nic*).

La formación de la horquilla en el ADN super-enrollado, producto de la presencia del motivo IR, es similar a la observada previamente en el origen de transferencia (*oriT*) del plásmido conjugativo R388.

En el *oriT* del plásmido movilizable pMV158, sin embargo, se pueden identificar dos pares de secuencias de bases IR. Estas secuencias, que se denominan IR-2 e IR-3, son las responsables de generar dos posibles conformaciones del origen de transferencia (*oriT*) de dicho plásmido (figura 1.13).

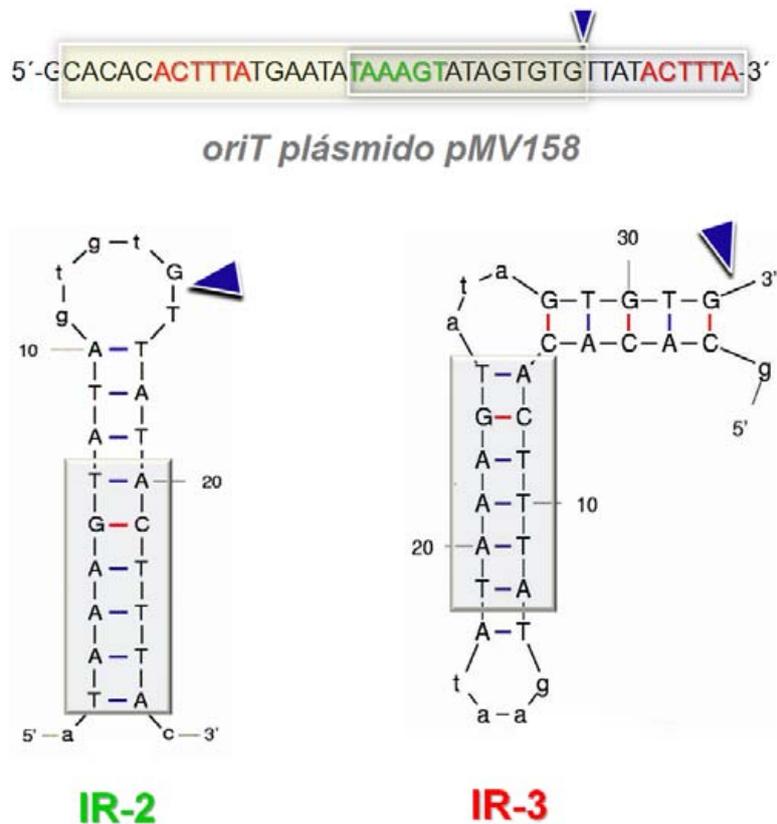


Figura 1.13. Secuencia del origen de transferencia (*oriT*) del plásmido pMV158, indicando los dos motivos IR presentes así como las posibles conformaciones que éstos puede adoptar.

La conformación IR-3 está favorecida termodinámicamente frente a la conformación IR-2 ( $\Delta G_{IR-2} = -4.40$  kcal/mol y  $\Delta G_{IR-3} = -6.88$  kcal/mol (DNA mfold server, Zuker M., Rensselaer Polytechnic Institute 1996-2009)).

Asimismo, los ensayos bioquímicos demuestran que la proteína MobM se une con mayor afinidad al motivo IR-3 que al IR-2 (resultados no publicados comunicados).



Se predice, por otra parte, la existencia de una estructura de tipo “coiled coil” entre los residuos 257-344 y 429-463 (COILS, Lupas *et al.*, 1991). Es en esta región que se localizan cinco residuos leucina que constituyen un motivo Leu-zipper responsable de la dimerización de la proteína.

Al igual que en el caso de TrwC y otras relaxasas conjugativas, es en el dominio N-terminal de MobM donde reside su actividad la relajación del ADN. Los alineamientos de secuencia, con relaxasas pertenecientes a la super-familia del plásmido pMV158, permiten reconocer la presencia de tres motivos conservados (figura 1.15).

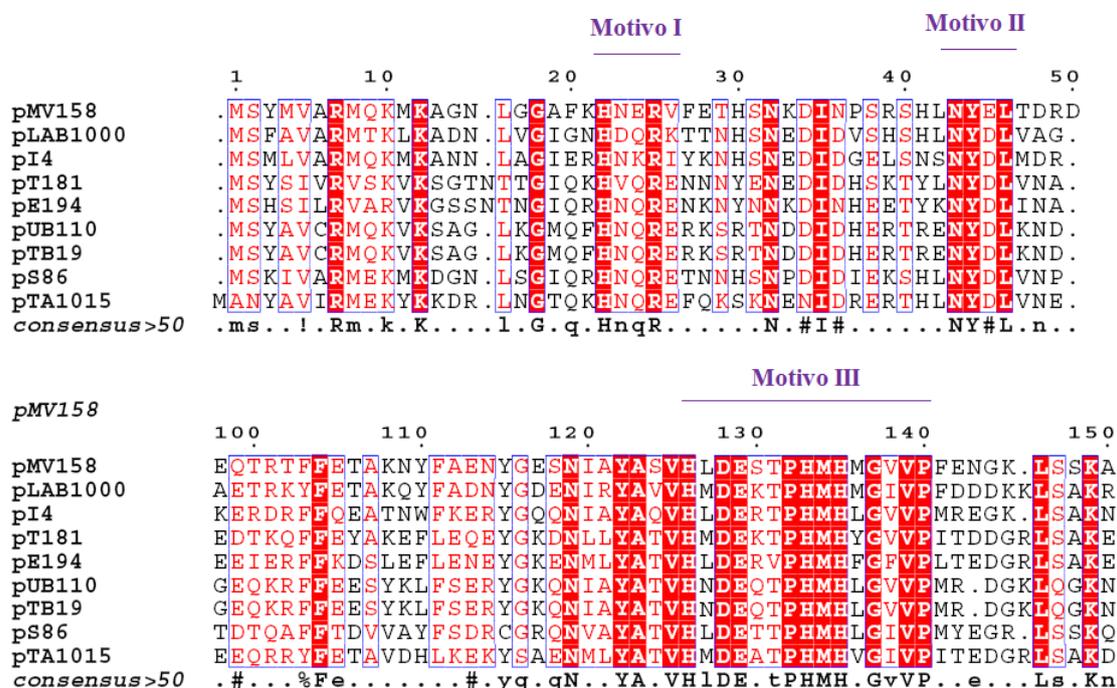


Figura 1.15. Alineamiento de secuencia de relaxasas pertenecientes a la super-familia del plásmido pMV158 en el que se muestra los motivos conservados I, II y III. (ESPrIPT 2.2, <http://espript.ibcp.fr/ESPrIPT/ESPrIPT/>)

Por homología con los sistemas conjugativos, en los sistemas movilizables el motivo I es el sitio donde se encuentra la tirosina catalítica encargada de corte de la cadena de ADN; el motivo II comprende los residuos involucrados en los contactos proteína-ADN en el extremo 3' de la región del sitio de corte (*nic*); y el motivo III contiene la tríada de histidinas que ayudan al ataque nucleofílico de la tirosina mediante la coordinación del ion metálico (Francia *et al.*, 2004).

Sin embargo, en la super-familia del plásmido pMV158 el motivo característico I se encuentra ausente. En su lugar, se ha sugerido que es en el motivo II el lugar donde se localiza la tirosina catalítica (Guzmán and Espinosa, 1997), aunque hasta el momento se carece de evidencia experimental que confirme esta hipótesis.

En el caso particular de la proteína MobM, y de acuerdo con esta predicción, las histidinas His126, His133 e His135 presentes en el motivo III constituirán la tríada encargada de coordinar el ión metálico indispensable para el desarrollo de la actividad relaxasa y la tirosina Tyr44, presente en el motivo II, sería el residuo catalítico responsable del ataque nucleofílico al grupo fosfato cuyo enlace fosfo-diéster cortará.

Aunque la naturaleza química del metal involucrado en el mecanismo de catálisis aún es materia de análisis, los ensayos bioquímicos sugieren al  $Mg^{2+}$  y al  $Mn^{2+}$  como los metales biológicamente activos. También se ha observado actividad relaxasa reducida o nula en presencia, respectivamente, de  $Ca^{2+}$  o  $Ba^{2+}$  y  $Zn^{2+}$  (de Antonio *et al.*, 2004; Guzmán and Espinosa, 1997).



## ***2. OBJETIVOS***



## 2. Objetivos

El objetivo principal del presente trabajo es estudiar, desde el punto de vista estructural, proteínas relaxasas que conforman parte de la maquinaria involucrada en el proceso de conjugación bacteriana.

Asimismo, se han definido los siguientes objetivos parciales:

- Identificar la naturaleza química y el rol del metal divalente que coordina la tríada de histidinas presente en el sitio activo de las relaxasas.
- Plantear un posible mecanismo catalítico para la reacción de corte de ADN o primer paso de la transferencia de cadena.
- Estudiar comparativamente dos diferentes relaxasas, de organismos Gram-negativos y Gram-positivos, de plásmidos conjugativos y movilizables.



### **3. *MATERIALES Y MÉTODOS***



## **3. Materiales y métodos**

### ***3.1. Preparación de la muestra***

#### ***3.1.1. Preparación de la muestra de proteína***

La concentración de proteína se determinó mediante la medición de la absorbancia a 280 nm. El coeficiente de extinción molar  $\epsilon$  puede ser estimado a partir de la secuencia de la proteína a través de utilidades disponibles en servidores web (ExpASY Proteomics Server, <http://www.expasy.ch/>).

La muestra se concentró utilizando unidades de filtrado Centricon (Amicon-Millipore) con membrana de corte de peso molecular de 10kD, a 4500 g y a 4°C, en una centrifuga MULTIFUGE 3 L-R Heraeus.

#### ***3.1.2. Preparación de los oligonucleótidos***

Los oligonucleótidos se adquirieron en la casa comercial MGW Biotech. Las muestras, con un grado de pureza HPLC según el fabricante, se utilizaron sin mediar pasos extra de purificación, sometiéndolas a un tratamiento específico para obtener la forma de horquilla reconocida por las relaxasas.

El protocolo empleado fue descrito previamente por Guasch y colaboradores (Guasch *et al.*, 2003) y consiste en disolver el ADN liofilizado en agua mili-Q y posteriormente someterlo a un proceso de templado, calentándolo durante 15 minutos en baño de agua a 80°C y enfriándolo en baño de hielo a 0°C y con agitación continua a 300 rpm durante 15 minutos.

#### ***3.1.3. Preparación y purificación de los complejos proteína-ADN***

Los complejos se prepararon, inmediatamente después del tratamiento al ADN, mezclando muy lentamente la proteína con el ADN en relación molar 1:1.2. Para

asegurar la completa formación del complejo, el ADN está 20% en exceso respecto a la proteína y la mezcla se deja en incubación durante 12 horas.

La purificación se llevó a cabo mediante cromatografía de exclusión molecular, utilizando una columna Superdex75 10/300 GL o Superdex200 10/300 (GE Healthcare), dependiendo del peso molecular de la muestra estudiada, y las fracciones correspondientes fueron concentradas por centrifugación con concentradores de membrana Centricon (Amicon-Millipore).

La concentración final de las muestras se determinó por el método de Bradford realizando la curva de calibración con albúmina sérica bovina (BSA) (Bradford, 1976).

### **3.1.4. Complejos TrwC-ADN**

#### **3.1.4.1. Preparación y purificación del complejo TrwC-ADN25**

Para la preparación del complejo se mezclaron una solución 0.179 mM de la proteína TrwC en 450 mM de NaCl y 5 mM de Hepes a pH=7.6, con una solución 0.179 mM del oligonucleótido 5'-GCGCACCGAAAGGTGCGTATTGTCT-3' disuelto en el mismo tampón y sometido previamente al proceso de templado.

La mezcla proteína-ADN se incubó durante 12 horas y el complejo así formado fue utilizado para los ensayos de cristalización. El paso de purificación por cromatografía de exclusión molecular se eliminó porque empíricamente se observó que esto no afecta ni a la cristalización ni a la calidad de los cristales.

#### **3.1.4.2. Preparación y purificación del complejo TrwC N293<sup>Y18F</sup>-ADN27**

Para la obtención del complejo se utilizó un mutante inactivo de la proteína TrwC en el que la tirosina catalítica Tyr18 se mutó por una fenilalanina. Al carecer de actividad relaxasa, es posible la formación y cristalización del complejo con un oligonucleótido cuya secuencia comprende el sitio de corte (*nic*) sin que la proteína sea capaz de cortarlo.

Una solución 0.23 mM del oligonucleótido en 450 mM de NaCl y 5 mM de Hepes a pH=7.6 del oligonucleótido 5'-GCGCACCGAAAGGTGCGTATTGTCTAT-3' se mezcló con una solución 0.23 mM de la proteína TrwC en el mismo tampón. La mezcla proteína-ADN se incubó durante 12 horas y, como en el caso del complejo TrwC-ADN25, se utilizó la muestra para los ensayos de cristalización sin posterior purificación.

### **3.1.5. Complejos *MobM-ADN***

#### **3.1.5.1. Preparación y purificación del complejo *MobM-ADN24 (IR-2)***

Una solución 2.5 mM en agua miliQ del oligonucleótido que comprende la secuencia de reconocimiento IR-2 (5'-TAAAGTATAGTGTGTTATACTTTA-3'), fue sometido al protocolo de templado preestablecido, mezclándolo posteriormente en relación 1.2:1 con una solución 0.079 mM de MobM en 20 mM de Tris pH=7.6, 500 mM de NaCl, 1% de glicerol, 1mM de EDTA y 1 mM de DTT, e incubando la mezcla durante 12 horas.

La purificación se realizó mediante cromatografía de exclusión molecular utilizando una columna Superdex 200 10/300 GL (figura 37) eluyendo la muestra con 20 mM de Tris pH=7.6, 500 mM de NaCl, 1% de glicerol, 1mM de EDTA y 1 mM de DTT.

Las fracciones de los tres picos principales se caracterizaron mediante SDS-PAGE revelado los geles con Coomasie y tinción de plata para la correcta visualización de la proteína y el oligonucleótido. El análisis de los geles permitió concluir que, dado que las fracciones 23 y 24 corresponden solamente a la muestra de proteína y la 29 al ADN la formación del complejo no tuvo lugar.

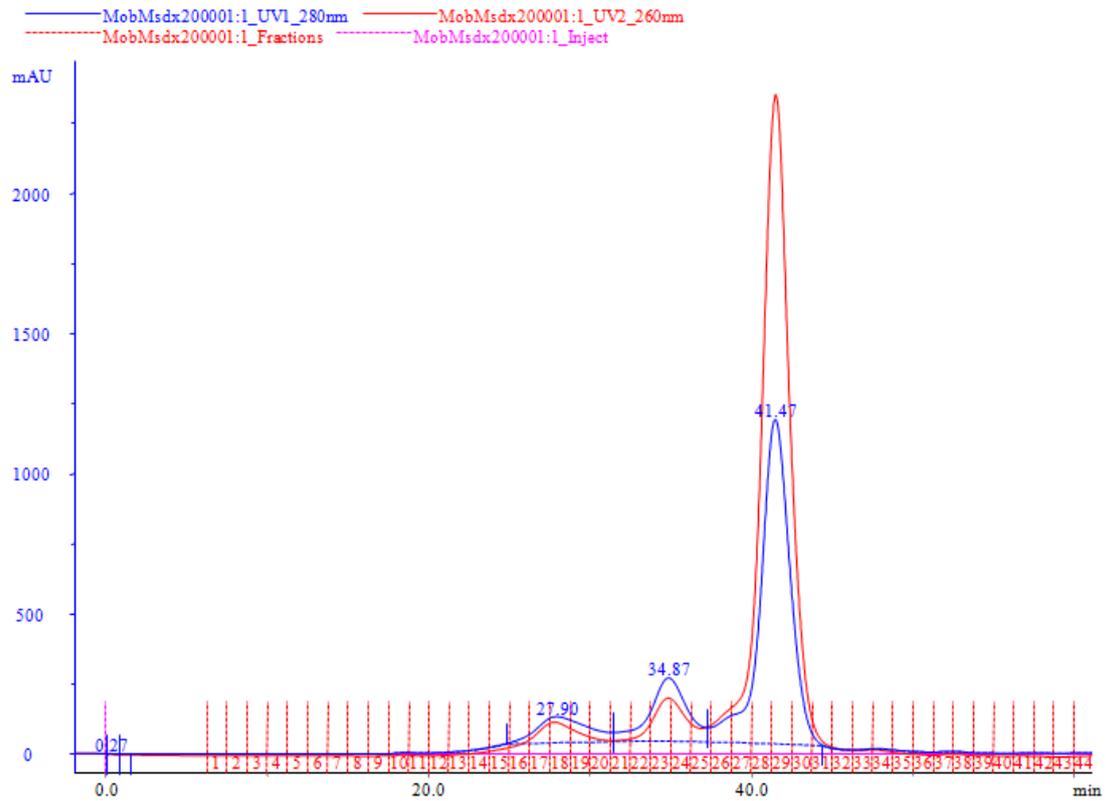


Figura 3.1. Purificación del complejo MobM-24ADN (IR-2) mediante cromatografía de exclusión molecular.

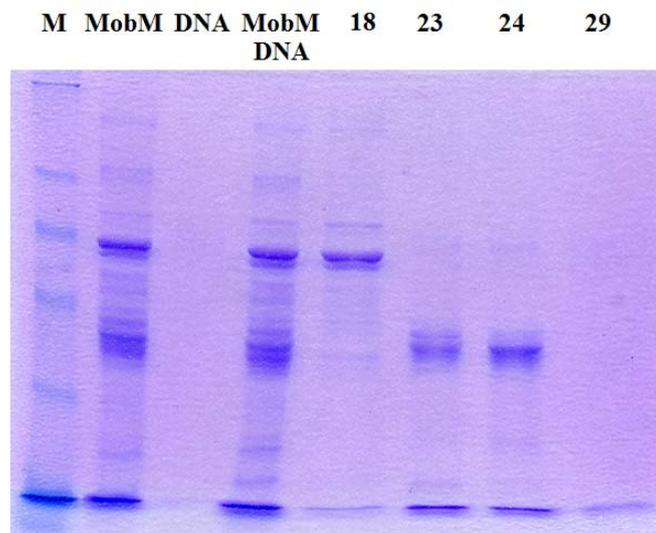


Figura 3.2. SDS-PAGE (15%) correspondiente a las fracciones 18, 23, 24 y 29 (Coomasie).

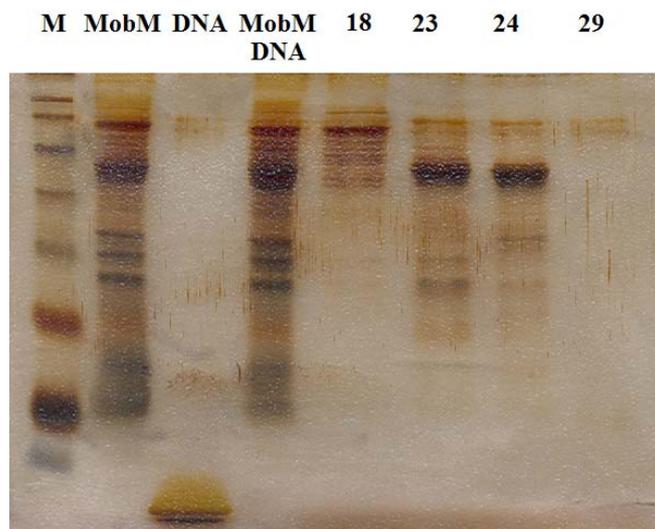


Figura 3.3. SDS-PAGE (15%) correspondiente a las fracciones 18, 23, 24 y 29 (tinción de plata).

Todos los esfuerzos posteriores dirigidos a favorecer la formación del complejo (disminución de la concentración de NaCl del tampón, incremento de la proporción de ADN, etc.) fueron infructuosos, demostrando la baja afinidad de la proteína por la conformación IR-2 a la hora de formar un complejo estable.

### 3.1.5.2. Preparación y purificación del complejo MobM-ADN32 (IR-3)

Una solución 1.7 mM del oligonucleótido 5'-CACACACTTTA TGAATATAAAGTATAGTGTGT-3' cuya secuencia corresponde al IR-3, se sometió al protocolo de templado. La mezcla proteína-ADN se realizó en una relación 1:1.2 adicionando al ADN una solución 0.074 mM de MobM en 20 mM de Tris pH=7.6, 500 mM de NaCl, 1% de glicerol, 1mM de EDTA y 1 mM de DTT, e incubando la mezcla durante 12 horas.

La purificación de la muestra se realizó mediante cromatografía de exclusión molecular utilizando una columna Superdex 200 10/300 GL, eluyéndola con el mismo tampón de la proteína.

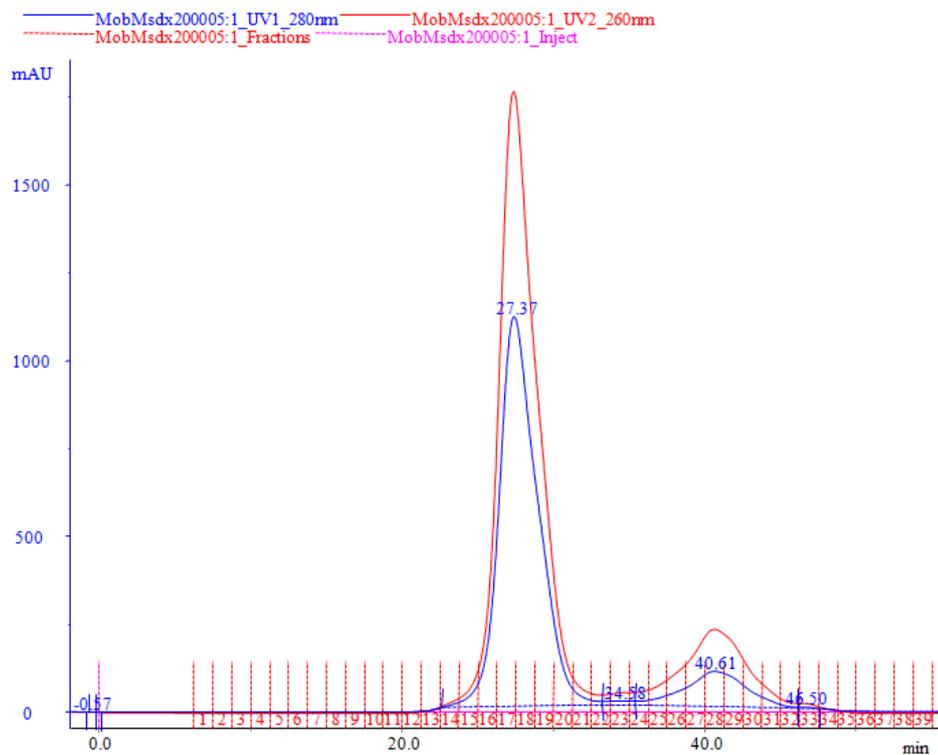


Figura 3.4. Purificación del complejo MobM-ADN32 (IR-3) mediante cromatografía de exclusión molecular.

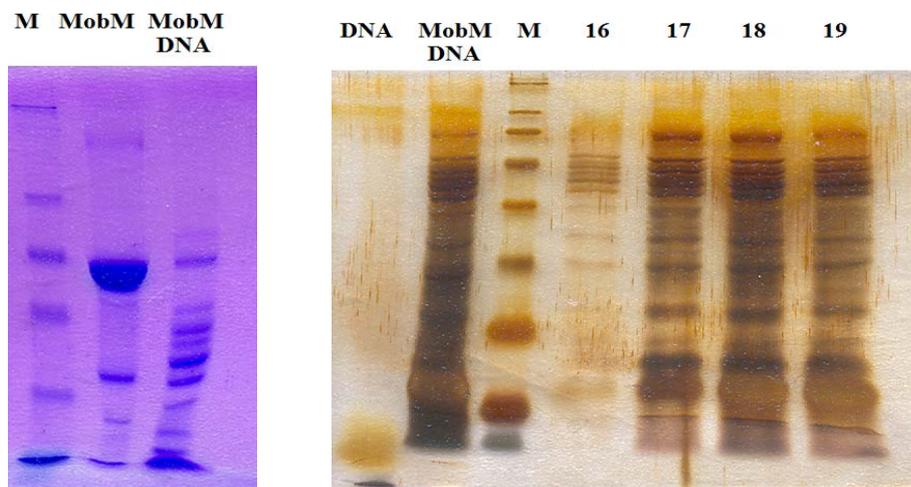


Figura 3.5. SDS-PAGE (15%) correspondiente a la proteína antes y después de la mezcla con el ADN (Coomasie) y las fracciones 16, 17, 18 y 19 de la cromatografía de exclusión molecular (tinción plata).

El elevado grado de degradación e inestabilidad del complejo obligaron a discontinuar el trabajo con el mismo.

### 3.1.5.3. Preparación y purificación del complejo MobM N198-ADN32

El oligonucleótido de 32 bases utilizado fue el mismo que en el caso anterior y para su preparación se procedió de manera idéntica a la ya descrita. La mezcla proteíno-ADN se realizó en la relación preestablecida, añadiendo al ADN una solución 0.102 mM de MobM N198 en 20 mM de Tris pH=7.6, 500 mM de NaCl, 1% de glicerol, 1mM de EDTA y 1 mM de DTT, e incubando la mezcla durante 12 horas.

La purificación del complejo se realizó mediante cromatografía de exclusión molecular utilizando una columna Superdex 200 10/300 GL y eluyendo la muestra con el tampón de la proteína.

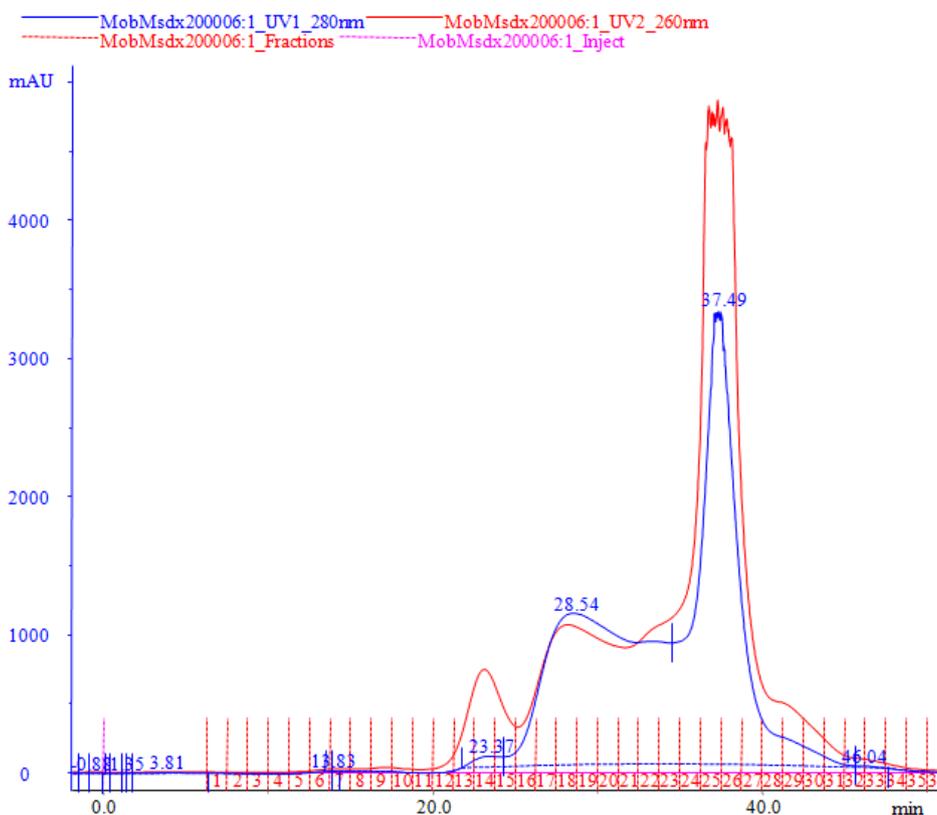


Figura 3.6. Purificación del complejo MobM N198-ADN32 (IR-3) mediante cromatografía de exclusión molecular.

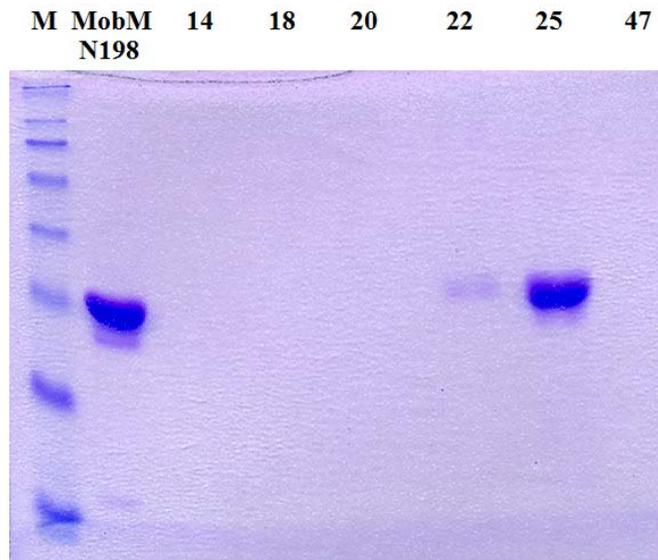


Figura 3.7. SDS-PAGE (15%) correspondiente a las fracciones 14, 18, 20, 22, 25 y 47 del cromatograma de la figura 3.6. (Coomasie).

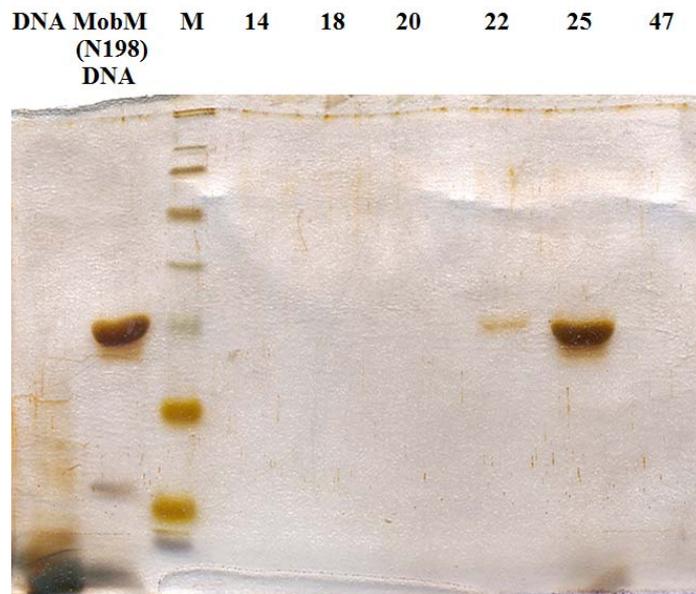


Figura 3.8. SDS-PAGE (15%) correspondiente a las fracciones 14, 18, 20, 22, 25 y 47 del cromatograma de la figura 3.6 (tinción plata).

Las fracciones 25, 26 y 27 correspondientes al complejo se concentraron, por centrifugación, hasta una concentración final de 5.3 mg/ml.

### 3.1.5.4. Preparación y purificación del complejo MobM N198-ADN26 (IR-3(-6))

Para el diseño del oligonucleótido de 26 bases, en las que fueron eliminadas seis bases de la secuencia IR-3 (5'-ACTTTATGAATATAAAGTATAGTGTG-3') se tuvo en cuenta la estructura del complejo TrwC-ADN. La eliminación de estas seis bases localizadas en el extremo 5' de la cadena de ADN, impide la formación de la segunda horquilla y convierte al extremo 3' en una cadena simple, tal como ocurre en el caso del oligonucleótido que une TrwC.

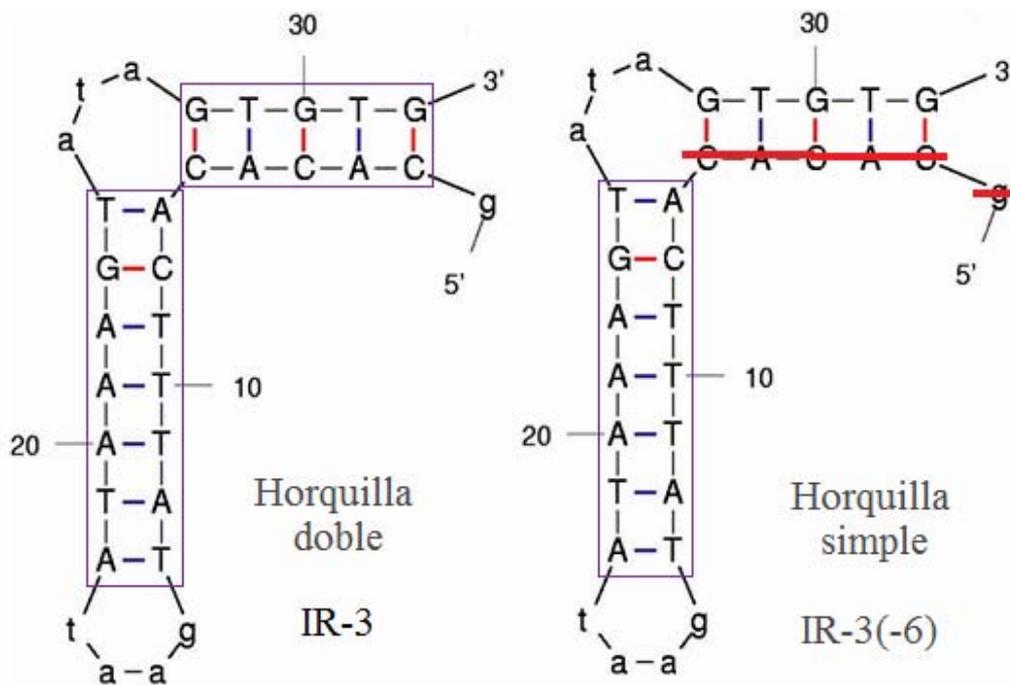


Figura 3.9. Esquema del diseño del oligonucleótido IR-3(-6) en que se muestra como la segunda horquilla es eliminada mediante la eliminación de las seis primeras bases de la secuencia IR-3.

La solución 1.7 mM del oligonucleótido se sometió al proceso de templado y luego se mezcló en relación 1:1.1 con una solución 0.17 mM de proteína en 500 mM de NaCl, 20 mM Tris pH=7.6, 1% de glicerol, 1mM EDTA y 1mM DTT.

La purificación se realizó mediante cromatografía de exclusión molecular y las fracciones 15, 16 y 17, correspondientes al complejo se concentraron a 8.5 mg/ml.

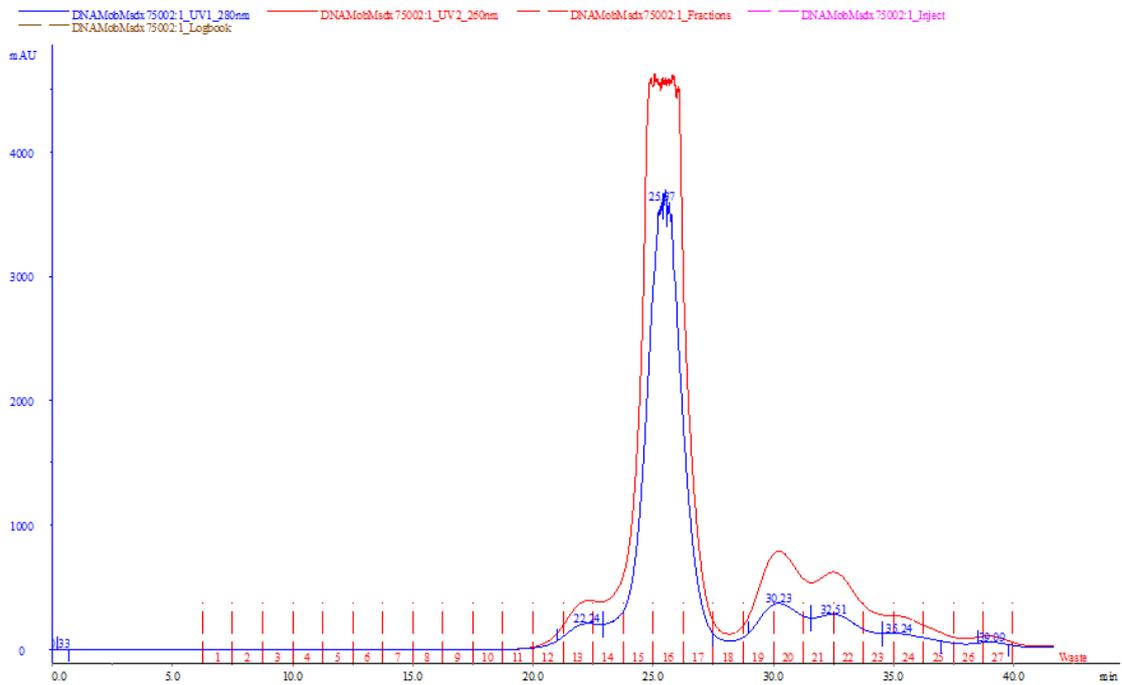


Figura 3.10. Purificación del complejo MobM N198-ADN26 (IR-3(-6)) mediante cromatografía de exclusión molecular.

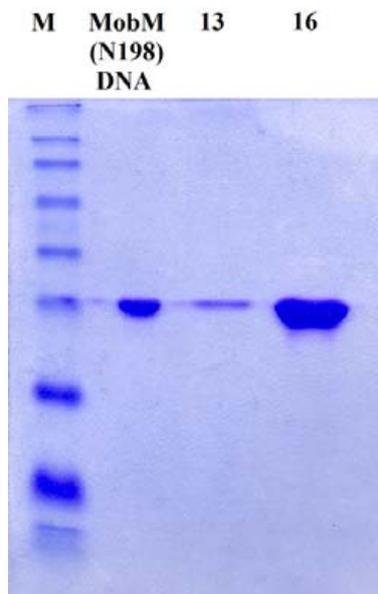


Figura 3.11. SDS-PAGE (15%) correspondiente a las fracciones 13 y 16 del cromatograma de la figura 3. 10.

### 3.1.5.5. Preparación y purificación del complejo SeMet-MobM N198-ADN26

Los complejos se prepararon y purificaron de manera idéntica a la descrita para la proteína nativa. Las fracciones 15, 16 y 17 de la cromatografía de exclusión molecular (figura 3.12), correspondientes al pico del complejo, se concentraron a 9 mg/ml.

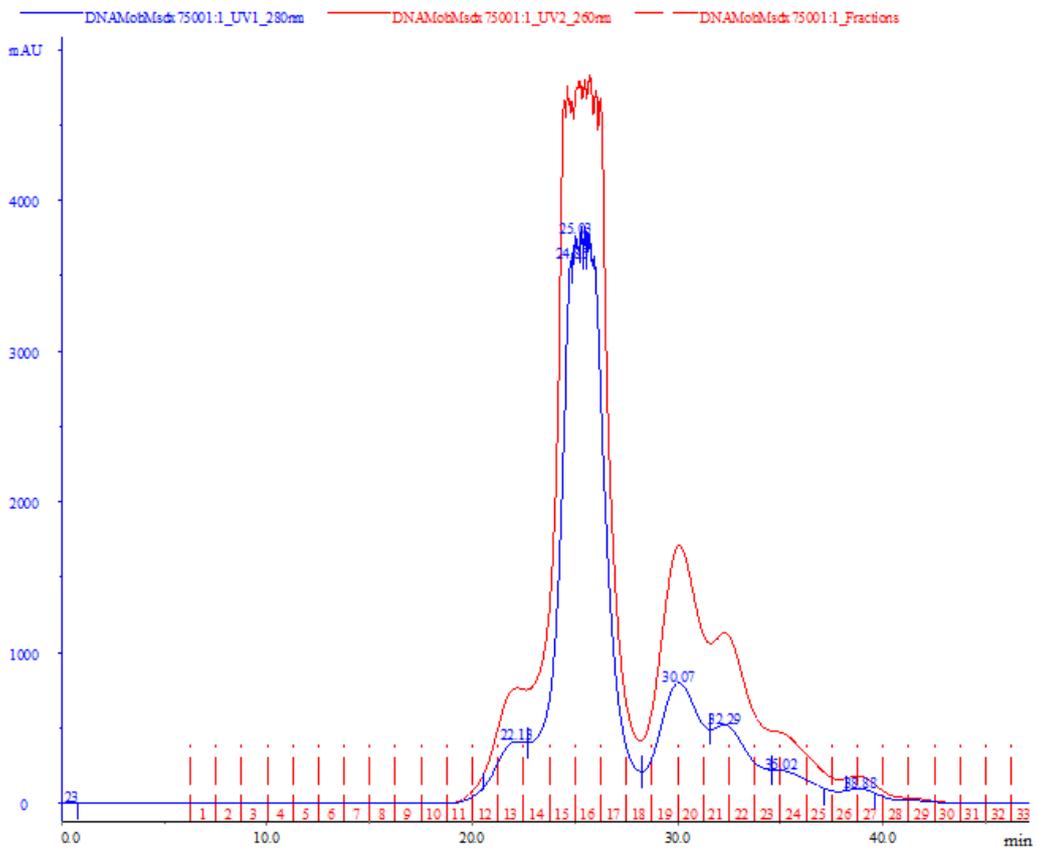
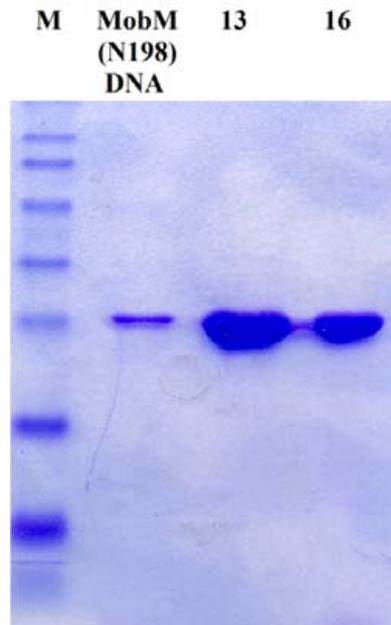
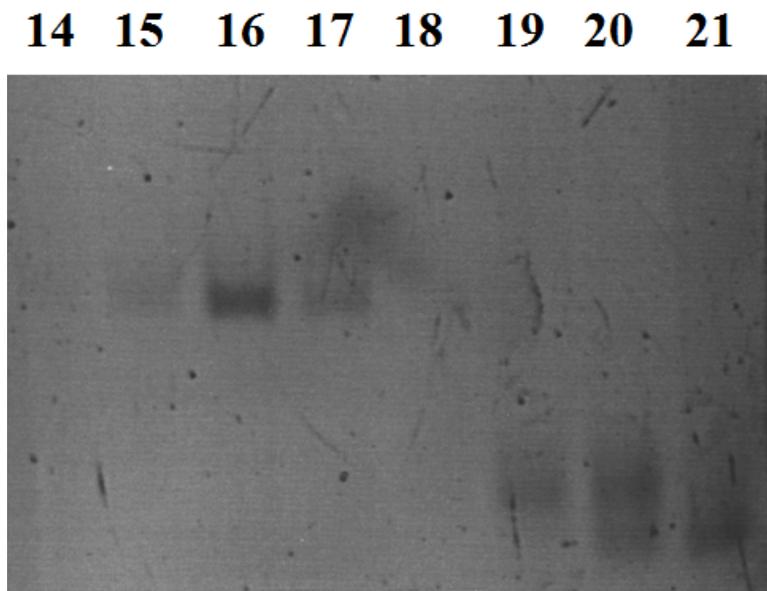


Figura 3.12. Purificación del complejo SeMet-MobM N198-ADN26 (IR-3(-6)) por cromatografía de exclusión molecular.



*Figura 3.13. SDS-PAGE (15%) correspondiente a las fracciones 13 y 16 de la cromatografía que se encuentra en la figura 3.12.*



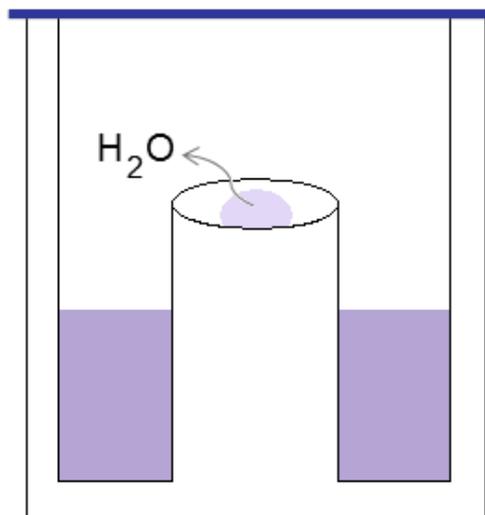
*Figura 3.14. Gel EMSA correspondiente a las fracciones 14 a la 21 del cromatograma de la figura 3.12.*

## 3.2. Obtención de los cristales

### 3.2.1. Cristalización

Los ensayos de cristalización se hicieron mediante la técnica de difusión de vapor. Esta técnica, que es una de las más empleadas a la hora de cristalizar macromoléculas, se basa en un principio físico-químico muy simple. Una gota, formada por la muestra a cristalizar y una solución conteniendo un agente precipitante, se deja equilibrar contra un reservorio que contiene solo la solución del agente precipitante (figura 3.15).

Debido a que la concentración de la solución del agente precipitante en la gota es menor que en el reservorio, la gota comienza a perder agua. La deshidratación de la gota produce un incremento en la concentración de la muestra a cristalizar y el equilibrio finalmente se alcanza cuando las concentraciones de precipitante del reservorio y de la gota son aproximadamente iguales.



$$[\text{ppte.}]_{\text{gota}} = \frac{[\text{ppte.}]_{\text{reservorio}}}{2}$$

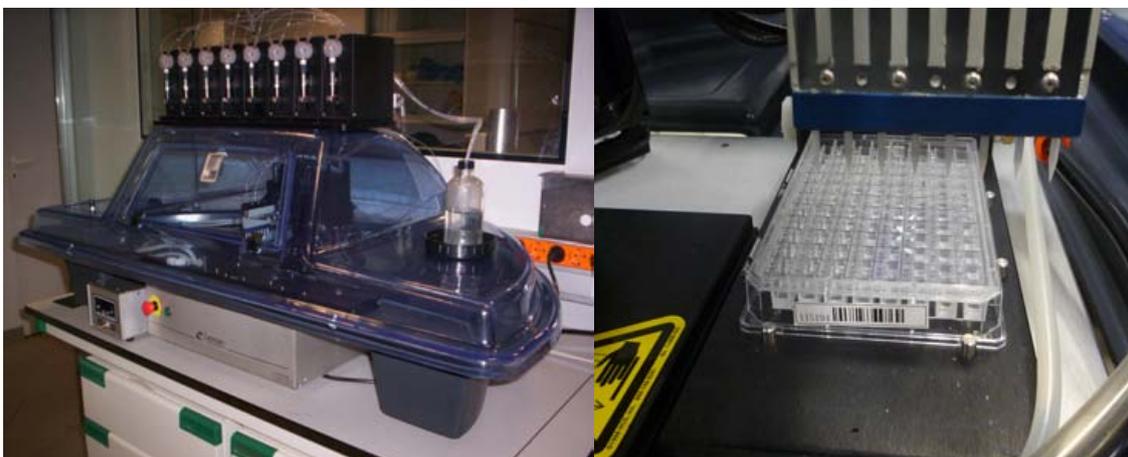
Figura 3.15. Dispositivo utilizado en el método de difusión de vapor para la cristalización de proteínas. En la medida que la gota tiende a alcanzar el equilibrio ésta va perdiendo gradualmente agua y aumenta la concentración de precipitante y proteína en la misma.

Dependiendo de la manera como se coloca la gota dentro del reservorio, la técnica de difusión de vapor tiene dos variantes: gota sentada o gota colgante. En el primer caso, el propio reservorio dispone de un soporte ubicado en el centro del mismo donde la gota se deposita y el reservorio se sella con una cinta adhesiva transparente. En el segundo caso, la gota se coloca en un cubre-objeto, a modo de tapa, sobre el reservorio y para que el sistema se mantenga sellado correctamente los bordes del reservorio se recubren con un material sellante.

### ***3.2.1.1. Búsqueda de las condiciones de cristalización***

La búsqueda de las condiciones iniciales se realizó en placas de cristalización de 96 pozos (CrystalQuick, Greiner), conteniendo en cada uno de ellos diferentes combinaciones de tampones, sales y agentes precipitantes. Estas soluciones fueron adquiridas en la casa comercial Hampton Research.

Las gotas de 200 nl (100 nl de proteína y 100 nl de buffer) fueron dispensadas automáticamente por el robot de cristalización Cartesian (Genomic Solutions) y las placas se hicieron por duplicado realizando los experimentos tanto a 20°C como a 4°C.



*Figura 3.16. (a) Robot Cartesian (Genomic Solutions) dispensador de nano-gotas y (b) detalle placa de cristalización Crystal Quick de 96 pozos (Greiner).*

El seguimiento de las gotas se efectuó mediante las imágenes tomadas por un robot Crystal Farm (Bruker AXS).



Figura 3.17. Robot de almacenaje y visualización de placas de cristalización Crystal Farm (Bruker AXS).

### 3.2.1.2. Optimización de los cristales

Una vez encontradas las condiciones iniciales de cristalización, y para obtener cristales de tamaño adecuado para las medidas de difracción de rayos X, se procedió al escalado de las gotas y a la optimización de las mismas.

Se utilizaron en este caso placas para gota sentada de 24 pozos (Cryschem Plate, Hampton Research). Las gotas fueron preparadas mezclando la proteína (1 $\mu$ l) y la solución del reservorio (1 $\mu$ ) en relación 1:1 y los experimentos se realizaron a 4°C o a 20°C de acuerdo a los resultados previos obtenidos en el paso anterior.

Las variables que se tuvieron en cuenta a la hora de la optimización de las condiciones de cristalización fueron: el pH del tampón, la concentración del agente precipitante, la concentración de las sales y/o de los aditivos.

### 3.2.2. Derivatización de los cristales con metales divalentes

Para lograr la incorporación de metales al sitio activo y permitir el estudio en detalle de su naturaleza y rol, se impregnaron los cristales en soluciones conteniendo diferentes metales divalentes de interés (Mg<sup>2+</sup>, Mn<sup>2+</sup>, Zn<sup>2+</sup>, Cu<sup>2+</sup>, Ca<sup>2+</sup>).

En primer lugar, se prepararon diluciones de sales de los metales en soluciones conteniendo idéntica formulación que las de la gota en las que los cristales crecieron. Una vez preparadas las soluciones, los cristales fueron retirados de la gota original y se transfirieron a una gota conteniendo las soluciones anteriormente mencionadas que se colocó en un nuevo pocillo conteniendo la condición de cristalización correspondiente y previamente equilibrado.

A manera de favorecer la incorporación de los metales en el cristal se hicieron, en paralelo, diferentes ensayos variando tanto el tiempo de inmersión de los cristales a la solución del metal (30 minutos, 1, 2, 5 y 12 horas), como la concentración de las soluciones de los mismos (1 mM, 5 mM y 10 mM).

### **3.2.3. Crio-protección**

Al ser expuestos a los rayos X los cristales sufren un importante daño, producto de la radiación, que afecta a su estructura interna y que causa la pérdida del poder de difracción, el aumento de la mosaicidad y una pérdida general de la calidad de los datos comprobable a través del aumento de los factores  $R_{\text{merge}}$  y  $B_{\text{Wilson}}$  (Murray and Garman, 2002). A nivel de la molécula de proteína, la radiación produce la ruptura de los puentes di-sulfuro, la pérdida de los grupos carboxilatos (Weik *et al.*, 2000) y la inducción de reacciones de oxido-reducción en los metal presentes.

Para minimizar estos efectos indeseados y alargar la vida útil del cristal para maximizar el tiempo total de exposición, los cristales son congelados en nitrógeno líquido. Durante el proceso de congelación es sumamente importante evitar la formación de micro-cristales de hielo, que interfieren con la difracción propia del cristal y producen la disrupción del mismo. Por este motivo, usualmente antes de ser congelado el cristal se sumerge en una solución que contiene un agente crio-protectante junto con el líquido madre. El agente crio-protectante se difundirá sobre la superficie del cristal e incluso en algunos casos penetrará al mismo, protegiéndolo y permitiendo el adecuado proceso de congelación.

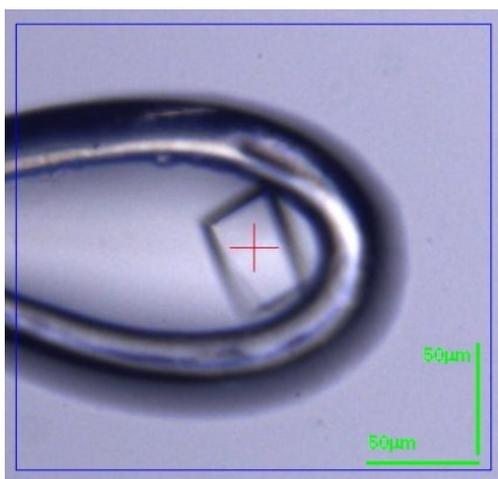
Pese a que existen diversos agentes crio-protectantes disponibles y ampliamente conocidos por su poder de crio-protección (glicerol, PEG, MPD, etc.), a priori no se

puede predecir cuál será el adecuado y en qué concentración añadirlo para evitar el deterioro de los cristales, por lo que esto debe testarse en cada caso particular.

En general, los crio-protectantes más comunes son efectivos en una concentración entre 15% y 25% (v/v). Las pruebas de crio-protección se realizan siempre primero a las concentraciones más bajas posibles, incrementando esta gradualmente a los efectos de utilizar la mínima concentración de agente crio-protectante.

### ***3.2.4. Preparación final de los cristales***

Para poder ser transportados y manipulados fácilmente los cristales son retirados (pescados) de la gota de cristalización con la ayuda de un pequeño lazo de nylon unido a una base metálica magnética conocido con el nombre de “loop” (figura 3.18).



*Figura 3.18. Imagen de un cristal de proteína montado en un loop de nylon.*

Luego de pescados los cristales son colocados durante breves segundos en una gota que contiene la solución del reservorio y el agente crio-protectante en la concentración adecuada. Por último, el “loop” se sumerge rápidamente en nitrógeno líquido y se lo cubre con un capuchón plástico que lo protege, manteniéndolos a temperatura de nitrógeno líquido en todo momento.

### 3.2.5. Complejos TrwC-ADN

#### 3.2.5.1. Cristalización del complejo TrwC-ADN25

Los cristales fueron obtenidos por difusión de vapor a 4°C con el método de la gota colgante utilizando como agente precipitante 20% (w/v) PEG MME 2000, 0.4 M de sulfato de amonio y 0.1 M de acetato de sodio a pH=4.6 (Guasch *et al.*, 2003). Las gotas se prepararon mezclando 2  $\mu$ l de solución del complejo (6.8 mg/ml) con 1  $\mu$ l de solución precipitante y los cristales alcanzaron su tamaño óptimo al cabo de aproximadamente 10 días.

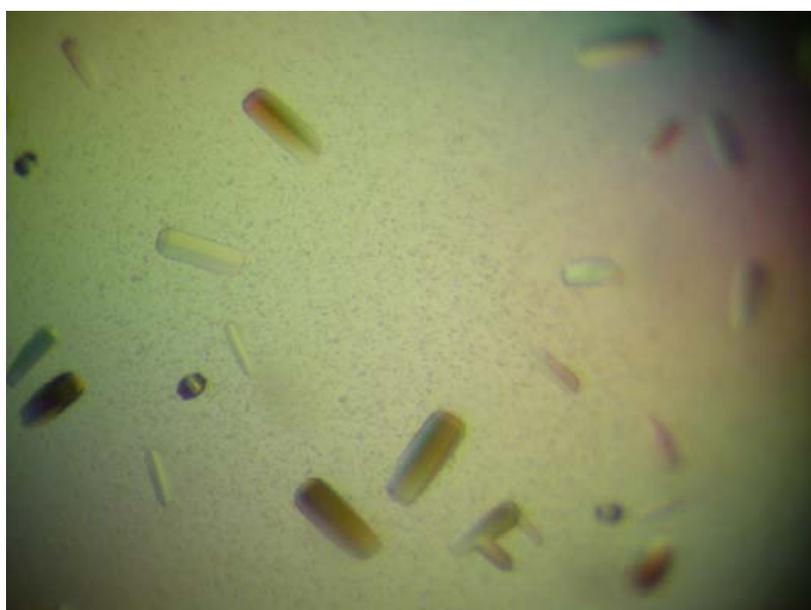


Figura 3.19. Cristales del complejo TrwC-ADN25.

#### 3.2.5.2. Derivatización del complejo TrwC-ADN25 con metales divalentes

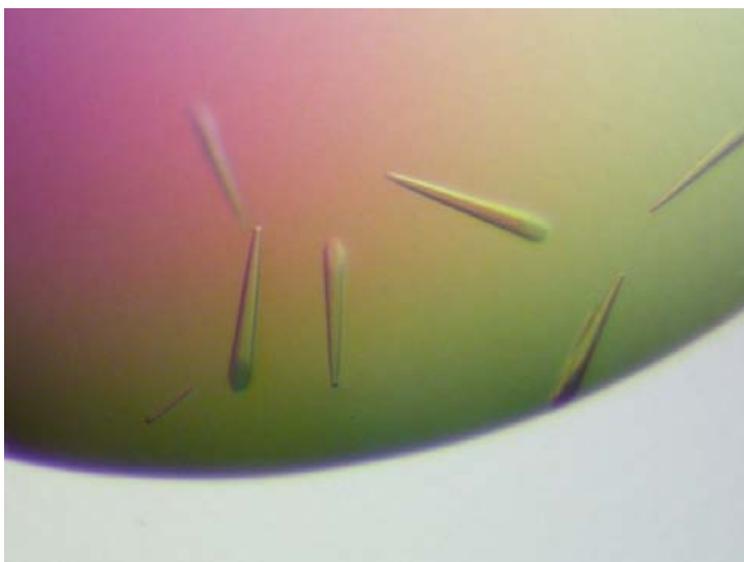
Los derivados de metales divalentes se prepararon colocando los cristales del complejo TrwCN293-ADN25 en gotas conteniendo la solución de precipitante y: (i) 5 mM de ZnSO<sub>4</sub> y 10 mM de MgCl<sub>2</sub>; (ii) 5 mM de ZnSO<sub>4</sub> y 10 mM de MnCl<sub>2</sub>; (iii) 5 mM MnCl<sub>2</sub>; (iv) 5 mM CuCl<sub>2</sub> y (v) 5 mM NiCl<sub>2</sub>, siendo las concentraciones indicadas las finales en la gota.

Los cristales se incubaron durante 24 horas en estas condiciones para luego pescarlos y congelarlos en nitrógeno líquido, utilizando 20% de glicerol como agente crio-protector.

### ***3.2.5.3. Cristalización del complejo TrwC N293<sup>Y18F</sup>-ADN27***

Los cristales fueron obtenidos por difusión de vapor a 4°C por el método de la gota colgante, mezclando 2 µl de solución del complejo (6.9 mg/ml) con 1 µl de solución precipitante conteniendo 32% (w/v) PEG MME 2000, 0.3 M de sulfato de amonio y 0.1 M de acetato de sodio a pH=4.6.

Al cabo de 27 días, cuando los cristales alcanzaron su máximo tamaño, se los pescó y congeló en nitrógeno líquido utilizando 20% de glicerol como agente crio-protectante.



*Figura 3.20. Cristales del complejo TrwC-ADN27.*

### **3.2.6. Complejos MobM N198-ADN**

Los intentos infructuosos de cristalizar la forma entera de la proteína MobM en complejo con las secuencias IR-2 e IR-3 de ADN conllevaron a la decisión de trabajar, tal como se hizo en el caso de la proteína TrwC, con una forma truncada de MobM que preservara la actividad relaxasa localizada en el dominio N-terminal de la misma. El constructo truncado, que comprende los residuos 2 al 198, se diseñó en base a la predicción de la secuencia secundaria de la proteína (figura 1.14), teniendo en cuenta las zonas de desorden y la eliminación del dominio “coiled-coil”.

Asimismo, dada la extremadamente baja afinidad que demostró tener la proteína por la cadena de ADN de secuencia IR-2, se decidió continuar el trabajo con oligonucleótidos cuya secuencia coincidiera total o parcialmente con la secuencia IR-3.

#### **3.2.6.1. Cristalización del complejo MobM N198-ADN32**

Las condiciones de cristalización se exploraron mediante los “screens” comerciales de la casa Hampton Research: Crystal Screen I, Crystal Screen II, Index, Natrix y Protein-DNA.

Las nano-gotas de 200 nl (relación proteína/precipitante 1:1) fueron dispensadas con el robot Cartesian y las placas se incubaron a 4°C. Las gotas fueron inspeccionadas durante tres meses de manera automática con el robot de visualización Crystal Farm y posteriormente de forma manual.

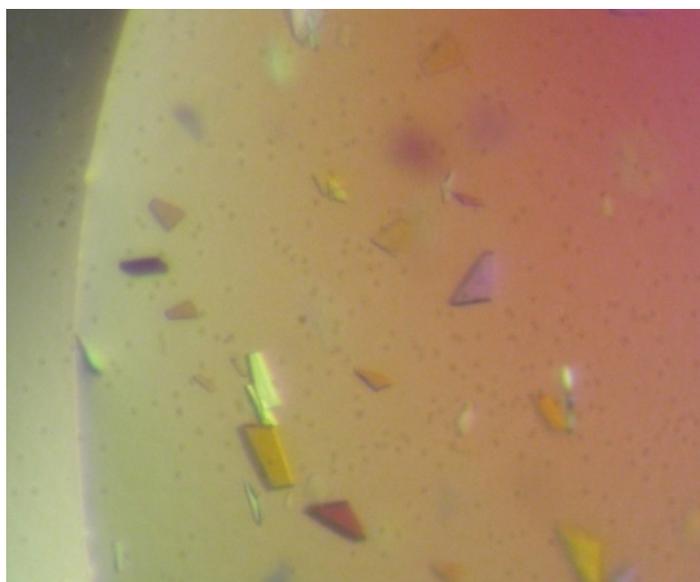
En ninguna de las condiciones testadas se obtuvieron cristales o precipitados cristalinos adecuados como para intentar una posterior optimización.

#### **3.2.6.2. Cristalización del complejo MobM N198-ADN26**

La búsqueda de las condiciones de cristalización se realizó con los “screens” comerciales de Hampton Research: Crystal Screen I, Crystal Screen II, Index, Natrix y Protein-DNA, dispensando las gotas con el robot Cartesian (200 nl, en relación 1:1) e incubando las placas a 4°C.

Al cabo de 15-20 días, diversas condiciones de cristalización fueron detectadas: 10-20% de PEG de diversos pesos moleculares (1000, 1500, 2000, 3350, 6000, 8000), en presencia de 0.1 M de acetato de sodio pH=4.6 y 10% de isopropanol en presencia de 0.1 M de acetato de sodio pH=4.6.

Estas condiciones fueron escaladas a micro-gota en placas de 24 pozos (1µl de proteína + 1µl de precipitante) e incubadas a 4°C. Su optimización final permitió obtener al cabo de 30-45 días varias condiciones de cristalización con cristales apropiados para realizar las medidas de difracción (figura 3.21).



*Figura 3.21. Cristales del complejo MobM NI98-ADN26.*

Las diferentes condiciones de cristalización obtenidas se detallan en la tabla 3.1 y se pueden agrupar en tres formas cristalinas diferentes.

	CONDICIONES DE CRISTALIZACIÓN
Cristales (i)	10% PEG 1500, 0.1 M NaAc pH=4.6 10% PEG 3000, 0.1 M NaAc pH=4.6 10% PEG 6000, 0.1 M NaAc pH=4.6
Cristales (ii)	10% Isopropanol, 0.1 M NaAc pH=4.6
Cristales (iii)	8% PEG 1000, 0.1 M NaAc pH=4.6 8% PEG 4000, 0.1 M NaAc pH=4.6 8% PEG 8000, 0.1 M NaAc pH=4.6

Tabla 3.1. Condiciones de cristalización del complejo MobM N198-ADN26.

Los cristales se prepararon para la recogida de los datos de difracción congelándolos en nitrógeno líquido, utilizando 25% de glicerol como agente crioprotector.

### 3.2.6.3. Cristalización del complejo SeMetMobM N198-ADN26

Los cristales de los complejos SeMetMobM N198-ADN26 se obtuvieron en las mismas condiciones que en el caso de la proteína nativa (tabla 3.1), encontrándose, asimismo, una nueva condición de cristalización: 60% MPD, 0.2M MgCl<sub>2</sub> y 0.1 M de NaAc a pH=4.6, que representa la cuarta forma cristalina (iv) obtenida.

## 3.3. Recogida y procesamiento de los datos de difracción

### 3.3.1. Recolección de datos

Las recolecciones de datos de difracción se realizaron en las diferentes líneas de macromoléculas, BM16, ID14-4 e ID23-1, del sincrotrón europeo ESRF.

La línea de macromolécula prototípica consta de tres sectores básicos: el sector donde se localizan los elementos relacionados con la óptica del sistema, la cabina donde se ubica el difractómetro y la sala de control, desde donde se opera mediante ordenadores los diferentes instrumentos (figura 3.22).

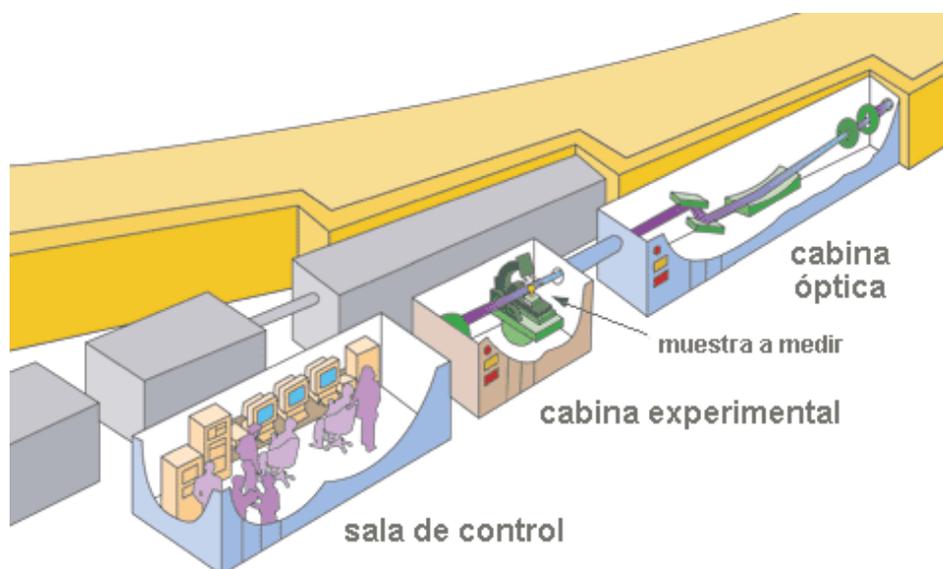
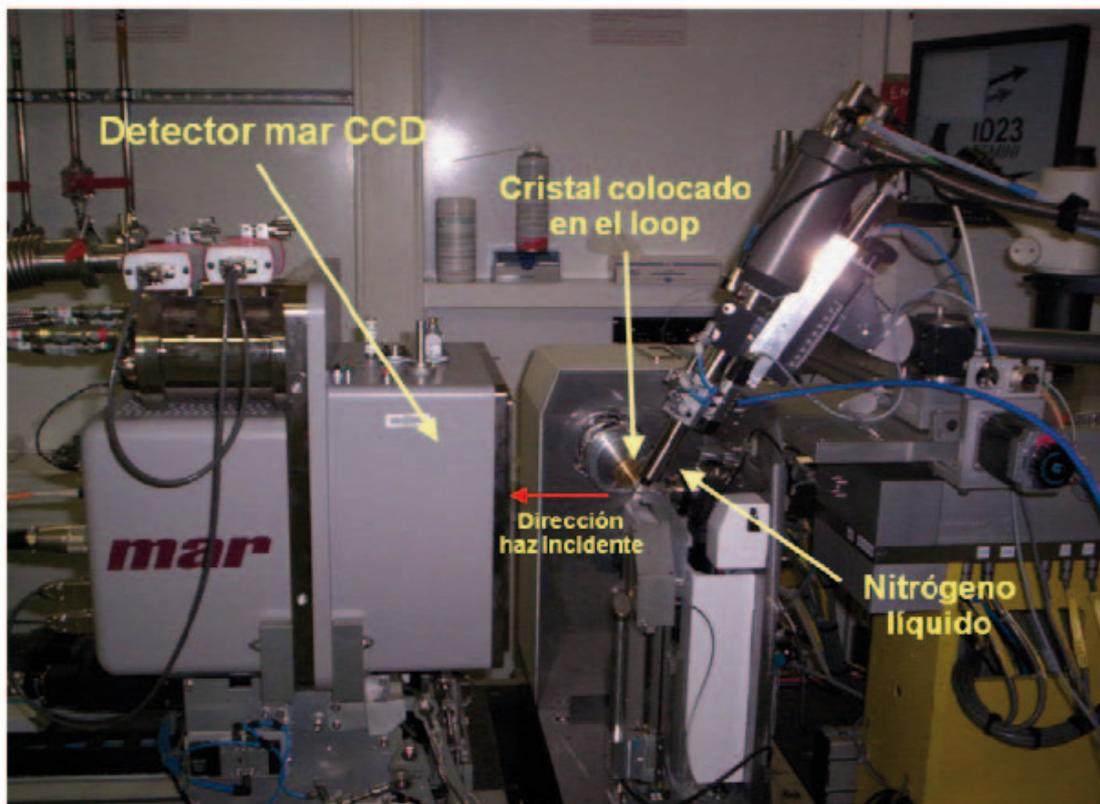


Figura 3.22. Línea de macromoléculas prototípica del ESRF (adaptación de la figura original de [www.esrf.eu](http://www.esrf.eu))

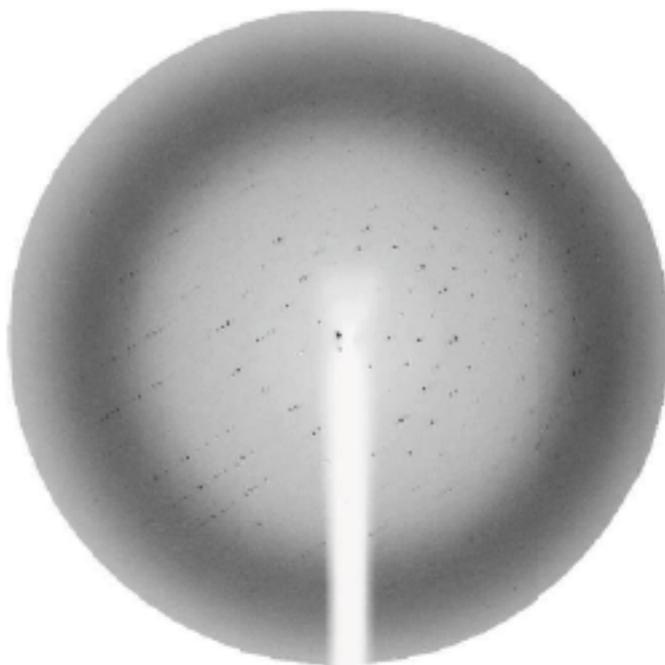
El cristal a medir se coloca en el goniómetro del difractómetro de manera tal que es rotado en la dirección perpendicular a la del haz de rayos X incidente ( $\Phi$ ) mientras es irradiado (figura 3.23).

La rotación típicamente se realiza en intervalos de  $1^\circ$  y los haces difractados son detectados y registrados por un detector de tipo CCD. El detector es colocado perpendicular a la dirección del haz de rayos X incidente, a una distancia de la muestra que depende de la resolución de los cristales.

Los patrones de difracción se almacenan como imágenes. Dichas imágenes presentan manchas de intensidad y posición características que dependen del contenido de la celda unidad y de las dimensiones de la celda y el grupo espacial del cristal, respectivamente (figura 3.24).



*Figura 3.23. Difractómetro utilizado en la línea de macromoléculas ID23-1 del ESRF. Se indica la posición de la muestra en relación con el haz incidente de rayos X y el detector de área (Mar CCD).*



*Figura 3.24. Imagen de difracción típica donde se observan como las reflexiones se distribuyen de acuerdo a un patrón que depende de la celda unidad y del grupo espacial en el cual ha cristalizado la proteína.*

El número total de grados que se debe girar el cristal, y por ende el número de imágenes de difracción a coleccionar, depende del grupo espacial. Así, los grupos espaciales de menor simetría (ej.  $P2$ ,  $P2_1$ ) requerirán recolecciones más extensas que los grupos espaciales de mayor simetría (ej.  $P2_12_12_1$ ).

### **3.3.2. Procesamiento de los datos**

El procesamiento de los datos consta de tres pasos: el indexado, la integración y el escalado, existiendo programas específicos que permiten realizar los cálculos requeridos en estos tres pasos.

El indexado es el primer paso para la extracción de la información contenida en las imágenes de difracción e implica asignar los correspondientes índices de Miller a cada una de las reflexiones observadas en la misma.

Los índices de Miller son tres números enteros  $hkl$  que identifican familias de planos paralelos que pueden definirse en el cristal y que de acuerdo a la construcción geométrica de W. H. y W. L. Bragg están asociados con la difracción de la red cristalina.

Si consideramos una familia de planos cuya distancia interplanar es  $d$ , si llamamos  $\theta$  al ángulo que forman los haces de rayos X incidentes (y los reflejados) respecto a los planos de la familia, para que exista una interferencia constructiva y se produzca difracción, la diferencia de caminos recorridos entre los haces incidentes (y los haces reflejados) tiene que ser un número entero de longitudes de onda  $\lambda$  (figura 3.25). Esto puede expresarse matemáticamente mediante la siguiente relación:  $n\lambda = 2d \sin\theta$ , que se conoce con el nombre de Ley de Bragg.

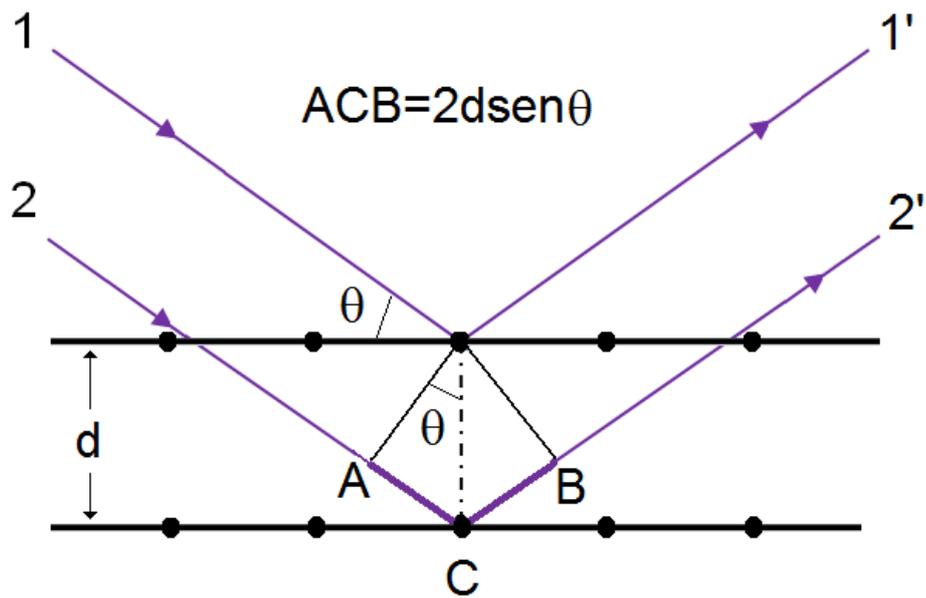


Figura 3.25. Reflexión de los rayos X en los planos cristalinos. La diferencia de caminos entre los haces 1-2 y 1'-2' está dada por la suma de las magnitudes de los segmentos AC y CB ( $2d\sin\theta$ ).

Para el indexado se utilizó el programa MOSFLM (Leslie, 1992) que calcula, en primer lugar, la matriz de orientación del cristal, y en función de ella establece los posibles grupos espaciales, la celda unidad y asigna los índices de Miller.

Una vez que es posible asociar cada una de las reflexiones con sus correspondientes índices de Miller, se procede a la integración. La integración permite calcular la intensidad de cada una de las reflexiones observadas. De esta manera, el archivo final contiene una lista que contiene la información básica necesaria para posteriores procesamientos: la intensidad y los índices de Miller asociados.

Debido a que durante la recogida de datos la intensidad del haz de rayos X varía, los datos deben ser corregidos y puestos en la misma escala. Este procedimiento se conoce con el nombre de escalado de los datos. Para el mismo se empleó el programa SCALA (Evans, 1993) de la suite de programas de CCP4 (Collaborative Computational Project, 1994).

### 3.3.3. Complejos TrwC-ADN

#### 3.3.3.1. Recolección de datos de los complejos TrwC-ADN25-M<sup>2+</sup>

Las medidas de difracción se realizaron en las líneas de macromoléculas del ESRF. Los cristales (i) fueron medidos en la línea ID14-4 a la energía correspondiente al pico de absorción del Zn (25ZNpkMG) y a 24 eV por debajo de éste (25ZNrmMg). El valor experimental de la energía de pico de absorción del Zn se determinó mediante un espectro de fluorescencia en un rango en torno al valor teórico de 9.661 keV (figura 3.26).

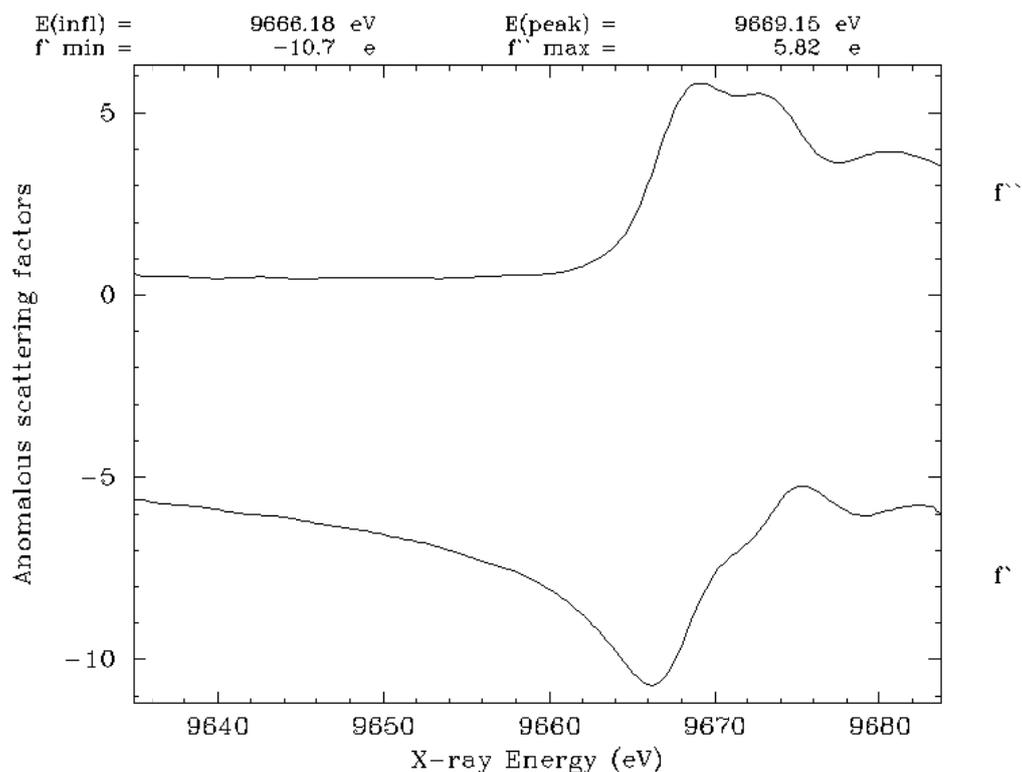


Figura 3.26. Perfil experimental de absorción del Zn en el que se indican los valores de los factores de dispersión anómala  $f'$  y  $f''$  a la energía del pico y del punto de inflexión.

Los datos de los cristales (ii) y (iii) se recogieron en la línea BM16, a la energía del pico de absorción del Zn (25ZNpkMN) y la del Mn (25ZNMNpk), los primeros; y a la energía del pico de absorción del Mn (set de datos 25MN), los segundos. Las energías de los picos de absorción se determinaron de manera análoga al caso anterior.

Los cristales (iv) se midieron en la línea ID23-1, a una energía mayor a la del valor teórico de pico de absorción del Cu (25CU) y los cristales (v) en la línea BM14, a una energía mayor al valor teórico del pico de absorción del Ni (25NI). Los detalles de las recolecciones se resumen en la tabla 3.2.

Conjunto de datos	25ZNpkMG	25ZNrmMG	25ZNpkMN	25ZNMNpk
Celda unidad (Å)	a=b=90.70 c=204.22	a=b=90.75 c=204.36	a=b=90.53 c=203.77	a=b=90.69 c=204.22
Grupo espacial	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22
Longitud onda (Å)	1.2820	1.2852	1.89223	1.2823
Rango resolución (Å) <sup>a</sup>	19.83-2.7 (2.85-2.7)	19.84-2.7 (2.85-2.7)	45.27-2.90 (3.06-2.90)	1.37-3.16 (3.16-3.0)
Nº reflexiones observadas <sup>a</sup>	169473 (24851)	146870 (21542)	130924 (18815)	120010 (17326)
Nº reflexiones independientes <sup>a</sup>	14370 (2037)	14393 (2039)	11442 (1598)	10633 (1496)
Compleitud <sup>a</sup>	99.9 (100)	99.8 (100)	98.9 (98.0)	100 (100)
Rmerge <sup>a,b</sup>	9.0 (22.4)	5.9 (15.6)	9.9 (36.1)	14.1 (42.7)
I/σ(I) <sup>a</sup>	5.7 (3.2)	9.1 (4.6)	7.1 (2.0)	4.8 (1.7)
Multiplicidad <sup>a</sup>	11.8 (12.2)	10.2 (10.6)	11.4 (11.8)	11.3 (11.6)
Compleitud anómala <sup>a</sup>	99.9 (100)	99.9 (100)	99.4 (98.6)	100 (100)
Multiplicidad anómala <sup>a</sup>	6.5 (6.6)	5.6 (5.6)	6.4 (6.4)	6.4 (6.3)
Conjunto de datos	25MN	25CU**	25NI***	
Celda unidad (Å)	a=b=90.24 c=202.08	a=b=91.09 c=205.55	a=b=92.73 c=208.86	
Grupo espacial	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22	
Longitud onda (Å)	1.8924	0.97855	1.00100	
Rango resolución (Å) <sup>a</sup>	43.90-2.6 (2.74-2.6)	24.68-2.5 (2.64-2.5)	43.9-2.29 (2.4-2.29)	
Nº reflexiones observadas <sup>a</sup>	171348 (25127)	386179 (56283)	114701 (13061)	
Nº reflexiones independientes <sup>a</sup>	15710 (2226)	18252 (2592)	28857 (3506)	
Compleitud <sup>a</sup>	99.9 (99.8)	99.9 (100)	98.8 (98.8)	
Rmerge <sup>a,b</sup>	12.7 (45.5)	10.4 (51.6)	4.9 (14.6)	
I/σ(I) <sup>a</sup>	4.6 (1.5)	5.6 (1.4)	10.2 (4.8)	
Multiplicidad <sup>a</sup>	10.9 (11.3)	21.2 (21.7)	4.7 (4.1)	
Compleitud anómala <sup>a</sup>	100 (100)	---	98.3 (88.9)	
Multiplicidad anómala <sup>a</sup>	5.9 (6.0)	---	2.5 (2.2)	

Tabla 3.2. Estadísticas de las recolecciones de datos de los complejos TrwC-ADN25. <sup>a</sup> Valores en paréntesis para el rango de mayor resolución. <sup>b</sup>  $R_{merge} = [\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)] \times 100$ .

### 3.3.3.2. Recolección de datos del complejo TrwC<sup>Y18F</sup>-ADN27

La recolección de datos se realizó en la línea de macromoléculas ID14-4 del ESRF. Los detalles se resumen en la tabla 3.3.

	TrwC-ADN27
Celda unidad (Å)	a=b=148.42 c=75.62
Grupo espacial	P6 <sub>5</sub>
Longitud onda (Å)	0.97730
Rango resolución (Å) <sup>a</sup>	25.0-2.7 (2.9-2.7)
Nº reflexiones observadas <sup>a</sup>	165382 (22961)
Nº reflexiones independientes <sup>a</sup>	27451 (3996)
Completitud <sup>a</sup>	99.7 (100)
Rmerge <sup>a,b</sup>	10.9 (41.3)
I/σ(I) <sup>a</sup>	5.4 (1.5)
Multiplicidad <sup>a</sup>	6.0 (5.7)
Completitud anómala <sup>a</sup>	99.0 (99.9)
Multiplicidad anómala <sup>a</sup>	3.1 (2.9)

Tabla 3.3. Recolección de datos del complejo TrwC-ADN27. <sup>a</sup> Valores en paréntesis corresponden al rango de mayor resolución. <sup>b</sup>  $R_{merge} = [\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)] \times 100$ .

El complejo cristaliza en una celda unidad hexagonal en el grupo espacial P6<sub>5</sub> con dos moléculas del complejo por unidad asimétrica.

### 3.3.4. Complejos MobM N198-ADN

#### 3.3.4.1. Recolección de datos del complejo MobM N198-ADN26

Las medidas de difracción se llevaron a cabo en las líneas de macromoléculas ID14-4 e ID23-2 del ESRF. En la tabla 3.4 se resume los principales parámetros para las tres formas cristalinas.

Conjunto de datos	Forma cristalina (i) nativa*	Forma cristalina (ii) nativa	Forma cristalina (iii) nativa **
Beamline	ID14-4	ID23-2	ID14-4
Celda unidad (Å)	a=54.40 b=65.25 c=76.16	a=65.79 b=54.48 c=77.23 $\beta=95.4^\circ$	a=53.00 b=55.53 c=56.25 $\beta=107.43^\circ$
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub>	P2 <sub>1</sub>
Longitud onda (Å)	1.0065	0.8726	1.07150
Rango resolución (Å) <sup>a</sup>	76.25-2.10 (2.21-2.10)	30.00- 2.50 (2.64-2.50)	55.56-2.30 (2.42-2.30)
Nº reflexiones observadas <sup>a</sup>	94475 (13927)	78658 (11506)	55788 (8226)
Nº reflexiones independientes <sup>a</sup>	16388 (2338)	19093 (2752)	13820 (1983)
Compleitud <sup>a</sup>	99.8 (100.0)	99.9 (100.0)	98.5 (97.5)
Rmerge <sup>a,b</sup>	8.0 (34.8)	19 (65.3)	8.0 (30.4)
I/σ(I) <sup>a</sup>	15.6 (5.2)	7.8 (1.9)	15.5 (6.2)
Multiplicidad <sup>a</sup>	5.8 (6.0)	4.1 (4.2)	4.0 (4.1)

Tabla 3.4. Estadística de la recolección de datos del complejo MobMN198-ADN26. <sup>a</sup> Valores en paréntesis para el rango de mayor resolución. <sup>b</sup>  $R_{merge} = [\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)] \times 100$ . \*Los datos corresponden a cristales obtenidos a partir de la condición de cristalización de PEG6000. \*\*Los datos corresponden a cristales obtenidos a partir de la condición de cristalización de PEG1000.

Los cristales de la forma (i) son ortorrómbicos y pertenecen al grupo espacial P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> con una molécula por unidad asimétrica, siendo el coeficiente de Matthews 2.11 Å<sup>3</sup>Da<sup>-1</sup> (Matthews, 1968) y el contenido de solvente 48%.

Los cristales de las formas (ii) y (iii) son monoclinicos y pertenecen al grupo

espacial  $P2_1$ . La primera de ellas cristaliza con dos moléculas de complejo por unidad asimétrica, un coeficiente de Matthews de  $2.15 \text{ \AA}^3 \text{ Da}^{-1}$  y un contenido de solvente del 49 %, mientras que la segunda lo hace con una sola molécula en la unidad asimétrica, un coeficiente de Matthews  $2.49 \text{ \AA}^3 \text{ Da}^{-1}$  y 56 % de solvente.

### 3.3.4.2. Recolección de datos del complejo *SeMet-MobM N198-ADN26*

La recolección de datos, que permitió la posterior resolución estructural, se realizó en la línea ID14-4 del ESRF con cristales ortorrómbicos correspondientes a la forma cristalina (i) midiendo un único conjunto de datos a la longitud de onda del pico de absorción del selenio.

La energía del pico de absorción del selenio se determinó experimentalmente a partir del espectro de fluorescencia de los cristales (figura 3.27), siendo  $12.6597 \text{ keV}$  y los factores de dispersión atómica  $f' -7.29$  y  $f'' 6.25$ .

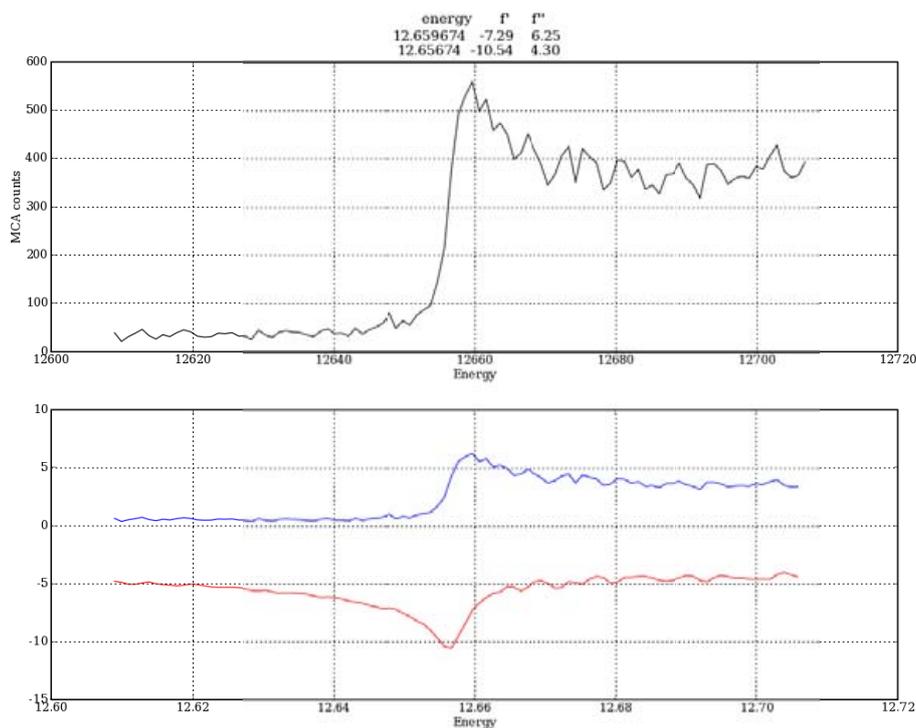


Figura 3.27. Perfil experimental de absorción del Se en el que se indican los valores de los factores de dispersión anómala  $f'$  y  $f''$  a la energía del pico y del punto de inflexión.

Asimismo se recogieron datos de la forma cristalina (iv) en la línea ID23-2. Los

principales parámetros relacionados con ambas colecciones de datos se resumen en la siguiente tabla (tabla 3.5).

Conjunto de datos	Forma cristalina (i) SeMet*	Forma cristalina (iv) SeMet
Celda unidad (Å)	a=54.50 b=65.19 c=76.99	a=43.57 b=52.85 c=56.04 $\beta=95.88^\circ$
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub>
Longitud onda (Å)	0.97950	0.8726
Rango resolución (Å) <sup>a</sup>	30.00-2.60 (2.74-2.60)	55.73-2.00 (2.11-2.00)
Nº reflexiones observadas <sup>a</sup>	51817 (7669)	70243 (7909)
Nº reflexiones independientes <sup>a</sup>	8878 (1270)	17056 (2268)
Completitud <sup>a</sup>	99.8 (100.0)	98.6 (90.7)
Rmerge <sup>a,b</sup>	7.7 (27.7)	9.3 (43.6)
I/σ(I) <sup>a</sup>	15.8 (5.9)	10.8 (2.4)
Multiplicidad <sup>a</sup>	5.8 (6.0)	4.1 (3.5)
Completitud anómala <sup>a</sup>	99.8 (100.0)	---
Multiplicidad anómala <sup>a</sup>	3.2 (3.2)	---

Tabla 3.5. Estadísticas de la recolección de datos de los complejos SeMetMobMN198-ADN26. <sup>a</sup> Valores en paréntesis para el rango de mayor resolución. <sup>b</sup>  $R_{merge} = [\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)] \times 100$ .  
\* Los datos corresponden a cristales ortorrómbicos obtenidos a partir de la condición de cristalización de PEG6000.

### 3.4. Resolución y refinamiento estructural

#### 3.4.1. Resolución estructural

El proceso de resolución estructural implica la determinación, a partir de las intensidades de difracción experimentalmente obtenidas, de la función densidad electrónica  $\rho(x,y,z)$  correspondiente. El conocimiento de esta función nos permite conocer las posiciones atómicas de cada uno de los átomos que constituyen la molécula proteica. La función densidad electrónica está dada por la siguiente relación:

$$\rho(xyz) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| \cos 2\pi[hx+ky+lz - \varphi(hkl)]$$

donde,  $V$  es el volumen de la celda unidad,  $x, y, z$  las coordenadas de cualquier punto del espacio,  $hkl$  los índices de Miller,  $|F_{hkl}|$  el módulo del factor de estructura y  $\varphi(hkl)$  su correspondiente fase.

Los factores de estructura representan, la contribución a la difracción, para cada familia de planos  $(hkl)$ , de los átomos contenidos en la celda unidad. Estos factores de estructura son la suma de una serie de términos que dependen básicamente de tres parámetros: el tipo de átomo, su poder de dispersión de los rayos X (descrito mediante el factor de dispersión atómica  $f_j$ ) y la posición del átomo en la celda unidad.

$$F_{hkl} = \sum_j f_j e^{2\pi i(hx+ky+lz)}$$

Se puede demostrar que las intensidades de los haces difractados, obtenidas experimentalmente cuando recogemos las imágenes de difracción, son proporcionales al cuadrado de los módulos de los factores de estructura. Sin embargo, la información de las fases  $\varphi(hkl)$  de los factores de estructura no se puede medir directamente. En cristalografía, este problema se conoce como el problema de las fases, y resolverlo, es el punto clave para lograr la determinación estructural.

### ***3.4.1.1. Métodos de resolución estructural utilizados***

Existen varios métodos diferentes que permiten abordar el problema de las fases: reemplazo molecular (MR, Molecular Replacement), reemplazo isomórfico simple (Single Isomorphus Replacement), reemplazo isomórfico múltiple (Multiple Isomorphous Replacement), difracción anómala simple (SAD, Single-wavelength Anomalous Diffraction), difracción anómala múltiple (MAD, Multiple-wavelength Anomalous Diffraction), e incluso combinación de ellos: SIRAS y MIRAS.

Para determinar la estructura tridimensional de las proteínas estudiadas en el presente trabajo se utilizaron dos de estas metodologías: la técnica de reemplazo molecular y la de difracción anómala simple (Single-wavelength Anomalous Diffraction), acordes cada una de ellas con las dificultades de cada caso en particular.

#### ***3.4.1.1.1. Reemplazo molecular***

El método de reemplazo molecular (Rossmann and Blow, 1962), es actualmente el método de resolución estructural más utilizado dado que presenta numerosas ventajas, en particular su simplicidad. Se basa en la utilización de la estructura de una proteína conocida como modelo para estimar las fases iniciales de la proteína desconocida cuya estructura se quiere determinar.

Evidentemente la eficacia del método depende de la similitud estructural entre la proteína de estructura desconocida y la que se va a utilizar como modelo. Cuanto mayor sea la homología, mayores serán las posibilidades de lograr determinar la estructura mediante este método. Si se parte de la suposición que la homología estructural está estrechamente vinculada a la homología de secuencia, es este parámetro el que se utiliza a la hora de buscar un modelo adecuado y en general debe ser igual o superior al 30%.

Una vez que se dispone del modelo, se determinan dos funciones, la función de rotación y función de traslación, que definen la correcta orientación y posición del modelo respecto a la estructura desconocida.

Las funciones de rotación y traslación consisten en la superposición de las funciones de Patterson del modelo y la estructura desconocida. Los nuevos programas incorporan métodos estadísticos donde se maximizan funciones de probabilidad (maximum likelihood) que facilitan encontrar las soluciones para ambas funciones.

Una vez orientado y posicionado correctamente el modelo, se calcula un mapa de densidad electrónica utilizando las intensidades experimentales y las fases que proporciona el modelo pudiendo iniciarse a partir de este momento el proceso de refinamiento.

#### **3.4.1.1.2. Difracción Anómala**

A longitudes de onda particulares los átomos son capaces de absorber los rayos X y como resultado de este efecto la Ley de Friedel  $F_{hkl} = F_{\bar{h}\bar{k}\bar{l}}$  debe cumplirse. Las reflexiones hkl ya no son iguales a las reflexiones -h-k-l, debido a que los factores de dispersión atómica  $f_j$ , a partir de los cuales se calculan los factores de estructura, se ven modificados con la introducción de dos nuevos términos,  $f''$  y  $f'''$ , tal que  $f_j = f_0 + f'' + if'''$ . Las diferencias generadas por el incumplimiento de la Ley de Friedel son aprovechadas entonces para obtener información sobre las fases.

Este fenómeno, que se conoce con el nombre de difracción o dispersión anómala es la base de los métodos de resolución estructural más ampliamente utilizados actualmente, MAD y SAD, especialmente en aquellos casos en que no se dispone de un buen modelo para aplicar el método de reemplazo molecular.

Las longitudes de onda a la cual se produce la difracción anómala son características de cada elemento y para los átomos livianos como el carbono, nitrógeno u oxígeno, quedan fuera del rango de los rayos X empleados en cristalografía, haciendo que la difracción anómala de estos elementos no sea utilizable. Sin embargo, la mayoría de los metales pesados con los que se suele derivatizar las proteínas (Se, Pt, Hg) o los metales presentes en las metalo-proteínas (Zn, Cu, Fe), presentan cantos de absorción fácilmente alcanzables en las líneas de macromoléculas de los sincrotrones actuales.

En la figura 3.28 se muestra el perfil de absorción para el selenio en el cual se identifican los dos puntos clave de las curvas: el pico (donde  $f''$  es máximo) y la inflexión (donde  $f'$  es mínimo).

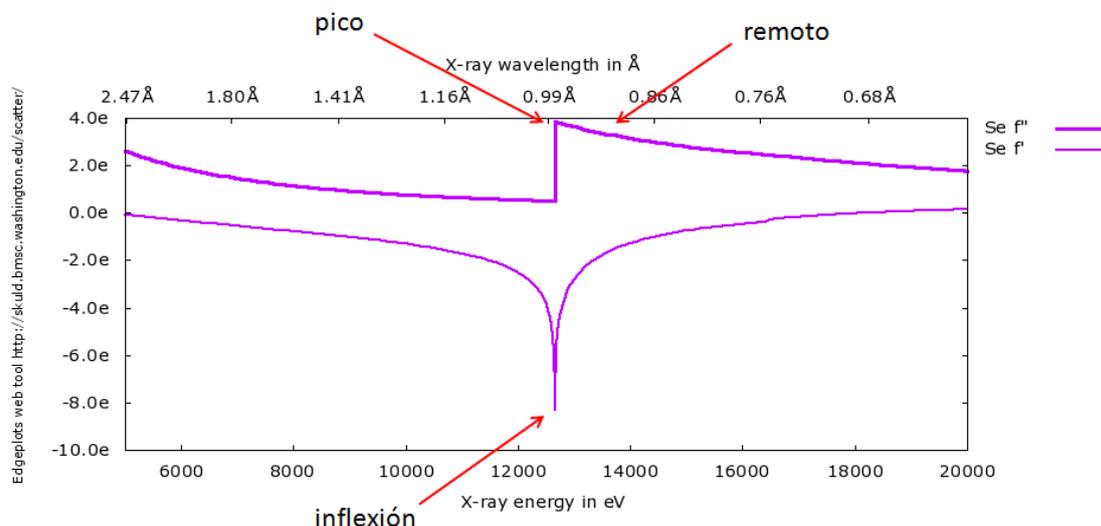


Figura 3.28. Coeficientes teóricos de dispersión anómala  $f'$  y  $f''$  en función de la longitud de onda para el átomo de selenio. En la figura se indican los puntos característicos: pico, inflexión y remoto, que corresponden a las longitudes de onda utilizadas habitualmente para la recogida de datos.

### 3.4.2. Construcción del modelo

La ubicación de los átomos pesados (Se) de la estructura a partir de datos de SAD recogidos en el pico de absorción del selenio se realizó mediante el programa SHELXD (Schneider and Sheldrick, 2002; Uson and Sheldrick, 1999), el cálculo inicial de las fases con el programa SHELXE (Sheldrick, 2002) y la mejora de la densidad (density modification) con PIRATE (Cowtan, 2002).

A partir del primer mapa experimental de densidad electrónica, el trazado inicial de la cadena polipeptídica se realizó de manera automática utilizando el programa RESOLVE (Terwilliger, 2004). El trazado posterior fue completado de forma manual mediante el programa gráfico COOT (Emsley and Cowtan, 2004).

En el caso de las estructuras que fueron resueltas por reemplazo molecular para la inspección de los mapas de densidad se utilizó el programa de visualización TURBO-FRODO (Roussel and Cambilleau, 1989).

### 3.4.3. Refinamiento

El proceso de refinamiento es un proceso iterativo que combina ciclos de cálculo con construcción y ajuste del modelo al mapa de densidad electrónica. El objetivo del refinamiento es lograr que el modelo se ajuste lo mejor posible a las amplitudes de los factores de estructura experimentalmente medidas.

La función básica a minimizar es el cuadrado de las diferencias entre las intensidades experimentales observadas y las calculadas a partir del modelo,

$$\Phi = \sum_{hkl} w_{hkl} (|F_{obs}| - |F_{calc}|)_{hkl}^2$$

donde  $w_{hkl}$  es un factor de peso que se introduce y que tiene en cuenta la precisión de la correspondiente intensidad medida. Esta función es una función de varias variables y todas ellas son tenidas en cuenta a la hora de realizar la minimización: las posiciones atómicas, los factores de temperatura  $B_j$  (relacionados con la movilidad de los átomos respecto a sus posiciones) y las ocupaciones de los átomos (que están relacionadas con la fracción en la que el átomo se encuentra en una misma posición en todas las celdas unidad del cristal).

Evidentemente el número de grados de libertad de este sistema es enorme y uno de los problemas que encontramos en las estructuras de proteínas es que las reflexiones medidas son insuficientes y el número de variables a calcular es mayor que las ecuaciones de las cuales disponemos.

Por este motivo es que en el refinamiento también se incluyen una serie de restricciones e imposiciones geométricas relacionadas con la estereoquímica de los aminoácidos, como son, distancias, ángulos de enlace, ángulos diedros, etc. El introducir estos parámetros durante el refinamiento disminuye los grados de libertad del sistema y posibilita la convergencia del mismo.

$$\Phi = \sum_{hkl} w_{hkl} (|F_{obs}| - |F_{calc}|)_{hkl}^2 + \sum_i w_i (d_i^{ideal} - d_i^{modelo})^2 + \sum_j w_j (\phi_j^{ideal} - \phi_j^{modelo})^2 + \dots$$

Durante todo el proceso de refinamiento se toman en cuenta, además, dos indicadores que nos aseguran la correcta convergencia del proceso, conocidos como los factores R y R-free. El factor R se define a partir de los módulos de los factores de estructura experimentales y calculados como:

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

y se relaciona con la bondad de ajuste del refinamiento. Cuanto menor sea el factor R mejor será el ajuste del modelo a los datos experimentales.

De manera análoga se calcula el factor R-free, pero a diferencia del R, para su cálculo solo se tienen en cuenta un pequeño porcentaje del total de las reflexiones (usualmente el 5%), que son excluidas de proceso de refinamiento. El factor R-free tiene la ventaja, por este motivo, de ser un parámetro que no se ve sesgado por el modelo ni por el proceso de refinamiento.

#### **3.4.4. Complejos TrwC-ADN**

##### **3.4.4.1. Resolución estructural y refinamiento del complejo TrwC-ADN25**

Los cristales de los complejos TrwC-ADN25-M<sup>2+</sup> provienen de la derivatización de cristales hexagonales del complejo TrwC-ADN25, y por lo tanto, pertenecen al mismo grupo espacial que éste (P6<sub>1</sub>22) con una sola molécula de complejo por unidad asimétrica.

	25ZNpkMG	25ZNrmMG	25ZNpkMN	25ZNMNpk
Átomos proteína <sup>a</sup>	2288		2288	
Átomos ADN <sup>a</sup>	512	No refinado <sup>c</sup>	512	No refinado <sup>c</sup>
Átomos metálicos <sup>a</sup>	Zn		Zn	
Aguas <sup>a</sup>	178		178	
Nº reflexiones utilizadas en el refinamiento <sup>b</sup>	13597		10055	
R factor (%) <sup>b</sup>	21.8			
R <sub>free</sub> (%) <sup>b</sup>	26.0	No refinado <sup>c</sup>		No refinado <sup>c</sup>
<i>Rmsd</i> respecto a valores típicos				
Enlaces (Å)	0.009		0.011	
Angulos (°)	1.31		1.44	
B factor promedio (Å <sup>2</sup> )	34.0		32.4	
	25MN	25CU**	25NI***	
Átomos proteína <sup>a</sup>	2226	2314	2275	
Átomos ADN <sup>a</sup>	512	512	512	
Átomos metálicos <sup>a</sup>	---	Cu	Ni	
Aguas <sup>a</sup>	256	207	162	
Nº reflexiones utilizadas en el refinamiento <sup>b</sup>	10614	18200	24598	
R factor (%) <sup>b</sup>	21.8	20.4	20.6	
R <sub>free</sub> (%) <sup>b</sup>	26.0	25.0	24.5	
<i>Rmsd</i> respecto a valores típicos				
Enlaces (Å)	0.010	0.005	0.013	
Angulos (°)	1.35	1.52	1.35	
B factor promedio (Å <sup>2</sup> )	31.6	37.1	38.2	

Tabla 3.6. Estadística del refinamiento de los complejos TrwC-ADN25-M<sup>2+</sup>. <sup>a</sup> Por unidad asimétrica. <sup>b</sup>  $R_{factor} = [\sum_{hkl} ||F_{obs}| - k|F^{calc}|] / \sum_{hkl} |F_{obs}| \times 100$ ;  $R_{free}$  igual que el anterior pero para un conjunto de reflexiones excluidas del refinamiento. <sup>c</sup> No se encontró evidencia de la presencia de Mg<sup>2+</sup> o Mn<sup>2+</sup>.

Las estructuras fueron resueltas por el método de reemplazo molecular utilizando el programa MolRep (Vagin and Teplyakov, 1997) y la estructura de la proteína libre de metal como modelo (código pdb 1OMH).

Los mapas de densidad electrónica se inspeccionaron con el programa TURBO-FRODO (Roussel and Cambilleau, 1989) para confirmar o descartar la presencia de los diferentes iones metálicos en el sitio activo de la proteína y en aquellos casos en que se constató la ausencia del mismo la estructura no fue refinada.

El refinamiento estructural se llevó a cabo mediante ciclos de “annealing”, con el programa CNS versión 1.1 (Brunger *et al.*, 1998), combinados con la inspección visual de los mapas (TURBO-FRODO) para corregir manualmente las posiciones de las cadenas laterales de los residuos y adicionar las moléculas de solvente.

Los principales parámetros y las estadísticas finales de todos los refinamientos se muestran en la tabla 3.6.

#### **3.4.4.2. Resolución estructural y refinamiento del complejo *TrwC<sup>Y18F</sup>-ADN27***

La estructura se resolvió por el método de reemplazo molecular con el programa MolRep (Vagin and Teplyakov, 1997), utilizando como modelo de partida la estructura de la proteína TrwC-ADN25 (código pdb 1OSB) y para el refinamiento posterior se empleó el programa REFMAC versión 5.2 (Murshudov *et al.*, 1997) que forma parte del conjunto de programas de la suite CCP4 (Collaborative Computational Project, 1994).

La base T25 de la cadena de ADN, que difería considerablemente respecto a la posición en el modelo, se eliminó. Luego de un ciclo de refinamiento y mediante la inspección de los mapas de densidad electrónica se pudo situar en la posición correcta. Otro ciclo de refinamiento y el posterior chequeo de los mapas de densidad, permitió la ubicación de la base A26. Finalmente, se intentó la localización de la base T27, pero la densidad electrónica extremadamente difusa no permitió su modelado.

	TrwC-ADN27
Átomos proteína <sup>a</sup>	2213, 2213
Átomos ADN <sup>a</sup>	492, 475
Átomos metálicos <sup>a</sup>	---
Aguas <sup>a</sup>	239
Nº reflexiones utilizadas en el refinamiento <sup>b</sup>	26050
R factor (%) <sup>b</sup>	20.5
R <sub>free</sub> (%) <sup>b</sup>	24.7
<i>Rmsd</i> respecto a valores típicos	
Enlaces (Å)	0.009
Angulos (°)	1.37
B factor promedio (Å <sup>2</sup> )	49.2

Tabla 3.7. Estadísticas del refinamiento del complejo TrwC-ADN27. <sup>a</sup> Por unidad asimétrica. <sup>b</sup>  $R_{factor} = [\sum_{hkl} ||F_{obs}| - k|F^{calc}|| / \sum_{hkl} |F_{obs}|] \times 100$ ;  $R_{free}$  igual que el anterior pero para un conjunto de reflexiones excluidas del refinamiento.

#### 3.4.4.3. Resolución estructural y refinamiento de los complejos MobM N198-ADN26

La estructura se intentó resolver por el método de reemplazo molecular utilizando el dominio N-terminal de la proteína TrwC como modelo (código pdb 1OMH). Sin embargo, la baja homología de secuencia existente entre ambas proteínas (menor del 10%) ya indicaba que las estructuras podían ser suficientemente diferentes como para imposibilitar la resolución estructural por este método, como así ocurrió. Se decidió entonces trabajar con la proteína derivatizada con seleno-metionina para obtener fases experimentales.

#### 3.4.4.4. Resolución estructural y refinamiento de los complejos SeMetMobM N198-ADN26

La estructura fue resuelta por el método de SAD (Dispersión Anómala Simple) utilizando la dispersión anómala del átomo de selenio, a partir de los datos recogidos con los cristales ortorrómbicos de la forma cristalina (i).

Las posiciones de siete de los ocho átomos de selenio se encontraron con el programa SHELXD (Schneider and Sheldrick, 2002).

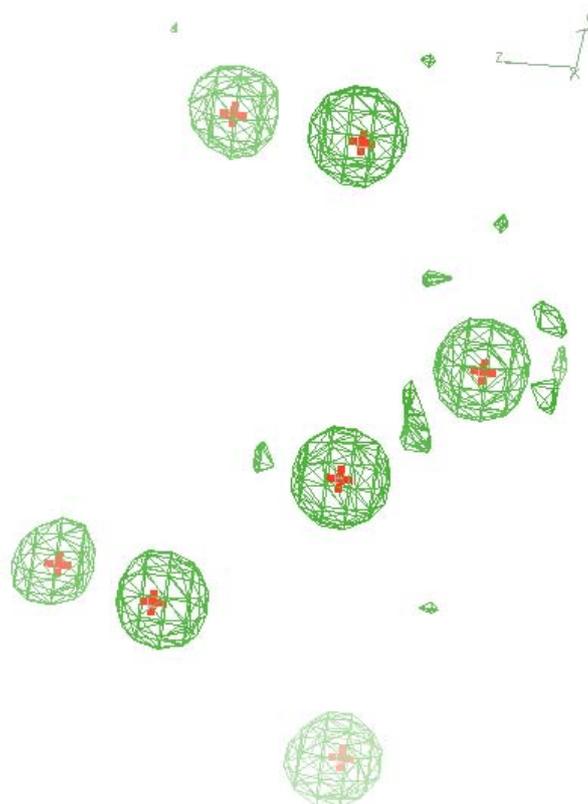
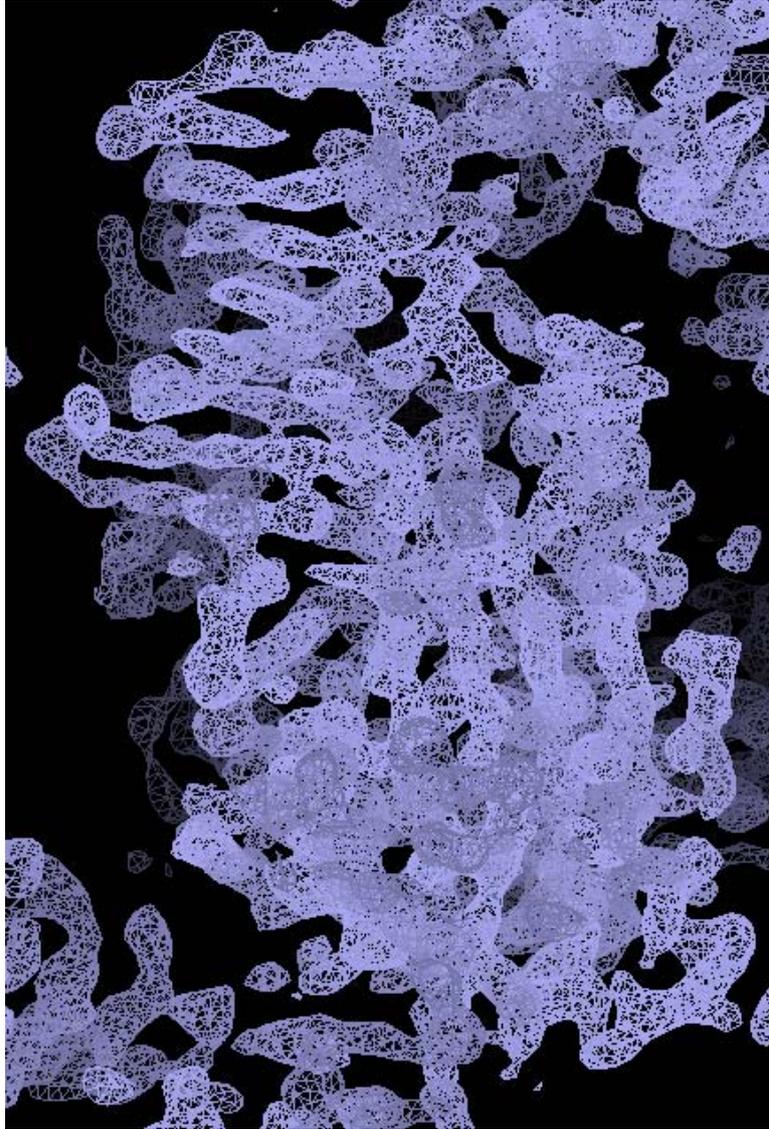


Figura 3.29. Mapa de Fourier anómalo dibujado a  $3.5\sigma$  en el que se han indicado la posición de los átomos pesados de selenio encontrados con el programa SHELXD.

El primer mapa de densidad electrónica se obtuvo a partir de las fases experimentales, mejorando la densidad con el programa PIRATE (Cowtan, 2002) (figura 3.30).



*Figura 3.30. Mapa de densidad electrónica del complejo SeMetMobM N98-ADN26 obtenido a partir de las fases experimentales. En él se reconocen varios elementos de estructura secundaria, como son la hoja beta que constituye el núcleo central de la estructura de la proteína. Asimismo también se puede identificar la cadena de ADN a la izquierda.*

El 30% de la cadena polipeptídica se construyó automáticamente con el programa RESOLVE (Terwilliger, 2004) y el 70% restante, junto con la cadena de ADN, se trazó manualmente con el programa Coot (Emsley and Cowtan, 2004).

La estructura se refinó a partir de datos nativos a 2.1 Å de resolución con el programa REFMAC 5.5 (Murshudov *et al.*, 1997) de la suite de programas del CCP4 (Collaborative Computational Project, 1994), combinando ciclos de refinamiento con la inspección visual de la densidad electrónica y el ajuste manual de la cadena

polipeptídica y de la de ADN a la densidad electrónica, utilizando para ello el programa Coot (Emsley and Cowtan, 2004).

Una vez que se obtuvo la estructura de la forma cristalina (i), las estructuras de las tres formas cristalinas restantes fueron determinadas mediante reemplazo molecular, usando como modelo de partida la estructura de la forma cristalina (i) y el programa PHASER (Read, 2001).

Los refinamientos se realizaron con el programa REFMAC 5.5 (Murshudov *et al.*, 1997) siguiendo el protocolo descrito. En la tabla 3.8 se muestran las estadísticas de los mismos.

	Forma cristalina (i)	Forma cristalina (iv)
Átomos proteína <sup>a</sup>	1389	1410
Átomos ADN <sup>a</sup>	535	535
Aguas <sup>a</sup>	98	125
Nº reflexiones utilizadas en el refinamiento <sup>b</sup>	15555	16174
R factor (%) <sup>b</sup>	21.2	20.7
R <sub>free</sub> (%) <sup>b</sup>	26.7	27.8
<i>Rmsd</i> respecto a valores típicos		
Enlaces (Å)	0.013	0.005
Ángulos (°)	1.76	0.92
B factor promedio (Å <sup>2</sup> )	28.5	26.9

Tabla 3.8. Estadísticas del refinamiento de los complejos MobM-ADN26. <sup>a</sup> Por unidad asimétrica. <sup>b</sup>  $R_{factor} = [\sum_{hkl} ||F_{obs}| - k|F^{calc}|| / \sum_{hkl} |F_{obs}|] \times 100$ ;  $R_{free}$  igual que el anterior pero para un conjunto de reflexiones excluidas del refinamiento.

### 3.5. Validación del modelo

Los modelos estructurales fueron validados con los programas PROCHECK (Laskowski *et al.*, 1993) de la suite de programas CCP4 (Collaborative Computational Project, 1994), MOLPROBITY (<http://molprobity.biochem.duke.edu/>) (Davis *et al.*, 2007) y las utilidades de validación del programa COOT (Emsley and Cowtan, 2004).

Se analizaron en todos los casos los parámetros geométricos más relevantes como son las distancias y los ángulos de enlace, las conformaciones de los posibles rotámeros y los diagramas de Ramachandran (Ramachandran and Sasisekharan, 1968).

Estos últimos, son una representación bidimensional de los ángulos diedros de los planos amida alrededor de los enlaces N-C $\alpha$  y C $\alpha$ -CO, llamados  $\phi$  (phi) y  $\psi$  (psi), respectivamente para cada uno de los residuos de la proteína (figura 3.31).

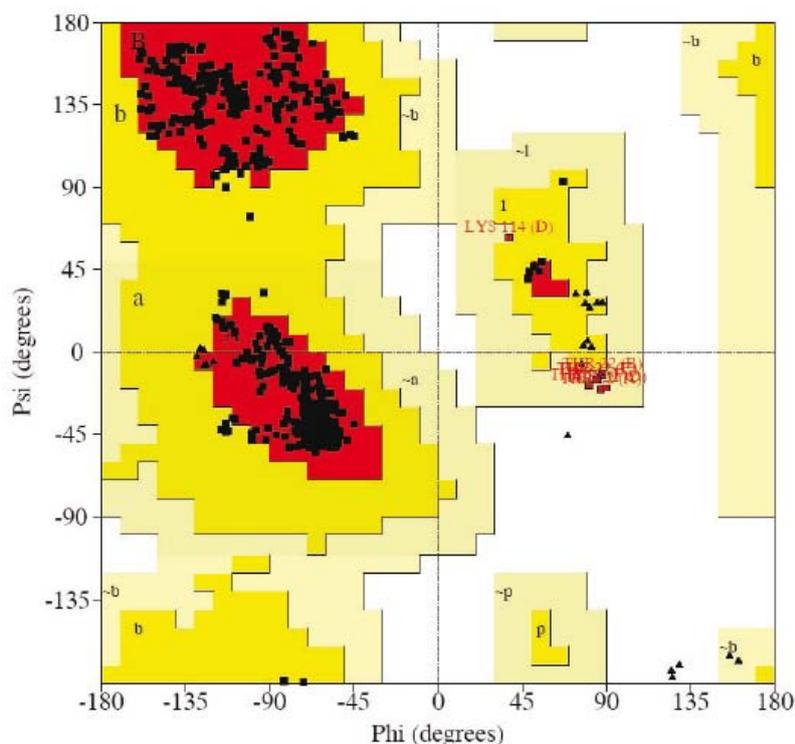


Figura 3.31. Ejemplo de un diagrama de Ramachandran de una proteína. Las zonas rojas representan las conformaciones permitidas más favorables.

Como estos ángulos adoptan valores particulares dependiendo de la estructura secundaria de la cadena polipeptídica donde se localiza el residuo, es posible chequear visualmente la validez del modelo desde el punto de vista estereo-químico a partir del diagrama. La localización de residuos en regiones prohibidas debe ser corregida o debidamente justificada.



## ***4. RESULTADOS Y DISCUSIÓN***



## **4. Resultados y discusión**

### **4.1. La relaxasa TrwC**

En 2003 fue descrita la estructura tridimensional del dominio N-terminal de TrwC en complejo con un oligonucleótido de 25 bases (Guasch *et al.*, 2003). En este trabajo se caracterizaba la unión de la proteína al ADN, las claves del reconocimiento de la secuencia específica del origen de transferencia (*oriT*), la arquitectura del sitio activo y la localización de una de las dos tirosinas catalíticas (Tyr18).

Sin embargo, una serie de interrogantes quedaron pendientes para posteriores estudios: la identidad y función del ión metálico que es coordinado por la tríada de histidinas en el centro activo, la localización del segundo residuo catalítico (Tyr26) y la posición del fosfato del enlace fosfo-diéster a ser cortado.

Con el fin de analizar y responder a estos interrogantes es que, en el presente trabajo, se realizó un estudio sistemático de la unión de la proteína a diferentes metales, caracterizándolos de acuerdo a su capacidad de difracción anómala de los rayos X a longitudes de onda típicas. Asimismo, se cristalizó y determinó la estructura tridimensional de la proteína con un oligonucleótido de 27 bases cuya secuencia comprende el sitio de corte que TrwC reconoce, utilizando para ello un mutante inactivo de la proteína (TrwC N293<sup>Y18F</sup>). Los resultados obtenidos se presentan y discuten en detalle a continuación.

#### **4.1.1. Detalles estructurales de los complejos TrwC-ADN25-M<sup>+2</sup>**

El plegamiento global de los complejos TrwC-ADN25-M<sup>2+</sup> es similar al observado en las estructuras de los complejos sin metal. No obstante, a nivel local, la unión del metal produce cambios conformacionales significativos en el del lazo  $\alpha_1$ - $\beta_1$ , que comprende los residuos 20 al 30 y la tirosina catalítica Tyr26.

En las estructuras con metal, este lazo que no pudo ser modelado en la estructura sin metal debido a su gran flexibilidad, se ordena parcial o totalmente y en mayor o

menor medida permite la localización de los residuos que lo constituyen (residuos 20 al 26 en las estructuras de  $\text{Ni}^{2+}$  y  $\text{Zn}^{2+}$  y 20 al 30 en la estructura de  $\text{Cu}^{2+}$ ).

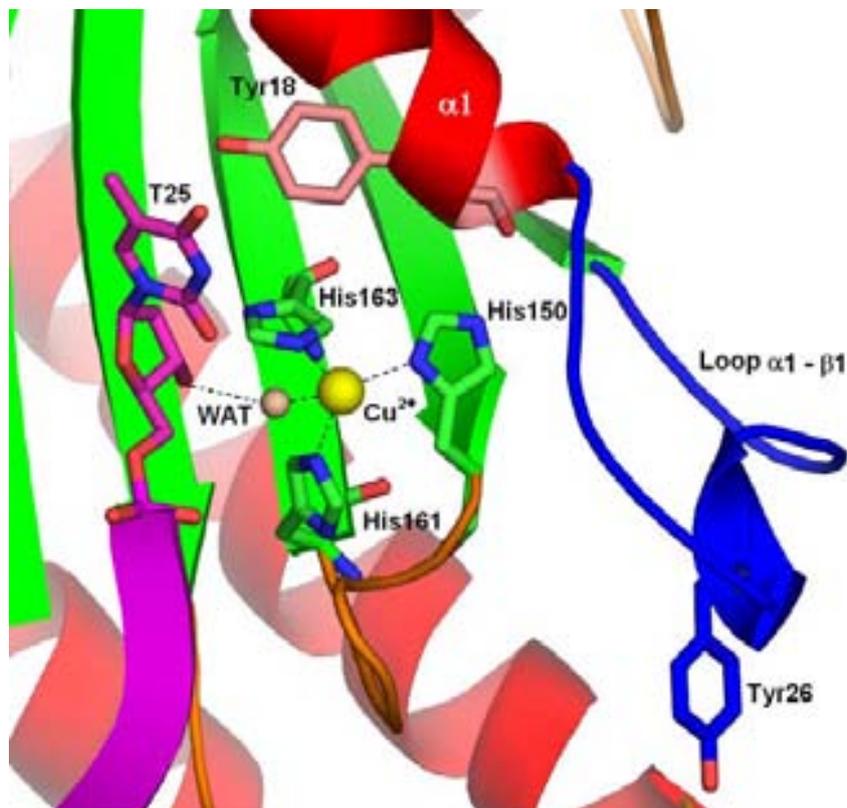


Figura 4.1. Centro activo del complejo  $\text{TrwC-ADN25-Cu}^{2+}$ . El ión metálico es coordinado por la triada de histidinas y una molécula de agua que establece enlace de hidrógeno con el oxígeno O3' del ADN. El lazo  $\alpha_1\text{-}\beta_1$  fue trazado completamente localizándose en él la tirosina Tyr26.

Las conformaciones de  $\alpha_1\text{-}\beta_1$  son idénticas en las estructuras de  $\text{Ni}^{2+}$  y  $\text{Cu}^{2+}$  difiriendo levemente en la estructura de  $\text{Zn}^{2+}$  entre los residuos 20 al 24. Pese a ello, en todas las estructuras la tirosina Tyr26 se ubica en la misma posición relativa y su grupo hidroxilo se dispone a 17 Å de distancia del sitio activo, orientado en dirección opuesta al grupo hidroxilo de la tirosina catalítica Tyr18 (figura 4.1).

La conformación de la tirosina Tyr18 muestra una significativa variación respecto a la estructura sin metal, dando cuenta de su gran flexibilidad la cadena lateral de la misma gira respecto al ángulo de torsión  $\chi_1$  desplazando al grupo hidroxilo 2.39 Å en la estructura de  $\text{Zn}^{2+}$ , 2.17 Å en la de  $\text{Ni}^{2+}$  y 0.68 Å en la de  $\text{Cu}^{2+}$  (figura 4.2).

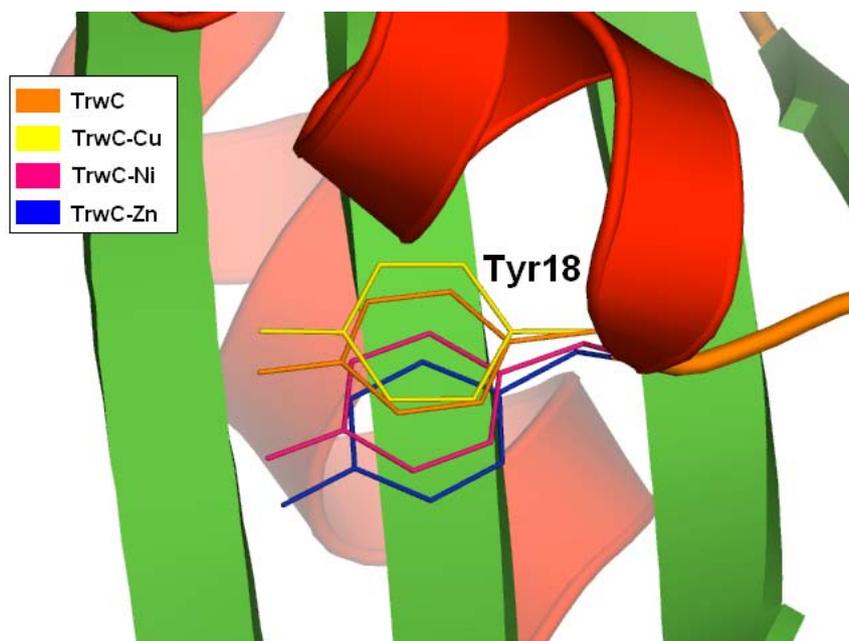


Figura 4.2. Flexibilidad conformacional del residuo catalítico Tyr18 en las estructuras con y sin metal.

La base T25 localizada en el extremo 3' de la cadena de ADN presenta, por su parte, también un alto grado de flexibilidad. Pese a que la densidad electrónica en los mapas es débil y algo dispersa, fue modelada en todos los casos sin ambigüedad y se orienta tal que el oxígeno O3' apunta en la dirección del metal en las estructuras de  $\text{Cu}^{2+}$  y  $\text{Ni}^{2+}$ , y para el lado contrario en la estructura de  $\text{Zn}^{2+}$  (figura 4.3).

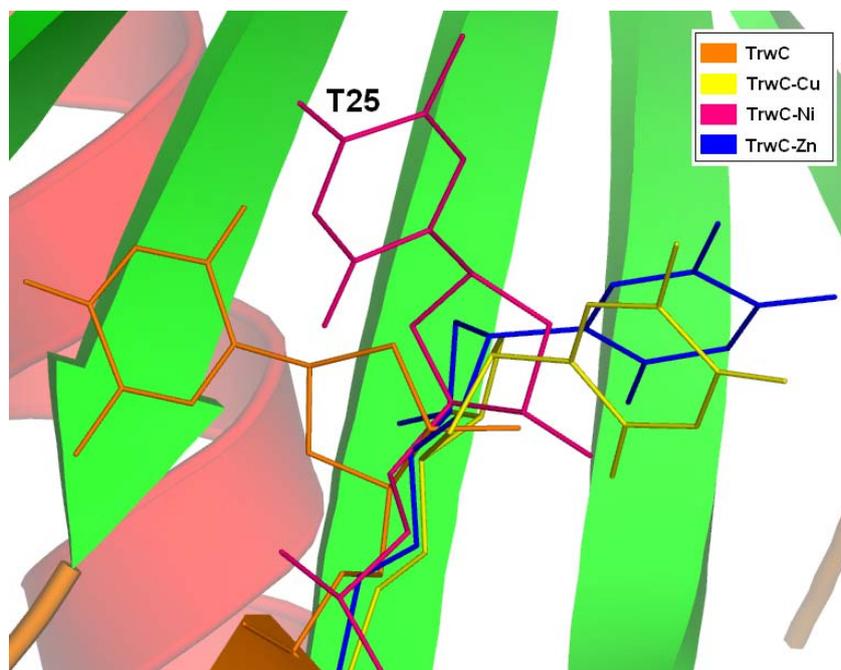


Figura 4.3. Conformaciones adoptadas por la base T25 en las diferentes estructuras con y sin metal.

La naturaleza química del metal, que la tríada de histidinas coordina en el sitio activo de la proteína, se analizó mediante un estudio sistemático por difracción de rayos X en el que se tuvo en cuenta la presencia o ausencia de difracción anómala del metal a la longitud de onda característica de cada uno de ellos.

La figura 4.4 muestra los mapas de diferencia anómala para cada uno de los metales estudiados. Los cristales (i), incubados en una solución conteniendo  $\text{ZnSO}_4$  y  $\text{MgCl}_2$ , fueron medidos a la longitud de onda correspondiente al pico de absorción del Zn y a una longitud de onda ligeramente mayor. El mapa de diferencia anómala muestra que los cristales que fueron medidos en el pico del Zn presentan una fuerte densidad electrónica (en amarillo) localizada en el sitio activo, mientras que aquellos que fueron medidos a una longitud de onda mayor (en azul) no presentan densidad alguna (figura 4.4 (a)). Es posible inferir, por tanto, que en este caso el  $\text{Zn}^{2+}$  es el metal que entra al sitio activo y es coordinado por la tríada de histidinas.

Las figuras 4.4 (b) y 4.4 (c) corresponden a los cristales (ii), tratados con  $\text{ZnSO}_4$  y  $\text{MnCl}_2$ . Como en el caso anterior, se observa una densidad electrónica perfectamente definida en el sitio del metal para los datos recogidos a la energía del pico del Zn, y ninguna para los recolectados a la energía del pico del Mn, por lo que también en este caso que el Zn es el metal que se une al sitio activo.

Los mapas de diferencia anómala de los cristales (iii), incubados en una solución conteniendo  $\text{MnCl}_2$  y medidos a la longitud de onda del pico del Mn, no muestran evidencia de la presencia de metal en el sitio activo (figura 4.4 (d)). Esto se ve reforzado por el hecho de que la orientación del residuo His150 es la misma que la adoptada en la estructura libre de metal y por lo tanto no es adecuada para la coordinación del mismo.

Por último, los mapas de diferencia anómala de las figuras 4.4 (e) y 4.4 (f) que corresponden a los cristales (iv y v) incubados en  $\text{CuCl}_2$  y  $\text{NiCl}_2$ , respectivamente, demuestran la presencia de  $\text{Cu}^{2+}$  y  $\text{Ni}^{2+}$  en los sitios activos correspondientes.

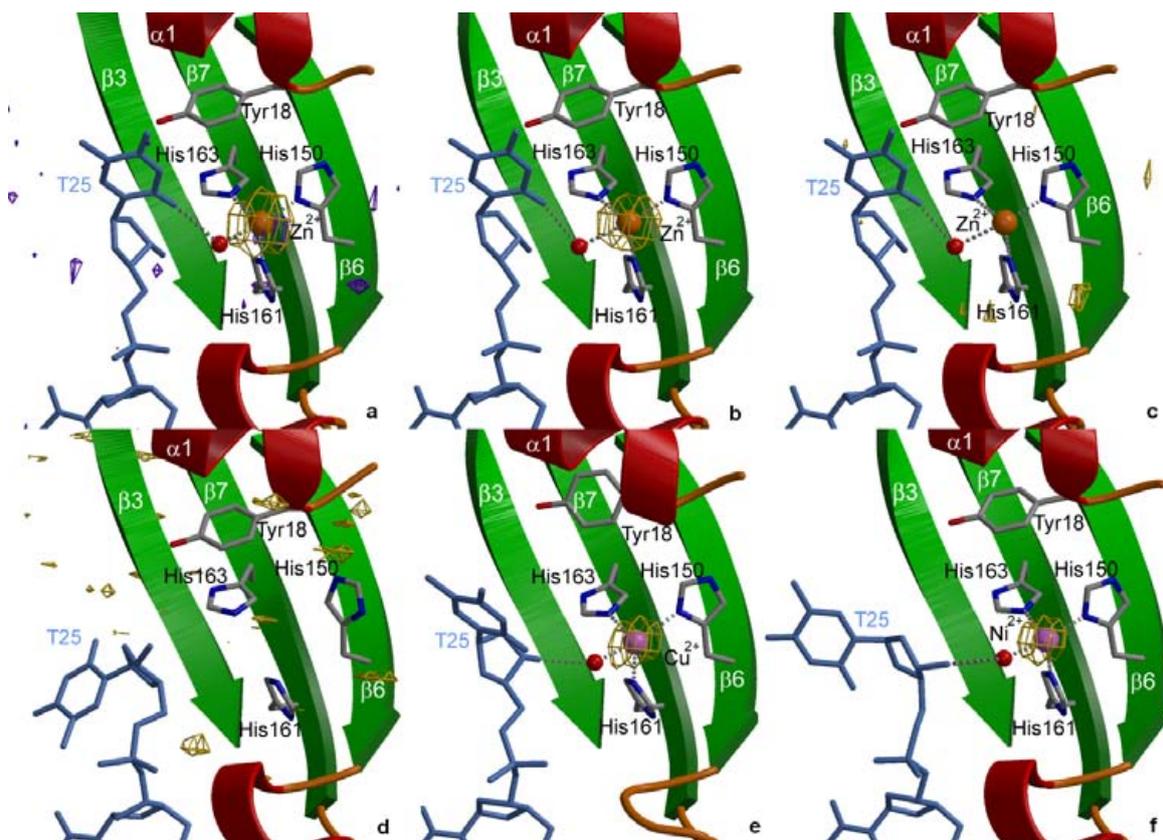


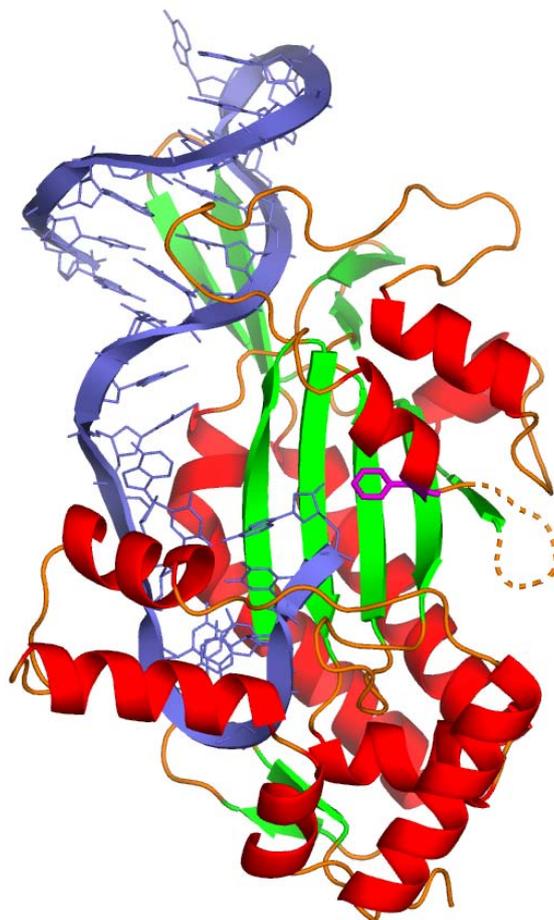
Figura 4.4. Mapas de diferencia anómala del centro activo de la TrwC con los diferentes metales estudiados. Cristales derivatizados con: (a)  $ZnSO_4$ , medidos en el pico de absorción del Zn (amarillo) y a una energía menor (azul), (b)  $ZnSO_4$  y  $MnCl_2$ , medidos en el pico de absorción del Zn y (c)  $ZnSO_4$  y  $MnCl_2$ , medidos en el pico de absorción del Mn, (d)  $MnCl_2$ , medidos en el pico de absorción del Mn, (e)  $CuCl_2$ , medidos en el pico de absorción del Cu y (f)  $NiCl_2$ , medidos en el pico de absorción del Ni.

En ningún caso se pudo verificar la presencia de  $Ca^{2+}$  o  $Mg^{2+}$  en el sitio activo. Si bien en los cristales incubados con  $MgCl_2$  se observó densidad electrónica en el sitio del metal, ésta fue asignada a una molécula de agua dado que la coordinación tetraédrica, y no octaédrica como es característica para el  $Mg^{2+}$  descartaron la presencia del mismo.

#### 4.1.2. Detalles estructurales del complejo $TrwCN293^{Y18F}$ -ADN27

De manera análoga a lo ocurrido con la estructura de la TrwC-ADN25 libre de metal (código pdb 1OSB), el complejo cristalizó en una celda hexagonal, en el grupo espacial  $P6_5$ , con dos moléculas de complejo por unidad asimétrica.

Las características estructurales generales son idénticas a las ya observadas en el complejo TrwC-ADN25 (figura 4.5).



*Figura 4.5. Complejo TrwC-ADN27, en magenta se indica la posición de la cadena lateral del residuo catalítico Tyr18 (mutado por fenilalanina) y en línea punteada la ubicación del lazo Glu20-Asp30 modelado en base a la estructura del complejo TrwC-ADN25-Cu<sup>2+</sup>.*

La cadena de ADN muestra la horquilla típica necesaria para el reconocimiento de la cadena de ADN por parte de la proteína y el giro en U de las bases T21, G22, T23 y C24 que permite la entrada al sitio activo de manera que, en la cavidad hidrofóbica que forman, queda atrapado el primer residuo de la proteína (SeMet1). La base T25 cambia de orientación respecto a la estructura 1OSB formando con la base G22 dos enlaces de hidrógeno, uno bifurcado de tres centros (O2 de T25 con N1 y N2 de G22) y otro simple (N3 de T25 con O6 de G22).

La conformación de la base T25 es tal que orienta el oxígeno O3' y al grupo

fosfato del residuo A26 hacia el sitio de unión del metal, que es ocupado en esta estructura por una molécula de agua. La base A26, que pudo ser localizada en los mapas de densidad electrónica y refinada en una sola de las dos moléculas presentes en la unidad asimétrica, establece interacciones con la base T25 y la SeMet5 (figura 4.6).

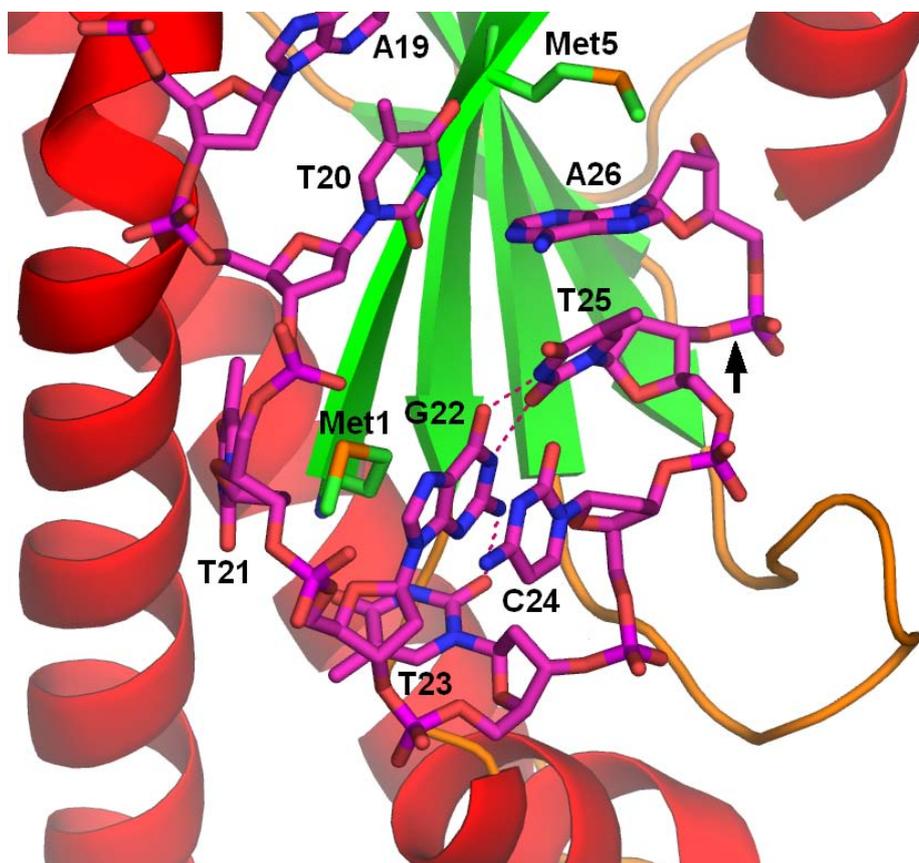


Figura 4.6. Giro en U de la cadena de ADN en el complejo TrwC-ADN27. A diferencia de los complejos TrwC-ADN25, la base T25 orienta al O3' en dirección del sitio de unión del metal estableciendo enlaces de hidrógeno con la base G22. La base A26 se sitúa paralela a T25 e interactúa con la SeMet5 cercana.

#### 4.1.3. La naturaleza del ión metálico en la relaxasa TrwC

Ensayos *in vitro* mostraron que en ausencia de ión metálico la TrwC no es capaz de cortar la cadena de ADN. El agregado de iones metálicos divalentes como  $Mg^{2+}$ ,  $Mn^{2+}$ ,  $Ca^{2+}$ ,  $Zn^{2+}$ ,  $Ni^{2+}$ ,  $Cu^{2+}$  o  $Co^{2+}$ , producen porcentajes de corte significativos. Sin embargo, la concentración necesaria de  $Mg^{2+}$  o  $Mn^{2+}$  para obtener porcentajes similares a los obtenidos con  $Zn^{2+}$ ,  $Ni^{2+}$  o  $Cu^{2+}$  es mucho mayor y la afinidad de la proteína por el  $Mg^{2+}$  es cuatro órdenes de magnitud menor que por el  $Zn^{2+}$  (Boer *et al.*, 2006).

El estudio estructural realizado con diferentes metales confirmó la alta afinidad de la proteína por el  $Zn^{2+}$ . La arquitectura del sitio activo provee el entorno de coordinación más comúnmente observado para el ión  $Zn^{2+}$ , en el que tres histidinas y una molécula de agua completan su esfera de coordinación.

De manera análoga también se observó la unión a los iones  $Cu^{2+}$  y  $Ni^{2+}$ , pero en ningún caso se encontró evidencia de unión a los iones  $Mg^{2+}$  o  $Mn^{2+}$ . La tríada de histidinas no provee el entorno característico de coordinación de estos metales, que coordinan usualmente con al menos un grupo carboxilo.

#### 4.1.4. Rol del ión metálico en el mecanismo catalítico

En la estructura del complejo TrwC-ADN27, una molécula de agua ocupa el lugar del ión metálico y mimetiza sus interacciones. El grupo fosfato que contiene el enlace fosfo-diéster a ser cortado orienta uno de sus átomos de oxígeno hacia el sitio del metal de manera tal que éste queda en la posición adecuada como para coordinar el ión metálico si este estuviera en el sitio activo.

Si se compara esta estructura con la de los complejos TrwC-ADN25-metal, se ve como la posición del oxígeno del fosfato se corresponde con la que ocupa la molécula de agua que completa la esfera de coordinación del ión metálico (figura 4.7).

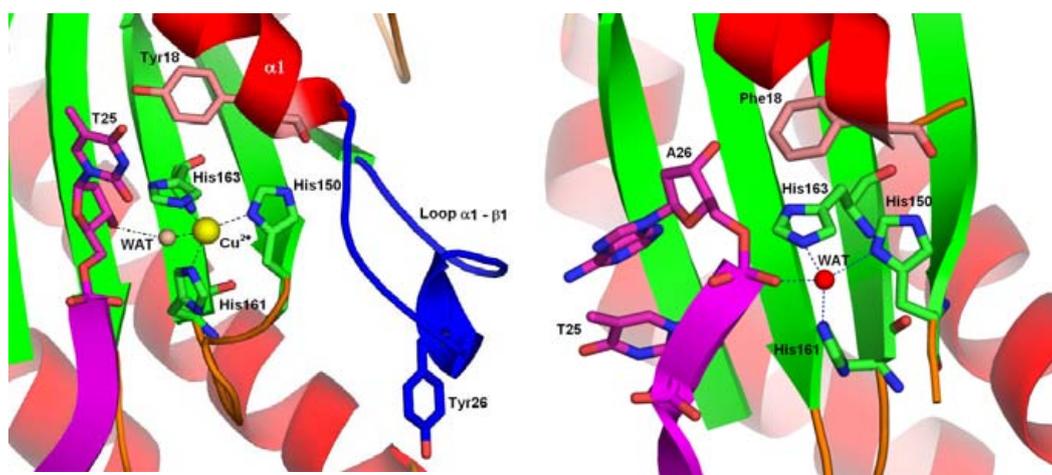


Figura 4.7. Comparación de los sitios activos de los complejos  $TrwC-ADN25-Cu^{2+}$  y  $TrwC-ADN27$ . La posición del metal y de la molécula de agua que completa su esfera de coordinación en (a), es ocupada por una molécula de agua y un átomo de oxígeno del grupo fosfato en (b).

Dado que la presencia del metal es esencial para que ocurra el corte del enlace fosfo-diéster de la cadena de ADN que inicia el proceso de conjugación, éste debe encontrarse estrechamente vinculado con el mecanismo catalítico de la reacción de trans-esterificación inicial.

La coordinación del metal mediante la tríada de histidinas y su posición respecto al grupo fosfato, mimetizada por una molécula de agua en la estructura del complejo con el oligonucleótidos de 27 bases, sugieren, tal como se esquematiza en la figura 5.8, un posible mecanismo para esta primera etapa.

La Tyr18, activada por el grupo carboxilo del aspártico Asp85 que se encuentra a distancia de enlace de hidrógeno del grupo hidroxilo de la tirosina Tyr18, ataca nucleofílicamente al fosfato. El metal coordina un oxígeno del grupo fosfato polarizando y debilitando el enlace a ser cortado, y favoreciendo de esta forma el ataque de la tirosina.

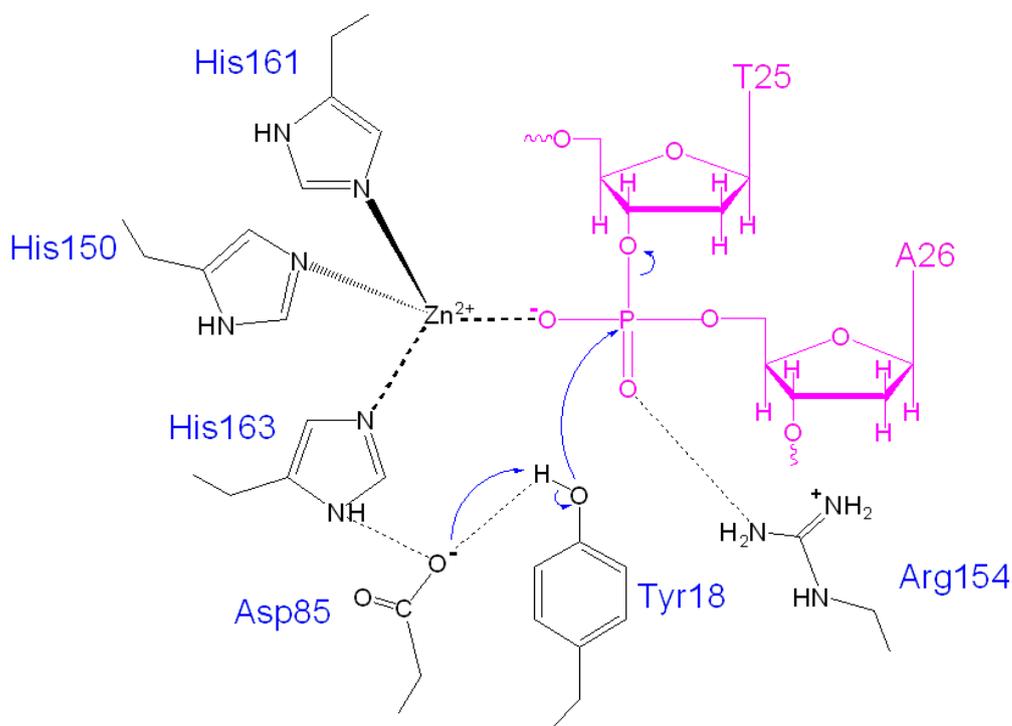


Figura 4.8. Esquema ilustrando el mecanismo catalítico propuesto para el primer corte del ADN.

El final de esta primera etapa ocurriría con la cesión de un protón al oxígeno O3' de la base T25. Los datos estructurales disponibles, sin embargo, no permiten establecer sin ambigüedad cuál es el residuo responsable de cumplir este papel.

#### ***4.1.5. La tirosina catalítica Tyr26***

El proceso de conjugación termina cuando se produce un segundo corte de la cadena de ADN de forma tal que se genera un nuevo extremo 3' y la reacción de religación de dicho extremo con el extremo 5' del fosfato tiene lugar.

La tirosina encargada de este segundo corte sería la tirosina catalítica Tyr26, que en las estructuras estudiadas se logró localizar en el lazo  $\alpha_1$ - $\beta_1$ . Este lazo, que comprende los residuos 20 al 30, se encuentra totalmente desordenado en las estructuras de los complejos de TrwC en los que no hay metal. Sin embargo, la presencia del metal lo ordena total o parcialmente y en el caso de la estructura TrwC-ADN25-Cu<sup>2+</sup> pudo ser trazado completamente.

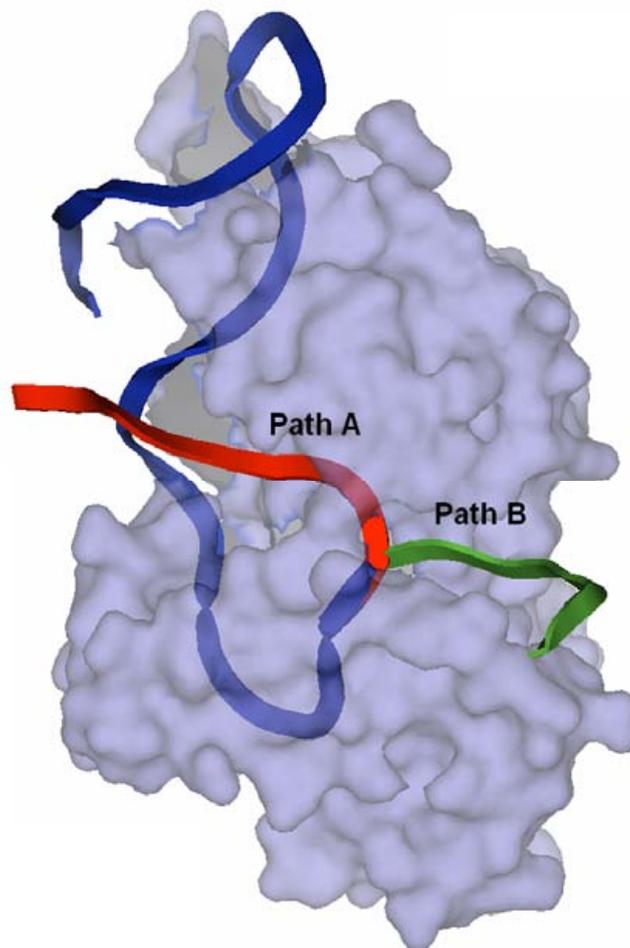
Si bien la localización de la tirosina Tyr26 en las estructuras determinadas es bastante lejana al sitio activo (17Å), es posible que, dada la extrema flexibilidad del lazo en la que esta se encuentra, un pronunciado movimiento del mismo la coloque próximo al centro activo permitiéndole realizar el corte.

Otra posibilidad es que, dada la ubicación de la tirosina Tyr26 en la superficie de la proteína, sea a través de una segunda molécula de TrwC que el corte se realice. Hasta el momento con la información bioquímica y estructural que se posee no es posible confirmar ni descartar ninguna de las dos posibilidades.

#### ***4.1.6. Las rutas de salida de la cadena de ADN***

En la estructura del complejo TrwC-ADN25 la posición de un par de iones sulfato, localizados cerca del sitio de corte, sugirió la posibilidad de que éstos podrían estar mimetizando la posición de dos grupos fosfato del ADN. Estos grupos fosfato estarían, por otra parte, indicando las direcciones de dos posibles rutas de salida que las cadenas de ADN utilizarían después del corte (figura 1.8).

En el complejo TrwC-ADN27 la posición de los fosfatos corriente abajo del sitio de corte muestran como la cadena de ADN continúa anti-paralela a la cadena simple que se extiende entre la horquilla y el giro en U. Pese a que el pasaje por encima del sitio activo se encuentra prácticamente cerrado por dos argininas, Arg14 y Arg154, el desplazamiento de las cadenas laterales de éstas permitiría fácilmente la salida de la cadena de ADN como se muestra en la figura 4.9 (camino A). Asimismo, un camino alternativo quedaría definido por la posición del fosfato presente en la estructura del complejo TrwC-ADN25 (camino B).



*Figura 4.9. Representación de la superficie del complejo TrwC N293<sup>Y18F</sup>-ADN27. El oligonucleótido de 27 bases es representado por la cadena de color azul. En rojo y verde se muestran las cadenas de ADN siguiendo las dos posibles rutas de salida (caminos A y B). Las cadenas de ADN modeladas fueron construidas manualmente y sus conformaciones determinadas a partir de minimizaciones energéticas con el programa CNS 1.1.*

Una hipótesis interesante es que ambas rutas de salida son utilizadas por la proteína. Luego del primer corte la cadena de ADN que permanece covalentemente unida a la Tyr18 seguiría el camino A (cadena roja). TrwC continuaría desarrollando su actividad helicasa 5'→3' y por acción de la ADN polimerasa que la acompaña, el sitio de corte se regeneraría, sería reconocido nuevamente por la relaxasa y la cadena de ADN se uniría a la proteína siguiendo el camino B.

## **4.2. La relaxasa MobM**

A efectos de caracterizar estructuralmente la proteína MobM en complejo con ADN mediante difracción de rayos X, se preparó y ensayó la cristalización de diferentes complejos, utilizando oligonucleótidos de diferente longitud y cuyas secuencias están contenidas en la secuencia del origen de transferencia (*oriT*) que MobM reconoce.

Por otra parte, en los estudios realizados se utilizó, tanto la proteína entera (MobM), como una forma troncada de la misma, que comprende los residuos 2 al 198 (MobM N198) y preserva su actividad relaxasa. Los resultados obtenidos y la discusión de los mismos se detallan a continuación.

### **4.2.1. Características estructurales generales de la relaxasa MobM**

La figura 4.10 muestra el plegamiento global del dominio N-terminal de la proteína MobM en complejo con el oligonucleótido de 26 bases. La secuencia del oligonucleótido corresponde a la secuencia del origen de transferencia (*oriT*) inmediatamente antes del sitio de corte del enlace fosfo-diéster y comprende el motivo IR, responsable de la formación de la horquilla.

La estructura I se determinó a partir de la forma cristalina ortorrómbica (i) y de las monoclinicas (ii) y (iii), presentando todas ellas una estructura idéntica en la conformación de la proteína y en la de la cadena de ADN. La estructura II se determinó a partir de la forma cristalina monoclinica (iv) y difiere, ligeramente de la I, en cuanto a la conformación del extremo 3' de la cadena de ADN y al trazado de la hélice  $\alpha 1$ ,

formada por los residuos Ala13  $\rightarrow$  Lys21, que no se pudo modelar en la estructura I debido a la ausencia de densidad en dicha zona.

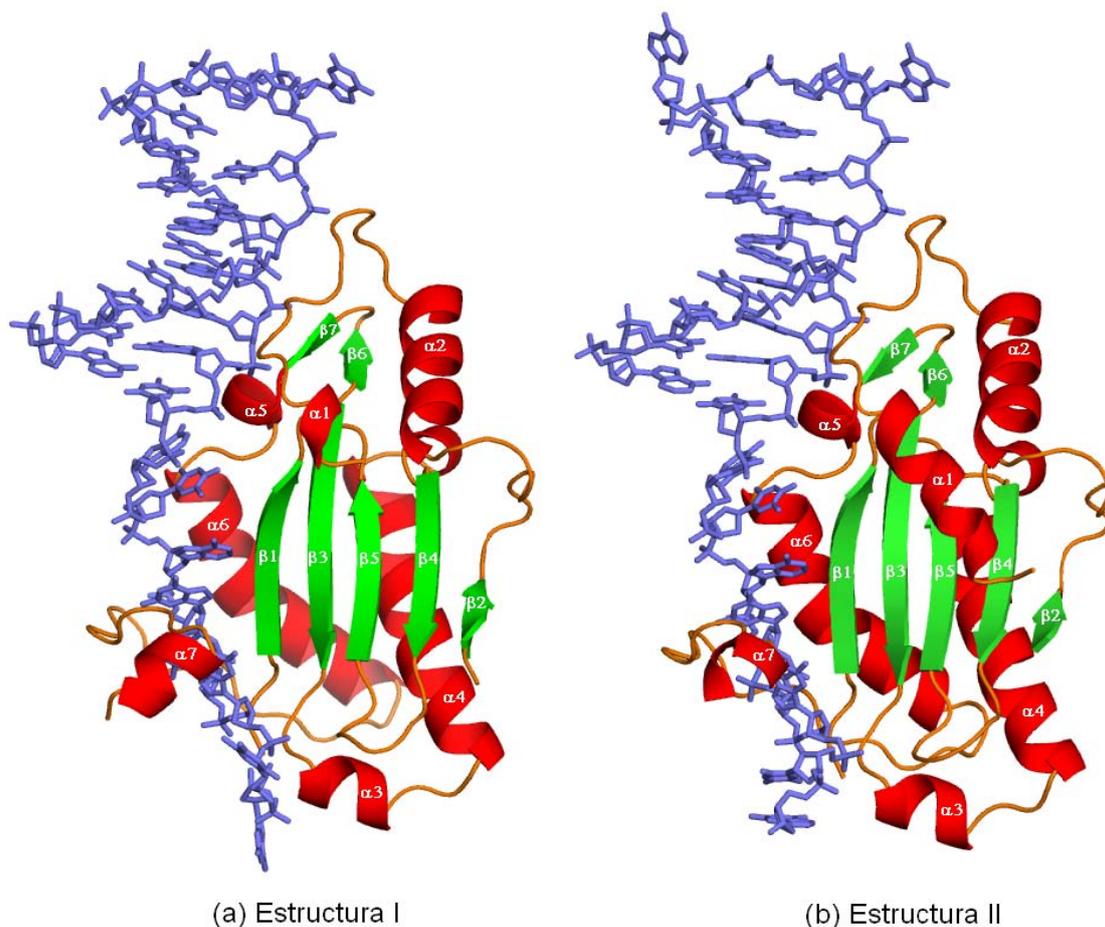


Figura 4.10. Representación global de las formas estructurales del complejo MobM N198-ADN26 en la que se indican los diferentes elementos de estructura secundaria. (a) Estructura I: correspondiente a los cristales ortorrómbicos (i) y monoclínicos (ii) y (iii) (b) Estructura II: correspondiente a los cristales monoclínicos (iv).

El núcleo central de la proteína está formado por cinco hebras beta anti-paralelas ( $\beta 1$ ,  $\beta 3$ ,  $\beta 5$ ,  $\beta 4$  y  $\beta 2$ ) que se encuentran flanqueadas por cuatro hélices alfa ( $\alpha 1$ ,  $\alpha 2$ ,  $\alpha 4$  y  $\alpha 6$ ). Las hélices alfa se disponen a ambos lados de la hoja beta quedando dos de ellas de un lado ( $\alpha 1$ ,  $\alpha 2$ ) y las otras dos del otro ( $\alpha 4$ ,  $\alpha 6$ ). La hélice  $\alpha 4$  es conectada a la hebra  $\beta 3$ , de la hoja beta central, mediante una pequeña hélice,  $\alpha 3$ . En el extremo opuesto, la horquilla beta formada por las hebras  $\beta 6$  y  $\beta 7$ , y que se encuentra unida a la hélice  $\alpha 5$ , conecta  $\beta 5$  con la hélice  $\alpha 6$ . Un extenso lazo ( $\alpha 6$ - $\alpha 7$ ), que comprende los residuos His171 $\rightarrow$ Thr188, rodea el extremo 3' de la cadena de ADN, uniendo la hélice  $\alpha 6$  con la hélice  $\alpha 7$  en el extremo C-terminal de la cadena polipeptídica.

La extensa cavidad que aloja a la cadena de ADN queda delimitada por el lazo  $\alpha 2\text{-}\beta 3$ , la horquilla beta  $\beta 6\text{-}\beta 7$ , la hélice  $\alpha 5$  y el lazo  $\alpha 5\text{-}\alpha 6$ , que contactan y participan en el reconocimiento de la horquilla y el lazo del ADN; y la hélice  $\alpha 6$ , la hebra  $\beta 1$ , el lazo  $\alpha 6\text{-}\alpha 7$  y la hélice  $\alpha 7$ , que reconocen el extremo de cadena simple del ADN.

En el alineamiento de secuencia de las relaxasas de la super-familia del plásmido pMV158 de la figura 4.11 se representan los elementos de estructura secundaria descritos. Cabe destacar el elevado grado de conservación de los residuos que constituyen la hoja beta central. En el caso particular de las hebras  $\beta 2$ ,  $\beta 4$  y  $\beta 5$  los residuos componentes se encuentran casi en su totalidad estrictamente conservados.

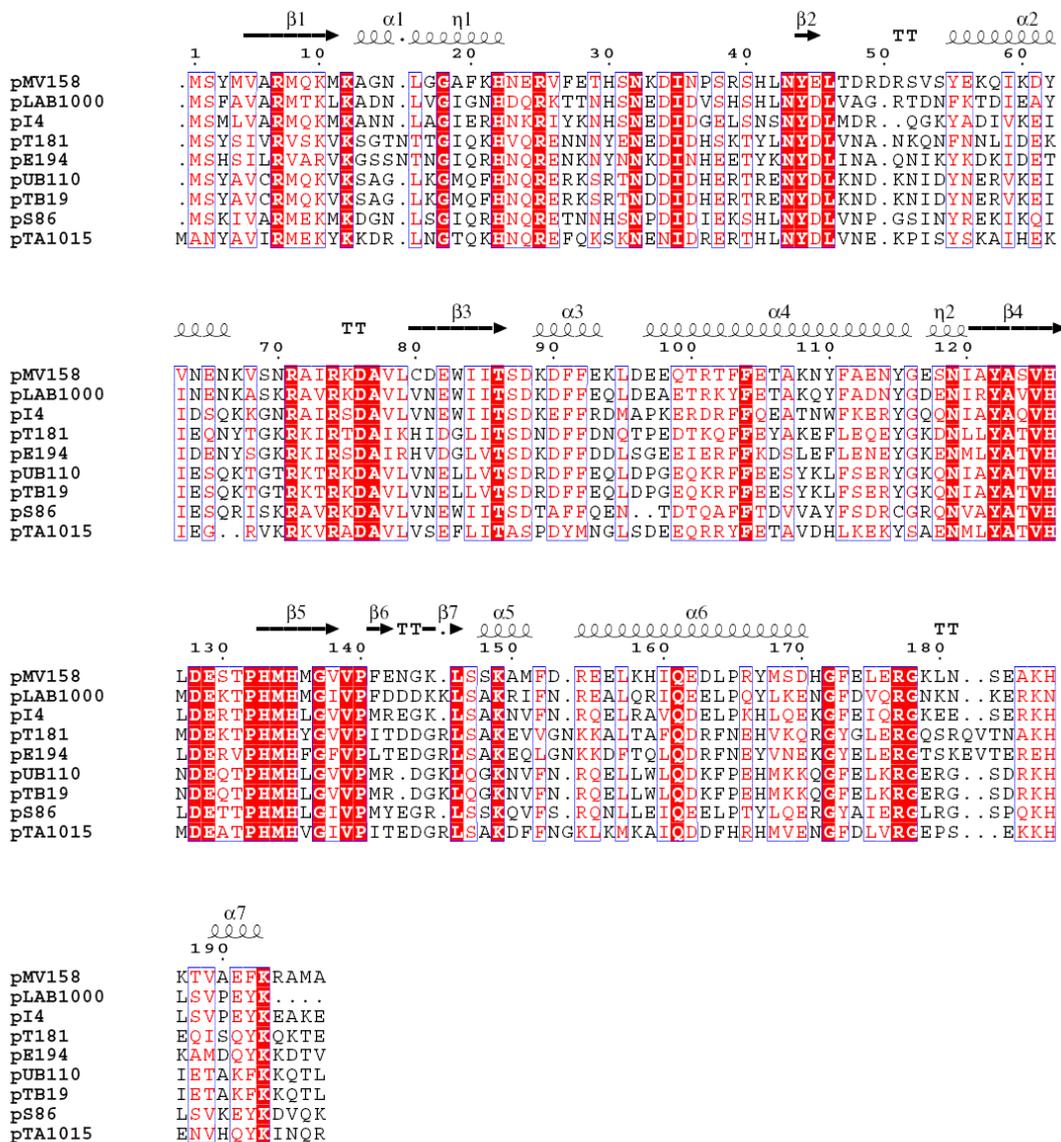
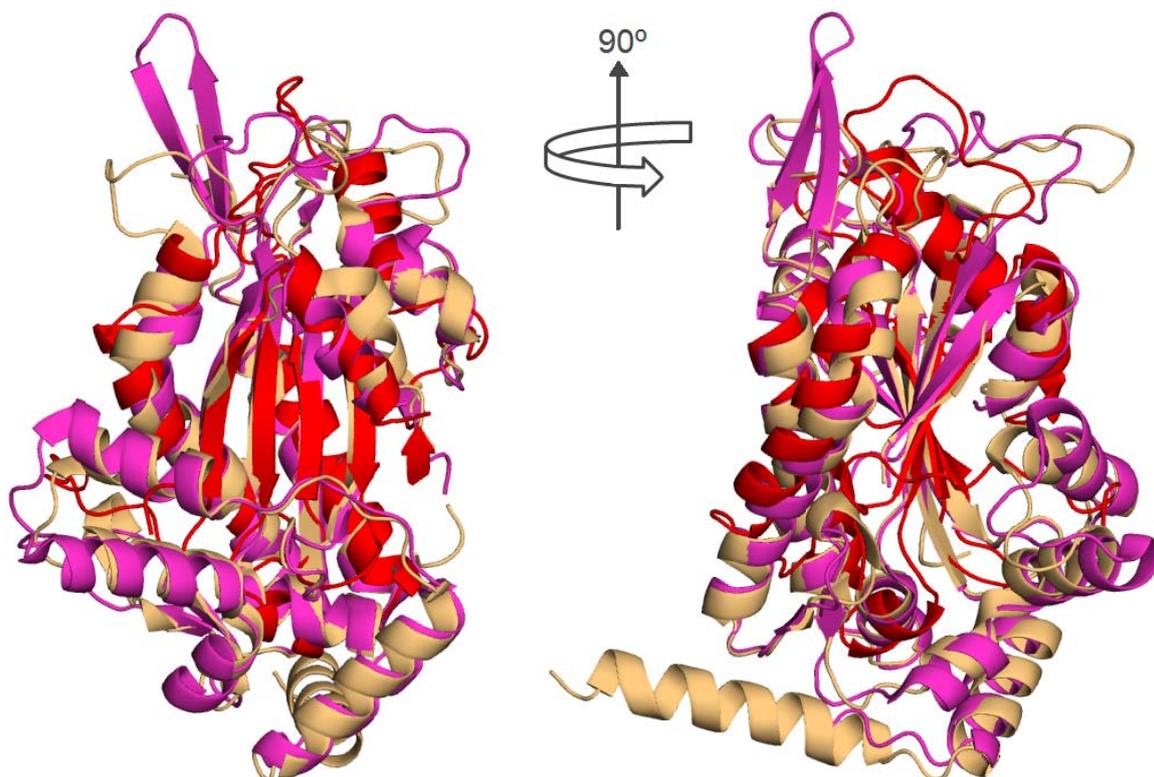


Figura 4.11. Estructura secundaria del dominio N-terminal de MobM (N198) y alineamiento de secuencias con otras relaxasas pertenecientes a la super-familia del plásmido pMV158.

El motivo estructural presente en el dominio N-terminal de la proteína MobM, constituido por las cinco hebras beta anti-paralelas que forman una hoja beta central flanqueada por cuatro hélices alfas, dos por cada lado, es el motivo estructural típico de las relaxasas. Si superponemos las estructuras de las relaxasas MobM, TrwC y TraI observamos como dicho motivo se encuentra conservado aún cuando la homología de secuencia entre ellas es extremadamente baja (figura 4.12).



*Figura 4.12. Superposición estructural de los dominios N-terminal de las relaxasas MobM del plásmido movilizable pMV158 (rojo), TrwC del plásmido conjugativo R388 (magenta) y TraI del plásmido conjugativo F (trigo). El núcleo de cinco hebras betas flanqueado por cuatro hélices alfa constituye el motivo estructural conservado en estas relaxasas.*

La figura 4.13 muestra con detalle el alineamiento de secuencia basado en la estructura para estas tres proteínas, resumiendo en la tabla 4.1 los parámetros más importantes de dicho alineamiento, como son: el número de residuos estructuralmente equivalentes ( $S$ ), el porcentaje de similitud relativa ( $\bar{s}$ ), el porcentaje de superposición de la primera proteína respecto a la segunda ( $c_q$ ), el porcentaje de superposición de la



Desde el punto de vista estructural MobM presenta un 66% de similitud con TrwC y un 62% de similitud con TraI, siendo la desviación media cuadrática 2.7 y 3.3 Å, respectivamente.

	S	$\bar{s}$	$C_q$	$C_t$	RMS (Å)	Identidad de secuencia (%)
TrwC→MobM	111	49	39	66	2.7	7
TraI→MobM	105	47	38	62	3.3	10

Tabla 4.1. Parámetros del alineamiento de secuencia basado en la estructura de los dominios N-terminales de las relaxasas MobM, TrwC (código 1omh) y TraI (código 1p4d) (TopMatch, Protein Structure Comparison, <http://topmatch.services.came.sbg.ac.at/>).

Pese al elevado grado de homología estructural de MobM respecto a TrwC/TraI, algunos de los elementos de estructura secundaria son claramente diferentes. Este es el caso del extenso lazo  $\alpha 6-\alpha 7$  y la hélice  $\alpha 7$ , ubicados en el extremo C-terminal de MobM, que no coinciden con ningún elemento estructural de las proteínas TrwC/TraI.

Asimismo se observa, en la estructura de MobM, la ausencia de la horquilla beta presente en la proteína TrwC (ver figura 4.12) y que, en dicha proteína es uno de los elementos de estructura secundaria que interviene en el reconocimiento de la cadena de ADN a través de su interacción con el surco mayor del ADN. Esta horquilla beta también se encuentra presente en la estructura de la relaxasa TraI, aunque en el complejo TraI-ADN esta región se encuentra completamente desordenada y no se observa en la estructura. La ausencia de este elemento estructural en MobM puede deberse a múltiples razones, pero probablemente esté relacionado con las características diferentes que presentan las cadenas de ADN que estas proteínas unen. Una de las diferencias más notorias es el tamaño considerablemente mayor del lazo generado por la horquilla de ADN en el caso de MobM.

#### 4.2.2. Las interacciones MobM-ADN

La unión de la cadena de ADN a la proteína involucra una extensa superficie de la misma, que representa  $2621 \text{ \AA}^2$  (30.2 % de la superficie total,  $8677 \text{ \AA}^2$ ) en la estructura I y  $2680 \text{ \AA}^2$  (30.2 % de la superficie total,  $8881 \text{ \AA}^2$ ) en la estructura II.

La superficie de interacción se encuentra cargada positivamente y define un profundo surco a lo largo del cual la cadena de ADN se sitúa.

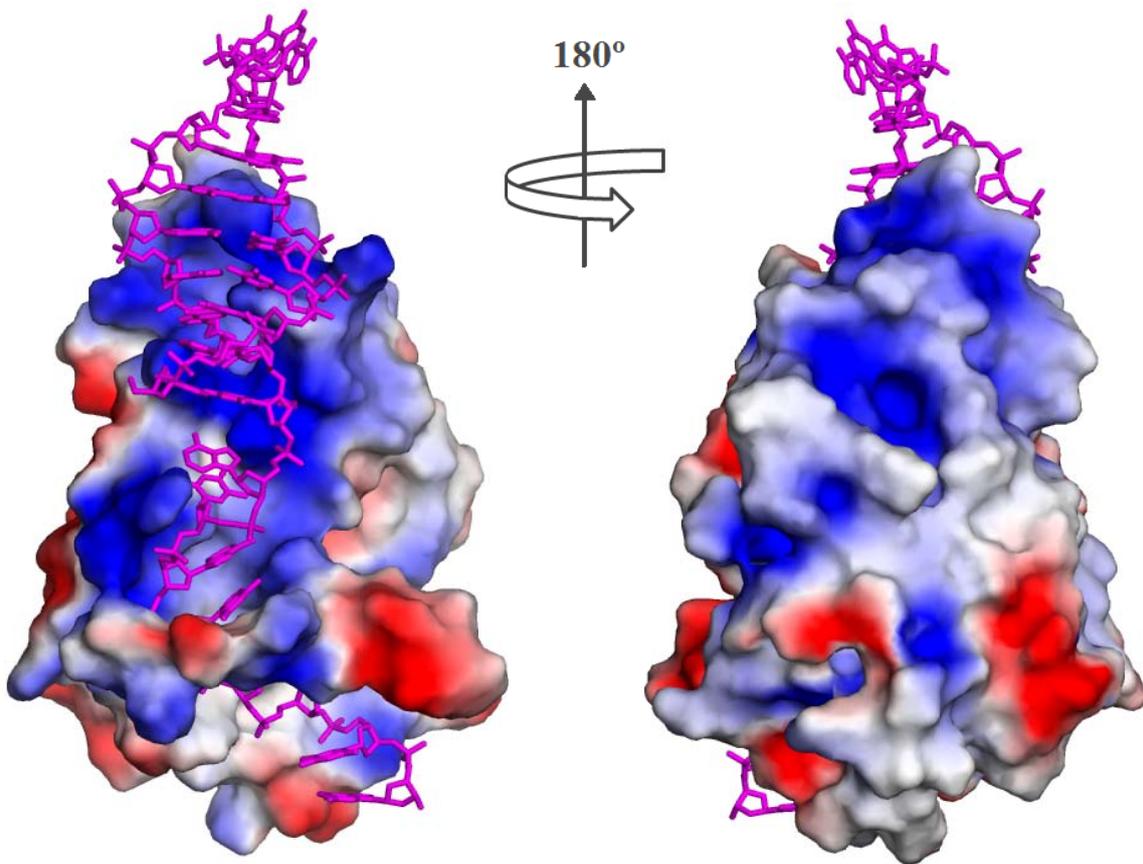


Figura 4.14. Representación de la superficie electrostática de la proteína MobM. La cadena de ADN se aloja a lo largo del extenso surco que se encuentra cargado positivamente.

Las interacciones del ADN con la cadena polipeptídica son numerosas y se producen mediante enlaces de hidrógeno, puentes salinos, interacciones de Van der Waals y apilamientos, que dan estabilidad al complejo en su conjunto (figura 4.15).

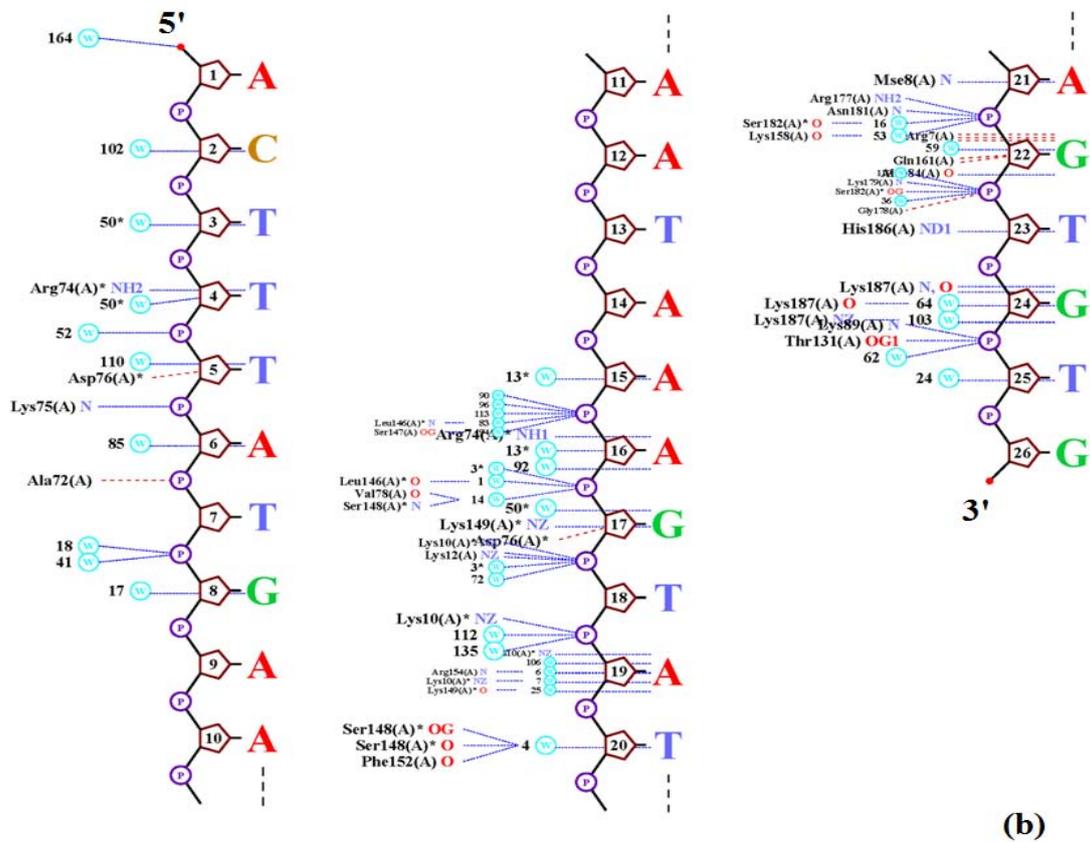
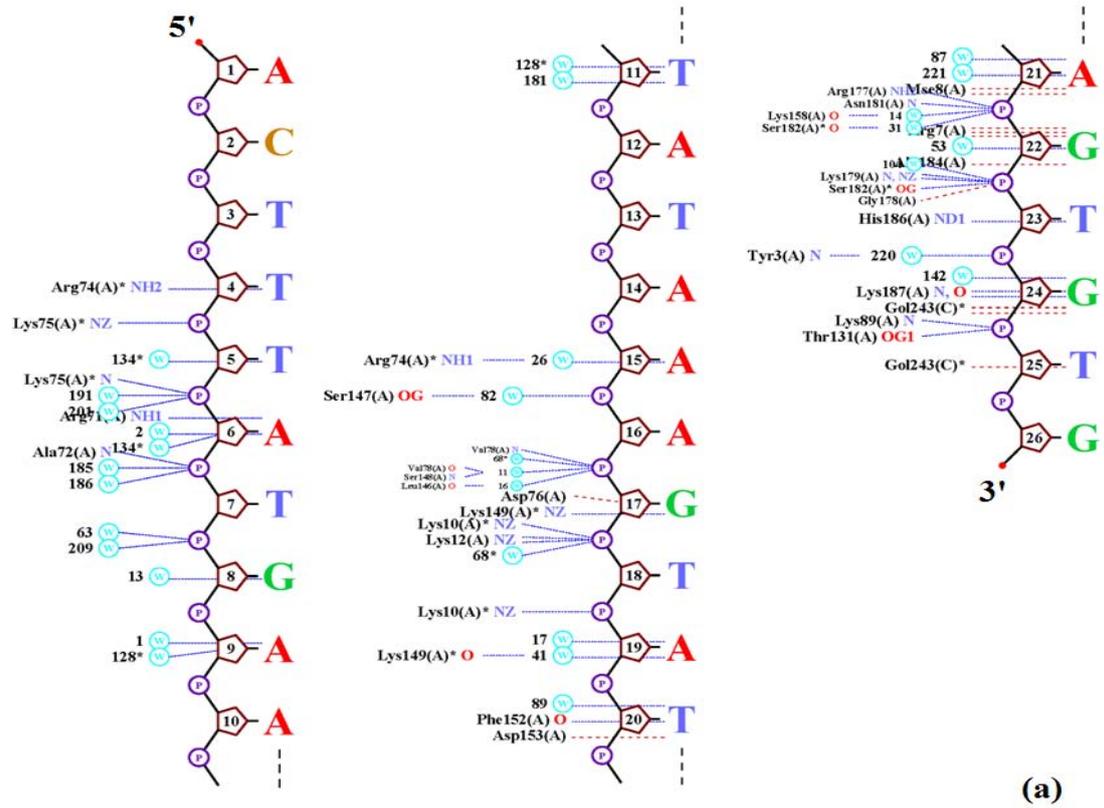
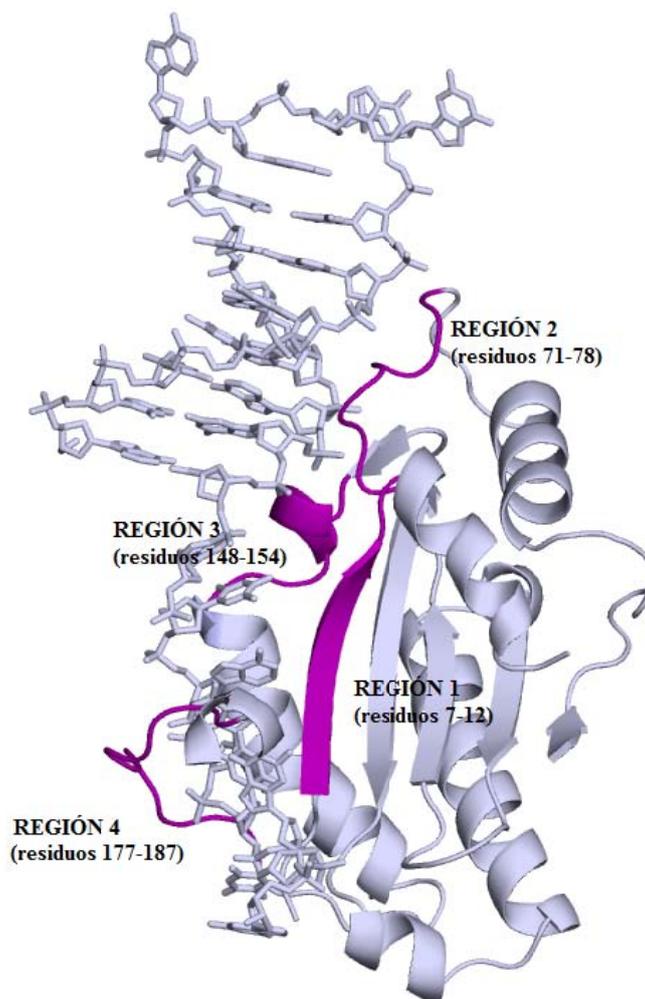


Figura 4.15. Representación esquemática de las interacciones establecidas por la cadena de ADN con la proteína y las aguas cercanas en las regiones expuestas al solvente en: (a) estructura I y (b) estructura II, del complejo MobM-ADN26 (NUCPLLOT, Luscombe et al., 1997).

Dichas interacciones están concentradas principalmente en cuatro regiones definidas de la proteína. La región 1 comprende los residuos Arg7, Met8, Lys10 y Lys12 de la hebra  $\beta$ 1 de la hoja beta central. La región 2 la forman los residuos Arg71, Ala72, Arg74, Lys75 y Val78 del lazo  $\alpha$ 2- $\beta$ 3. La región 3 está constituida por los residuos Ser148, Lys149, Phe152 y Arg154 de la hélice  $\alpha$ 5 y el lazo  $\alpha$ 5- $\alpha$ 6 y, finalmente, la región 4, que es la más extensa de todas, comprende los residuos Arg177, Lys179, Asn181, Ser182, Ala184, His186 y Lys187 del lazo  $\alpha$ 6-  $\alpha$ 7 (figura 4.16).



*Figura 4.16. Regiones de la proteína donde se concentran, principalmente, los residuos que establecen contactos con la cadena de ADN (púrpura).*

Comparando en detalle los contactos que establece la proteína con la cadena de ADN en las estructuras I y II se observa que, a pesar de la similitud estructural, algunos residuos forman interacciones que difieren entre ambas estructuras (tabla 4.2).

Interacciones proteína-ADN				Distancia (Å)	
				Estructura I	Estructura II
Arg7	NE	G22	N7	3.06	3.27
	NE	G22	O6	3.33	3.24
	NH2	G22	O6	3.18	---
Met8	N	A21	N1	3.02	2.91
	O	A21	N6	3.02	3.06
Lys10	NZ	A19	N7	---	2.93
	NZ	A19	O2P	2.93	2.93
	NZ	T18	O2P	2.77	2.91
	NZ	A19	N1	2.77	---
Lys12	NZ	A19	N3	3.22	---
	NZ	G17	O3*	2.96	2.72
	NZ	T18	P	---	3.3
Arg71	NZ	T18	O1P	2.82	2.62
	NH1	A6	N3	2.50	2.34
	NH2	T7	O4*	3.30	---
Ala72	N	T7	O1P	2.82	3.12
Arg74	NH1	A16	N3	3.11	2.98
	NH2	G17	O4*	3.27	3.22
	NH2	T4	O2	2.79	2.93
	NH2	T5	O4*	3.33	3.27
Lys75	N	A6	O1P	2.77	2.86
	NZ	T5	P	3.19	---
	NZ	T5	O1P	2.90	---
	NZ	T5	O5*	2.60	---
Val78	N	G17	O1P	2.96	3.03
Lys89	N	T25	O2P	2.96	2.99
Asp90	N	G26	O4*	3.28	---
	OD2	G26	N7	3.16	---
Thr131	OG1	T25	O1P	2.46	2.73
Ser148	N	G17	O2P	---	3.30
Lys149	NZ	G17	O6	2.73	2.65
	N	G17	O2P	3.33	---
Phe152	O	T20	N3	2.92	3.23
Arg154	N	T20	O2	3.01	2.99
	NH1	A21	O1P	2.85	---
Gln161	OE1	G22	O4*	3.31	3.34
Arg177	NH2	G22	O1P	2.75	2.70
	NH2	G22	O5*	3.25	3.29
Lys179	N	T23	O1P	2.86	2.83
	NZ	T23	O1P	2.84	---
Asn181	N	G22	O2P	2.96	2.92
Ser182	OG	T23	O2P	2.57	2.48
	O	G22	N2	3.29	---
Ala184	O	G22	N2	3.10	2.87
His186	ND1	T23	O4	2.80	2.92
Lys187	N	G24	O6	2.95	3.11
	NZ	G24	N2	---	3.24
	O	G24	N1	2.69	2.71
	O	G24	N2	3.32	---

Tabla 4.2. Interacciones proteína-ADN en las estructuras I y II en la que se han indicado los átomos que participan de la interacción y las distancias correspondientes.

Las cadenas laterales de los residuos básicos Arg7, Lys10, Lys12, Arg71, Arg74, Lys149 y Arg177, establecen contactos con las mismas bases de la cadena de ADN en ambas estructuras. Sin embargo, Lys75, Arg154, Lys179 y Lys187 establecen interacciones diferentes dependiendo de la estructura: Lys75, interactúa con la base T5 a través del N de su cadena principal en ambas estructuras, pero forma un enlace de hidrógeno con el NZ de su cadena lateral sólo en la estructura I; Arg154 forma, a través de su grupo guanidinio, un fuerte enlace de hidrógeno con la base A21 en la estructura I, pero no en la estructura II; el grupo amino de la Lys187 establece una interacción directa con la base G24 del ADN en la estructura II, que no se observa en la I.

Otros residuos, de diferente naturaleza química a los anteriormente mencionados contribuyen a la unión de la proteína a la molécula de ADN mediante la formación de fuertes enlaces de hidrógeno. Dichos residuos, Thr131, Ser182 e His186, establecen interacciones similares en ambas estructuras y representan algunas de las interacciones más potentes observadas, con distancias de enlace comprendidas entre 2.46 y 2.92 Å.

Los átomos de la cadena principal de la proteína también participan en la unión de la proteína a la cadena de ADN. Estas uniones son, sin embargo, más débiles respecto a las establecidas por las cadenas laterales, siendo sus distancias en los enlaces de hidrógeno comparativamente mayores.

Existen también interacciones entre la cadena polipeptídica y la cadena de ADN mediadas por moléculas de agua (tabla 4.3). En dichas interacciones participan, en general, los átomos de la cadena principal de la proteína, aunque también se observan con menor frecuencia, enlaces con las cadenas laterales. Es aquí donde se observan las mayores diferencias entre las estructuras I y II. Las únicas interacciones conservadas incluyen los átomos de oxígeno de la cadena principal de los residuos Lys158 y Ser182 y el grupo hidroxilo de Ser147.

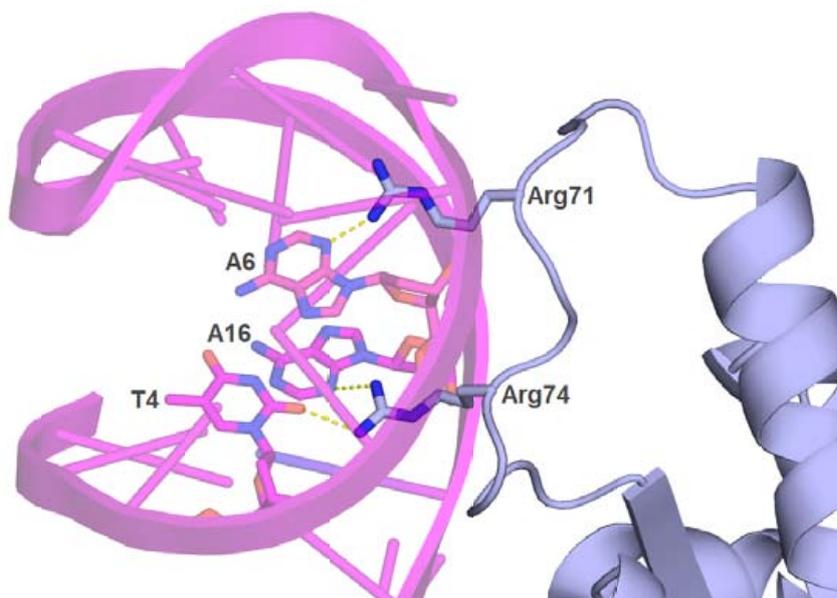
Interacciones proteína-ADN mediadas por aguas						Distancia (Å)	
						Estructura I	Estructura II
Tyr3	N	HOH220	---	G24	O1P	2.48	---
Arg74	NH1	HOH26	---	A16	O4*	3.10	---
				A15	N3	2.77	---
Val78	O	---	HOH14	G17	O1P	---	2.75
Leu146	O	---	HOH1	A16	O3*	---	3.12
	O	---	---	G17	O1P	---	2.81
	N	---	HOH83	A16	O1P	---	2.79
Ser147	OG	HOH82	HOH94	A16	O5*	3.27	3.16
				A16	O2P	2.76	2.86
Ser148	OG	---	HOH4	T20	O4	---	2.92
	N	---	HOH14	G17	O1P	---	2.75
Lys149	O	HOH4	---	A19	N6	2.71	---
Phe152	O	---	HOH4	T20	O4	---	2.92
Lys158	O	HOH14	HOH53	G22	O1P	2.55	2.59
Ser182	O	HOH31	HOH16	G22	O2P	2.70	3.00
Lys187	NZ	---	HOH103	G24	N7	---	2.67

Tabla 4.3. Interacciones proteína-ADN mediadas por moléculas de agua. En la tabla se indican, para la estructura I y II, la distancia entre la molécula de agua y el átomo de la molécula de ADN con la cual interactúa.

Algunas de las diferencias observadas en ambas estructuras son debidas a la flexibilidad inherente de la cadena polipeptídica. Por ejemplo, Tyr3 se ubica en el extremo N-terminal de la proteína, Val78 se localiza en el lazo  $\alpha 2$ - $\beta 3$ , Leu146 se encuentra en el límite entre la hoja  $\beta 7$  y la hélice  $\alpha 5$ . Por otro lado, la diferencia de conformación del extremo 3' de la cadena de ADN en las estructuras genera interacciones diferentes con las bases T23 o G24, dependiendo de la estructura que estemos considerando. Este es el caso, por ejemplo, de los contactos que establece la Lys187 con la base G24 en la estructura I y que se encuentran ausentes en la estructura II.

De las interacciones ya mencionadas, es de especial interés la interacción que establece el lazo  $\alpha 2$ - $\beta 3$  a través de las Arg71 y Arg74, Estos residuos se encuentran estrictamente conservados en las relaxasas de la super-familia del plásmido pMV158 y

juegan un papel clave en el reconocimiento y anclaje de la cadena de ADN a la proteína. Dos giros betas consecutivos del lazo  $\alpha 2$ - $\beta 3$  en el que se encuentran, orienta a estos residuos hacia el estrecho surco menor de la doble cadena de ADN, donde penetran y establecen enlaces de hidrógeno con las bases T4, A6 y A16 (figura 4.17).



*Figura 4.17. Reconocimiento de la horquilla de la cadena de ADN por medio de las argininas Arg71 y Arg74, que penetran en el surco menor del ADN y anclan al mismo a la proteína.*

Este doble giro beta, mediante el cual el lazo penetra en el surco menor de la cadena de ADN, también se encuentra presente en la relaxasa TrwC, y representa un nuevo motivo de unión al ADN que permite que las cadenas laterales penetren en el surco menor y alcancen las bases para el reconocimiento específico de secuencia. En ella, el lazo  $\alpha 3$ - $\beta 3$  entra en el surco menor situando a los residuos Ala73 y Arg75 en contacto con las bases G13 y A5, respectivamente, de manera análoga a lo observado en la estructura de MobM.

Luego de la horquilla, la cadena de ADN continúa a lo largo de la proteína en dirección al sitio activo. Si se superpone la estructura del complejo TrwC-ADN25 a la de los complejos MobM-ADN26, se observa que la posición del fosfato 23 de la cadena de ADN de la TrwC coincide con la del fosfato 24 de la cadena de ADN de la MobM, sin embargo, a partir de allí las dos bases restantes adoptan posiciones muy diferentes.

En la estructura del complejo TrwC-ADN25 un giro en U de la cadena de ADN posiciona a éste en el sitio activo de la proteína, mientras que en el complejo MobM-ADN26 este giro no se produce y, como consecuencia de esto, el extremo 3' de la cadena de ADN no ingresa al sitio activo (figura 4.18).

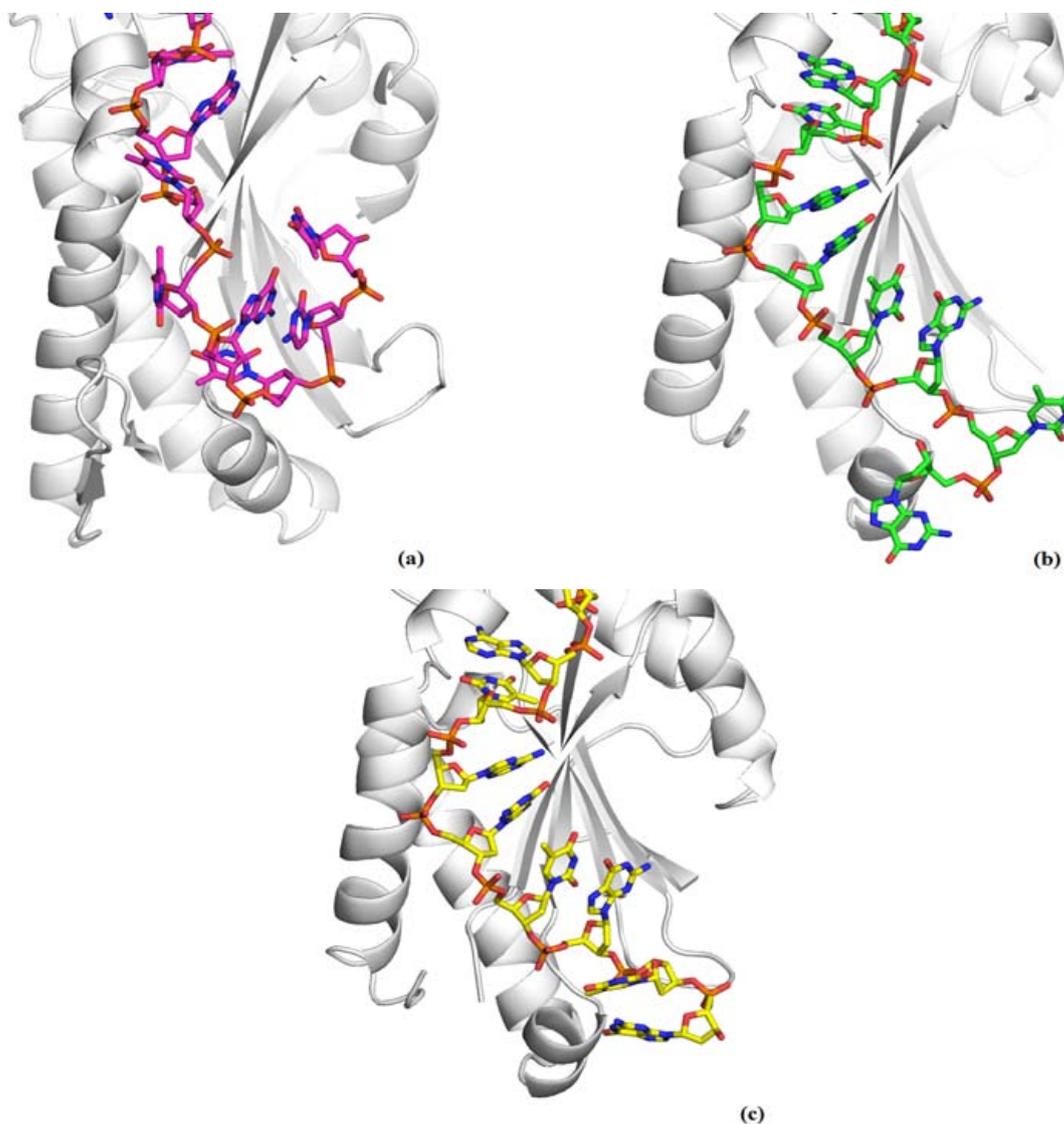
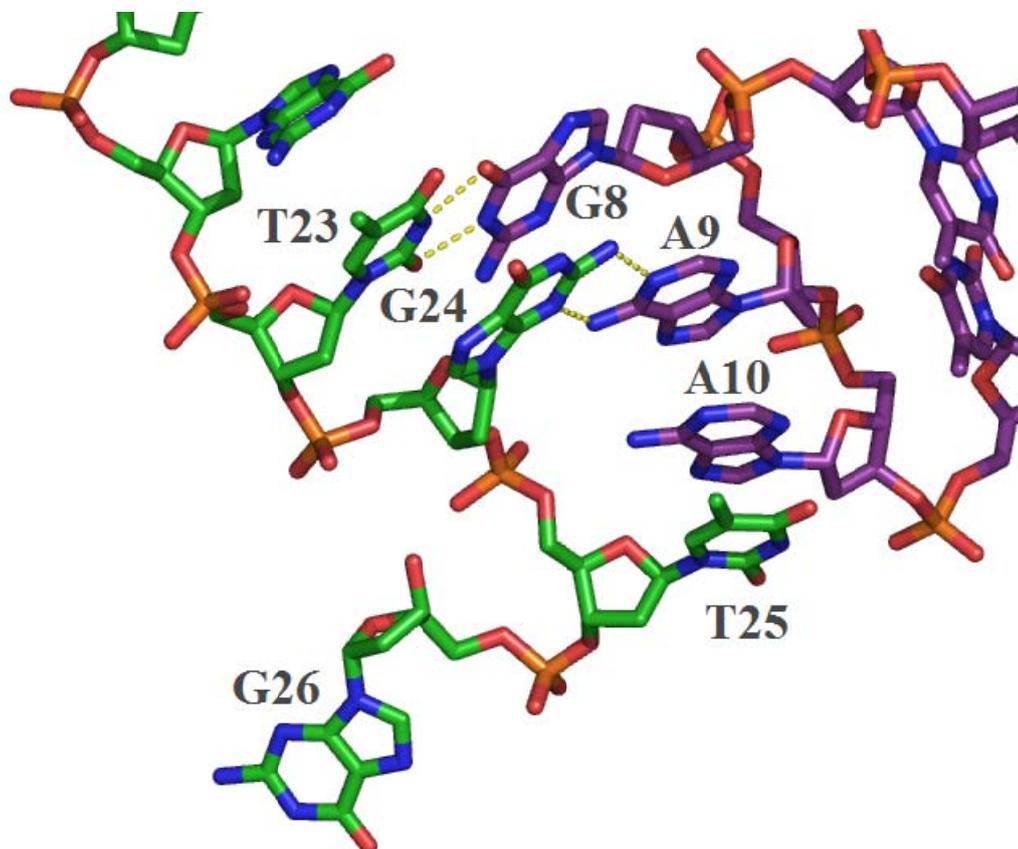


Figura 4.18. Conformación del extremo 3' de las cadenas de ADN en: (a) TrwC-ADN25, (b) MobM-ADN26 estructura I, (c) MobM-ADN26 estructura II.

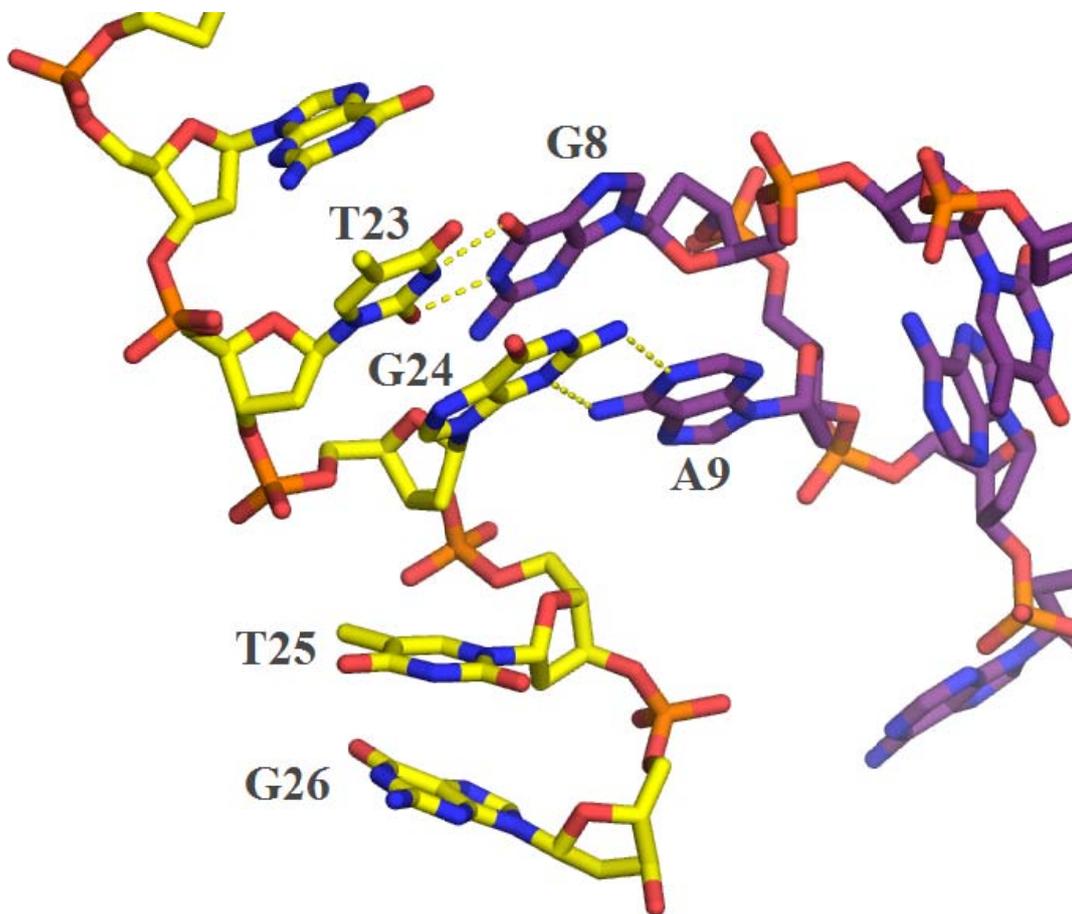
En su lugar, las bases T23, G24, T25 y G26 se acomodan de manera tal que establecen contactos mediante enlaces de hidrógeno o apilamiento con las bases de una molécula vecina de ADN.

En la estructura I del complejo MobM-ADN26, T23 y G24 forman enlaces de hidrógeno con las bases G8 y A9, mientras que la base T25 se apila paralela a la base A10 (figura 4.19).



*Figura 4.19. Interacción del extremo 3' del ADN (verde) con una molécula vecina (púrpura) en la estructura I del complejo MobM-ADN26. Las bases T23 y G24 establecen enlaces de hidrógeno con las bases G8 y A9, respectivamente. La base T25 se apila paralela a la base A10.*

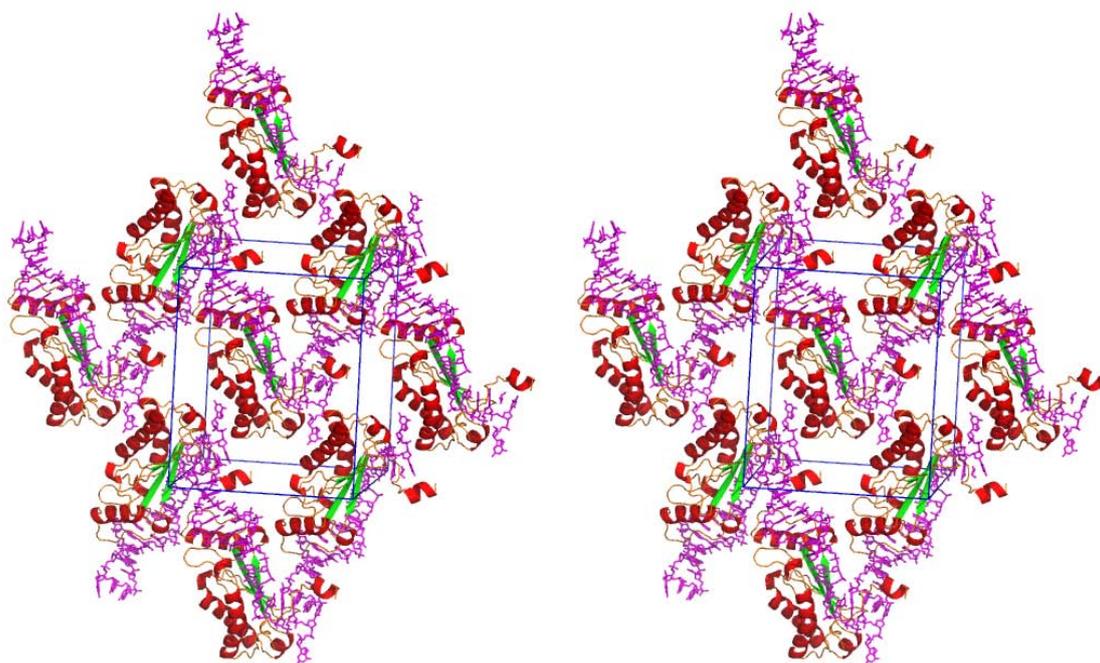
Una situación similar ocurre en la estructura II, en la que las bases T23 y G24 interactúan mediante enlaces de hidrógeno con las bases G8 y A9 de la molécula vecina de ADN y la base T25 se apila paralela a la base G26 (figura 4.20).



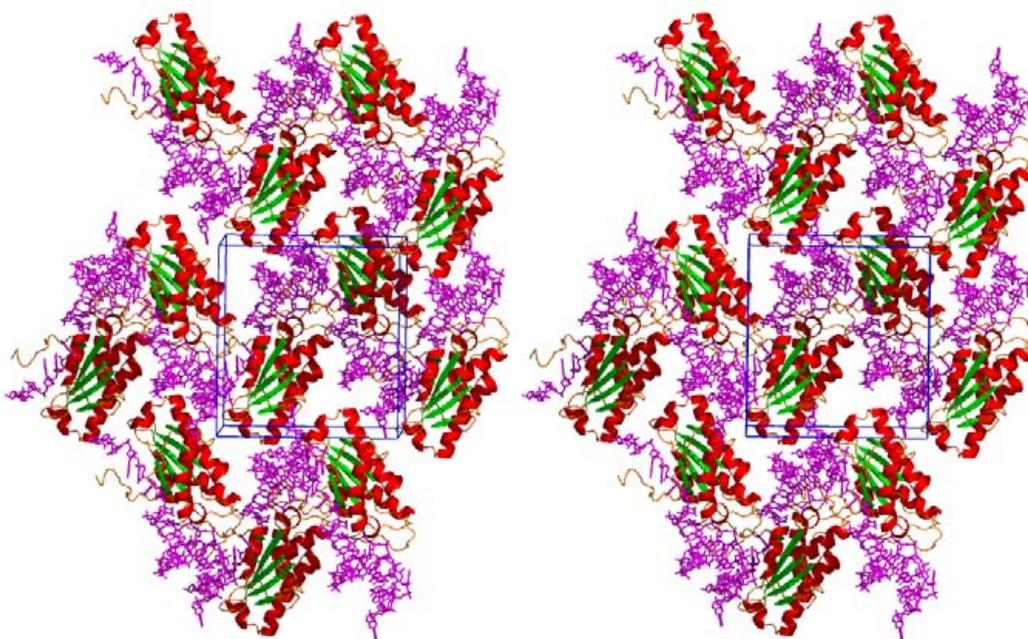
*Figura 4.20. Interacción del extremo 3' del ADN (amarillo) con una molécula vecina (púrpura) en la estructura II del complejo MobM-ADN26. Las bases T23 y G24 forman enlaces de hidrógeno con G8 y A9, respectivamente. La base T25 se apila paralela a la G26.*

Esta interacción entre cadenas de ADN vecinas rigen el empaquetamiento de las moléculas del complejo (figuras 4.21 y 4.22).

En ambas estructuras, la cadena de ADN se coloca de manera tal que el lazo que define la horquilla y que está formado por las bases G8, A9 y A10, interactúa con el extremo 3' de la cadena de ADN vecina (bases T23, G24 y T25) quedando de esta forma posicionado cerca del centro activo.



*Figura 4.21. Diagrama estéreo de empaquetamiento de las moléculas del complejo en la estructura I.*



*Figura 4.22. Diagrama estéreo de empaquetamiento de las moléculas del complejo en la estructura II.*

#### 4.2.3. La arquitectura del sitio activo

La figura 4.23 muestra en detalle los residuos situados en el centro activo de la proteína y la ubicación relativa de la cadena de ADN, que corresponde a la de la molécula vecina del complejo.

La tríada de histidinas, encargada de coordinar el metal necesario para la catálisis, está formada por las histidinas His126 (localizada en la hebra  $\beta$ 4) y las His133 e His135 (localizadas en la hebra  $\beta$ 5). La conformación de las histidinas His126 y His133 en la estructura no es la apropiada para la coordinación del metal. Sin embargo, esto puede deberse a la ausencia del mismo en el sitio activo, ya que desde el punto de vista conformacional ambos residuos tienen la flexibilidad necesaria como para orientarse adecuadamente.

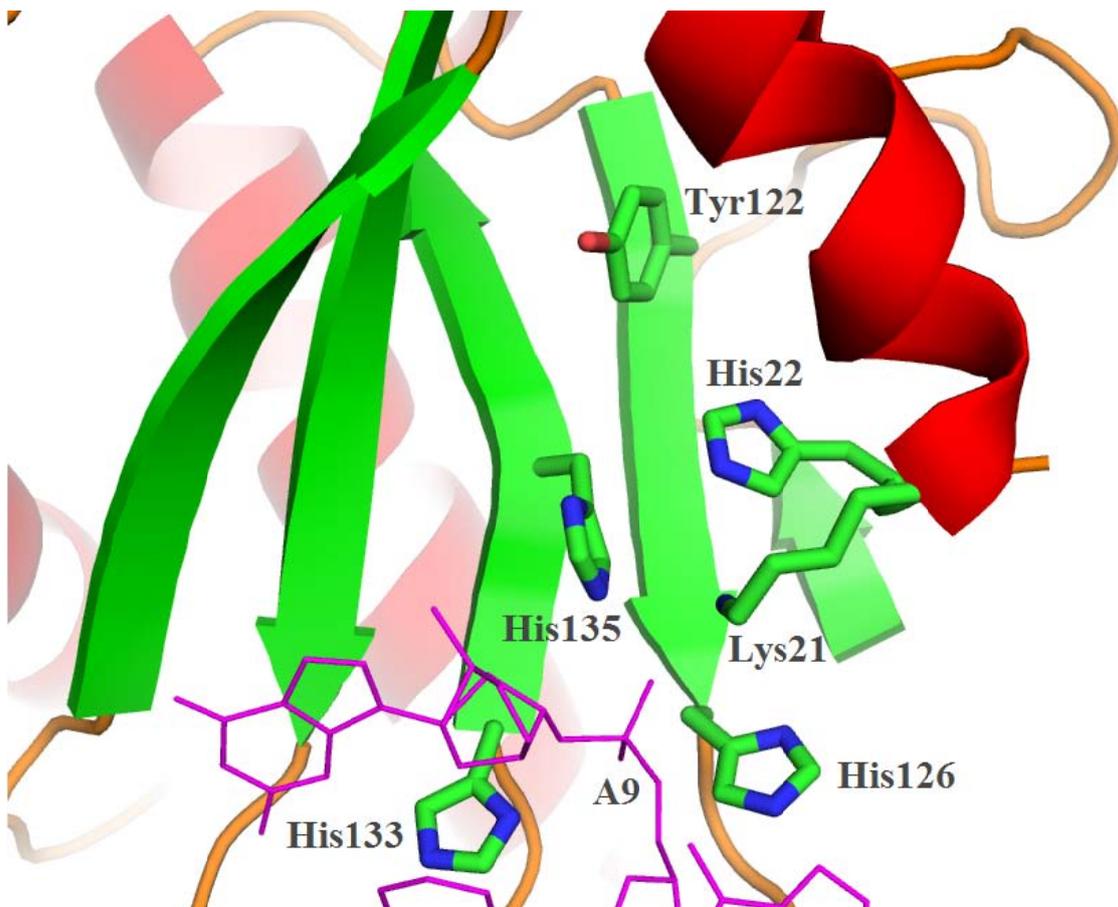


Figura 4.23. Centro activo de la relaxasa MobM. La cadena de ADN que se representa en color magenta corresponde a la de una molécula vecina del complejo. La tríada de histidinas

Junto con la tríada de histidinas, también se encuentran ubicados en el sitio activo la tirosina Tyr122 (posible residuo catalítico), la lisina Lys21 y la histidina His22.

#### **4.2.4. El residuo catalítico**

La identidad del residuo catalítico que efectúa el ataque nucleofílico sobre el fosfato del ADN y, por tanto, es el residuo esencial para catalizar el corte del enlace fosfo-diéster en la familia de las relaxasas pertenecientes a la super-familia de plásmidos movilizables pMV158, no se ha podido determinar hasta el momento y continúa siendo motivo de debate.

Los estudios de mutagénesis selectiva, realizados con la proteína Orf1 del plásmido pBBR1, perteneciente a la super-familia de los plásmidos pMV158, demostraron que, aún cuando todas las tirosinas de Orf1 son mutadas, la capacidad de movilización del plásmido no se ve afectada. Esto implicaría que el residuo catalítico involucrado en el mecanismo de movilización no es una tirosina (Szpirer *et al.*, 2001). Los autores del trabajo, sin embargo, no pudieron determinar sin ambigüedad cual sería entonces dicho residuo.

Por otra parte, las predicciones basadas en los alineamientos de secuencia de las relaxasas de esta super-familia, postulan que es la tirosina Tyr44 el residuo catalítico que utiliza MobM para iniciar el proceso de conjugación mediante el corte del enlace fosfo-diéster específico.

En las estructuras obtenidas se puede localizar esta tirosina formando parte de la hebra  $\beta$ 2, encontrándose su grupo hidroxilo a una distancia de aproximadamente 15.8 Å del grupo fosfato que contiene el enlace a cortar, cuya posición es mimetizada por el grupo fosfato de la base A9 (figura 4.24). La lejanía de este residuo, inadecuada para el ataque nucleofílico, no obstante podría ser reducida dado que su flexibilidad le permitiría adoptar una conformación más próxima situándose a 11 Å del grupo fosfato.

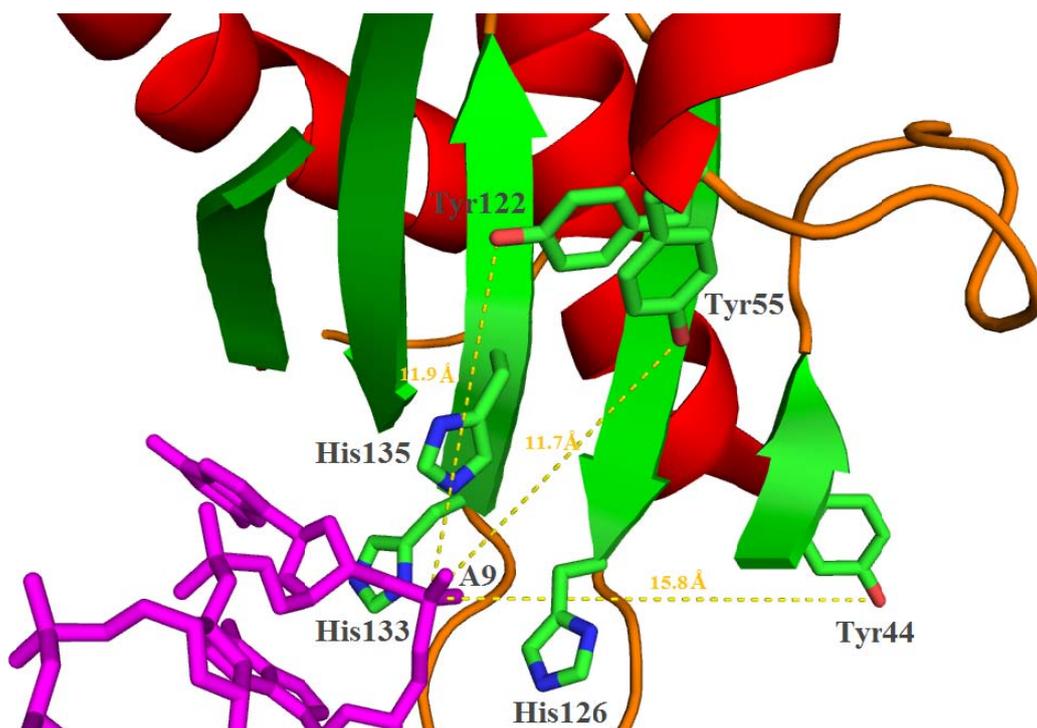


Figura 4.24. Sitio activo del complejo MobM-ADN26 en el que se han indicado para las tres tirosinas próximas, Tyr44, Tyr55 y Tyr122, las distancias al grupo fosfato de la base A9.

En las inmediaciones del sitio activo se encuentran también otras dos tirosinas, Tyr122 y Tyr55, que se ubican aproximadamente a la misma distancia del grupo fosfato. La Tyr122, se localiza en la hebra beta  $\beta$ 4, y su grupo hidroxilo se orienta hacia el sitio activo a una distancia de 11.9 Å del grupo fosfato, mientras que el grupo hidroxilo de la Tyr55, ubicada en la hélice alfa  $\alpha$ 2, se encuentra a una distancia de 11.7 Å de dicho grupo (figura 4.25).

La Tyr55, que es la más próxima de las tres, es sin embargo un residuo que no se encuentra conservado en la familia de las relaxasas homólogas a la MobM y que es sustituido por una fenilalanina (figura 4.11), por lo que es poco probable que dicha tirosina sea el residuo catalítico en la MobM.

Si se comparan los centros activos de los complejos MobM-ADN26 y TrwC-ADN25-Cu<sup>2+</sup> superponiendo ambas estructuras (figura 4.24), se observa que la tríada de histidinas se localizan en la misma posición relativa, difiriendo únicamente en la conformación (His126 e His133).

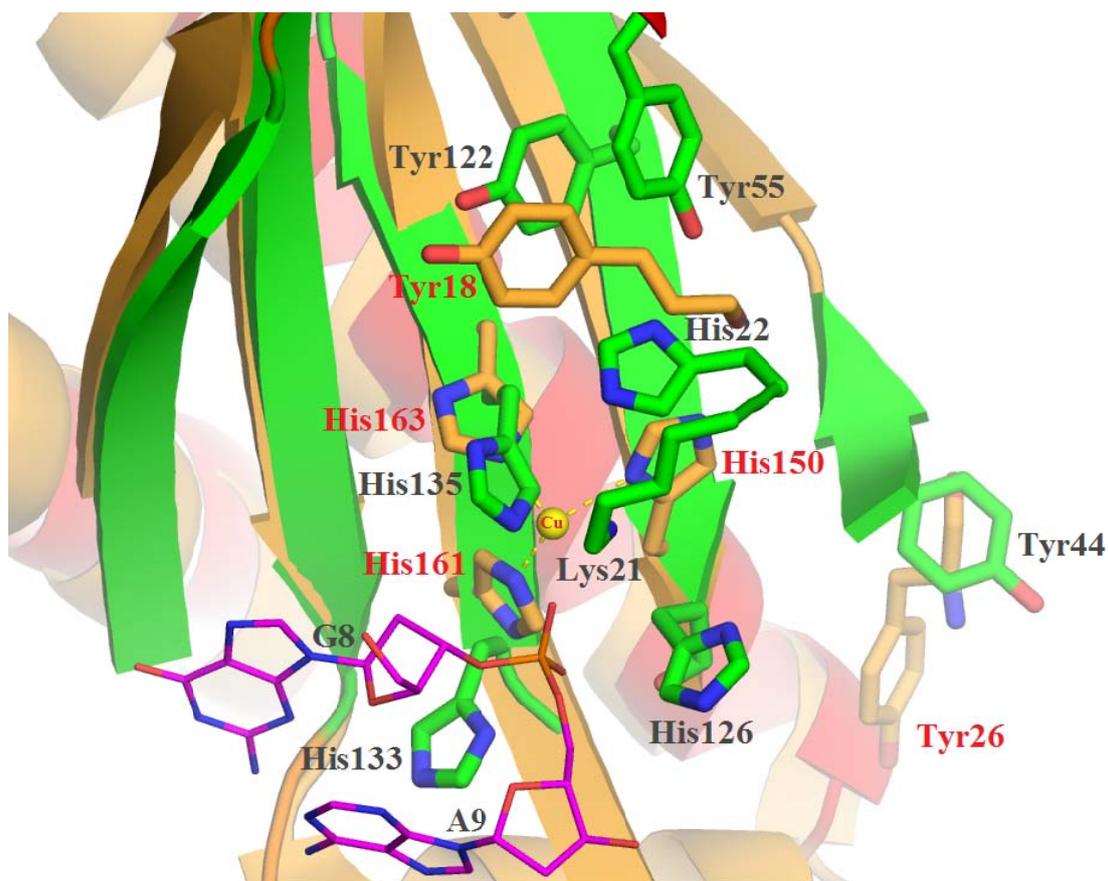


Figura 4.25. Superposición de los sitios activos de MobM y TrwC. En magenta se representan los residuos catalíticos de TrwC (His150, His161, His163, Tyr18Phe) y en verde los de MobM (His126, His133, His135, Tyr122). La cadena de ADN en color magenta corresponde a la de la molécula vecina.

Pese a que ninguna de las tres tirosinas anteriormente mencionadas se dispone exactamente en el lugar de la tirosina catalítica Tyr18, la Tyr122 es la que adopta la localización más próxima y semejante a ella. La tirosina Tyr44, por otra parte, parece ocupar un sitio estructuralmente análogo al de la Tyr26 de la proteína TrwC.

Muy próximos al sitio activo, se encuentran además una histidina, His22, y una lisina, Lys21. La función de dichos residuos se desconoce aún y existe incluso la posibilidad de que participen de manera activa en la catálisis.

Una posibilidad es que la His22 actúe como base general quitándole el protón a la tirosina para lograr su activación. En la estructura se observa la proximidad de la Tyr55 a la His22, pero no se puede descartar que un cambio de conformacional sitúe a

la Tyr122 próxima a la His22.

#### ***4.2.5. El rol y la naturaleza del metal en la actividad catalítica***

Como en el caso de TrwC, la presencia de un metal divalente es imprescindible para que MobM pueda desarrollar su actividad relaxasa. Los ensayos bioquímicos realizados, señalan al  $Mg^{2+}$  y/o  $Mn^{2+}$  como posibles metales divalentes utilizados por MobM (de Antonio *et al.*, 2004).

Para estudiar desde el punto de vista estructural el papel del metal se siguieron diferentes estrategias para la obtención de cristales del complejo MobM-ADN con diferentes metales divalentes, en particular  $Mg^{2+}$  y  $Mn^{2+}$ . Por un lado se intentó derivatizar los cristales, en forma análoga a como se hizo en el caso de TrwC y, por otro, se intentó adicionar los metales al tampón de la proteína. En ambos casos los esfuerzos fueron infructuosos y todas las estructuras obtenidas hasta el momento carecen de metal en el sitio activo.

La ausencia de metal es muy probablemente la causa de la distorsión parcial de las conformaciones de las histidinas His126 e His133. Basándonos en la superposición estructural de los sitios activos de los complejos TrwC-ADN25 y MobM-ADN26, es posible especular que una vez que el metal entra al sitio activo las histidinas se ordenan de manera tal que coordinan el metal.

#### ***4.2.6. La conformación del extremo 3' de la cadena de ADN***

La conformación que adopta el extremo 3' de la cadena de ADN en los complejos MobM-ADN26 no es acorde con lo que se esperaba, teniendo en cuenta lo observado en las estructuras previas de las relaxasas TrwC y TraI. En estas dos últimas estructuras, un pronunciado giro en U de la cadena de ADN coloca el extremo 3' en el centro activo de la proteína. En MobM, en cambio, el extremo 3' de la cadena de ADN no produce este giro y en el lugar que le correspondería se sitúa el grupo fosfato de una base de una cadena vecina que orienta el lazo determinado por la horquilla hacia el sitio activo.

La razón de que esto ocurra puede deberse a una serie de factores y existen al menos dos diferentes hipótesis para explicar lo observado. La primera de ellas es que lo que se observa en la estructura es simplemente un artefacto cristalográfico producido por el empaquetamiento de las moléculas del complejo. Sin embargo, esta hipótesis no parece ser probable, porque en todas las formas cristalinas obtenidas, a partir de condiciones de cristalización muy diferentes, el comportamiento es el mismo.

La segunda posibilidad es de origen estructural y se relaciona con la conformación que adopta el extremo C-terminal. Como puede observarse en la figura 4.26, en la estructura del complejo TrwC-ADN25 tres hélices alfa ( $\alpha 8$ ,  $\alpha 9$  y  $\alpha 10$ ) son las responsables de que la cadena de ADN encuentre el camino correcto hacia el sitio activo, sin embargo, en el complejo MobM-ADN26 estas tres hélices alfa finales no están y en su lugar se localiza un lazo flexible en cuyo extremo se encuentra una pequeña hélice alfa  $\alpha 7$ .

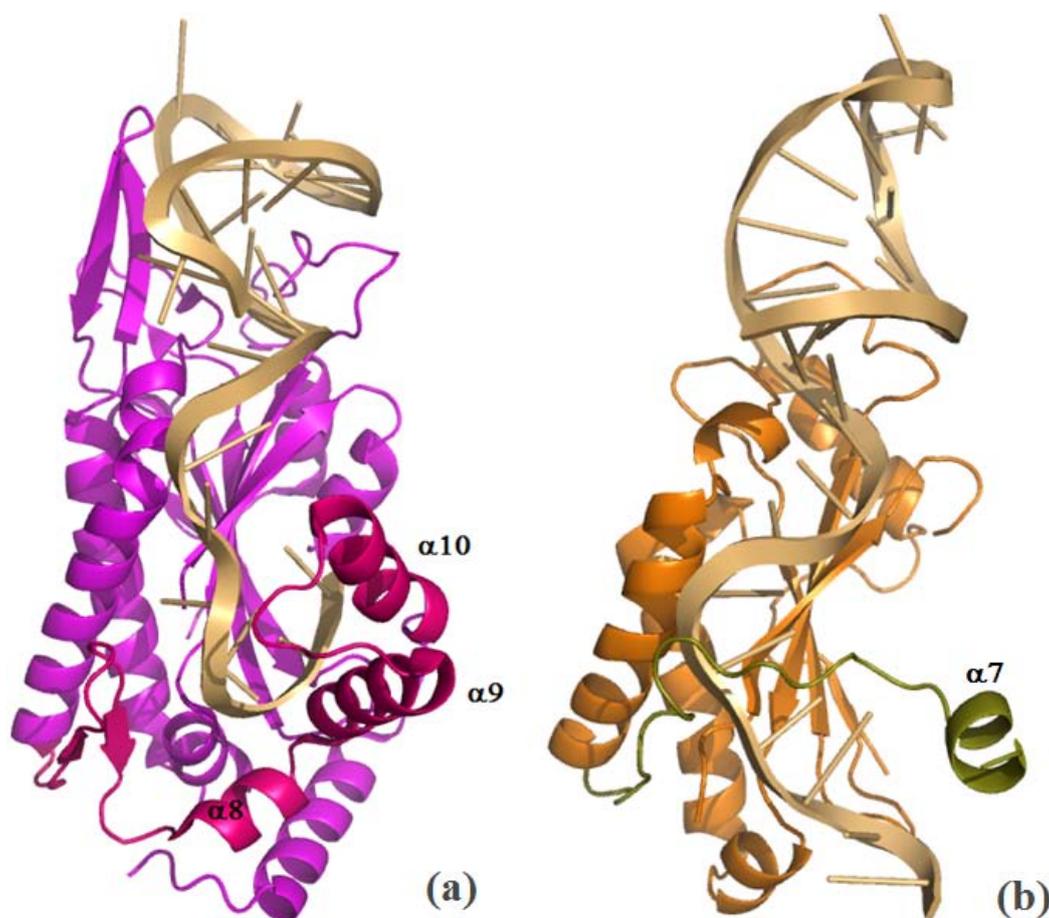


Figura 4.26. (a) Complejo TrwC-ADN25. (b) Complejo MobM-ADN26.

Cuando analizamos los motivos por el cual estas dos hélices alfa no están presentes en la estructura de la MobM y estudiamos nuevamente la predicción de estructura secundaria (figura 1.14), vemos que las hélices ausentes se encuentran definidas por los residuos 217 al 222 y 226 al 243 que fueron eliminadas al diseñar el constructo.

Si la relevancia de estos dos elementos estructurales en la correcta conformación del extremo 3' de la cadena de ADN guiándola hacia el sitio activo es tal, lo que observamos en la estructura es entonces el producto de la eliminación de dichos residuos que afecta, por otra parte, también la conformación del lazo y la hélice  $\alpha 7$ .

#### ***4.2.7. El mecanismo catalítico***

Una de las características estructurales notoriamente diferente, cuando comparamos la estructura del complejo con ADN del extremo N-terminal de la proteína MobM con otras relaxasas de estructura conocida, es la forma como el lazo de la horquilla de ADN se sitúa en el sitio activo.

Si bien es cierto que esta particular conformación puede deberse simplemente al hecho de que el extremo 3' de la cadena de ADN no entra al sitio activo, por alguno de los motivos descritos anteriormente, y por lo tanto el sitio que queda libre es ocupado por el lazo que define la horquilla, también cabe la posibilidad de que esto tenga relación con la manera como MobM desarrolla su actividad catalítica.

La posición que adopta el lazo de la horquilla, situando el grupo fosfato de la base A9 en el lugar que debería tener el fosfato que contiene el enlace a cortar por la MobM, parece estar mimetizando la conformación que adoptaría la cadena de ADN en la conformación IR-2 a punto de ser cortada por la MobM. En esta conformación, la base G26 que se ubica en el lazo define el sitio de corte que la proteína reconoce (figura 4.27). De manera análoga, en la estructura se observa que una base G (la base G8) precede a la base A9 que expone su fosfato para ser cortado. Como la proteína no puede cortar en este punto, el lazo queda inalterado y lo que vemos en la estructura podría representar una imagen instantánea del momento en que el ADN en la conformación IR-2 expone el enlace fosfo-diéster a cortar al centro activo de la proteína.

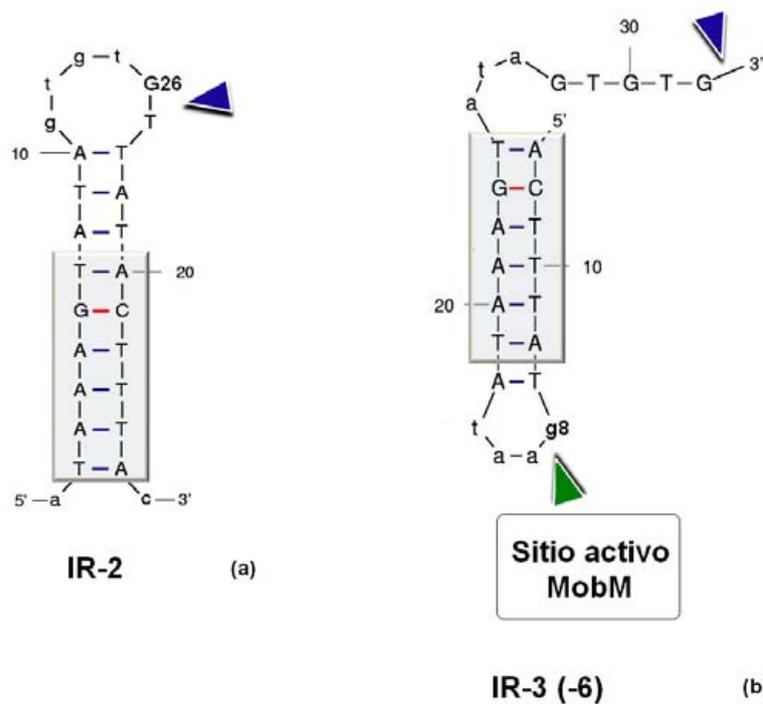


Figura 4.27. (a) Conformaciones IR-2 (a) e IR-3(6) (b) del ADN, donde se ha indicado la base G26 que se localiza inmediatamente antes del sitio de corte que reconoce MobM. Asimismo se ha representado en (b) la localización relativa de la base G8, observada en la estructura, respecto al centro activo.

Para que esto sea posible, una vez que MobM se ha unido a la cadena de ADN super-enrollado en la conformación IR-3, debe ocurrir un cambio conformacional de manera tal que la cadena adopta la conformación IR-2 y deja el sitio de corte (*nic*) como cadena simple haciéndolo accesible y permitiendo, de esta manera, el corte.

Si tenemos en cuenta que la proteína MobM sin trincar actúa como dímero en solución, podríamos relacionar lo observado en la estructura cristalina con lo propuesto por Espinosa y colaboradores, suponiendo que en la estructura se está mimetizando de alguna manera este estado de transición.

Una hipótesis posible es que mientras un monómero permanece unido a la cadena de ADN que se encuentra en la conformación IR-3, en el otro el cambio conformacional de la cadena de ADN se ha producido pasando a estar en la conformación IR-2 y orientando el lazo de la horquilla que contiene el sitio *nic* hacia el sitio activo del otro monómero para que la reacción de corte se produzca.

Al no poseer datos estructurales del dímero de MobM, ni de los complejos con oligonucleótidos de secuencia IR-2, lo anterior no pasa de una simple especulación y no puede ser fehacientemente demostrado en base a la información con la que se cuenta.