Universidad de Cantabria

Facultad de Medicina

Departamento de Biología Molecular

Las proteínas RHH confieren especificidad a la conjugación bacteriana

Memoria presentada para optar al grado de Doctor por

Yera Tena Primo

Santander, junio de 2013

El presente trabajo ha sido realizado en el Departamento de Biología Molecular de la Universidad de Cantabria, bajo la dirección del Catedrático Fernando de la Cruz Calahorra y el Profesor Gabriel Moncalián Montes, gracias a una beca predoctoral de la Fundación Marqués de Valdecilla (2007-2010).

Durante este periodo se realizó una estancia de tres meses en el laboratorio de la Dra. Phoebe Rice (Universidad de Chicago, EEUU), gracias a una ayuda de la Fundación Marqués de Valdecilla.

D. Fernando de la Cruz Calahorra, Catedrático de Genética, y **D. Gabriel Moncalián Montes**, Profesor Titular de Genética, ambos pertenecientes al departamento de Biología Molecular de la Universidad de Cantabria,

CERTIFICAN: Que Dña. Yera Tena Primo ha realizado bajo nuestra dirección el trabajo que lleva por título "Las proteínas RHH confieren especificidad a la conjugación bacteriana".

Consideramos que dicho trabajo se encuentra terminado y reúne los requisitos necesarios para su presentación como memoria de Doctorado, al objeto de poder optar al grado de Doctor por la Universidad de Cantabria.

Santander, a 19 de junio de 2013,

Fdo. Fernando de la Cruz Calahorra Gabriel Moncalián Montes

Aunque todos los que me habéis acompañado durante este tiempo sabéis que parte de esta Tesis es tan mía como vuestra quería aprovechar esta oportunidad que se me brinda para agradeceros a todos, guías, compañeros y amigos todo el apoyo que me habéis dado, ya que sin vosotros no me encontraría ahora escribiendo estas palabras.

Agradecer en primer lugar a Fernando y a Gabi la oportunidad que me brindaron, así como las indicaciones y consejos que me han servido para llegar hasta aquí. Agradecer también a todos mis compañeros, pasados y presentes por todo lo que he aprendido con ellos y de ellos. A Iñaki y Elena, por resolver dudas varias sobre proteínas. A Dani y Blanca, por facilitar mis primeros pasos y hacer que lograra entender aquellas conversaciones científicas que al principio me sonaban a algo abstracto. A Wiki-Mapi por compartir su sabiduría y estar siempre dispuesta a ayudar, por esos alineamientos sin igual. A Ruli, otro pozo de sabiduría, por tantos momentos humeantes que ayudaban a desconectar y ver las cosas desde otra perspectiva. A Mati y a Sandra, ¿qué sería de mí sin vosotras? A los chicos Alejandro, Jorge, Lillo, Juantxo, David…por echarme una mano siempre, por hacerme reir y sentirme a gusto en el laboratorio y fuera de él, se os echa de menos. A MC, gran amiga, que aún no siendo parte del laboratorio has tenido que aguantar nuestros lamentos. A Inma, Leti, Carlos, Ana, Sheila, María, Val, Raquel, Sandrita, Andrés, Delfi…de todos he aprendido algo.

A Irene (y a Gonzalo), Esther y Héctor, ¿qué puedo decir? Sin duda, mi mejor recuerdo de Santander. Por tantos momentos juntos, por compartir penas y alegrías, por estar siempre ahí, por ayudarme en todo, por acogerme. Gracias a vosotros he podido acabar esta etapa, y espero teneros presentes en todas las que están por llegar.

A mi familia, por apoyarme siempre y por ayudar en los momentos más difíciles. Por supuesto, a Juan, por su paciencia infinita y por muchas cosas más…entre ellas Telmo, mi mejor resultado, mi motivación y alegría, mamá promete atenderte como mereces ahora que doy fin a este documento.

A todos, gracias.

A mis chicos.

Índice

ABREVIATURAS 1

1. INTRODUCCIÓN 3

1. Transferencia genética horizontal 4
2. Conjugación bacteriana 7
    2.1 Modelo de conjugación 9
3. El plásmido R388 12
    3.1 Región Mpf 14
    3.2 Región Dtr 15
        3.2.1 *oriT* de R388 17
        3.2.2 Orígenes de transferencia de otros sistemas 19
            a) *oriT* MOB_P 19
            b) *oriT* MOB_Q 20
            c) *oriT* MOB_{F12} 21
        3.2.3 Proteína acopladora 22
        3.2.4 La relaxasa 25
        3.2.5 El relaxosoma 31
4. Proteínas auxiliares 31
    4.1 Proteínas auxiliares de la conjugación codificadas por el cromosoma huésped 31
        4.1.1 IHF en la conjugación 33
        4.1.2 IHF en R388 35
    4.2 Proteínas auxiliares de la conjugación codificadas por el Plásmido 37
        4.2.1 Sistema MOB_P 37
        4.2.2 Sistema MOB_Q 39
        4.2.3 Sistema MOB_F 40
        4.2.4 Sistema MOB_{HEN} 43
        4.2.5 TrwA de R388 43
5. Superfamilia de proteínas RHH 45

2. OBJETIVOS 50

3. MATERIALES Y MÉTODOS 54

MATERIALES 56
1. Cepas bacterianas 56
2. Medios de cultivo 56
3. Plásmidos utilizados 57

3.1 Construcción de plásmidos concretos 59
4. Oligonucleótidos 61
MÉTODOS 62
1. Métodos generales de Biología Molecular 62
2. Métodos microbiológicos 68
3. Métodos bioquímicos 71
4. Métodos *in silico* 79

4. RESULTADOS 82

1. La deleción de los elementos de estructura secundaria del *oriT* restan eficiencia de movilización 84
2. Los plásmidos R388 y R7K se diferencian casi exclusivamente en el gen trwA y los sitios de unión de la proteína en el *oriT* 89
3. Cada TrwA reconoce específicamente sus sitios de unión 91
4. La unión al ADN y la movilización se ven afectadas cuando alteramos los sitios de unión de TrwA 98
5. Efecto de las alteraciones de los sitios de unión de TrwA en la actividad del promotor *trwABC* 103
6. Determinación de los contactos específicos de TrwA con el ADN 109
7. La especificidad de TrwA no depende sólo de los aminoácidos polares de su hoja beta 113
8. Cristalización TrwA-ADN 118
8.1 Predicción *in silico* de la estructura de TrwA unida al ADN 121
9. Interacciones de TrwA con otras proteínas del relaxosoma 124
9.1 Interacción TrwA-TrwC 128
9.2 Efecto de IHF en la unión de TrwA 129
10. Nuevas proteínas accesorias similares a TrwA 133
10.1 Pseudomonas stutzeri 135
10.2 Bordetella parapertussis 137
10.3 Otras proteínas RHH 138

5. DISCUSIÓN 144

1. Diferencias puntuales en las proteínas TrwA y sus sitios de unión en R388 y R7K son responsables de dotar de especificidad al sistema 147
2. La disposición que adoptan los sitios de unión de TrwA en $oriT_{R388}$ es vital para asegurar el correcto funcionamiento de la proteína 148

3. La distribución característica de los sitios de unión de TrwA de R7K y R388 también se encuentra en otros plásmidos 152
4. Posible interacción de TrwA con otras proteínas del relaxosoma 156
5. Determinación de los contactos específicos de TrwA con el ADN 159

6. CONCLUSIONES 164

7. BIBLIOGRAFÍA 168

Abreviaturas

RHH	Ribbon hélix hélix
TGH	Transferencia genética horizontal
EGM	Elementos genéticos móviles
ADN	Ácido desoxirribonucléico
oriT	Origen de transferencia
T4SS	Sistema de secreción tipo IV
Mpf	Formación del par conjugativo (maiting pair formation)
Dtr	Transferencia y replicación del ADN (DNA transfer and replication)
orf	Marco de lectura abierto (open reading frame)
ATP	Adenosin trifosfato
aa	Aminoácido
Kb	Kilobase
pb	Par de bases
A, T, G, C	Adenina, timina, guanina, citosina
IR	Repetición invertida (inverted repeat)
DR	Repetición directa (direct repeat)
IHF	Factor de integración del huésped (integration host factor)
Å	Angstrom
RCR	Replicación por círculo rodante (rolling circle replication)
SDS	Dodecilsulfato sódico
N-terminal	Amino terminal
C-terminal	Carboxilo terminal
T-DNA	ADN transferido
mer	Medida de longitud de un oligonucleótido
TrwCr	Dominio relaxasa de TrwC
$TrwC_H$	Dominio helicasa de TrwC
ICE	Elemento conjugativo e integrativo (integrative and conjugative element)
KDa	Kilodalton
PDB	Banco de datos de proteínas (protein data bank)
LB	Luria-Bertani
g	gramo
g/l	Gramo por litro
p/v	Concentración en peso por volumen
v/v	Concentración en volumen por volumen
min	Minutos
µg	Microgramo
pg	Picogramo
ml	Mililitro
PCR	Reacción en cadena de la polimerasa (polymerase chain reaction)
DO	Densidad óptica
mM	Milimolar
cm	Centímetro
Tm	Temperatura de desnaturalización (melting temperatura)
U	Unidades de actividad enzimática
EDTA	Ácido etileno-diamino-tetracético
TBE	Tris-boro-EDTA

r.p.m	Revoluciones por minuto
PAGE	Electroforesis en gel de poliacrilamida (Poliacrilamide gel electrophoresis)
PEG	Polietilenglicol
IPTG	isopropil-β-D-tiogalactopiranósido
PMSF	Fenil-metil-sulfonil-fluoruro
Tris	Trihidroximetil-amino-metano
MES	Ácido 2-morfolino etano sulfónico
HEPES	Ácido N-2-hidroxietilpiperacina-N`-2-etanosulfónico
GFP	Proteína fluorescente verde (green fluorescent protein)
ESRF	Instalaciones europeas de radiación sincrotrón (European synchrotron radiation facilities)

1.

INTRODUCCIÓN

1. INTRODUCCIÓN

1. TRANSFERENCIA GENÉTICA HORIZONTAL

La transferencia genética horizontal se define como la transmisión no-genealógica de material genético de un organismo a otro (Goldenfeld and Woese, 2007). La transferencia genética horizontal es un evento más versátil que la propia reproducción, al producirse entre organismos no estrechamente emparentados evolutivamente hablando. Se puede producir entre distintos reinos en una u otra dirección, habiéndose descrito este proceso entre bacterias, desde bacterias a arqueas (Rest and Mindell, 2003) y viceversa (Gophna *et al.*, 2004), desde arqueas a eucariotas (Andersson *et al.*, 2005), desde bacterias a eucariotas (Watkins and Gray, 2006), desde eucariotas a bacterias (Guljamow *et al.*, 2007) e incluso entre eucariotas (Nedelcu *et al.*, 2008). Aun así, ha sido más ampliamente aceptada y documentada en la evolución de las bacterias y arqueas, estimándose que entre un 1 y 32% de los genes de los diversos genomas microbianos han sido adquiridos por transferencia genética horizontal (TGH) (Koonin *et al.*, 2001).

Al ser el flujo de la información genética tan abundante entre procariotas, permite que las bacterias y arqueas tengan un genoma "core" pequeño conservado dentro de una especie, y un gran pangenoma muy variable (Tettelin *et al.*, 2008). La transferencia genética horizontal presenta varias implicaciones evolutivas (de la Cruz and Davies, 2000; Levin and Bergstrom, 2000; Ochman *et al.*, 2000): permite la recombinación homóloga entre cepas o especies emparentadas y la integración de nuevo material genético, ayudando en la adaptación a nuevos ambientes o a condiciones de estrés, y por último, se ha planteado que la movilidad génica dirige los procesos cooperativos microbianos (Nogueira *et al.*, 2009).

Es importante resaltar que estos mecanismos de TGH son responsables de la diseminación de los genes de resistencia a antibióticos, un problema de salud mundial, y ha dotado a este campo de un interés adicional (de la Cruz and Davies, 2000).

Existen tres mecanismos principales por los que se produce la TGH en el mundo bacteriano. Estos mecanismos esquematizados en la figura I1 son los siguientes:

I. Transformación. Es la incorporación de ADN libre del entorno que se produce de forma natural en muchas especies bacterianas (*Acinetobacter*, *Synechocystis*, *Bacillus*, *Neisseria*, o *Helicobacter*). Las bacterias capaces de ser transformadas se denominan competentes. En condiciones de laboratorio se ha conseguido la transformación de tipos celulares carentes de competencia natural mediante diversos métodos como la electroporación o la transformación por choque térmico. Puede ocurrir recombinación homóloga entre el ADN incorporado y el genoma de la célula que lo capta, dando lugar a un nuevo fenotipo.

II. Conjugación. Se denomina así a la transferencia de una molécula de ADN entre dos células que establecen contacto por medio de una estructura proteica conocida como *pilus*. Este mecanismo se explica con más detalle en el apartado 2 de esta introducción, ya que es el principal proceso estudiado en esta Tesis.

III. Transducción. Es la transferencia de ADN mediante la infección por bacteriófagos. Generalmente, el fago contiene exclusivamente ADN vírico, aunque hay ocasiones en las que incorpora ADN bacteriano y lo introduce en otra célula al infectarla. Este nuevo ADN puede establecerse en la célula huésped mediante un proceso de recombinación homóloga

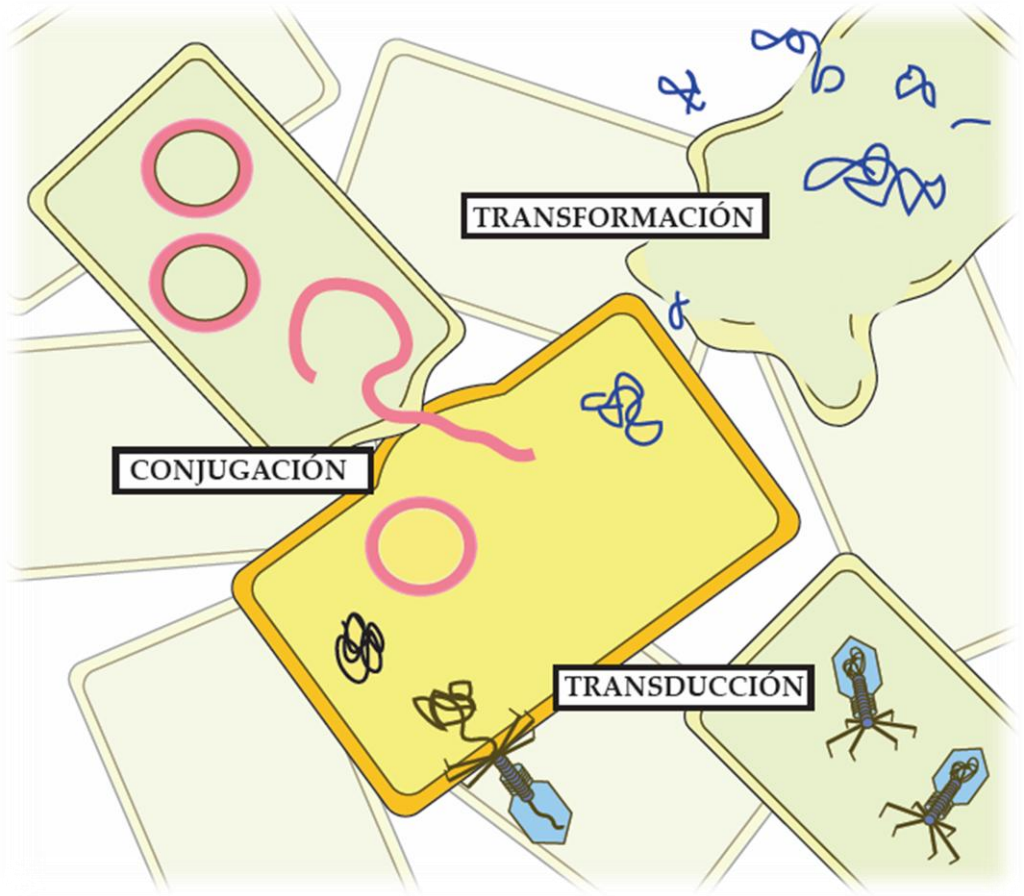


Figura I1. Imagen donde se muestran los tres mecanismos por los cuáles se lleva a cabo la transferencia genética horizontal: transformación, conjugación y transducción.

Esta transferencia de información se produce generalmente mediante los elementos genéticos móviles (EGMs), entre los que se distinguen distintos tipos de vectores:

I. Elementos transponibles. Engloban una variedad de segmentos de ADN que puede moverse dentro del genoma. Fueron nombrados como "genes saltarines" tras su descubrimiento en los años 40 por Barbara McClintock. Se incluyen en este grupo los transposones, los integrones y las secuencias de inserción. Contienen, en muchos casos, genes de resistencia a antibióticos, a metales pesados o genes codificantes para proteínas de degradación de compuestos xenobióticos.

II. Plásmidos. Un plásmido es una molécula de ADN separada del cromosoma y que puede replicarse independientemente del mismo. Aproximadamente la mitad de los plásmidos descritos hasta la fecha son capaces de transferirse de unas bacterias a otras, siendo la mitad de ellos conjugativos (son capaces de movilizarse por si mismos) y la otra mitad movilizables (necesitan de la maquinaria de los plásmidos conjugativos para transferirse)(Smillie *et al.*, 2010).

III. Fagos. Son los virus que infectan a las bacterias, y responsables, por lo tanto, de la transducción. Aún siendo ubicuos, representan los elementos que menos aportación hacen a los fenómenos de transferencia.

2. CONJUGACIÓN BACTERIANA

Es un mecanismo de transferencia genética horizontal en el cual se transfiere un EGM desde una célula donadora a una receptora mediante un contacto célula-célula, mediado por un complejo proteico conocido como aparato conjugativo. La promiscuidad de este proceso va más allá del mundo procariota, pudiendo ocurrir incluso entre especies muy alejadas filogenéticamente. En condiciones de laboratorio se ha observado transferencia de ADN conjugativo de bacterias a levaduras, a células de plantas e incluso a células animales (Heinemann and Sprague, 1989; Waters and Guiney, 1993). Un ejemplo de la versatilidad de este proceso en la naturaleza es el sistema de transferencia de ADN del género bacteriano *Agrobacterium* al núcleo de células vegetales, mediante una maquinaria de transferencia altamente similar al aparato conjugativo de las bacterias Gram-negativas (Lessl and Lanka, 1994).

La conjugación es el mecanismo de transferencia genética horizontal que contribuye en mayor medida al reservorio de genes móviles en el mundo procariota (Halary *et al.*, 2010). El ADN que se transfiere es generalmente autónomo y no requiere de la existencia de un fenómeno de recombinación posterior para que se estabilice en la

célula receptora, aumentando así la eficiencia de la conjugación con respecto a otros mecanismos de transferencia genética horizontal.

Las bases moleculares del proceso conjugativo han sido ampliamente estudiadas, ya que es un mecanismo que genera variabilidad genética al contribuir a la diseminación de factores de virulencia (Ziebuhr *et al.*, 1999), y es el principal mecanismo por el que se expanden los genes de resistencia a antibióticos (Mazel and Davies, 1999). Además, debido al peculiar procesamiento del ADN para este proceso, cabe la posibilidad de utilizar los elementos que participa en este proceso como una herramienta biotecnológica (Llosa and de la Cruz, 2005).

Lederberg y Tatum descubrieron la conjugación bacteriana en 1947, cuando al realizar estudios en *E.coli* con diversas auxotrofías observaron que los caracteres podían transmitirse de unas a otras (Lederberg and Tatum, 1946). El mecanismo subyacente no se descubrió hasta años después, cuando en 1953 se propuso que esta transmisión estaba asociada a un factor que se denominó F (de fertilidad) (Cavalli *et al.*, 1953; Hayes, 1953). Este fue el primer plásmido conjugativo que se descubrió, y se considera el prototipo de los estudios de conjugación. Pasaron muchos años hasta que se descubrió que la conjugación es un mecanismo extendido en la naturaleza, y que los plásmidos conjugativos están presentes en prácticamente todos los microorganismos conocidos. El número de plásmidos conjugativos o movilizables que se descubren sigue creciendo día a día.

Algunos de estos plásmidos denominados plásmidos de amplio rango de huésped están presentes en un amplio rango de bacterias. Otros plásmidos denominados plásmidos de estrecho rango de huéspedes son, en cambio, estables en un número limitado de bacterias. Otra diferencia que presentan los plásmidos es el tipo de *pilus* producido por la maquinaria de conjugación. Algunos plásmidos tienen un *pilus* flexible que les permite conjugar tanto en medio líquido como en sólido, y otros producen un *pilus* rígido que les permite conjugar únicamente en medio sólido (Bradley *et al.*, 1980).

A pesar de las diferencias que muestran los distintos tipos de plásmidos respecto a la conjugación, todos comparten los mismos requerimientos, como son la necesidad de un *pilus* u otro mecanismo para establecer contacto célula-célula, la presencia de

enzimas para el procesamiento del ADN, un mecanismo para asegurar el establecimiento del plásmido entrante en la célula receptora, un sistema regulador para el control y la inducción de la transferencia, y la presencia de una región específica en el ADN a transferir denominada origen de transferencia.

2.1. Modelo de la conjugación

Hoy en día, se contempla la conjugación como la asociación de dos mecanismos comunes en el entorno bacteriano, como son la replicación por círculo rodante y la presencia de un sistema de secreción de macromoléculas tipo IV (Llosa *et al.*, 2002). Este modelo conocido como "shoot and pump" se ha convertido en poco tiempo en el modelo aceptado para explicar este proceso biológico (figura I2).

El modelo propone que el ADN es transportado a la célula receptora en dos pasos diferenciados. Comienza con el reconocimiento de una secuencia de corte en una cadena del origen de transferencia por la proteína conocida como relaxasa. Esta proteína corta el ADN y se queda unida covalentemente a él, y es el sistema de secreción tipo IV del plásmido el encargado de reconocer y transportar la relaxasa, y pasivamente, el ADN que lleva unido. El segundo paso del mecanismo lo lleva a cabo la proteína acopladora, que actúa como una bomba de ADN empujándolo a la célula receptora a través del sistema de secreción.

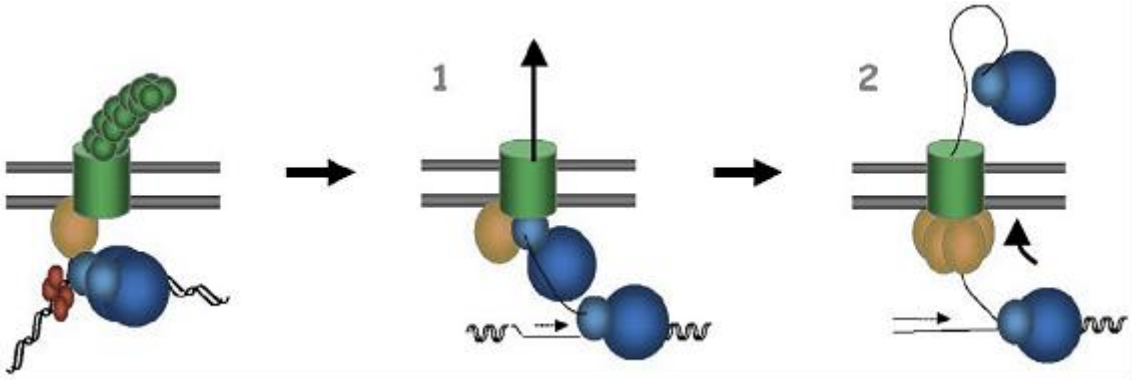


Figura I2. Esquema del modelo "shoot and pump" para la transferencia del ADN conjugativo en el plásmido R388. El complejo de procesamiento del ADN se compone de un dímero de la relaxasa TrwC (esferas azules), la proteína accesoria TrwA (esferas rojas) y el origen de transferencia representado como hebras negras. La proteína acopladora TrwB se muestra como esferas naranjas. El ADN que se sintetiza por RCR se muestra como una flecha discontinua, y la flecha curva indica la fuerza motora dependiente de ATP requerida para que se produzca el bombeo del ADN. El T4SS se representa como un cilindro verde. Paso 1: la relaxasa es secretada por el T4SS junto con la cola de ADN unida covalentemente. Paso 2: La proteína acopladora bombea el ADN de la célula donadora a la receptora, mientras el ADN se sigue replicando en la donadora. Tomado de (Llosa and de la Cruz, 2005).

Las relaxasas están relacionadas con proteínas de iniciación de la replicación por círculo rodante, y sus secuencias diana en el *oriT* están relacionadas estructuralmente con los orígenes de replicación por círculo rodante *oriV* (Waters and Guiney, 1993). Por otro lado, las proteínas conjugativas que se ensamblan para dar lugar al conducto transmembrana pertenecen a la familia de transportadores de secreción tipo IV (T4SS). Este tipo de transportador es el empleado por *A. tumefaciens* para transferir el T-ADN a células de plantas (Christie, 2001), y también es utilizado por patógenos animales para inyectar factores de virulencia en células diana (Ding *et al.*, 2003). Estos dos módulos (el encargado del procesamiento del ADN y el transportador), se conectan mediante una proteína acopladora presente en los sistemas conjugativos.

Un esquema general propuesto para la conjugación bacteriana sería el mostrado en la figura I3.

Figura I3. Esquema del proceso conjugativo tal y como se detalla en el texto.

(1, 2) Se forma el par conjugativo. La célula donadora y la célula receptora entran en contacto por medio del *pilus* conjugativo sintetizado por la primera. En este primer paso, intervienen las proteínas del sistema Mpf encargadas de la formación del *pilus* y de la estabilización del par conjugativo.

(3, 4) La relaxasa, o proteína piloto localiza su secuencia diana en el *oriT*, corta la cadena que se va a transferir y se queda covalentemente unida al extremo 5´. Se produce un desenrollamiento del ADN dando lugar a un ADN de cadena

sencilla que es transferido. En algunos plásmidos, esta actividad helicasa es llevada a cabo por la propia relaxasa, y en otros, por una helicasa codificada por el cromosoma de la bacteria. Se cree que la relaxasa unida al ADN es en primer lugar, transferida mediante el T4SS, y posteriormente, es la proteína acopladora la que bombearía el ADN para que el complejo relaxasa-ADN llegue a la célula receptora.

(5) El ADN transferido debe establecerse en la célula receptora. Para ello, la molécula de ADN circular de cadena sencilla se convierte en ADN de cadena doble mediante replicación y posteriormente, es superenrollada.

(6) Una vez acabado el proceso conjugativo, las células se desagregan y en la célula receptora se expresan los genes que dan lugar a la exclusión de superficie. De esta manera, se impide que entren más copias del mismo plásmido. La célula receptora, puede ahora actuar como donadora en un nuevo ciclo de transferencia.

3. EL PLÁSMIDO R388

El plásmido conjugativo R388 es un plásmido de amplio rango de huéspedes, que se aisló originalmente en *E. coli* (Datta and Hedges, 1972). Se ha comprobado la transferencia de R388 a diferentes especies de proteobacterias como *Salmonella typhimuriun*, *Shigella flexneri*, o *Pseudomonas aeruginosa* (Bradley and Cohen, 1976) y a *A. tumefaciens* (Loper and Kado, 1979). R388 codifica para las resistencias a sulfanilamida y trimetropim (Datta and Hedges, 1972) y produce unos *pili* finos y rígidos (Bradley, 1980) que hacen que se transfiera mejor en medio sólido que en líquido.

En el año 1988 los plásmidos fueron divididos en grupos de incompatibilidad en función de la especificidad de sus sistemas de replicación, ya que cuando dos plásmidos comparten el mismo sistema de replicación, son incompatibles en una misma línea

celular (Couturier *et al.*, 1988). Se formaron 19 grupos. R388 pertenece al grupo de incompatibilidad IncW, junto con otros plásmidos como pSA y R7K (Ward and Grinsted, 1982). Los plásmidos pertenecientes a este grupo son los plásmidos conjugativos más pequeños estudiados hasta el momento (30-40Kb). R388 tiene un tamaño de 33.923 pb y la obtención de la secuencia completa ha permitido el estudio detallado de su organización genética. Se ha asignado una función a la mayoría de los 43 marcos de lectura identificados (Fernandez-Lopez *et al.*, 2006).

Casi la mitad del plásmido está implicado en conjugación (región TRAw), abarca 14.9 Kb y se divide en dos regiones conocidas como *Dtr* (DNA transfer and replication) y *Mpf* (mating pair formation). La región *Dtr* incluye los ORFs que participan en el procesamiento conjugativo del ADN, mientras en la región *Mpf* se engloban aquellos responsables de la formación y el ensamblaje del T4SS y del *pilus* conjugativo (figura I4). Estas dos regiones pueden separarse y añadirse independientemente en *trans*, restituyendo la capacidad de conjugación (Bolland *et al.*, 1990). Ambas regiones se describen en detalle en las siguientes secciones.

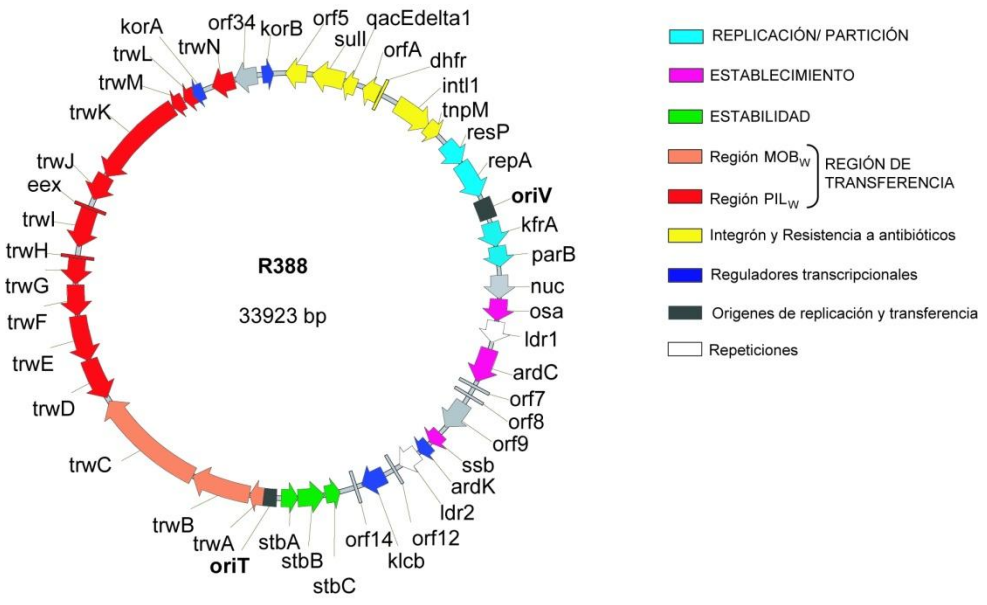


Figura I4. Mapa genético del plásmido R388**.** Se señala el proceso en el que están implicados cada uno de los genes mediante un código de colores. Tomada de (Fernandez-Lopez et al., 2006).

3.1. Región Mpf

Esta región está compuesta por los genes *trwD, E, F, G, H, I, J, K, M, L, N*, que están implicados en la formación del T4SS y el *pilus* conjugativo. Se ha visto que todos estos genes son esenciales para la conjugación excepto *trwN*. En esta región también se engloban los genes *korA* y *korB* que regulan la expresión de los genes *trw*, el gen *eex* que es responsable de la exclusión de superficie y el *orf34* que es un homólogo de *kikA* que produce una proteína letal en *Klebsiella*. El gen *trwH* es también un gen letal de un sistema kil-kor que no permite el mantenimiento del plásmido en la célula hospedadora a menos que se encuentren presentes los genes *kor* que regulan su expresión (Bolland, 1991).

Las proteínas Trw de esta región forman el sistema de secreción tipo IV (T4SS) de R388. Los T4SS son sistemas que translocan ADN y proteínas a través de envueltas celulares generalmente mediante un mecanismo que requiere contacto directo con una célula receptora. Se han descrito tres tipos de T4SS: a) los de sistemas conjugativos, en los que funcionan como maquinarias que translocan sustratos de ADN a células receptoras, b) los de sistemas de translocación de efectores que transportan proteínas u otros efectores moleculares a células diana eucariotas, y c) los de sistemas de expulsión o captación de ADN que translocan ADN desde o hacia el medio extracelular (Cascales and Christie, 2003).

Los T4SS conjugativos conforman la subfamilia más extensa y ampliamente distribuida de los T4SS, habiendo sido descritos estos T4SS conjugativos para la mayoría de especies bacterianas y para algunos miembros de las arqueas. El sistema más estudiado, en cambio, es el de las proteínas VirB del plásmido Ti de *A. tumefaciens*, por el cual *Agrobacterium* transfiere su T-ADN a células de plantas (Christie and Cascales, 2005). Las proteínas Trw de la región *Mpf* del plásmido R388 muestran una similitud de secuencia de entre un 42 y un 58% con las del T4SS de *A. tumefaciens*, y la organización genética es similar (Bolland, 1991). Por homología, el operón *trwHGFED* correspondería a VirB7, 8, 9, 10, 11 y el operón *trwLMKJI* a VirB2, 3, 4, 5, 6. La proteína VirD4 sería un homólogo de la proteína acopladora TrwB, que se tratará ampliamente en el apartado 3.2.3 de esta introducción.

Las proteínas VirB se han dividido en tres grupos funcionales. El primer grupo está formado por las proteínas que suministran la energía en forma de ATP (VirB4, VirB11 y VirD4). De sus homólogos en R388, TrwD, TrwB y TrwK respectivamente, se ha probado que los tres muestran actividad ATPasa (Arechaga *et al.*, 2008; Rivas *et al.*, 1997; Tato *et al.*, 2005). Un segundo grupo de proteínas forma el canal que atraviesa tanto la membrana interna como la externa (VirB6, VirB8, VirB10, VirB3, VirB7 y VirB9 en *A. tumefaciens* y TrwI, TrwG, TrwE, TrwM, TrwH y TrwF en R388). Finalmente, estarían los componentes del *pilus*, formado por la asociación de la pilina (VirB2 o TrwL) con VirB5, VirB7 y VirB1 (Cascales and Christie, 2003), figura I5.

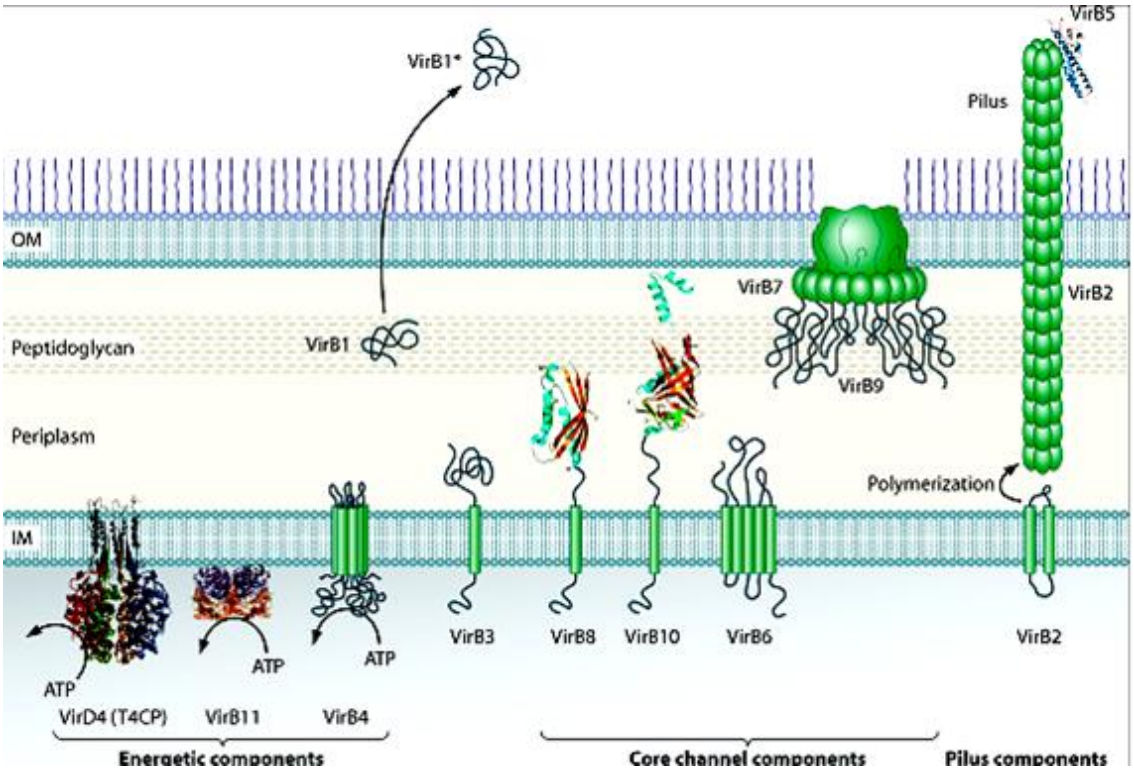


Figura I5. Localización de las subunidades VirB/VirD4 de *A. tumefaciens*. La proteína acopladora VirD4 y los componentes Mpf (VirB1 a VirB11) se representan según la supuesta función que desempeñan, pudiendo ser energética, de formación del canal o de componente del pilus. IM: membrana interna, OM: membrana externa. Se muestran las estructuras cristalinas de los homólogos de VirD4 (TrwB_R388), VirB5 (TraC_pKM101), VirB8 (dominio soluble de VirB8Bs) y VirB10 (dominio soluble de ComB10Hp).Tomado de (Alvarez-Martinez and Christie, 2009).

Se ha estudiado la posible complementación de T4SS de distintos plásmidos. El T4SS de R388 puede ser intercambiado por el del plásmido pCU1 (IncN) o R6K (IncI), pero no por el de RP4 (IncP) (Bolland *et al.*, 1990; Llosa *et al.*, 2003). Se demostró que la eficiencia de la movilización heteróloga está relacionada con la fuerza de la interacción que se da entre la proteína acopladora y el T4SS (Llosa et al., 2003). Se ha observado también, que el componente VirB10 del T4SS implicado en virulencia de *Bartonella tribocorum*, *A. tumefaciens* y *Brucella suis* puede interaccionar con la proteína acopladora TrwB que participa en la conjugación de R388. Más aun, la proteína TrwE (VirB10) de R388 puede ser parcialmente sustituida funcionalmente por la TrwE de *Bartonella tribocorum* (de Paz et al., 2005).

3.2. Región Dtr

La region Dtr es la encargada del procesamiento del ADN previo a su transporte. Es una región que ocupa 5,2Kb en R388 y codifica para los genes *trwA*, *trwB* y *trwC* que se encuentran agrupados en un mismo operón. Contiene también el origen de transferencia del plásmido. A continuación, se detalla cada componente de esta región.

El origen de transferencia, es la única secuencia de ADN requerida en *cis* para la conjugación, y su incorporación en un plásmido no transmisible lo convierte en movilizable (Lanka and Wilkins, 1995). Abarca regiones intergénicas de entre 38 y 500pb.

La secuencia y la longitud del *oriT* varía entre los distintos plásmidos, pero todos comparten algunas características comunes:

- Contienen el sitio *nic* en el que se realiza el corte específico por la relaxasa.
- Están ubicados en regiones intergénicas, no codificantes en posición asimétrica con respecto a los genes *tra*, de manera que estos genes son los últimos en entrar a la célula receptora.
- Presentan elementos de estructuras secundarias como repeticiones directas e invertidas, siendo muchos de ellos sitios de unión de proteínas.
- Son regiones ricas en A+T para facilitar la fusión de la doble cadena.

- Incluyen en su secuencia promotores para la expresión de los genes *tra*.

3.2.1 *oriT* de R388

El *oriT* de R388 fue inicialmente descrito como una secuencia de 402 pb (Llosa *et al.*, 1991), y más tarde se delimitó a 264 pb (60-330) (Fernandez R, resultados sin publicar). Su contenido en G+C es del 49%, frente al 57% del plásmido completo. Contiene en su secuencia una serie de estructuras secundarias que se detallan a continuación y se muestran en la figura I6:

- Hay cuatro repeticiones del hexanucleotido CACTAC, tres de ellas de manera directa y una de manera indirecta (IR4), que forman los sitios de unión de TrwA (Moncalian *et al.*, 1997).
- Se han descrito cuatro repeticiones invertidas más:
  - IR_1: Repetición invertida de 10 nucleótidos (CCCAATGCGC) separados por 3 nucleótidos. Se ha comprobado por *footprinting* con $KMnO_4$ que esta IR es la que mayor tendencia tiene a formar cruciformes (los oligonucleótidos centrales de las demás IRs también muestran cierta sensibilidad) (Moncalián, 2000).
  - IR_{11}: Repetición invertida de 7 nucleótidos (TGACCCT) separados por 4 nucleótidos.
  - IR_2: Repetición invertida de 6 nucleótidos separados por 4 nucleótidos y que constituye el sitio de unión de TrwC (CGCACC).
  - IR_3: Repetición invertida de 9 nucleótidos (AAGTCATTC) separados por 5 nucleótidos.
- En el *oriT* de R388 hay además dos sitios de unión de IHF de 13 nucleótidos, cuya presencia se comprobó por ensayos de footprinting y retardo (Moncalian *et al.*, 1999b). El sitio nombrado como *ihfA* está situado en el IR_3 entre los pb 203-215 (figura I6) y se ajusta a la secuencia consenso de unión de IHF 5' WATCAANNNNTTR 3' (Friedman, 1988), mientras que el sitio *ihfB* contiene dos desviaciones de la secuencia consenso y se encuentra entre los dos sitios de unión de TrwA (IR_4 y las últimas repeticiones directas, 278-290 pb). La afinidad que muestra IHF por la secuencia consenso en *ihfA* es mayor.

- En la región promotora del operón *stbABC* hay dos bloques de cinco repeticiones directas cada una. Estas repeticiones tienen 9 nucleótidos, están espaciadas por dos nucleótidos, y se ha demostrado que son sitios de unión para StbA (Guynet *et al.*, 2011).
- El sitio de corte *nic* se sitúa entre los nucleótidos 176 y 177 (figura I6). El corte es único y específico de cadena, y debido a que la cadena que se transfiere es la cortada y se transfiere primero el extremo 5´, los genes *trw* serán los últimos en entrar a la célula receptora (Llosa *et al.*, 1995).

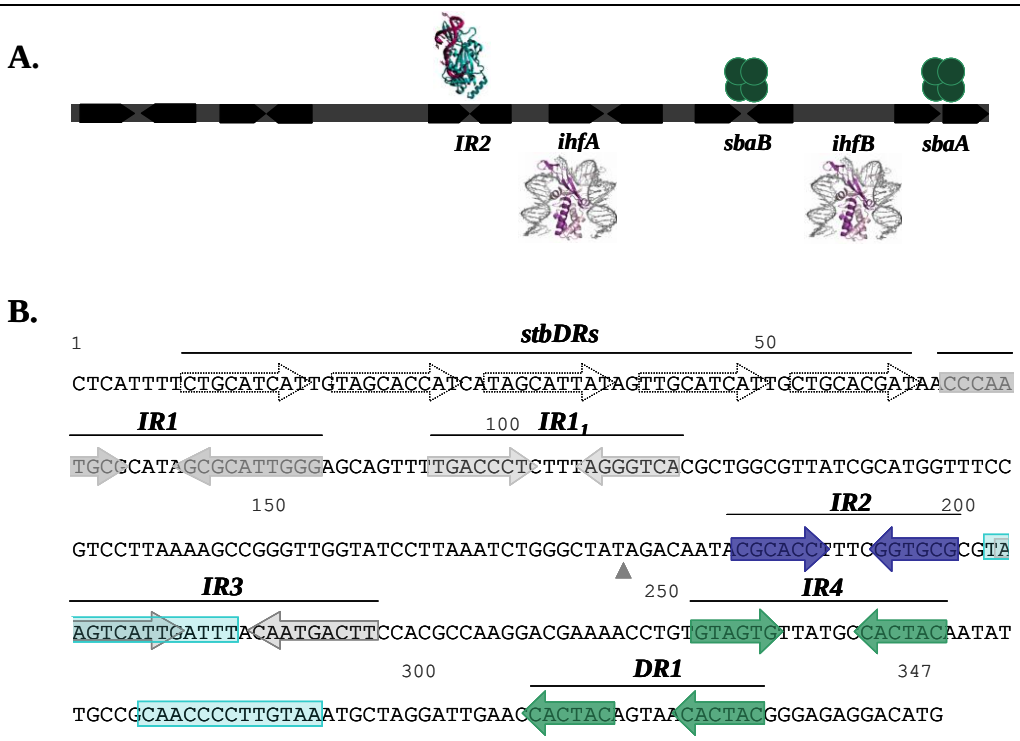


Figura I6. Origen de transferencia de R388. **A.** Representación esquemática de los sitios de unión de proteínas conocidos. **B.** Secuencia nucleotídica del *oriT* donde se destacan los elementos de estructura secundaria definidos hasta la fecha.

Las relaxasas de un mismo grupo de incompatibilidad guardan homología entre sí y lo mismo ocurre con sus *oriT*s asociados que muestran también cierta conservación de secuencia (Francia *et al.*, 2004; Parker *et al.*, 2005).

3.2.2. Orígenes de transferencia de otros sistemas

a) *oriT* MOB_P

Los análisis de las secuencias de los *oriT* de RP4 y del plásmido estrechamente emparentado R751 (Furste *et al.*, 1989; Pansegrau *et al.*, 1988) proveyeron las bases para el estudio descriptivo de todos los orígenes de transferencia en general. El sitio *nic* está bordeado por promotores divergentes que transcriben el operón de la relaxasa y los operones líder, siendo *traJ* y *traK* respectivamente los primeros genes flanqueantes del *oriT* en RP4 (figura I7). El operón líder se denomina así porque codifica los primeros genes que entrarán en la célula receptora. Este operón posee, en la mayoría de los plásmidos, tres genes conservados (*traK*, *traL* y *traM* en RP4, *stbA*, *stbB* y *stbC* en R388) que participan en los procesos de conjugación y partición (Guynet *et al.*, 2011). El *oriT* de RP4 abarca aproximadamente 350 pb y contiene repeticiones simétricas que son reconocidas por las proteínas TraJ y TraK, que definen la especificidad de transferencia del plásmido ya que no son intercambiables ni las variantes producidas por RP4 y R751 (Ziegelin et al., 1991).

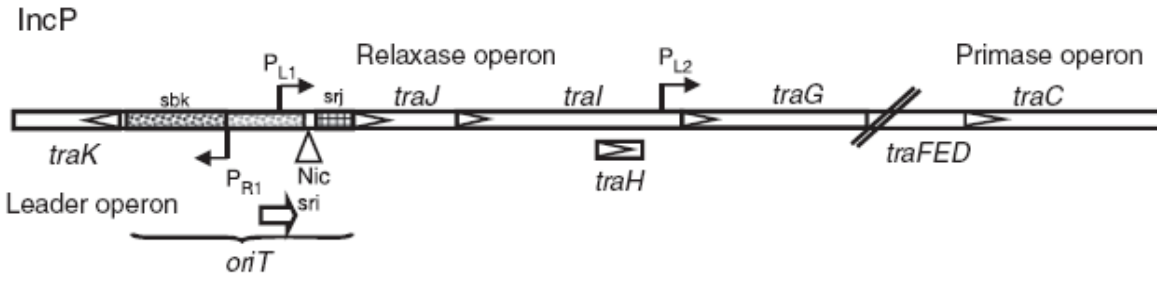


Figura I7. Esquema del *oriT* de RP4 (basado en (Pansegrau *et al.*, 1994a). Los tres promotores implicados en la expresión de componentes del relaxosoma se indican con flechas salientes. Los sitios de unión de TraK se representan como rectángulos moteados, siendo más oscuros cuanta mayor afinidad muestran. Imagen tomada de (de la Cruz *et al.*, 2010).

b) *oriT* MOB_Q

El *oriT* de R1162/RSF1010/R300B es un locus compacto compuesto por unas 38 pb (Brasch and Meyer, 1987). A diferencia de R388 y F (ver sección 3.2.1 y 3.2.2 c), no tiene sitios de unión para proteínas codificadas por el cromosoma, lo que le aporta independencia para ser funcional en diversos citoplasmas.

El *oriT* consiste en una repetición invertida con 10 pb por brazo, una región adyacente rica en AT donde comienza la separación de hebras en el relaxosoma, y una región rica en GC que incluye el sitio *nic* (figura I8). Esta organización del *oriT* se encuentra altamente conservada entre los *oriT*s asociados a la amplia familia de las relaxasas MOB_Q, siendo el tamaño y la secuencia de la repetición invertida lo que muestra más variación (Becker and Meyer, 2003). El brazo distal de la repetición respecto al *nic* no es necesario para iniciar la transferencia, aunque la relaxasa se une mejor a un *oriT* que contenga la repetición invertida intacta (Scherzinger *et al.*, 1993). En las 19 pb del *oriT* que son necesarias para el inicio de la transferencia, se toleran cambios de bases en varias posiciones. Como consecuencia de esta especificidad laxa, la presencia de secuencias consenso en el cromosoma puede servir como origen de transferencia cuando el plásmido R1162 está presente en la célula (Meyer, 2009).

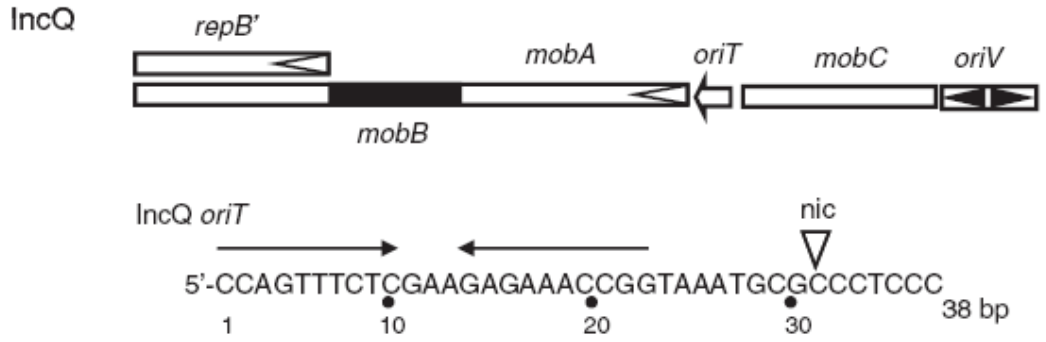


Figura I8. Representación esquemática de la organización de las regiones de movilización y replicación de R1162 (Scholz *et al.*, 1989). En la imagen inferior se muestra la secuencia de la cadena que sufre el corte del *oriT*. La repetición invertida indica el sitio de unión de la relaxasa MobA. Imagen tomada de (de la Cruz *et al.*, 2010).

c) *oriT* MOB_{F12}

La región del *oriT* de los plásmidos tipo F ocupa unas 350 pb y está bordeada por los promotores del gen *ygfA* en F, que codifica para un homólogo de una transglicosilasa lítica, y de TraM, una proteína del relaxosoma (Frost *et al.*, 1994; Koraimann, 2003)(figura I9). El *oriT* de F contiene los elementos comunes nombrados anteriormente como son una serie de repeticiones que definen los sitios de unión de la relaxasa y de las proteínas accesorias TraM y TraY, así como una región rica en AT relevante para la apertura de la doble cadena. El *oriT* de F contiene también curvaturas intrínsecas y sitios de unión de IHF (Tsai *et al.*, 1990), y se han descrito tres sitios de unión para H-NS. En dirección 5´ al sitio de unión de IHF hay un sitio de unión de TraY en F y R100 (Inamoto and Ohtsubo, 1990; Nelson *et al.*, 1993). El sitio *nic* se localiza entre los nucleótidos 140 y 141 en la hebra T. Inmediatamente después del *nic* en dirección 5´ aparece una secuencia de 9pb que determina la especificidad de corte de la relaxasa.

Otros orígenes de transferencia pertenecientes a plásmidos tipo F también han sido caracterizados en detalle, entre ellos el *oriT* de R100 (Inamoto *et al.*, 1991), R1 (Schwab *et al.*, 1991), y pED208 (Di Laurenzio *et al.*, 1991). Aunque el sitio *nic* se encuentra conservado, hay pequeños cambios en la secuencia de las 9pb adyacentes y en la repetición invertida que es reconocida por TraI (Harley and Schildbach, 2003). La región rica en AT, y los sitios de unión de TraM, TraY e IHF están presentes en todos los *oriT*s de la familia del plásmido F, aunque el orden y el número de estos sitios de unión son variables. Como ocurre en F, estos plásmidos contienen curvaturas intrínsecas adyacentes al sitio *nic*, seguidas de un sitio de unión de IHF que ha sido comprobado en R100 (Dempsey and Fee, 1990) y peD208 (Di Laurenzio *et al.*, 1995).

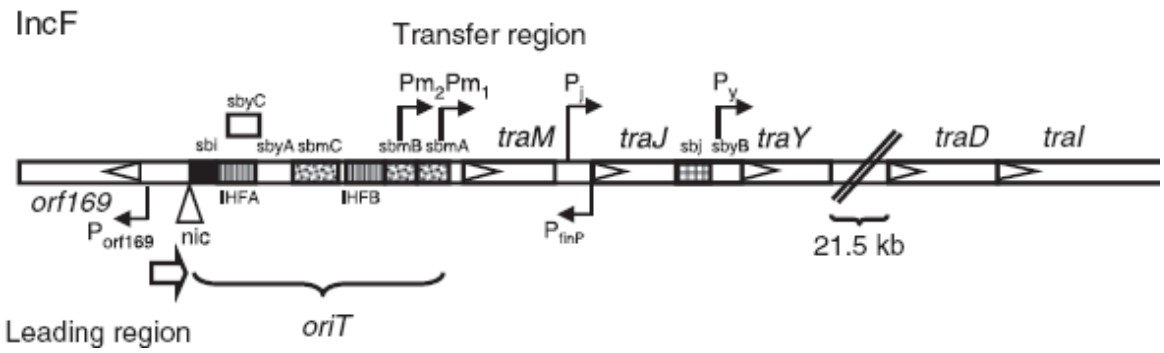


Figura I9. Representación abreviada del *oriT* de F basado en (Frost et al., 1994). Los promotores principales se indican como flechas salientes. Los sitios de unión de TraI (*sbi*), TraY (*sbyA, -B, -C*), TraM (*sbmA, -B, -C*) y TraJ (*sbj*) se muestran como rectángulos negros, blancos, moteados y a cuadros respectivamente. La doble barra diagonal indica un corte en la secuencia de 21,5Kb. Imagen tomada de (de la Cruz *et al.*, 2010)

3.2.3. Proteína acopladora

La proteína acopladora funciona como nexo de unión entre el relaxosoma y el T4SS estando presente en todos los sistemas conjugativos estudiados, así como en muchos T4SS implicados en virulencia bacteriana (Garcillan-Barcia *et al.*, 2009). En el plásmido R388 se ha podido demostrar la interacción de su proteína acopladora TrwB con proteínas pertenecientes tanto al relaxosoma como al T4SS. TrwB interacciona con las proteínas TrwA y TrwC del relaxosoma, y con el componente del T4SS TrwE (Llosa et al., 2003). La interacción con las proteínas del relaxosoma es muy específica, mientras que la interacción con el T4SS no lo es tanto, indicando que una proteína acopladora puede llegar a interaccionar funcionalmente con varios T4SS conjugativos.

TrwB es una proteína de 507 aa y 56 KDa que se encuentra anclada en la cara interna de la membrana citoplasmática mediante dos segmentos transmembrana en el N-terminal. Para facilitar su estudio *in vitro*, se purificó el dominio citoplasmático, TrwBΔN70, monómero en solución, el cuál ha sido ampliamente caracterizado bioquímicamente.

Esta proteína TrwBΔN70 une inespecíficamente ADN de cadena sencilla o doble, aunque recientemente se ha demostrado que muestra más afinidad por ADN que contiene estructuras G-cuádruplex en su secuencia (Matilla *et al.*, 2010).

En la parte citosólica de TrwB, se localizan un motivo Walker A y un motivo Walker B de unión a nucleótidos (Llosa *et al.*, 1994a). Este dominio tiene capacidad de unión de NTPs, y la mutación puntual en el sitio WalkerA inhibe la conjugación (Moncalian *et al.*, 1999a). Como puede indicar la presencia de este dominio, se ha demostrado que TrwB muestra actividad ATPasa *in vitro* dependiente de ADN. Esta actividad ATPasa se encuentra modulada por TrwA, con la que interacciona por su dominio C-terminal (Tato *et al.*, 2007). El dominio citosólico de TrwB muestra una similitud estructural con la F_1 ATPasa y otras proteínas que bombean o se desplazan sobre el ADN (SpoIIIE, FtsK o la helicasa Gen4 del fago T7), sugiriendo que TrwB podría mostrar un mecanismo parecido, como se propone en el modelo "shoot and pump" (Llosa et al., 2002).

La estructura cristalina de TrwBΔN70 ha sido resuelta por difracción de rayos X (Gomis-Ruth *et al.*, 2001). Se observa cómo una asociación de seis monómeros equivalentes se disponen de manera que dejan un canal central con un diámetro de aproximadamente 20Å (figura I10). Cada monómero presenta dos dominios diferenciados, uno formado por hélices α y orientado al citosol, y el otro, el dominio de unión a nucleótidos formado por estructuras secundarias α/β y próximo a la membrana.

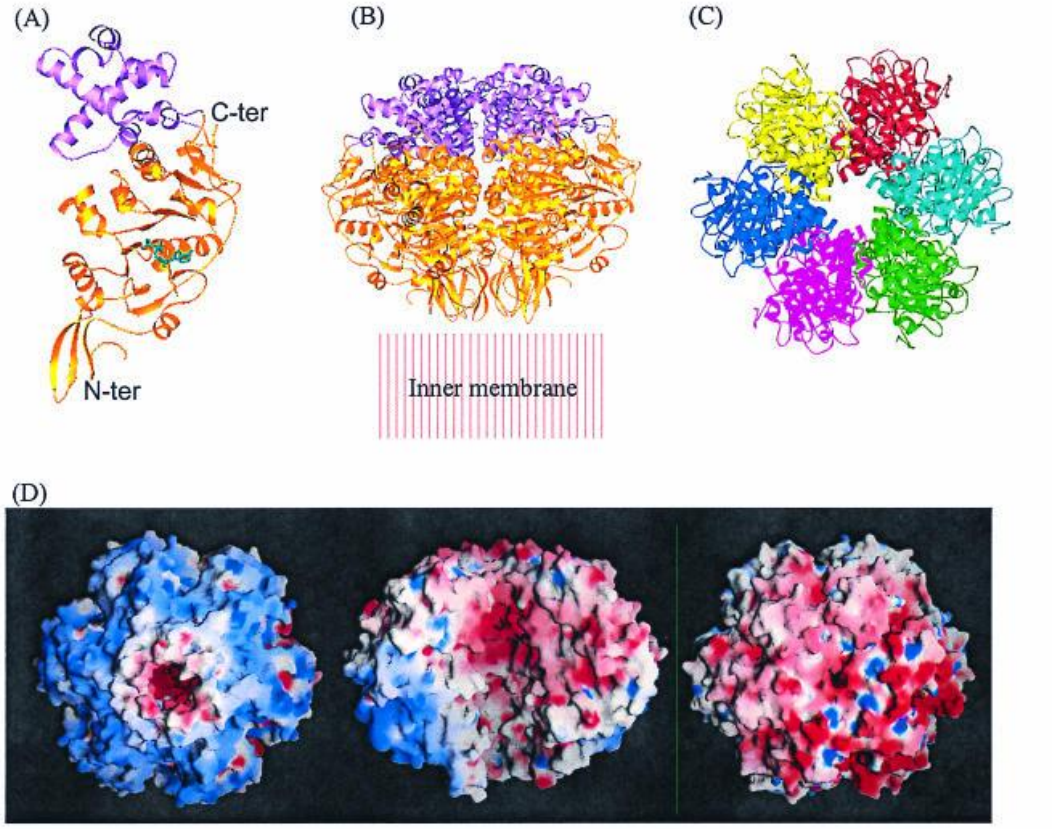


Figura I10. Características estructurales de TrwB. A. Monómero de TrwBΔN70 unida a ADP. Se observan el dominio de hélices α en morado y el dominio de unión a nucleótidos en dorado, dentro del cual se aprecia la molécula de ADP en azul. B. Hexámero de TrwBΔN70 visto de lado. C. Hexámero visto desde el canal en la base del hexámero formado por el dominio de hélices α, cada monómero muestra un color. D. Superficie electrostática del hexámero de TrwBΔN70, coloreado según la carga (en azul la positiva y en rojo la negativa). En la imagen de la izquierda se observa el hexámero desde la cara de la membrana. En el centro, se observa el corte del hexámero en el que se aprecia el canal central. En la imagen de la derecha, se muestra una imagen del hexámero visto desde el lado citoplasmático. Imagen tomada de (Yeo and Waksman, 2004).

En un estudio reciente, se ha realizado una disección funcional de TrwBΔN70. Mediante el análisis de una colección de mutantes puntuales, se han delimitado los dominios implicados en la interacción tanto con el T4SS como con el relaxosoma, en la unión a ADN y en la actividad ATPasa (de Paz *et al.*, 2010).

3.2.4. La relaxasa

La relaxasa es el único elemento indispensable de la maquinaria MOB, siendo la proteína responsable del inicio y de la terminación del procesamiento conjugativo del ADN.

Las relaxasas se denominan así porque se observó que al incubar estas proteínas con ADN superenrollado se formaba un complejo nucleoprotéico que relajaba el ADN al tratarlo con proteinasa K y SDS (Clewell and Helinski, 1969).

Son endonucleasas específicas de sitio y hebra que reconocen y cortan una hebra concreta del *oriT* en un sitio concreto denominado *nic*, mediante una reacción de transesterificación (Byrd and Matson, 1997; Lanka and Wilkins, 1995). Las relaxasas comparten tres motivos comunes en su secuencia (Francia *et al.*, 2004; Pansegrau *et al.*, 1994b)(figura I11):

a) Motivo I. Contiene la tirosina o tirosinas catalíticas, que están implicada/s en la actividad de corte y ligación de ADN.

b) Motivo II. Está implicado en interacciones de la proteína con el ADN, y contiene una serina conservada.

c) Motivo III. Posee una tríada de histidinas conservadas que coordinan el catión divalente requerido para el ataque nucleofílico por la tirosina catalítica.

Estos motivos se encuentran conservados sobre todo en relaxasas de bacterias gram negativas, siendo la conservación menor en bacterias gram positivas (Guzman and Espinosa, 1997). Los motivos I y III se encuentran también conservados en las proteínas Rep implicadas en la iniciación de la RCR, aunque en estas los motivos no se encuentran en el mismo orden que en la estructura primaria de las relaxasas (Guasch *et al.*, 2003; Ilyina and Koonin, 1992).

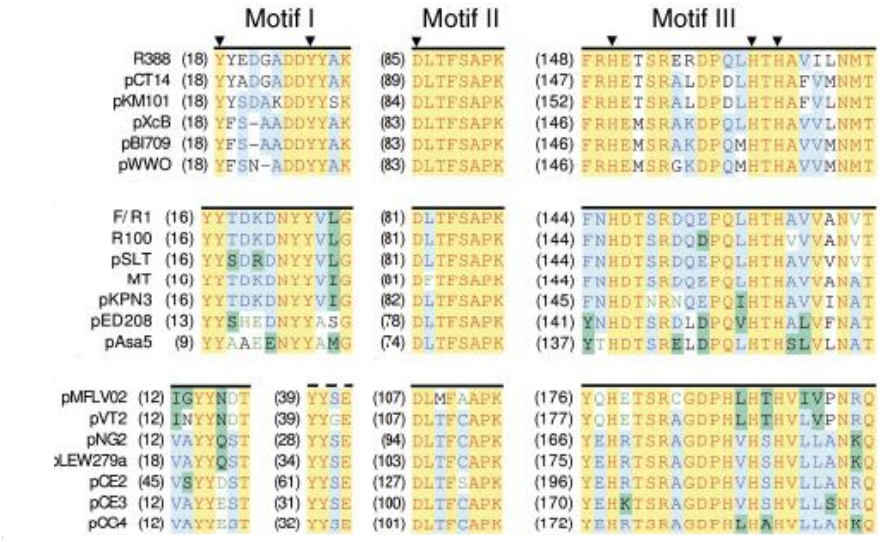


Figura I11. Motivos de secuencia conservados en las relaxasas representativas de la familia MOB_F. En el alineamiento se tomaron en cuenta los primeros 300 residuos N-terminales que contienen el dominio relaxasa. Las puntas de flecha verticales apuntan los residuos importantes para su función en $TrwC_{R388}$ (Guasch et al., 2003). En sombreado amarillo se muestran los aminoácidos idénticos que se mantienen, en azul los altamente conservados, y en verde los que muestran similitud. Sin sombrear, los residuos en verde indican poca similitud, y los residuos en negro son los no conservados. Tomado de (Garcillan-Barcia et al., 2009).

Las relaxasas se han clasificado en cuatro grupos según la conservación de secuencia de la proteína y del sitio *nic* que reconocen (Lanka and Wilkins, 1995). La familia MOB_F abarca las relaxasas de los plásmidos IncF, IncN e IncW, en MOB_P se engloban las relaxasas de IncP, IncI, IncX y VirD2 de pTi de *Agrobacterium*, en MOB_Q se encuentran las relaxasa pertenecientes a IncQ y TraA_pTI de *Agrobacterium*, y en MOB_V la relaxasa de pMV158 y otros homólogos. Una clasificación reciente de plásmidos movilizables ha propuesto la división de relaxasas en otras dos familias. Estas familias son MOB_C en la que se encuentra la relaxasa de CloDF13 (Francia et al., 2004), y MOB_{HEN}, a la que pertenece la relaxasa de ColE1. La relaxasa de CloDF13 no contiene motivo 3H, ni ha sido identificada la presencia de tirosinas catalíticas (Nunez and De La Cruz, 2001).

Recientemente, se ha publicado un trabajo en el que se utiliza la secuencia de la relaxasa como método para la clasificación de plásmidos. La mayor diferencia respecto a la revisión de Francia et al 2004 es la adición de dos familias nuevas de relaxasas que

están presentes predominantemente en plásmidos conjugativos de gran tamaño pertenecientes a MOB_F y MOB_H. Además, se han añadido nuevas relaxasas a la familia MOB_C en al que antes se incluían solamente ClodF13 y tres relaxasas más. Así, se ha propuesto la clasificación de los sistemas conjugativos en 6 familias que han adoptado la nomenclatura MOB de las relaxasas (MOB_F, MOB_H, MOB_C, MOB_Q, MOB_P, MOB_V) (Garcillan-Barcia et al., 2009).

La reacción que lleva a cabo la relaxasa conlleva la formación de un intermediario covalente en todos los casos estudiados hasta la fecha, excepto en el plásmido CloDF13 (Nunez and De La Cruz, 2001). La primera etapa consiste en la reacción de corte de un enlace fosfodiéster concreto en el sitio *nic* mediante una reacción de transesterificación tras al cuál la relaxasa se queda unida covalentemente al extremo 5´del ADN por medio de un enlace fosfotirosina. Esta reacción es reversible, ya que el extremo 3´-OH libre del ADN tiene la capacidad de actuar como grupo nucleofílico, pudiendo lanzar un ataque al extremo 5´unido covalentemente, religando así el ADN cortado y restableciendo el original. Para que se produzca el corte es necesario un desapareamiento previo de las bases que rodean el sitio de corte, pudiendo producirse esta fusión parcial del ADN por la unión de la relaxasa o de proteínas accesorias.

En el caso concreto de TrwC, esta es una proteína perteneciente a la familia de relaxasas MOB_F. Tiene 966 aminoácidos, con un peso molecular teórico de 180KDa, y forma dímeros en solución (Grandoso *et al.*, 1994).

Las relaxasas son proteínas grandes formadas por dos o más dominios. La actividad relaxasa se localiza en el dominio N-terminal y en el C-terminal, se encuentran dominios helicasa, primasa u otros de función desconocida. En el caso de TrwC, está formada por dos dominios, un dominio relaxasa N-terminal (1-300), y un dominio helicasa C-terminal (332-966, homólogo a helicasas 5'-3' tipo RecD). La región C-terminal contiene dos actividades relacionadas, una actividad ATPasa y una actividad helicasa. La actividad ATPasa es dependiente de ssADN e iones Mg^{2+}, y la actividad helicasa depende de ATP e iones Mg^{2+} (Grandoso *et al.*, 1994).

Además de estas actividades, TrwC es capaz de producir recombinación sitio-específica entre secuencias de *oriT in vivo* (Llosa et al., 1994b), y se ha descrito que es capaz de integrar el ADN transferido de manera sitio-específica en un *oriT* localizado en la célula receptora (Draper *et al.*, 2005).

Dominio relaxasa:

En TrwC el dominio relaxasa propiamente dicho contiene los tres motivos característicos de las relaxasas. Tiene capacidad para cortar por si solo tanto ADN superenrollado como ADN de cadena sencilla aunque, sin embargo, no es capaz de cortar ADN lineal de doble cadena. El fragmento mínimo que realiza el corte y transferencia de cadenas en TrwC es $TrwC_{N293}$ (1-293).

La estructura del dominio relaxasa $TrwC_{N293}$ fue la primera estructura de una relaxasa en ser resuelta (Guasch et al., 2003). La estructura de $TrwC_{N293}$ fue resuelta junto con un oligonucleótido de 25mer que contiene la secuencia reconocida por TrwC (Figura I12). Posteriormente, se han obtenido las estructuras de TraI de F y MobA de R1162, y se ha observado que tienen plegamientos similares a las ADN polimerasas y a las proteínas de iniciación de RCR.

Se dispone de la estructura en complejo con ADN y con varios metales, pero los intentos para obtener la estructura de la proteína completa que incluiría el dominio helicasa han resultado infructuosos hasta la fecha.

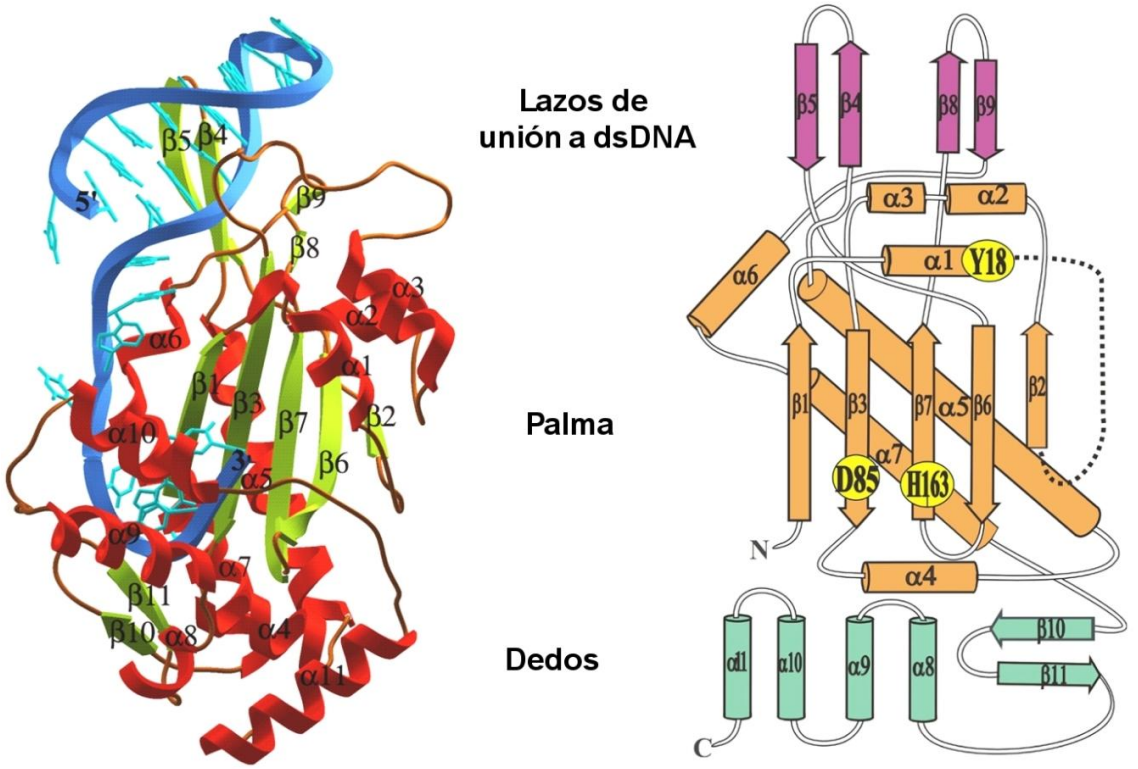


Figura I12. Estructura del complejo TrwCr:R(25+0). En la imagen de la izquierda se representa la estructura tridimensional del complejo: en verde las hojas beta, en rojo las hélices alfa y en azul el ADN. En la imagen de la derecha se muestra un esquema de la topología de la proteína, donde se resaltan los residuos clave en el mecanismo catalítico de la proteína en amarillo (Guasch et al., 2003)..

La estructura que adopta TrwCr se compara con una mano izquierda por analogía con las polimerasas de ADN (Hubscher *et al.*, 2002). La parte central de la proteína en la que se apoya el ADN sería la palma, mientras que los dedos estarían formados por las hélices 8-11 y las hebras 10-11 que se pliegan sobre el centro activo.

La estructura general de las tres relaxasas cuya estructura ha sido resuelta es similar. TrwCr y TraI36 (incluye los primeros 330 aminoácidos de TraI) (Datta *et al.*, 2003) muestran una estructura general muy parecida. El fragmento relaxasa de MobA de R1162 (minMobA) también es similar, aunque carece del dominio "dedos" y solo tiene una tirosina catalítica (Monzingo *et al.*, 2007) (Figura I13).

TrwC y TraI pertenecen a la misma familia de relaxasas y su similitud de secuencia es aproximadamente de un 42%, en cambio esta similitud es de solo un 10 y

un 12% respectivamente cuando se comparan con minMobA. Al comparar las estructuras, sin embargo, esta similitud es mucho mayor. El parecido en cuanto al plegamiento y a la función hace pensar que estas proteínas descienden de un mismo ancestro, aunque minMobA quede filogenéticamente alejada de las otras.

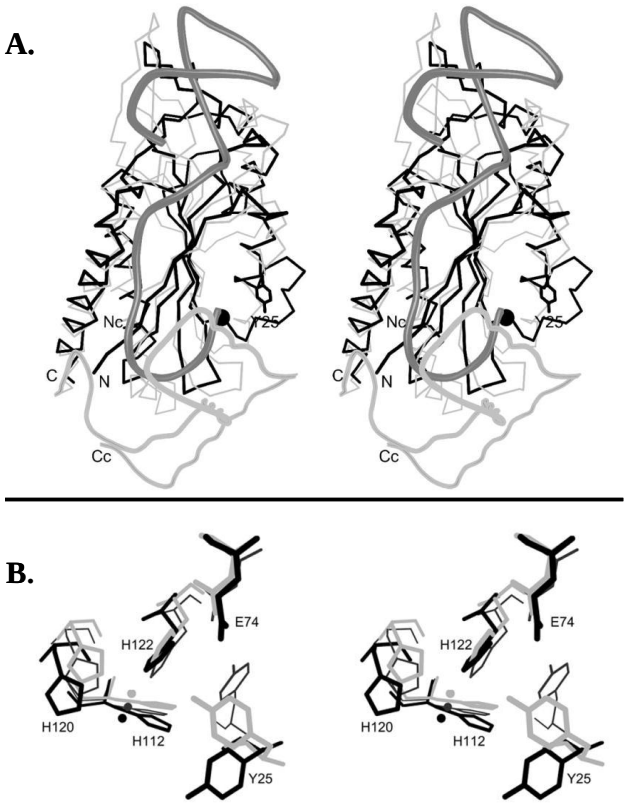


Figura I13. **A.** Superposición de minMobA y TrwC con ADN. El Cα de minMobA se muestra en trazos negros, el N y el C-terminal aparecen señalados, asi como el átomo de metal del sitio activo y la cadena lateral de la Tyr25. La región N-terminal de TrwC que tiene un plegamiento similar a minMobA se muestra con trazos grises finos (Nc). La parte C-terminal de TrwC que no tiene homólogo en minMobA se ilustra con trazos grises más anchos. **B.** Superposición de los residuos activos y el átomo de metal de minMobA con sus equivalentes en TrwC y TraI. Los residuos de minMobA y TraI se muestran en negro, siendo el trazo más fino en el último, y los residuos de TrwC, en gris. Tomado de (Monzingo et al., 2007).

3.2.5. El relaxosoma

Se denomina relaxosoma al complejo nucleoproteico responsable de la iniciación de la conjugación bacteriana. Está constituido por el origen de transferencia, la relaxasa, y proteínas accesorias que facilitan la acción de la relaxasa.

En R388, el relaxosoma está formado por el *oriT*, la relaxasa TrwC y las proteínas auxiliares TrwA e IHF que se detallarán a continuación.

4. PROTEINAS AUXILIARES

Las proteínas auxiliares que participan en la formación del relaxosoma, pueden estar codificadas por el plásmido y también por el cromosoma de la bacteria huésped. En general son proteínas que, por su unión al *oriT*, facilitan la fusión parcial del sitio nic que permite a la relaxasa acceder a la cadena sencilla sobre la que actúa.

4.1. Proteínas auxiliares de la conjugación codificadas por el cromosoma huésped

Aunque los elementos transmisibles como los plásmidos codifican muchos de los factores necesarios para su propio mantenimiento y diseminación, también dependen de proteínas codificadas por el huésped para la regulación de algunas de sus funciones clave. Algunas de estas proteínas se denominan proteínas “nucleoides”, grupo al que pertenecen un abundante número de proteínas que modulan la estructura del cromosoma. Estas proteínas también funcionan como reguladores tanto positivos como negativos de la transcripción, replicación y recombinación.

IHF es una proteína codificada por el cromosoma de *E. coli* y bacterias similares. Es una proteína pequeña, que muestra especificidad de secuencia por el ADN. Fue identificada en un principio como una proteína del huésped que participaba en la integración del bacteriófago λ y de ahí su nombre (de sus siglas en inglés Integration

Host Factor), aunque después se haya demostrado su participación en muchos procesos que implican complejos proteína-ADN de mayor orden de como pueden ser la replicación (se une al *oriC*), en la regulación transcripcional (reconoce secuencias que se encuentran en dirección 5´de promotores específicos de la subunidad σ54 de la polimerasa), y en una variedad de sistemas de recombinación sitio-específica (Goosen and van de Putte, 1995; Nash, 1996). La función primaria de IHF parece ser estructural, ya que introduce una curvatura en el ADN que facilita la interacción de otros componentes en un complejo nucleoprotéico.

IHF es un heterodímero codificado por los genes *himA* y *hip*. Sus dos subunidades muestran un peso molecular de 10KDa y son 30% idénticas en secuencia.
La proteína se une a una secuencia de ADN de 30 a 35 pb que contiene una secuencia consenso de 13 pb, 5´ WATCAANNNNTTR 3´, donde W puede ser A/T y R puede ser A/G (Friedman, 1988). Esta secuencia de reconocimiento es bastante variable, con únicamente dos de las pb conservadas en todos los sitios de unión de IHF. Así y todo la especificidad de secuencia de la proteína es considerable, ya que reconoce estas secuencias sobre otras al azar con una afinidad 10^3 o 10^4 veces mayor (Mengeritsky *et al.*, 1993; Wang *et al.*, 1995). Se ha propuesto, que la especificidad de secuencia de IHF viene dada tanto por el reconocimiento específico de nucleótidos, como por el reconocimiento "indirecto" por una conformación concreta que adopta el ADN. Según se observa en la estructura tridimensional de la proteína, IHF interacciona con el surco menor del ADN produciendo una curvatura de aproximadamente 160º (Rice *et al.*, 1996), figura I14.

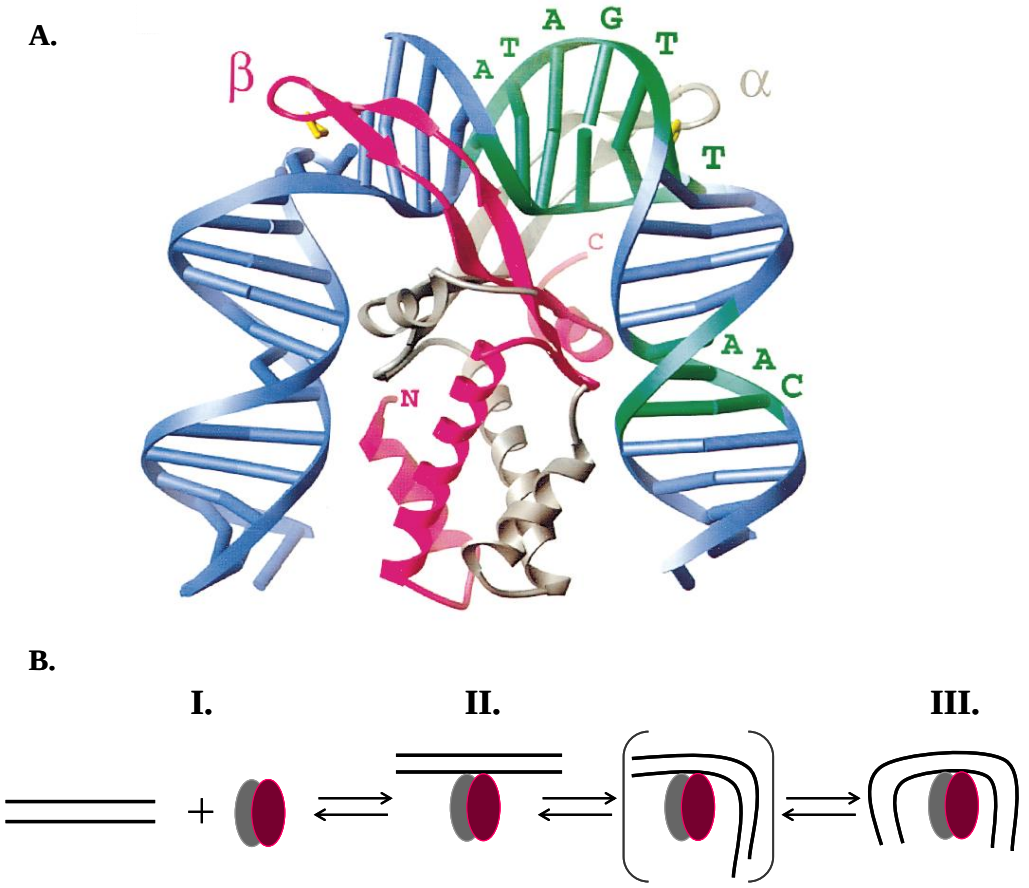


Figura I14. A. Estructura tridimensional de IHF unida a ADN. La subunidad α está coloreada en gris, y la β, en rosa. La secuencia consenso se muestra en verde. Tomado de (Rice et al., 1996). B. Mecanismo propuesto para la formación del complejo IHF-ADN. IHF se representa con óvalos fucsias y grises, y el dúplex de ADN con líneas paralelas. En un comienzo, IHF une ADN lineal (I.) y forma un intermediario de ADN lineal-proteína (II.), finalmente curva el ADN formando el complejo final (III.). Entre los pasos II. y III. puede existir un complejo en el que solo un lado del ADN esté curvado (mostrado en paréntesis). Modificado de (Sugimura and Crothers, 2006).

4.1.1. IHF en la conjugación

IHF es un componente de los relaxosomas de MOB_F. Se han descrito sitios de unión de IHF en los orígenes de transferencia de los plásmidos R388, F, R1, R100 y pED208, entre otros, aunque pueden diferir en número y en localización. Plásmidos pertenecientes a otras familias como MOB_P o MOB_Q no tienen sitios de unión descritos para IHF u otras proteínas codificadas por el cromosoma huésped.

La función que se le atribuye es modificar la estructura del ADN facilitando así el acceso a otras proteínas que requieran esa conformación para realizar su función.

Estudios realizados en F sugieren que TraI es reclutada en su sitio de unión por la acción de las proteínas accesorias TraY e IHF. Ambas proteínas se unen al *oriT* y curvan el ADN, provocando cambios conformacionales que pueden propagarse hasta el sitio *nic*, generando una región de cadena sencilla a la que puede unirse TraI (Nelson et al 95). Se ha observado que aunque la unión de estas proteínas no es cooperativa, sus sitios de unión deben estar correctamente espaciados y orientados el uno respecto al otro para que se de una movilización eficiente (Williams and Schildbach, 2007). IHF junto a las otras dos proteínas accesorias de F, TraM y TraY, se unen a secuencias cercanas al *nic* y estimulan la actividad de corte de TraI en el *oriT* superenrollado o lineal tanto *in vivo* como *in vitro* (Kupelwieser *et al.*, 1998).

De los dos sitios de unión que tiene IHF en el *oriT* de F, solo *ihfA* parece ser necesario para la reacción de corte (Fu *et al.*, 1991; Howard *et al.*, 1995), mientras que la deleción de *ihfB* no parece tener efecto en el corte *in vitro* o *in vivo*. Este dato contrasta con los resultados obtenidos en R100, donde se demostró que la unión de IHF a su sitio de unión *ihfB* es importante para la transferencia *in vivo* (Abo and Ohtsubo, 1995) y se ha sugerido que podría formar un complejo con la proteína TraM que mejoraría la transferencia del ADN (Abo and Ohtsubo, 1993).

Por otro lado se ha estudiado el papel de IHF en la transferencia de ICEs (de sus siglas en inglés *Integrative and Conjugative Element*). Los ICE son un grupo diverso de elementos genéticos móviles procarióticos que se integran en el cromosoma del huésped y son transmisibles por conjugación. Hasta la fecha no se ha estudiado en detalle el papel que pueden tener los factores codificados por el huésped en la transmisión de estos elementos. SXT es un ICE que fue identificado por primera vez en un aislado clínico de *Vibrio cholerae*. Su genoma de 100 kb contiene secuencias que codifican elementos para la transferencia por conjugación y maquinarias de integración en el cromosoma, así como genes que confieren resistencia a múltiples antibióticos (Beaber *et al.*, 2002). Se estudió el efecto que podía tener IHF sobre este elemento, y se demostró que este factor es necesario para que *V. cholerae* actúe como donador de SXT. La integración y la escisión de SXT, en cambio, no se veían influenciadas por la ausencia

de IHF. Este efecto parece ser específico del huésped, ya que mutantes de *E. coli* que carecen de IHF eran capaces de llevar a cabo una transferencia eficiente de SXT. El mutante IHF$^-$ en *V. cholerae* se puede considerar defectuoso para la transferencia conjugativa, ya que la transferencia del plásmido RP4 también se ve reducida al utilizar esta cepa (McLeod *et al.*, 2006).

4.1.2. IHF en R388

En el caso de R388 no se han observado diferencias en la conjugación al utilizarse donadoras o receptoras deficientes en IHF.

Sí se demostró un efecto sobre el corte realizado en el *nic* por la relaxasa, ya que en presencia de IHF, este corte se ve inhibido (Moncalian *et al.*, 1999b) aunque no produce ningún efecto en la unión de la relaxasa al ADN superenrollado. Al examinar el ADN extraído de cepas IHF$^-$, el ADN que se obtuvo se encontraba principalmente en una forma relajada, mientras que el extraído de cepas IHF$^+$ aparecía de manera superenrollada en su mayoría. Se concluye de este dato que IHF podría tener un rol inhibitorio de la reacción de corte también *in vivo*. Este efecto es contrario al que ocurre en F, donde IHF estimula el corte realizado por la relaxasa del sistema TraI (Nelson *et al.*, 1995). Un análisis detallado de los sitios de unión en los dos plásmidos señaló dos diferencias que podrían ser cruciales para explicar este efecto antagonista. En primer lugar, el sitio *ihfA* de R388 se encuentra 10 pb más cerca del sitio *nic* que en F, de tal manera que quedaría pegado al sitio IR2 de unión a TrwC, mientras que en F está situado una vuelta de hélice más alejado. La otra diferencia es que el sitio *ihfA* se ubica en distintas orientaciones relativas respecto al *nic* en los dos plásmidos (figura I15). Además de esto, en presencia de IHF se observa un sitio sensible a DNasa I en IR_2 del $oriT_{R388}$, lo que sugiere que IHF estaría induciendo un cambio conformacional en esta región (Moncalian et al., 1999b).

Figura I15. Comparación de la región IR2-*nic*-*ihfA* de F y R388 donde se observa la diferencia en la localización relativa que adoptan. Imagen tomada de (Moncalian et al., 1999b).

Se ha propuesto, que la unión de IHF al $oriT_{R388}$ induce un cambio estructural en la topología del ADN que provoca una rigidez alrededor del sitio *nic*, creando un relaxosoma "cerrado" que impide la separación de las hebras, y por lo tanto el corte por la relaxasa.

Se ha observado, a su vez, un efecto de IHF en la recombinación sitio específica mediada por TrwC (Cesar and Llosa, 2007). En ausencia de TrwA, la unión de IHF al *oriT* mantiene los niveles de recombinación bajos, probablemente reteniendo al relaxosoma en esa conformación cerrada mencionada anteriormente. Se propone que IHF y TrwA tienen funciones antagonistas en la recombinación, IHF estaría ejerciendo una inhibición mientras que TrwA contrarrestaría esa inhibición, siendo necesaria para que el proceso de recombinación se de correctamente.

Se podría afirmar, por lo tanto, que TrwA sería la responsable de activar tanto la recombinación como la conjugación modificando la conformación del relaxosoma a una forma "abierta" que dejaría expuesto el ADN de cadena sencilla en los alrededores del sitio *nic*, facilitando el corte inicial que lleva a cabo la relaxasa.

Con los datos obtenidos en los distintos experimentos realizados, se ha especulado con que la activación del procesamiento del ADN durante la conjugación de R388 puede estar mediada por la liberación de IHF del *oriT*.

4.2. Proteínas auxiliares de la conjugación codificadas por el plásmido

4.2.1. Sistema MOB_P

El plásmido RP4 no depende de proteínas auxiliares codificadas por el huésped, pero a cambio tiene tres proteínas accesorias (TraH, TraJ, TraK) que aportan la conformación y especificidad que requiere el proceso de transferencia (figura I16). TraJ es una proteína de 13,2KDa que se une al *oriT*. Concretamente reconoce de modo específico el brazo derecho de una repetición de 19 pb que se encuentra a 8 pb del *nic* (Furste et al., 1989). Un análisis más detallado de la secuencia del brazo derecho de la repetición muestra la presencia de una repetición invertida más corta (G/AGTG/A). El reconocimiento de esta secuencia por TraJ se ha propuesto como el paso inicial de la cascada que genera la formación del relaxosoma (Pansegrau *et al.*, 1990). Se piensa que la unión de TraJ altera la estructura local del ADN de forma que permite el acceso de TraI a su sitio de reconocimiento.

TraH también es necesaria para la formación de un relaxosoma estable. Es una proteína ácida oligomérica de 12,9 KDa que no une ADN, pero se cree que se une tanto a TraI como a TraJ estabilizando el relaxosoma, siendo este complejo visible en ADN superenrollado utilizando un microscopio electrónico (Pansegrau et al., 1990). Se ha propuesto que esta proteína pudiera tener una actividad tipo chaperona.

La tercera proteína que participa en la formación del relaxosoma es TraK (14,6KDa). TraK está implicada en la formación tanto *in vitro* como *in vivo* del relaxosoma, y en el aumento en la obtención de plásmido en forma circular abierta que se aísla del mismo. Junto con TraJ, parece ser un determinante de la especificidad de transferencia (Furste *et al.*, 1989; Waters *et al.*, 1991). Aunque se observó que la presencia de una repetición invertida de 49 pb estaba implicada en el inicio de la unión de TraK, no resultó ser un sustrato para la unión *in vitro*. El análisis *in vitro* de los complejos nucleoproteícos TraJ-*oriT* revela que TraK hace que se envuelva a su alrededor una región de unas 180 pb del origen de transferencia (Ziegelin *et al.*, 1992). Esto produce un cambio local en la superhelicidad que ayudaría a desenrollar la región de mella, facilitando el acceso a TraI. Recientemente nuestro grupo ha mostrado cierta

homología de secuencia y función entre TraK de RP4 y StbA de R388, habiendo demostrado que StbA está implicada tanto en conjugación como en estabilidad del plásmido (Guynet et al.)

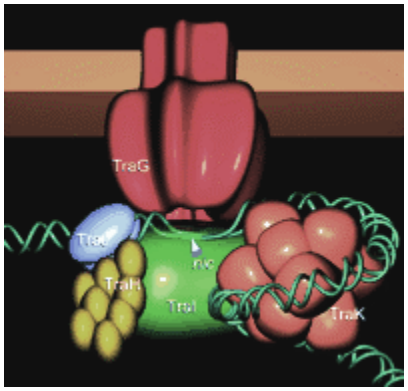


Figura I16. Modelo propuesto para el relaxosoma de RP4. TraG (en rojo) es la proteína acopladora que podría funcionar como canal para la translocación de ADN de cadena sencilla. La relaxasa TraI, que se muestra en verde, estaría interaccionando con el *oriT* y TraG. TraJ se muestra en azul, el homomultímero de TraH en amarillo, y TraK, en naranja. Tomado de (Schroder *et al.*, 2002).

La proteína NikA del plásmido R64 del grupo MOB_P también ha sido caracterizada bioquímicamente. NikA se une a una repetición invertida próxima al sitio *nic* e induce una curvatura en el ADN (Furuya and Komano, 1995). Recientemente se ha obtenido la estructura tridimensional por resonancia magnética nuclear del fragmento N-terminal de NikA, NikA(1-51). Este fragmento es un homodímero compuesto por dos dominios idénticos RHH, mientras que la proteína entera (NikA(1-109)) se encuentra como homotetrámero en solución. Las hojas beta de los dos dominios diméricos que formarían el tetrámero se unen a dos repeticiones internas homólogas de 5 pb dentro de la repetición de 17 pb descrita como su sitio de unión. Como tetrámero, la proteína completa muestra mayor afinidad por la repetición A (situada a 8 nucleótidos del sitio nic) y es capaz de producir una mayor curvatura en el *oriT*. Los autores de este trabajo proponen que el descubrimiento del plegamiento de tipo RHH en NikA implica no sólo su unión a ADN, sino también la unión a la relaxasa NikB. Basan su hipótesis incapacidad de NikB para unirse al *oriT* en ausencia de NikA y citan la existencia de

otras proteínas RHH que participan en la interacción con otras proteínas como pueden ser ParG, FitA o MetJ. (Yoshida *et al.*, 2008), figura I18.

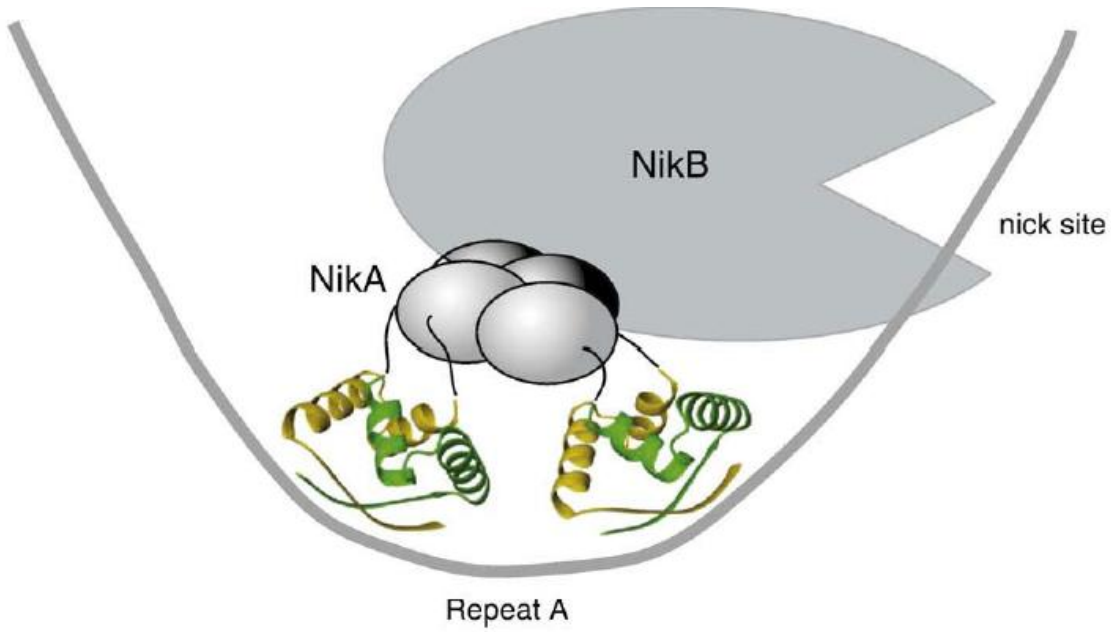


Figura I18. Modelo del mecanismo de acción de NikA en la conjugación bacteriana. Los dos dominios RHH de NikA se muestran en verde y amarillo. La región terminal tetramérica se representa con elipses grises, y NikB se muestra como una gran elipse (Yoshida et al., 2008).

4.2.2 Sistema MOB_Q

Para la transferencia eficiente del plásmido R1162, se necesitan dos pequeñas proteínas además de la relaxasa. Estas proteínas son MobB (15KDa) y MobC (11KDa). Ambas aparecen en el complejo en altas cantidades estequiométricas, y no se ha delimitado un sitio de unión claro para ellas en el *oriT*. Esta característica, y la falta de sitios de unión para proteínas codificadas por el huésped, pueden ser los motivos por los que el *oriT* de R1162 (35pb) es tan pequeño en comparación con el de F (350pb).

MobC es necesaria para la extensión de la región de cadena sencilla que se genera en el *nic* (Zhang and Meyer, 1997). Puede que envuelva el ADN, induciendo un estrés torsional que se libera mediante la separación de hebras.

MobB es necesaria para que el relaxosoma sea completamente funcional a la hora de producir el corte *in vitro*, pero cuando está ausente del complejo los efectos no son dramáticos (Scherzinger *et al.*, 1992), de hecho, los plásmidos carentes de MobB se movilizan a una frecuencia de 0,1% respecto al salvaje. Se ha observado que MobB participa en el transporte de la relaxasa a la célula receptora, también en ausencia de ADN plasmídico (Vergunst *et al.*, 2000). MobB está asociada con la membrana, y tiene un dominio putativo de unión a membrana que es importante para la transferencia (Parker and Meyer, 2007). Aunque los detalles de su funcionamiento no se conocen, se propone que MobB mediante el anclaje de la relaxasa a la membrana aumenta la interacción de la misma con la proteína acopladora.

4.2.3. Sistema MOB_F

A diferencia de las relaxasas de IncP, las relaxasas de IncF son capaces de realizar el corte de ADN de cadena doble linear o relajado, para lo que se requiere la presencia de dos proteínas accesorias, una codificada por el huésped que ya se ha comentado anteriormente (IHF), y TraY codificada por el plásmido (Inamoto *et al.*, 1994; Nelson *et al.*, 1995). Para el corte de sustratos de ADN superenrollados, la relaxasa no necesita de estas proteínas, aunque la eficiencia que se obtiene es más baja (Inamoto et al., 1994).

TraY es una proteína pequeña necesaria para la unión y el corte de la relaxasa en F (Karl *et al.*, 2001; Nelson *et al.*, 1995) y R100 (Inamoto et al., 1994). Es una proteína de unión a ADN relacionada estructuralmente con las proteínas RHH que se explicarán más en detalle en el siguiente apartado. Se une a homopurina-homopirimidina [d(GA)n : d(CT)n] (Krasilnikova *et al.*, 2001) y curva el ADN con un ángulo de entre 50-55º. Es un monómero en solución, que se une a dos sitios en el *oriT* (*sbyA* y *sbyC*), y a un sitio en el promotor Py (*sbyB*) (Luo *et al.*, 1994; Nelson *et al.*, 1993). Mediante su unión al promotor, regula el operón *tra* de F positivamente (Silverman and Sholl, 1996), o

negativamente en R100 (Taki *et al.*, 1998). Estimula además los dos promotores de TraM de F (Penfold *et al.*, 1996).

El sitio de unión de la relaxasa y el sitio *sbyA* están demasiado lejos uno de otro como para permitir interacciones directas de ambas proteínas. No obstante, al encontrarse separados por el sitio de unión *ihfA*, un giro inducido por IHF al unirse a este sitio podría permitir un contacto específico entre estas proteínas (Byrd and Matson, 1997). Los factores que favorecen la existencia de regiones con carácter de cadena sencilla en el ADN favorecen también la reacción catalizada por TraI (Nelson et al., 1995), por lo que un paso crucial en el ensamblaje del relaxosoma es la presentación de la región *nic* en una forma que asemeje ADN de cadena sencilla.

TraM de F es una proteína de 127 aminoácidos que muestra conservación de secuencia en otros sistemas de transferencia de IncF (Frost et al., 1994). Realiza funciones cruciales en la conjugación, ya que no solo promueve la formación del relaxosoma y autorregula su expresión, sino que se ha descrito su interacción con la proteína acopladora del sistema TraD (Lu *et al.*, 2008) y con la relaxasa TraI, estimulando el corte en el *oriT* y las reacciones de separación de hebras llevadas a cabo por la última (Ragonese *et al.*, 2007; Sut *et al.*, 2009). Es una proteína de unión a ADN que se une a regiones ricas en GC rodeadas de curvaturas intrínsecas y sitios de unión a IHF en el *oriT* (Frost et al., 1994). Hay de dos a cuatro sitios de unión de TraM (*sbmA*-D) que pueden ser descritos como repeticiones directas o invertidas que contienen pequeños palíndromos. El sitio de unión más cercano al *nic*, *sbmC*, es importante para la transferencia del plásmido, mientras que los sitios *sbmA* y *sbmB* están implicados principalmente en la regulación de los dos promotores de TraM (Fu et al., 1991). En fragmentos de ADN lineal, TraM se une a *sbmA* y *sbmB* con mayor afinidad que a *sbmC*, siendo esta unión cooperativa, ya que la afinidad por el sitio *sbmC* aumenta diez veces en presencia de *sbamAB*. Al igual que TraY, es capaz de producir una curvatura en el ADN (Fekete and Frost, 2002).

Se ha propuesto que TraM cumple una función en la señalización del inicio de la transferencia de ADN, por lo que sería la encargada de regular la conjugación en F (Willetts and Wilkins, 1984).

En un estudio reciente, se ha obtenido la estructura tridimensional de TraM de pED208 unida a *sbmA* por cristalografía de rayos X (Wong JJ, 2011)(figura I19). El dominio de unión a ADN está formado por una hoja beta y dos hélices alfa (RHH), por lo que se engloba en la misma familia que TraY. Se aprecia en la estructura como dos tetrámeros de ADN se unen a un sitio *sbmA*. Los dos módulos RHH de dos tetrámeros de TraM contactan con la diana GANTC dentro de *sbmA*. Cada tetrámero, reconoce un par de motivos GANTC separados por 12 pb. La cooperatividad de unión a ADN se consigue mediante la torsión y el desenrollamiento del ADN para posicionar los motivos GANTC alternos en la misma cara de la hélice.

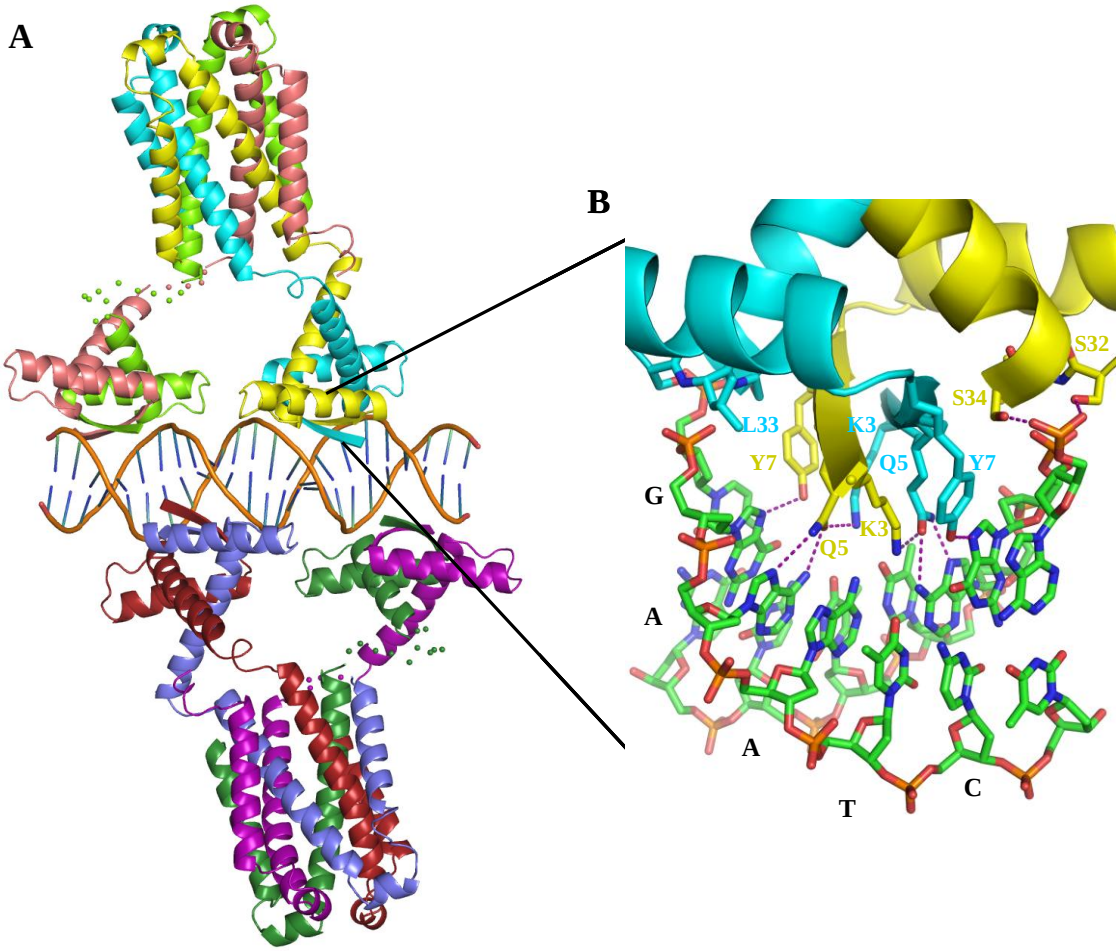


Figura I19. A. Vista ortogonal de la estructura general del complejo TraM-*sbmA*. Los conectores desordenados se muestran como esferas, una por cada Cα que no puedo ser refinado. **B.** Interacciones entre el dominio N-terminal de TraM y el ADN. Los puentes de hidrógeno se muestran como líneas punteadas, y se destacan los aminoácidos que establecen contacto directo con el ADN. Tomado de (Wong JJ, 2011).

4.2.4. Sistema MOB_{HEN}

En un estudio reciente, se ha analizado la posible función de la proteína MbeC de ColE1 como proteína accesoria. Mediante la comparación de secuencia aminoacídica y predicción de estructura secundaria se ha determinado que MbeC está relacionada con la familia de proteínas RHH (Figura I20). Por proteólisis limitada se fragmentó la proteína en dos dominios estables: el dominio N-terminal que contiene un motivo RHH, y el dominio C-terminal que es similar al de otras proteínas auxiliares. En el *oriT* de ColE1 hay una repetición invertida que se creía que podía ser el sitio de unión de MbeC, pero mediante ensayos de unión se observó que esta proteína no reconocía el fragmento que llevaba esta repetición, sino que se unía a la región *nic* del *oriT*. Los resultados obtenidos sugieren que MbeC cumple una función similar en la conjugación al de TraY y TrwA de F y R388 respectivamente (Varsaki *et al.*, 2009).

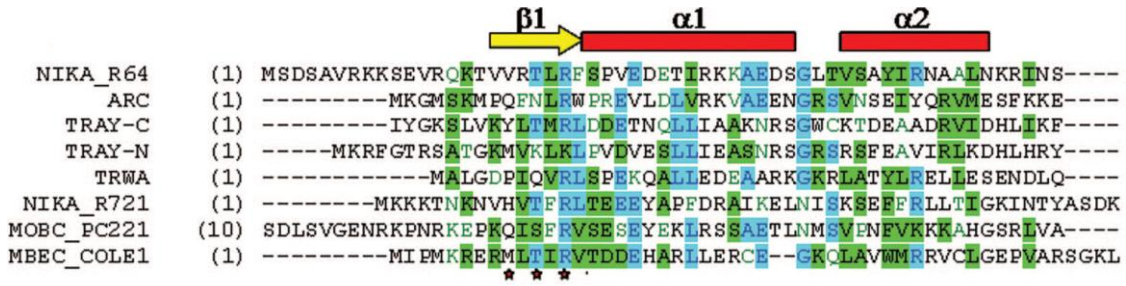


Figura I20. Análisis mediante CLUSTALW de los 50 primeros aminoácidos de MbeC junto con otras proteínas auxiliares RHH. Tomado de (Varsaki et al., 2009).

4.2.5. TrwA de R388

El gen *trwA* está localizado en el plásmido R388 al principio del operón *trwABC* y justo después del *oriT*, por lo que sería el último gen en entrar a la célula receptora. La proteína para la que codifica es una proteína de 121 aminoácidos y 13,4KDa, que tetrameriza en solución formando un complejo de 53KDa. Una cremallera de leucinas

que podría situarse entre los aminoácidos 34-67 puede ser la responsable de las interacciones proteína-proteína que se llevan a cabo en la oligomerización (Llosa et al., 1994a).

Mediante proteólisis limitada con tripsina, se han obtenido dos complejos estables de TrwA, un dominio N-terminal implicado en la unión a ADN, y un dominio C-terminal responsable de la tetramerización de la proteína. El dominio N-terminal contiene un motivo RHH y mantiene intacta la capacidad de unión a ADN de la proteína. El dominio C-terminal de TrwA (36-121), en cambio, no tiene capacidad de unión a ADN aunque sigue siendo tetrámero en solución (Moncalian and de la Cruz, 2004).

TrwA se une a dos sitios en el *oriT*, formados por una repetición directa (*sbaA*) y otra invertida (*sbaB*) del hexanucleótido CACTAC, siendo mayor la afinidad *in vitro* por el sitio *sbaA* (Moncalian et al., 1997). TrwA realiza varias funciones necesarias para una conjugación eficiente. Actúa como represor transcripcional de su propio operón *trwABC*, cuyo promotor se encuentra solapando los sitios de unión de TrwA. Por otra parte, la unión de TrwA al ADN estabiliza y aumenta la unión de la relaxasa TrwC al *oriT*, siendo este el motivo más probable por el cuál aumenta el corte que realiza TrwC en el sitio *nic* (Moncalian et al., 1997). TrwA incrementa, además, la actividad de recombinación sitio específica mediada por TrwC (Cesar *et al.*, 2006), así como la actividad ATPasa de la proteína acopladora del sistema TrwB (Tato et al., 2007).

Mediante estudios de predicción de estructura secundaria, y los resultados de alineamientos que destacan la similitud con represores transcripcionales ya caracterizados como Arc y Mnt, u otras proteínas accesorias como NikA o TraY, se engloba a TrwA en la familia de proteínas RHH. Se ha demostrado, además, que establece contactos específicos con el ADN mediante los residuos polares de la lámina beta de este dominio RHH, ya que las mutaciones en los aminoácidos Q8, R10 y S12 afectan a su capacidad de unión al *oriT* (Moncalian and de la Cruz, 2004).

5. SUPERFAMILIA DE PROTEÍNAS RHH

La familia RHH de factores transcripcionales utiliza un motivo estructural tridimensional conservado para unirse al ADN de manera específica de secuencia. Esta familia de proteínas funcionalmente diversa regula la transcripción de genes implicados en la captación de metales, biosíntesis de aminoácidos, división celular, el control de número de copias de plásmidos, el ciclo lítico de los bacteriófagos, y puede que en muchos otros procesos celulares (Schreiter and Drennan, 2007).

Es una familia creciente de proteínas, habiendo sido identificadas secuencias de factores de transcripción RHH en bacterias, arqueas y bacteriófagos, auunque no así en eucariotas. La unidad funcional del dominio RHH es un dímero en el que dos motivos RHH se encuentran estrechamente entrelazados para formar un dominio estable que tiene simetría dos y es capaz de unirse al ADN. El dominio de simetría dos generalmente se forma por homodimerización, excepto en el caso de la proteína TraY de F, que contiene dos repeticiones de la secuencia del motivo RHH en una única cadena polipeptídica (Lum and Schildbach, 1999).

La hebra beta presente en el N-terminal de cada monómero RHH forma una hoja beta antiparalela. La unión del dímero RHH al ADN posiciona esta hoja beta en el surco mayor del ADN (figura I21). Tres cadenas laterales de aminoácidos de cada hebra se posicionan hacia el interior del surco y forman contactos específicos de secuencia cruciales con las bases de los nucleótidos. La unión específica por los residuos de la hoja beta es una característica propia de los miembros de la familia RHH, y distingue a estas proteínas de la superfamilia ubicua HTH, que inserta una hélice alfa en el surco mayor del ADN. También se llevan a cabo una serie de contactos no específicos conservados con el esqueleto fosfato del ADN mediante dos amidas del N-terminal de la segunda hélice del dominio RHH, en ambos lados del surco mayor.

Aunque la identidad de secuencia primaria de las secuencias RHH es baja y no hay ninguna posición de aminoácidos que se encuentre totalmente conservada, un análisis más detallado de la secuencia y alineamientos estructurales sacan a relucir la presencia de varios motivos conservados. En la hebra beta del extremo N-terminal, entre

las posiciones 2 y 7 se repite un patrón de residuos hidrofóbicos y polares alternos, en los que las cadenas laterales hidrofóbicas que se encuentran en las posiciones 3, 5 y 7 apuntan al núcleo del dominio y las polares de las posiciones 2, 4 y 6 apuntan hacia fuera del dominio para establecer la conexión con los nucleótidos. Un alineamiento de la secuencia primaria de las proteínas RHH conocidas muestra la presencia de una arginina o lisina en la posición 2 de prácticamente todas ellas. Un segundo motivo estructural estaría compuesto por las dos hélices alfa situadas después de la hebra beta. En estas dos hélices hay cuatro residuos hidrofóbicos conservados que mantienen la integridad del núcleo hidrofóbico, siendo esta una característica esencial para mantener la integridad del dominio.

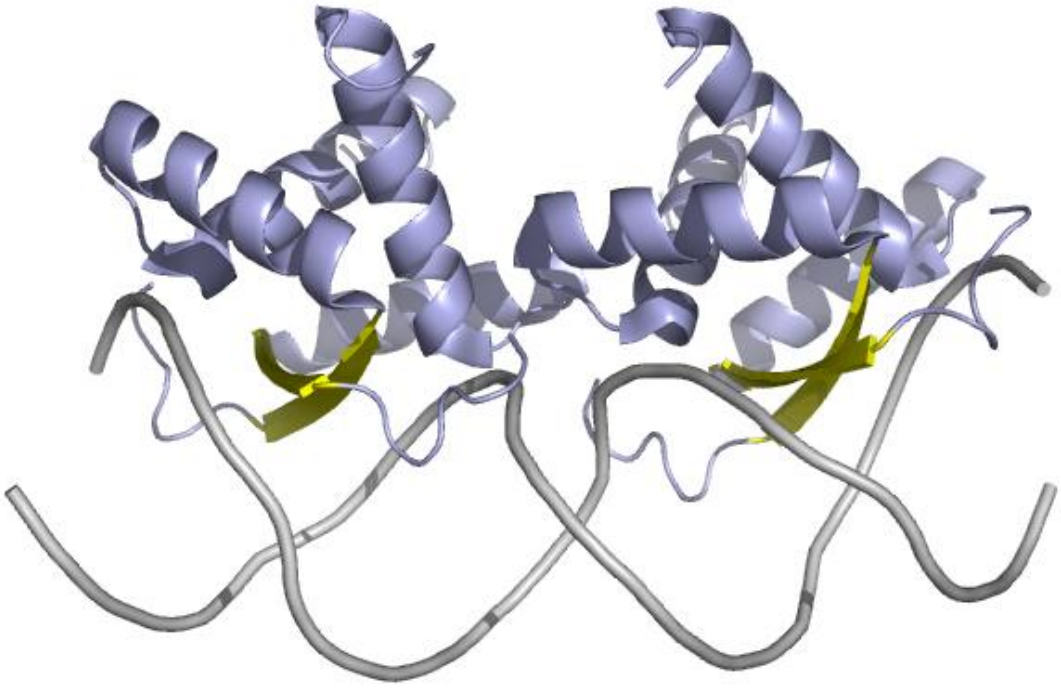


Figura I21. Represor Arc (RHH) unido a su operador. En azul se muestran las hélices alfa y en amarillo las hojas beta que establecen contactos específicos con el ADN que se muestra en gris (PDB 1PAR).

Aunque el modo de unión a ADN de cada dominio RHH es prácticamente el mismo, la manera en la que se da el contacto específico con las bases del ADN no parece estar conservada entre los distintos miembros de esta familia. Cada miembro de esta familia de proteínas reconoce un único operador (u operadores). Proteínas como MetJ pueden unirse a una secuencia simétrica de operador, mientras que otras proteínas RHH reconocen sub-sitios asimétricos. Repeticiones en tándem de los sitios permiten la unión cooperativa de varios dominios RHH. En el represor omega, se pueden encontrar hasta 10 repeticiones en tándem, y además, al reconocer una secuencia no palindrómica, hay tres posibilidades para la orientación de de dos sub-sitios adyacentes, los cuáles podrían estar formando una repetición directa, o dos repeticiones invertidas, dependiendo de si los dos brazos se encuentran enfrentados o apuntando hacia fuera. La orientación de estos brazos modula la afinidad del represor omega por el operador (de la Hoz *et al.*, 2004).

Actualmente no existe un código que relacione directamente la secuencia aminoacídica de la hoja beta con una secuencia específica de ADN, lo que hace que la predicción de nuevos operadores RHH resulte difícil.

Las proteínas RHH no sólo están implicadas en el reconocimiento específico del ADN, sino que también participan en interacciones proteína-proteína. Algunos ejemplos son el repressor ParG que interacciona con la proteína de partición ParF tanto en presencia como en ausencia de ADN, el repressor MetJ que se une a varios operadores con el co-represor S-adenosilmetionina, o la proteína del sistema de partición ParD, que se une a la toxina ParE probablemente por su extremo C-terminal (Yoshida et al., 2008).

En este trabajo caracterizando en detalle TrwA vamos a ahondar en el papel de las RHH en conjugación y cómo siendo capaces de reconocer IR y DR, la disposición de estos sitios es importante en el *oriT*. Esta caracterización será muy útil a la hora de diseñar *oriT* mínimos o nuevos *oriT* con diferentes requerimientos para ser transferidos.

2.

OBJETIVOS

Objetivos

El principal objetivo que nos planteamos al inicio de esta tesis fue el diseño y construcción de un sistema conjugativo eficiente, fácilmente manipulable y no tan dependiente de la regulación por parte de distintas proteínas. Para ello necesitamos conocer en detalle el papel de los distintos elementos estructurales del *oriT*. Diseñamos un *oriT* sintético que nos permitiría estudiar la movilización del plásmido en presencia y en ausencia de cada elemento de estructura secundaria, definiendo el rol que desempeña cada uno en la conjugación. Por otro lado, necesitábamos estudiar en profundidad el papel de las proteínas auxiliares (TrwA en R388), centrándonos en su especificidad de unión al ADN y en el motivo de la peculiar distribución que muestran sus sitios de unión en el *oriT*. Ya que TrwA desempeña varias funciones durante la conjugación del plásmido R388, con este estudio pretendíamos obtener más detalles del procesamiento inicial del ADN, y por otro lado, aportar nuevos datos para facilitar el estudio de la familia de represores transcripcionales RHH.

A continuación se detallan los objetivos concretos planteados:

1. Definir los elementos mínimos que debe disponer un *oriT* funcional.

2. Ahondar en el estudio de la especificidad de la proteína auxiliar de la conjugación TrwA.

3. Descubrir las razón por la cuál TrwA presenta una distribución de sus sitios de unión a modo de repetición directa e invertida. Comprobar si esta distribución es única en R388 o se encuentra también en otros sistemas.

4. Estudiar las posibles relaciones de TrwA con otras proteínas del relaxosoma como TrwC o IHF.

3.

MATERIALES Y MÉTODOS

3. MATERIALES Y MÉTODOS

MATERIALES:

1. CEPAS BACTERIANAS

Cepa	Genotipo	Referencia
DH5α	F^- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17(r_K^- m_K^+), λ–	(Grant *et al.*, 1990)
C41 (DE3)	F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm* (DE3)	(Miroux and Walker, 1996)
UB1637	F^- *recA56 his lys trp rpsL*	(de la Cruz and Grinsted, 1982)
BW27783	*lacI*^{q}*rrnB3* Δ*lacZ4787 hsdR514* Δ(*araBAD*)*567* Δ(*rhaBAD*)*568* Δ(*araFGH*) Φ(Δ*araEp* P_{CP8}-*araE*)	(Khlebnikov *et al.*, 2001)

Tabla M1. Cepas utilizadas en este trabajo.

2. MEDIOS DE CULTIVO

Para el cultivo líquido de bacterias se utilizó el medio Luria-Bertani (LB) (triptona 10 g/l, extracto de levadura 5 g, NaCl 5 g/l, Pronadisa). En medio sólido se utilizó LB-agar (se suplementa el LB con agar 1,5 % (p/v), Pronadisa). Todos los medios de cultivo se esterilizaron autoclavándolos a 120 ˚C durante 20 min. Para conservar las cepas, se centrifugaron los cultivos en fase estacionaria y se resuspendieron en 50% de peptona 1,5% (p/v) y 50% de glicerol. Las cepas se guardaron a -20 ˚C para su uso frecuente y a -70 ˚C para su almacenaje a largo plazo.

Se añadieron antibióticos a los medios de cultivo para hacerlos selectivos. Las concentraciones a las que se utilizaron son las que se muestran a continuación: 100 μg/ml de ampicilina sódica (Apollo), 25 μg/ml de cloranfenicol (Sigma), 50 μg/ml de sulfato de kanamicina (Sigma), 25 μg/ml de ácido nalidíxico (Sigma), 300 μg/ml de sulfato de streptomicina (Apollo).

El cloranfenicol se preparó disuelto en etanol. En el resto de los casos, los antibióticos se prepararon disueltos en agua a una concentración 1000 veces mayor de la concentración indicada.

3. PLÁSMIDOS UTILIZADOS

Se muestran a continuación dos tablas que contienen respectivamente los plásmidos ajenos utilizados (Tabla M2) y los construidos en este trabajo (Tabla M3).

Plásmido	**Descripción**	**Fenotipo**	**Referencia**
pAR12	pBAD33:*trwA*	Cm^R	(Fernández-Lopez R, 2011)
pET29C	Vector de expresión	Kn^R	Novagen
pSU19	Vector de clonación	Cm Rep(p15A)	(Bartolome *et al.*, 1991)
pSU711Δ*oriT:aac(3)-IV*	Derivado de R388 sin *oriT*	Gm^RKn^R (IncW)	(Demarre *et al.*, 2005)
pSU1547	pET22a:*trwA*R388	Ap^R, Rep (pMB8)	(Moncalian and de la Cruz, 2004)
pSU1548	pET22::trwA(2– 73)	Ap^R, Rep (pMB8)	(Moncalian and de la Cruz, 2004)
pSU1550	pET22::trwA(36–121)	Ap^R, Rep (pMB8)	(Moncalian and de la Cruz, 2004)
pSU1621	pET3a::*trwC*	Ap^R, Rep (pMB8)	(Guasch *et al.*, 2003)
pSU1588	pET3a::*trwC*-N293	Ap^R, Rep (pMB8)	(Hernando, 2000)
pSUMC29	pET3a::$trwC_{HEL}$	Ap^R, Rep (pMB8)	Cabezas, M (sin publicar)
pSU2007	Derivado Km^R de R388	Cm^RRep(p15A)	(Martinez and de la Cruz, 1988)
pSU4910	pSU18::*oriT*	Cm^RRep(p15A)	Fernandez-Lopez R. (sin publicar)
pUA66	Plásmido reportero	Kn^RRep (pSC101)	(Zaslaver *et al.*, 2004)
R388	Salvaje	Su^R, TpR, Tra1, IncW	(Datta and Hedges, 1972)
R7K	Salvaje	IncW	(Coetzee *et al.*, 1972)

Tabla M2. Plasmidos ajenos utilizados en este trabajo.

Plásmido	Descripción	Vector	Enzimas/Oligonucleótidos 5' 3'
pYTP1	*oriT*$_{R7K}$	pSU19	*Hind*III, *Bam*HI/ oriTR7K388SH, oriTR7KEBN
pYTP2	*oriT_trwA*$_{R7K}$	pSU19	*Hind*III, *Bam*HI/ oriTR7K388SH, oriTAR7K388EB
pYTP3	*oriT_trwA*$_{R388}$	pSU19	*Hind*III, *Eco*RI/ oriTR7K388SH, TrwaEEco
pYTP4	*oriT*synth	pSU19	*EcoR*I *Hind*III (ver M3.1)
pYTP5	pYTP4ΔIR1	pSU19	*Pst*I
pYTP6	pYTP4Δbeg	pSU19	*Pst*I *Kpn*I
pYTP7	pYTP4ΔIR2	pSU19	*Bss*HII
pYTP8	pYTP4ΔIR3	pSU19	Ver M3.1
pYTP9	pYTP4ΔsbaA	pSU19	*Bgl*II
pYTP10	pYTP4ΔsbaB	pSU19	*Bam*HI
pYTP11	pYTP4ΔsbaAB	pSU19	*Bgl*II *Bam*HI
pYTP12	pYTP42IR	pSU19	*Bgl*II / sbaAIR1, sbaAIR2
pYTP13	pYTP42DR	pSU19	*Bam*HI / sbaBDR1, sbaBDR2
pYTP14	pYTP4sbaA,sbaBinv	pSU19	Ver M3.1
pYTP16	pYTP4IR2Flip	pSU19	FliporiTA, FliporiTB
pYTP17	pYTP4_GFP	pUA66	BspHIoriBeg, XhoIoriEnd
pYTP18	pYTP17QC-35	pUA66	trwAprmQuickA, trwAprmQuickB
pYTP19	pYTP12_GFP	pUA66	Los utilizados en pYTP17
pYTP20	pYTP19QC-35	pUA66	Los utilizados en pYTP18
pYTP21	pYTP13_GFP	pUA66	Los utilizados en pYTP17
pYTP22	pYTP21QC-35	pUA66	Los utilizados en pYTP18
pYTP23	TrwA$_{R7K}$K8Q,T12S	pET29C	TrwAR7KE, AR7K288S
pYTP24	pYTP23F9V	pET29C	TrwAR7KE, 7K88V
pYTP25	pYTP23E5D	pET29C	TrwAR7KE, 7K88D
pYTP26	pYTP24E5D	pET29C	TrwAR7KE, 7K88DV
pYTP27	pET29C::*trwA*	pET29C	*Xho*I, *Nde*I

Tabla M3. Plásmidos construidos en este trabajo. Los plásmidos marcados con asteriscos se detallan en el siguiente sub-apartado 3.1.

3.1 Construcción de plásmidos concretos

- pYTP4 y el diseño del origen de transferencia sintético

Muchos de los plásmidos construidos en este trabajo y enumerados en la tabla M3 tienen como base el plásmido pYTP4 que contiene el *oriT* denóminado sintético.

Se denomina síntesis de ADN artificial al proceso de sintetizar un gen o una secuencia de ADN cualquiera sin que sea necesario para ello un ADN molde inicial. Hoy en día existen compañías especializadas en la síntesis de fragmentos desde pocos cientos de pares de bases hasta algunas kilobases de longitud. La compañía elegida en nuestro caso fue GeneArt (Alemania).

Nos interesaba desarrollar un estudio del origen de transferencia de R388, por lo que diseñamos un *oriT* en el que cada módulo de estructura secundaria estuviera delimitado por sitios de restricción que permiten su eliminación independientemente del resto de la secuencia (figura M1). No solo podemos eliminar estos elementos, sino que podemos introducir en su lugar cualquier otra secuencia de interés siempre que mantengamos los sitios de restricción en los extremos. Mediante la digestión, y posterior religación del *oriT*, obtenemos una serie de mutantes que carecerán del sitio elegido, lo que nos permite el estudio de su papel en el proceso de la conjugación.

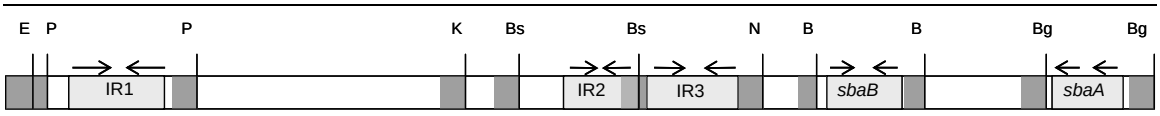


Figura M1. Esquema del origen de transferencia sintético con los elementos de estructura secundaria representados como rectángulos gris claro, y los sitios de restricción como bloques en gris más oscuro. Las letras mostradas en la parte superior indican el sitio de restricción insertado (E; *Ecor*I, P; *Pst*I, K; *Kpn*I, Bs; *Bss*HII, N; *Nco*I, B; *Bam*HI, Bg; *Bgl*II).

- pYTP8

Tuvimos que abordar la construcción del mutante que carece del elemento IR_3 de una manera secuencial al carecer esta región de dos sitios de restricción que lo delimitaran (figura M1). En primer lugar realizamos una doble digestión con *Pau*I y *Nco*I del plásmido pYTP4. Los extremos resultantes de la digestión no son compatibles, por lo que se llevó a cabo una ligación utilizando el fragmento klenow de la polimerasa I de ADN (Fermentas) para rellenar los extremos no cohesivos. Al llevar a cabo esta digestión, la secuencia que abarca la región del *nic* y el sitio de unión de la relaxasa también desaparece, por lo que insertamos de nuevo este fragmento mediante la hibridación de oligonucleótidos comerciales previamente fosforilados (aquellos utilizados en la construcción del plásmido pYTP16). Para la inserción, previamente se digirió el plásmido de nuevo con *Pau*I, ya que después de su religación con klenow, se había restablecido un sitio de restricción para dicha enzima (figura M2).

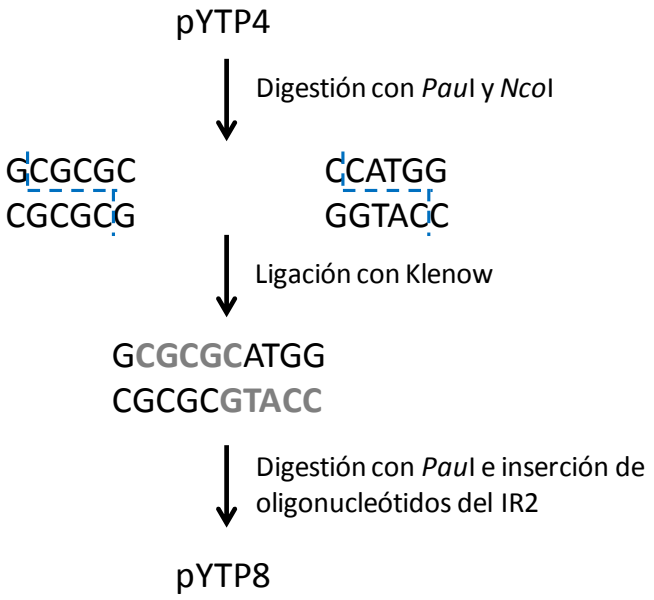


Figura M2. Esquema de la construcción del plásmido pYTP8.

- p YTP14

Para conseguir el mutante con los sitios *sbaB* y *sbaA* en posiciones invertidas, se partió del clon p YTP13 que contiene dos sitios *sbaA* y se procedió a digerirlo con la enzima *Bgl*II, de modo que se eliminaría el sitio *sbaA* más cercano al inicio de la proteína. Una vez digerido el plásmido, se ligó con la hibridación de los oligonucleótidos utilizados para construir el plásmido pYTP12, introduciendo así un sitio *sbaB* donde antes había un sitio *sbaA*.

4. OLIGONUCLEÓTIDOS

Se muestran dos tablas con los oligonucleótidos utilizados para métodos de Biología Molecular (tabla M4) y métodos bioquímicos (tabla M5).

Oligonucleótido	**5'** **Secuencia** **3'**
oriTR7K388SH	AAAAAaagcttCGTTTTAGCTCATTTTC
oriTR7KEBN	AAAAAggatccCATATGCGCTTGTGGTCCTAGTTGT
oriTAR7K388EB	AAAAAggatccTCAATCCTCCTTCCCCT
TrwaEEco	AAAAAgaattcTCAATCCTCCTTCCCCT
FliporiTA	CGCGCTATAGACAATACGCACCTTTCGGTG
FliporiTB	GCGCGCACCGAAAGGTGCGTATTGTCTATAG
sbaAIR1	GATCTGTGTAGTGTTATGCCACTACAA
sbaAIR2	GATCTTGTAGTGGCATAACACTACACA
sbaBDR1	GATCCGAACCACTACAGTAACA
sbaBDR2	GATCCCCGTAGTGTTACTGTAGTGGTTCG
BspHIoriBeg	AAAAATCATGAATTCCTGCAGGATAACC
XhoIoriEnd	AAAAACTCGAGCTAGGGCCATGAGATCT
TrwAR7KE	AAAAACTCGAGATCCTCCTTCCCCTCCGG
AR7K288S	AAAAAACATATGGCACTTGGCGAACCTATCCAATTCCGG CTCAGCCCGGAGAAC
7K88D	AAAAACATATGGCACTTGGCGACCCTATCCAATTCCGGC
7K88V	AAAAACATATGGCACTTGGCGAACCTATCCAAGTACGGC TCAGCCC
7K88DV	AAAAACATATGGCACTTGGCGACCCTATCCAAGTACG
trwAprmQuickA	GTTATGCCACTACAGGATTGCCGCAACCCCTTGTAAATG
trwAprmQuickB	CATTTACAAGGGGTTGCGGCAATCCTGTAGTGGCATAAC

Tabla M4. Oligonucleótidos utilizados para clonaje y mutagénesis.

Oligonucleótido	5’ Secuencia 3’
sbaA$_{R388}$	AACCACTACAGTAACACTACGGGA
sbaB$_{R388}$	ACCTGTGTAGTGTTATGCCACTACAATATT
sbaA$_{R7K}$	TTGGGACCACAACTAGGACCACAAAGCG
sbaB$_{R7K}$	AAACATCCGTGGTCTACAGGGACCACAGAT
sbaA$^{1,-1}$	AAC**T**ACTACAGTAA**T**ACTACGGGA
sbaA$^{2,-2}$	AACC**G**CTACAGTAAC**G**CTACGGGA
sbaA$^{3,-3}$	AACCA**T**TACAGTAACA**T**TACGGGA
sbaA$^{4,-4}$	AACCAC**C**ACAGTAACAC**C**ACGGGA
sbaA$^{5,-5}$	AACCACT**G**CAGTAACACT**G**CGGGA
sbaA$^{6,-6}$	AACCACTA**T**AGTAACACTA**T**GGGA

Tabla M5. Oligonucleótidos utilizados para los ensayos de retardo en gel.
- En todos los casos estos oligonucleótidos fueron hibridados con sus complementarios, que no han sido mostrados por simplificación.
- Las mutaciones .introducidas en los oligonucleótidos aparecen sombreadas y subrayadas.
- En los experimentos de cristalización, se utilizaron variantes en longitud de los primeros cuatro oligonucleótidos.

MÉTODOS:

1. METODOS GENERALES DE BIOLOGIA MOLECULAR

1.1 Extracciones de ADN

Para las extracciones de ADN se utilizaron kits comerciales. Para la extracción de ADN plasmídico se utilizó el kit *GenElute Plasmid Miniprep Kit* de Sigma. Para la extracción de bandas de gel de agarosa, se utilizó el *GenElute Gel Extraction Kit* y, por último, para limpiar ADN o cambiarle el tampón, se utilizó el *GenElute PCR Clean-Up Kit* (ambos de Sigma). En todos los casos se siguieron las instrucciones dadas por el fabricante.

Para determinar la concentración y la pureza de las muestras de DNA se midió la absorbancia A260 y la relación A260/280 utilizando un espectrofotómetro *Nano-Drop ND-1000* (Thermo Scientific).

1.2. Transformaciones

1.2.1. Transformaciones por TSB

La transformación de ADN por medio de TSB (LB pH 6,1, con 10% de PEG, 5% DMSO, 10 mM de $MgCl_2$ y 10 mM de $MgSO_4$)(Chung and Miller, 1988), es un método sencillo con el que se consiguen frecuencias de hasta $2x10^8$ transformantes por μg de ADN. Además, las células que han adquirido competencia por este método pueden ser congeladas para ser transformadas en otro momento sin una pérdida significativa de eficiencia.

Para la transformación, se crece en LB un cultivo bacteriano de la cepa que se quiere transformar hasta una DO_{600} de 0,3-0,6, y se recogen las células por centrifugación a 4,000 g. Las células se resuspenden en un volumen de TSB equivalente a 1/10 del volumen inicial (típicamente se resuspenden en 1 ml de TSB a 4˚C) y se incuban en hielo durante 10 min. Se hacen alícuotas de 0,1 ml en tubos fríos, se añaden 10-100pg del plásmido que se quiere transformar, y se mantienen en hielo durante 5-30 min. Acto seguido, se ponen a crecer las células en TSB-glucosa (0,9 ml de TSB con 20 mM de glucosa) a 37 ˚C con agitación durante una hora para expresar la resistencia del antibiótico. Se siembran en placas de LB-Agar con el antibiótico adecuado para seleccionar.

1.2.2. Transformaciones por electroporación

Éste fue el método elegido cuando la frecuencia de transformación necesaria era muy alta, ya que con él se consiguen frecuencias de hasta 10^{10} transformantes por μg de ADN (Dower *et al.*, 1988). Para las células competentes para electroporar se crecieron los cultivos hasta una DO_{600} de 0,5-0,7 y se incubaron durante 30 min en hielo. Se centrifugaron a 3000 g durante 10 min y se lavaron dos veces con agua destilada milliQ estéril a 4 ˚C. Las células se resuspendieron en un 2% del volumen inicial con glicerol al 10% (p/v). Estas células se repartieron en alícuotas de 60 μl y se conservaron a -80 ˚C para su posterior utilización.

Para las transformaciones por electroporación se dializó el ADN a transformar en filtros Millipore GS de 0,05 μM de tamaño de poro en una placa de Petri con agua MilliQ durante 30 min. Se recogió el ADN de este filtro y se añadió de 1 a 50 pg de este ADN a una alícuota de las células competentes. Se depositaron las células junto al ADN en una cubeta de electroporación de 0,2 cm *Gene Pulser* de BioRad (Hercules, Carlifornia) manteniendo todo a 4 ˚C. Se utilizó un electroporador Micropulser™ de BioRad y la electroporación se realizó en las condiciones recomendadas para la transformación de *E. coli*. Inmediatamente después, se añadió 1 ml de LB estéril y se resuspendieron las células en él. Después, se incubaron a 37 ˚C con agitación para la expresión de la resistencia al antibiótico (diferentes tiempos en función del antibiótico que se iba a utilizar para seleccionar), y se sembraron en placas de LB-Agar con el antibiótico correspondiente.

1.3. Amplificación de ADN: PCR

Para la amplificación de fragmentos de ADN, siempre que se requirió una alta fidelidad en la polimerización, como en el caso de las mutagénesis o en los procesos de clonación, se utilizó la polimerasa Vent (BioLabs, Massachussets). En el resto de las amplificaciones, como en el caso de la comprobación de colonias, se utilizó la polimerasa Biotaq (Bioline, Londres). Las reacciones se hicieron en un volumen final de 50 μl, y las concentraciones de cada reactivo se eligieron según la recomendación de cada fabricante.

Las reacciones se llevaron a cabo en un termociclador UNO II (Biometra, Alemania). El programa incluyó en todos los casos un primer paso de desnaturalización de 5 min a 94˚C. Después se llevaron a cabo entre 25 y 30 ciclos compuestos por 30 segundos a 94˚C para desnaturalizar la doble hebra del ADN, 30 segundos a la temperatura de hibridación adecuada para la reacción en cuestión, y 1 min por kb de ADN a 72˚C para la elongación del ADN. Después de estos ciclos, se añadió un último paso de 10 min a 72˚C, con el que se completa la elongación de los fragmentos. Una vez terminada la amplificación las muestras se conservaron a 4˚C. Para calcular la temperatura de hibridación óptima se escogió la Tm más baja de los dos cebadores, y se restaron 5˚C a esta temperatura.

1.4. Mutagénesis dirigida

Para la introducción de mutaciones puntuales en el DNA se siguieron dos estrategias.

I. Se diseñaron cebadores que contenían la mutación deseada, y mediante una reacción normal de amplificación se obtuvieron productos que contenían en su secuencia dicha mutación. Esta estrategia se eligió cuando los nucleótidos a mutar se encontraban cerca de alguno de los extremos del fragmento a clonar. Los productos previamente secuenciados fueron clonados en un vector de expresión.

II. Se utilizó el kit de mutación sitio-específica Quickchange de Stratagene. En este caso, la mutación que deseamos se introduce en dos cebadores complementarios, y se utiliza la polimerasa PfuTurbo para replicar las dos hebras del plásmido. El producto final de la reacción de amplificación es un plásmido mutado que contiene mellas. Seguidamente, el producto de la amplificación es digerido con la endonucleasa *Dpn*I, la cual muestra especificidad por el DNA metilado y hemimetilado, y es utilizada para digerir el ADN molde parental y así seleccionar el ADN sintetizado que contiene la mutación. El ADN mellado que contiene la mutación deseada, se transforma ahora a células competentes, y será extraído de las colonias resultantes que ya habrán reparado las mellas en el plásmido. Un esquema del proceso se presenta en la figura M3.

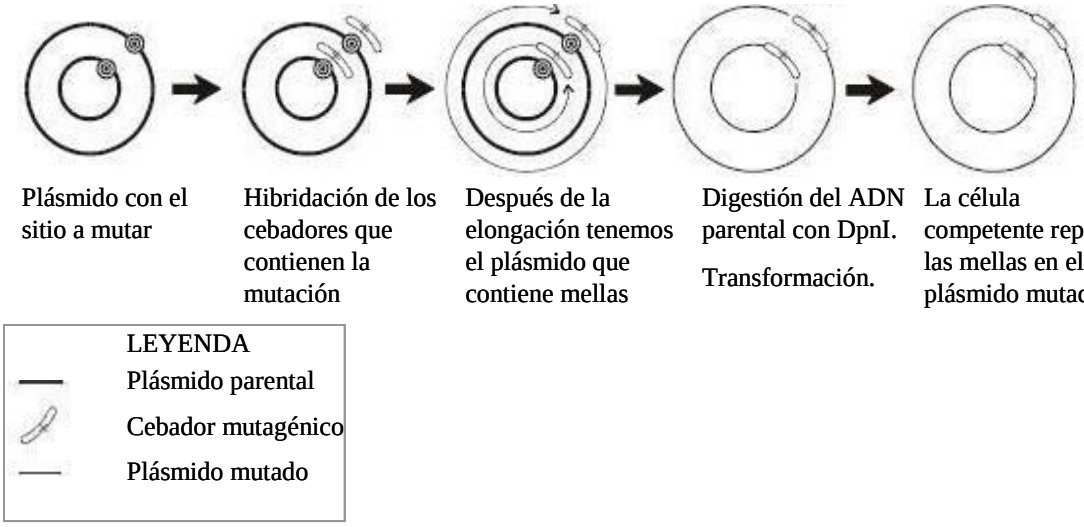


Figura M3. Representación esquemática de la reacción de mutación sitio-específica realizada mediante el método de Quickchange (modificada del protocolo proporcionado por la marca).

1.5. Digestiones enzimáticas

Las digestiones enzimáticas se llevaron a cabo con enzimas de restricción de Fermentas (Ontario, Canadá), siguiendo las instrucciones dadas por la casa comercial. Los volúmenes típicos de estas reacciones podían variar entre 20 y 50µl. El tiempo de reacción fue de media unas dos horas y la temperatura a la que se llevaron a cabo 37 ˚C (siempre que no se indique lo contrario). Para inactivar la reacción se siguieron las recomendaciones específicas para cada enzima (típicamente calentamiento a 65˚C o 80ºC durante 20 min).

1.6. Ligaciones

Los vectores de clonación y los fragmentos de PCR fueron digeridos con los enzimas correspondientes según el método que se ha indicado anteriormente. En el caso de los clones construidos mediante hibridación de oligonucleótidos comerciales, éstos se diseñaron con colas laterales correspondientes a los sitios de restricción elegidos, por lo que no requerían digestión por enzimas de restricción. Si se procedió, en cambio, a su

fosforilación mediante la T4 polinucleotido kinasa (PNK, Fermentas). Para ello se mezclaron 40pmoles de los oligonucleótidos complementarios en un volumen de 20µl con 2µl de ATP 10mM, el tampón requerido por la enzima, y 10 unidades de la PNK, añadiendo el volumen restante de agua hasta llegar a los 20µl.

Por otro lado se eliminaron los fosfatos terminales del vector digerido para aumentar la eficiencia de la ligación y disminuir el número de falsos positivos. Para ello, se utilizó el enzima fosfatasa alcalina (Roche) según las indicaciones del fabricante. Se utilizaron 10 U por reacción de desfosforilación, y se mantuvo 1 hora a 37˚C. Esta reacción se inactivó por medio de incubación a 65˚C durante 10 min.

Las reacciones de ligación se llevaron a cabo poniendo una relación 1:5 vector:inserto implicando típicamente 0,5-10ng de ADN en 20µl de volumen final. Se utilizó para estas ligaciones la ligasa del fago T4 (de Fermentas). En cada caso se utilizaron 4U por reacción. Las reacciones de ligación se llevaron a cabo a 22˚C durante toda la noche. Como control negativo de las ligaciones se utilizó la misma mezcla que en la reacción positiva sin inserto, añadiendo el volumen equivalente de agua en este caso hasta completar los 20µl.

1.7. Secuenciación de ADN

Para comprobar la secuencia de los clones obtenidos, las muestras de ADN se enviaron al "Servicio de Secuenciación de ADN del CIB" (SECUGEN S.L. Madrid) y más adelante al servicio de secuenciación MACROGEN Inc. DNA Sequencing Service (Seul, Korea).

1.8. Electroforesis de ADN

Para la observación y manipulación del ADN se realizaron electroforesis en geles de agarosa al 1, 1,5 ó 2% (p/v) según el tamaño del ADN que se pretendiese resolver. La agarosa de disolvió en tampón TBE 0,5X (Tris-HCl 45 mM, ácido bórico 45 mM

EDTA 0,5 mM pH 8,2). Las muestras de ADN a resolver se mezclaron antes de la eletroforesis con 1/10 de su volumen de tampón de carga (azul de bromofenol 0,25% (p/v) y glicerol 30% (v/v)). Para la visualización del ADN se añadió a la agarosa bromuro de etidio a 0,25 mg/L, o en la última etapa de la tesis, RedSafe 0,4x (Intron Biotechnology). Como marcador se utilizó el patrón HyperLadder I (Biolabs) o 100bp para pesos moleculares bajos (Amersham), en función del tamaño del ADN que se estaba analizando.

La electroforesis se llevó a cabo en cubetas horizontales (BioRad) en tampón TBE 0.5X a voltajes entre 80-100 V. Las imágenes de los geles se obtuvieron en un transiluminador de luz ultravioleta GelDoc 2000 y se analizaron con el programa Quantity One (BioRad).

2. METODOS MICROBIOLÓGICOS

2.1. Conjugaciones

Las conjugaciones se realizaron según se describe previamente en (Llosa *et al.*, 1991). En todos los experimentos de conjugación se utilizó DH5α como cepa donadora, y UB1637 como cepa receptora.

Los cultivos se crecieron hasta fase estacionaria con los antibióticos correspondientes. Se midió la densidad óptica de los cultivos a 600nm y se añadió el mismo número de células donadoras y receptoras a un único tubo. Para eliminar el antibiótico, las células se lavaron dos veces con LB estéril, se centrifugaron y se resuspendieron en el mismo volumen de LB limpio. El pellet resultante se resuspendió en aproximadamente 50µl. Se ha descrito que tanto R388 como sus derivados conjugan mejor en sólido que en líquido (Bradley and Cohen, 1976), por lo que las conjugaciones se realizaron sobre placas de LB-Agar utilizando filtros de nitrocelulosa (GS Millipore de 0,22µm de tamaño de poro y 2mm de diámetro). Se extiende la mezcla conjugativa sobre este filtro y se mantiene durante 60 minutos a 37ºC. Pasado este tiempo, el filtro

se retira con pinzas estériles y se introduce en un tubo que contiene 2ml de LB estéril, resuspendiéndose las células por medio de vórtex. Estos dos mililitros que contienen la mezcla conjugativa constituirán la dilución cero. En las últimas etapas de la tesis, las conjugaciones se hicieron directamente sobre el LG-Agar sin utilizar filtros de nitrocelulosa en placas de 24 pocillos. En estas placas se había añade previamente un mililitro de LB-Agar en cada pocillo, el cual sirve de soporte para la mezcla de conjugación. Después de una hora a 37ºC, la conjugación se para añadiendo un ml de LB estéril al pocillo, en el que se resuspende la muestra mediante la pipeta y se hacen las correspondientes diluciones como en el protocolo utilizado tradicionalmente.

Para el cálculo de la frecuencia de conjugación, se hacen las diluciones adecuadas y se siembran en placas de LB-Agar con el antibiótico adecuado para seleccionar las células donadoras por un lado y las células transconjugantes por otro. Las frecuencias de conjugación fueron determinadas calculando el número de transconjugantes por donador. Teniendo en cuenta que las frecuencias de conjugación siguen una distribución de Poisson, se han obtenido los logaritmos en base 10 de las frecuencias y se han calculado las desviaciones estándar de los valores logarítmicos. El rango de las frecuencias se ha obtenido sumando y restando la desviación estándar a la media logarítmica y calculando a continuación el antilogaritmo de los valores resultantes.

2.2. Medición de la actividad del promotor

La actividad transcripcional de un promotor se determinó mediante el perfil de expresión de GFP siguiendo la metodología utilizada en (Zaslaver *et al.*, 2006) con algunas modificaciones introducidas en (Fernández-Lopez R, 2011). Mediante esta metodología, se mide directamente la actividad del promotor, no la concentración de ARN producido, lo que hace que esta técnica sea insensible a las diferencias en el procesamiento o la degradación del ADN.

Para la medición de la actividad, se clonaron las variantes del *oriT* sintético a estudiar que contienen el promotor de TrwA en el vector reportero pUA66 (Tabla M2), utilizando las enzimas de restricción *Pag*I, *Xho*I (Fermentas). Este vector contiene el

gen GFP-mut2 en sentido 3´ a un fuerte sitio de unión al ribosoma. Esta variante de la GFP madura rápidamente (Kalir *et al.*, 2001), es estable y no resulta tóxica para *E. coli*. Al ser tan estable, se podría considerar que la tasa de desaparición de la misma es debida únicamente a la dilución derivada del crecimiento celular, característica que se tomará en cuenta al procesar los datos de las mediciones.

Las mediciones se realizaron paralelamente en presencia y en ausencia del regulador, el cual era directamente expresado desde el plásmido R388 que se había introducido en la cepa por conjugación. Las células fueron puestas a crecer en medio M9 complementado con casaminoácidos (0,2%) y glicerol (0,5%), además del antibiótico correspondiente. Después de crecer las células o/n los cultivos se diluyeron 10.000 veces en el mismo medio en placas de 96 pocillos y se incubaron en un fluorímetro Victor3 a 37ºC con agitación orbital. Ya que es posible que las modificaciones post-traduccionales y las tasas de degradación de proteínas tengan un efecto importante sobre las concentraciones de reguladores, las mediciones se llevan a cabo en estado estacionario. En esta fase de crecimiento bacteriano, se puede asumir que la concentración media de regulador no varía.

El lector de placas es capaz de detectar la DO y la fluorescencia al mismo tiempo, y ambas se determinan para cada pocillo cada 5 minutos. Para contrarrestar la evaporación de agua del medio, el programa está diseñado para que se inyecte agua después de tres pasos de medida de absorbancia y fluorescencia. El análisis de los datos se realizó siguiendo el procedimiento utilizado en (Zaslaver *et al.*, 2006) aunque se introdujeron algunas modificaciones que se mencionan en (Fernández-Lopez R, 2011). A los datos de DO se les restaron los valores obtenidos de medio de cultivo sin células, mientras que los valores de fluorescencia se obtuvieron restando los valores observados con la cepa vacía a la misma DO. La actividad del promotor fue determinada midiendo la fluorescencia en fase estacionaria por unidad de DO.

En los casos en los que la medición se ha hecho en presencia únicamente del regulador TrwA (sin introducir el plásmido R388), se utilizó el plásmido pAR12 que lleva el gen *trwA* clonado bajo el promotor P*ara*BAD (pAR12), de manera que mediante la adición de arabinosa se induce la expresión del regulador transcripcional de manera controlada y medible. Este sistema desarrollado por Keasling y colaboradores

(Khlebnikov *et al.*, 2001) utiliza una cepa de *E. col*i (BW27783) que expresa constitutivamente el transportador de arabinosa araE. El promotor P*ara*BAD se induce añadiendo diferentes concentraciones de arabinosa al medio (figura M4). Como el transportador no se encuentra bajo el control del metabolito transportado, se evita así una variable que puede influir negativamente a la hora de registrar las mediciones. La fluorescencia por unidad de DO se obtuvo después de restarle los valores obtenidos con la cepa que contenía el vector pBAD33 vacío.

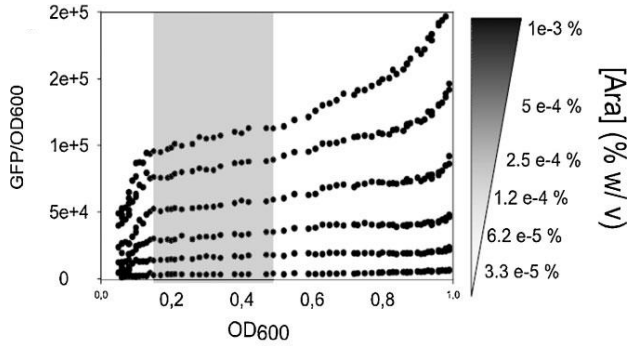


Figura M4. Perfiles de expresión de cultivos que contiene el pBAD33::GFPmut2 a distintas concentraciones de arabinosa. Cada línea de puntos corresponde a una concentración de arabinosa, que se indica en la escala triangular de la derecha. Después de varias rondas de replicación, las células alcanzan un estado estacionario (rectángulo sombreado). Tomado de (Fernandez-Lopez *et al.*, 2010).

3. METODOS BIOQUIMICOS

3.1. Purificación de proteínas

3.1.1. Sobreexpresiones

Las sobreexpresiones se realizaron en un volumen de uno o dos litros de LB divididos en dos matraces. En todos los casos se utilizó la cepa C41 de *E.coli* con un plásmido derivado del vector pET29C o pET22a con el gen que codifica para la proteína que se quiere purificar clonado en el.

La tarde anterior a la sobreexpresión, se inocula una colonia puntual de la cepa indicada en un matraz de 50ml de LB estéril en el que se dejará crecer en presencia de antibiótico durante toda la noche en agitación a 37ºC. Este matraz se utiliza como preinóculo para crecer 2l de sobreexpresión (dilución 1:40). Antes de añadir las células, añadimos el antibiótico al medio, y ponemos los matraces en agitación a 130r.p.m y 37ºC en un agitador Certomat BS-T (B. Braun Biotech International). Cuando la DO_{600} alcanza un valor de entre 0,6-0,8, añadimos 500µl de isopropil-β-D-tiogalactopiranósido (IPTG) 1M a cada litro de cultivo (C_{final}=0,5mM). Cogemos una alícuota de células en este momento (t_0), que se centrifugan y después de descartar el sobrenadante las resuspendemos en 100µl de azul de carga (Tris-HCl pH 6,8 250 mM, SDS 5 % (p/v), glicerol 50 % (p/v), azul de bromofenol 0,05 % (p/v), β-mercaptoetanol 250 mM). Dejamos que la sobreexpresión tenga lugar durante tres horas a partir de la inducción con IPTG, y después de este tiempo, recogemos las células por centrifugación, descartando el sobrenadante, y guardamos este pellet a -80ºC hasta que realicemos la purificación. El cultivo sobreexpresado se centrifuga en tubos de 400ml, 12 minutos a 5000 r.p.m y 4ºC. Antes de centrifugar las células, recogeremos otra alícuota de células (t_1) que, resuspendidas en azul de carga, correremos en un gel desnaturalizante de SDS-PAGE junto al t_0 recogido previamente para comprobar si se ha producido la sobreexpresión.

3.1.2. Purificación de TrwA y sus variantes

Tanto $TrwA_{R388}$ como $TrwA_{R7K}$ y sus derivados y mutantes utilizados en este trabajo han sido clonados en un vector de expresión de tal forma que añade una cola de histidinas en el extremo carboxilo terminal de la proteína. La purificación se ha realizado de la misma manera en todos los casos, siendo necesaria la utilización de dos pasos de purificación para obtener una fracción de proteína pura.

El pellet celular conservado a -80ºC se descongela a temperatura ambiente para acelerar la lisis de las células. Una vez descongelado el pellet, se resuspende en 50ml de tampón A (50mM Tris pH=7.5, 1M NaCl) al que se le han añadido 0,0002% del inhibidor de proteasas PMSF. La lisis se lleva a cabo por sonicación, mediante 3 ciclos de 2 minutos cada uno, dejando entre cada ciclo 5 minutos para que la muestra no se caliente en exceso. El lisado celular se centrifuga a 40000r.p.m durante 40 minutos,

tiempo en el que las membranas celulares y organelas se depositan en el fondo del tubo en forma de pellet. El sobrenadante que contiene la fracción proteica soluble de la célula se carga en una columna HisTrap de 5ml (GE Healthcare) previamente equilibrada con tampón A. Las proteínas unidas a la columna se eluyen con un gradiente de tampón B (50mM Tris pH=7.5, 1M NaCl, 0,5M Imidazol) utilizando un sistema AKTA (GE). Tras la elución, las fracciones que contienen la proteína que nos interesa son concentradas mediante centrifugación a 4000r.p.m a 4ºC durante el tiempo necesario utilizando un dispositivo Amicon Ultra Centrifugal (Millipore). A continuación se centrifugó la muestra durante 5 minutos a 13.000r.p.m para eliminar agregados o precipitados y se cargó en una columna de filtración en gel Superdex 75 HR 10/30 (Pharmacia) que previamente ha sido equilibrada con tampón C (50mM Tris pH=7.5, 150mM NaCl, 1mM EDTA). La proteína fue eluida utilizando el mismo tampón C, y las fracciones que contenían la proteína de interés se mezclaron y cuantificaron y se almacenaron en alícuotas a -80ºC para su uso posterior.

3.2. Cuantificación de proteínas

Para medir la concentración de proteínas en solución, se utilizan técnicas espectroscópicas. En este trabajo se ha utilizado principalmente el método de Bradford.

Al tratarse de un método rápido y sencillo, esta técnica (Bradford, 1976) es de uso generalizado en los laboratorios de bioquímica. Para medir la concentración de proteínas se utiliza el reactivo Protein Assay (BioRad), que contiene Azul Brillante Coomassie G-250 en ácido fosfórico y metanol. El ensayo se basa en que el agente cambia de color al unirse a las proteínas. Bajo condiciones ácidas el colorante Coomassie es más estable en la forma doblemente protonada, de color rojo. Sin embargo, al unirse a las proteínas, las interacciones hidrofóbicas e iónicas estabilizan la forma aniónica del Coomassie, que es azul. De esta forma, se produce un cambio de color visible que es proporcional a la cantidad de proteína presenten la muestra y que puede ser fácilmente cuantificado por colorimetría midiendo la absorbancia a 595 nm.

Para llevarlo a cabo, se hace una dilución 1:5 del reactivo y se añaden 800µl de esta dilución a 200µl de cada una de las muestras. Para la recta patrón se preparan 6 muestras de 200µl cada una con concentraciones crecientes de seroalbúmina bobina (BSA, Sigma), de 0 a 9µg de proteína. Se mezcla y se dejan pasar 10 minutos en oscuridad antes de medir la A_{595}. Al mismo tiempo que se han preparado las muestras de la recta, se preparan también las muestras problema. Las concentraciones se han calculado como la media de 3 muestras.

3.3. Electroforesis de proteínas

Para la visualización de proteínas, estas se separaron electroforéticamente en geles de SDS-poliacrilamida a distintos porcentajes de entre el 10 y el 15% dependiendo del tamaño de la proteína. Los geles se prepararon con acrilamida:bisacrilamida 29:1 de BioRad. Las muestras se mezclaron con 1/5 de su volumen del tampón de carga (SAB 5X: Tris-HCl pH 6,8 250 mM, SDS 5 % (p/v), glicerol 50 % (p/v), azul de bromofenol 0,05 % (p/v), DTT 250 mM) y se calentaroN durante 2min a 95ºC. Se realizó la electroforesis en un tampón Tris 25mM, glicina 192mM, SDS 0,1% (p/v). Una vez realizada la electroforesis, los geles se tiñeron con una solución de Azul Brillante Coomassie R250 0.1%(p/v), metanol 40% (v/v) y ácido acético glacial al 10% (v/v) durante 15 min con agitación a temperatura ambiente. Una vez teñidos, los geles se destiñeron en una solución que contenía metanol al 40%(v/v) y ácido acético glacial al 10%(v/v), nuevamente a temperatura ambiente y con agitación. Las imágenes de los geles fueron tomadas con un escáner convencional.

3.4. Ensayos de unión ADN-proteína

3.4.1. Marcaje de oligonucleótidos en 5'

Para el marcaje radiactivo de oligonucleótidos de cadena sencilla, en un volumen final de 20µl se incubaron 20pmol de oligonucleótido, 3µl de [γ-^{32}P]-ATP (10µCi/µl) y 10U de T4 polinucleótido kinasa (Fermentas) en el tampón provisto para la

enzima. Tras 45 min a 37ºC la mezcla se llevó a 70ºC para inactivar el enzima y se eliminó el exceso de ATP no incorporado mediante la utilización de columnas Micro Bio-Spin 6 Chromatography (BioRad).

Se utilizó el mismo protocolo para marcar radiactivamente los fragmentos de doble cadena obtenidos por PCR.

3.4.2. Ensayo de retardo en gel

Todos los retardos en gel mostrados en este trabajo se realizaron con ADN de cadena doble obtenido por PCR o por hibridación de oligonucleotidos comerciales. Para la obtención de fragmentos de ADN de doble cadena por hibridación de oligonucleótidos comerciales, el oligonucleótido marcado radiactivamente se mezcló con el oligonucleótido complementario frío a concentraciones iguales, se calentaron a 95ºC durante 5 min y la mezcla se dejó enfriar lentamente hasta alcanzar la temperatura ambiente. Una vez hibridados, los fragmentos se conservaron a 4ºC o a -20ºC si no iban a ser utilizados por un tiempo.

Aproximadamente 2fmol del fragmento de doble cadena se incubaron en un volumen final de 20µl con cantidades crecientes de TrwA en tampón Tris 50mM (pH=7.5), NaCl 50mM, glicerol 5%, EDTA 0,1mM, DTT 1mM. En todos los casos se añadió ADN de esperma de salmón a una concentración final de 4µg/ml para evitar uniones no específicas. Tras incubar 10 min a temperatura ambiente, se añadieron 4µl de azul de carga que se encontraba a 5x y se cargaron 15µl en un gel de acrilamida de porcentaje variable entre el 5% y el 8% dependiendo del tamaño del fragmento de ADN. La electroforesis se desarrolló a 100V durante 45 min en tampón TBE. Después de la electroforesis, los geles se secaron media hora a 80ºC y la radioactividad se cuantificó en un sistema FLA-5100 Fluorescent Image Analyzing System (Fujifilm).

En los casos en los que se realizó retardo con ADN competidor frío, el protocolo fue el mismo, pero se añadieron cantidades crecientes de ADN competidor en las siguientes proporciones: 100x, 1000x, 2500x, 7500x, 15000x, 25000x.

Mediante la figura M5 se pretende ilustrar la apariencia que tendrían los dos tipos de ensayos de retardo en gel.

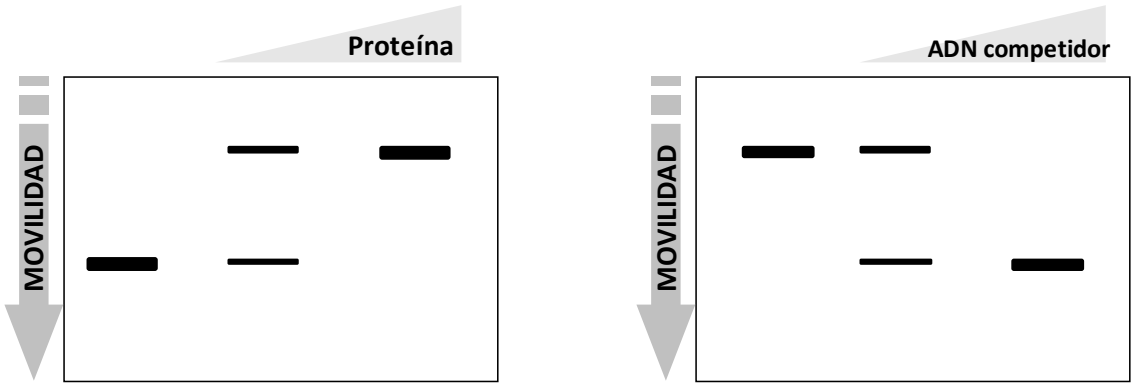


Figura M5. Representación esquemática del principio de ensayo de movilidad en gel para el análisis de unión de la proteína al ADN. En el primer dibujo por la izquierda se representa lo que sería un retardo normal en la que se aprecia un descenso de movilidad en la banda de ADN a mayor concentración de proteína. En el segundo, se representa lo que se esperaría obtener en un retardo con ADN competidor, en el que la movilidad crece a medida que el ADN frío compite para unirse a la proteína.

En los experimentos de retardo que se realizaron con mutantes del origen de transferencia, el ADN se detectó mediante tinción con SYBR Green I (Sigma) en vez de con fósforo radioactivo. El ensayo se llevó a cabo de la misma manera pero con ambos oligonucleótidos a hibridar sin marcar. Después de la electroforesis, los geles se tiñeron durante 10 minutos en agitación en un recipiente que contenía una concentración de 1X de SYBR Green disuelto en TBE. Después de la tinción, se sustituyó la solución de SYBR Green por agua y se mantuvo en agitación durante 15 minutos más. Los geles se visualizaron en el sistema FLA-5100 (Fujifilm) utilizando un láser con excitación a 473 nanómetros.

La constante de disociación (Kd) es una constante de equilibrio termodinámico de la unión de un ligando a un receptor. En el caso de la unión proteína-ADN, esta constante se puede expresar como la concentración que permite el 50% de la unión máxima cuando la concentración del fragmento de ADN utilizado es despreciable frente

a la concentración de proteína. Así, gracias a un ensayo EMSA, el cálculo de la constante de disociación se realiza mediante la medición de la intensidad de ADN no unido a la proteína en condiciones de concentración creciente de proteína.

3.5. Cristalización de proteínas

Para obtener los complejos proteína-ADN para su cristalización, se mezcló la proteína purificada TrwA con la hibridación de oligonucleótidos que conformaban el sitio *sbaA* o el *sbaB* en una relación proteína:ADN 1:1,3. Para optimizar la formación de cristales, se utilizaron oligonucleótidos de distintas longitudes con o sin colas en los extremos.

Las pruebas de cristalización se llevaron a cabo por el método de difusión de vapor en gota sentada almacenando las placas a 22ºC y a 4ºC. Los cristales se obtuvieron a una concentración de proteína de 10mg/ml, en las siguientes condiciones:

$TrwA_{R388}$_*sbaA*: 0,15M $(NH_4)_2SO_4$, MES pH=5.5, glicerol 10%

$TrwA_{R7K}$_*sbaA*: 0,2M $(NH_4)_2SO_4$, MES pH=5.5, PEG 2K 5%

TrwAR7K_*sbaB*: $MgCl_2$ 0.2M, HEPES pH=7.5, PEG 400 20%

Los experimentos de difracción de los cristales obtenidos se llevaron a cabo en la línea BM16 del sincrotrón ESRF de Grenoble (Francia).

3.6 Estudio de interacción de proteínas

El estudio de la interacción de proteínas se realizó en un principio partiendo de proteínas purificadas, y posteriormente se llevó a cabo una coexpresión *in vivo* de las dos proteínas a analizar para comprobar si se repetían los resultados obtenidos *in vitro*.

3.6.1. Interacción en columna con proteínas purificadas.

En todos los experimentos realizados tanto con proteínas enteras como con los distintos fragmentos, se calculó su concentración y se mezcló una de las proteínas a analizar con cola de histidinas con la otra proteína sin cola de histidinas en una proporción 1:1 en un tampón que contenía 100mM de NaCl, 50mM de Tris a pH=7,2, 70mM de imidazol y 10% de glicerol. La mezcla de las dos proteínas se incubó a temperatura ambiente durante 10 minutos antes de incorporarla a una columna Mini Bio-Spin Chromatography de Bio-Rad a la que previamente se le habían añadido 100µl de una resina Ni-NTA agarose de Qiagen que se equilibró con el mismo tampón en el que se encuentran las proteínas. Una vez se cargó la mezcla que contenía las proteínas a analizar, esperamos 5 minutos antes de comenzar una sucesión de lavados de un volumen de columna cada uno con el tampón mencionado anteriormente para eliminar las proteínas no unidas a la resina. Se realizaron 10 lavados, de los cuáles se tomaron alícuotas de los cinco primeros lavados y del último para su posterior análisis. Una vez realizados los lavados, procedimos a la elución de las proteínas aún unidas a la resina utilizando un tampón con la misma composición que el anterior pero esta vez con 500mM de imidazol. Las fracciones derivadas de la elución, junto con las correspondientes a los distintos lavados y a la mezcla inicial que se cargó en la columna, se mezclaron con tampón de carga para su análisis en geles SDS-PAGE de distinto porcentaje según el tamaño de las proteínas a analizar.

3.6.2. Co-purificación TrwA-TrwC.

Se transformaron al mismo tiempo en la cepa C41 (DE3) los plásmidos para la sobreexpresión de TrwA (pYTP27, Km^R) y TrwC (pSU1621, Ap^R), asegurándonos la presencia de ambos plásmidos al mantener la presión selectiva de los antibióticos Km y Ap en todo momento. Se probaron distintos tiempos de sobreexpresión para elegir el tiempo óptimo en el que se obtuviera un buen nivel de expresión de ambas proteínas, y se eligieron 3 horas de sobreexpresión. La inducción y sobreexpresión se llevó a cabo como si de una sola proteína se tratase, siguiendo el protocolo mencionado en el apartado 2.3.1.1 hasta la recogida de las células.

Los primeros pasos previos a la purificación se realizaron como se detalla en 2.3.1.3, variando únicamente el tampón de lisis, que contiene esta vez 500mM de NaCl en vez de 1M. Posteriormente a la lisis por sonicación, se introdujeron algunos cambios en el protocolo en lo que concierne a la composición de los tampones y la carrera utilizada en el HPLC para la elución de las proteínas.

El lisado celular se diluyó hasta una concentración final de NaCl de 100mM y 70mM de imidazol. La mezcla resultante se cargó a un flujo de 3ml/min en una columna de afinidad Ni-NTA de 5ml (GE Healthcare) previamente equilibrada con tampón A (100mM NaCl, 50mM Tris pH=7,5, 70mM Imidazol). A continuación se lava la columna con el mismo tampón el que tenemos nuestro lisado con al menos 10 volúmenes de columna, para eliminar el exceso de proteínas no adheridas a la misma. Conectamos la columna al sistema FPLC y seguimos pasando tampón A hasta que obtenemos una línea base completamente plana, lo que nos indica que todas las proteínas unidas no específicamente a la columna han sido ya eliminadas. En este momento, aplicamos un gradiente rápido de un volumen de columna con el tampón B (100mM NaCl, 50mM Tris pH=7,5, 0,5M Imidazol), de tal manera que las proteínas antes unidas a la columna quedarán eluidas en las fracciones de un mililitro que recogemos.

4. MÉTODOS *IN SILICO*

4.1 Modelado unión TrwA-ADN

El modelado de la estructura terciaria de TrwA unida al ADN se llevó a cabo en varios pasos. En primer lugar, se lanzó la secuencia al servidor Phyre2 (Kelley and Sternberg, 2009) para obtener la homología estructural con proteínas ya resueltas. El servidor nos devolvió la homología con varias proteínas de las que se seleccionaron las tres primeras para su análisis más detallado. Estos modelos corresponden a una subunidad de la proteína libre sin ADN, por lo que se llevó a cabo un alineamiento con la proteína Arc (paradigma de las proteínas RHH) unida a ADN para poder obtener una

visión aproximada de la unión de TrwA al ADN. Estos alineamientos se realizaron mediante el programa Matras (Kawabata, 2003), que facilita el alineamiento estructural de dos cadenas. Cada una de las cuatro cadenas de TrwA (una por cada monómero) se alineó con las cuatro cadenas correspondientes de Arc, obteniendo cuatro alineamientos que se unieron en una sola molécula para obtener el tetrámero de TrwA. La conversión a imagen, la visualización y la modificación de detalles se llevó a cabo mediante Pymol (Pymol Molecular Graphics System, Versión 1.2r3pre, Schrödinger, LLC).

4.

RESULTADOS

4. RESULTADOS

1. LA DELECIÓN DE LOS ELEMENTOS DE ESTRUCTURA SECUNDARIA DEL *oriT* RESTAN EFICIENCIA DE MOVILIZACIÓN

Los orígenes de transferencia contienen elementos de estructura secundaria, que suelen albergar sitios de unión para proteínas del relaxosoma que participan activamente en el procesamiento conjugativo del ADN o factores de transcripción que, aún no participando en este complejo nucleoprotéico también están implicados en la conjugación. El *oriT* de R388 contiene varios sitios de unión a proteínas ya descritos como los sitios *sbaAB* de unión de TrwA, el sitio de unión de la relaxasa (IR2) o la repetición IR3 que contiene en su secuencia un sitio de unión no consenso para IHF (Moncalian *et al.,* 1999b). Además se encuentran en el *oriT* de R388 otras repeticiones que se cree que juegan un papel estructural en el mantenimiento del *oriT* como puede ser la repetición IR1. En la figura I6 del apartado de introducción se muestran todas estas repeticiones en el *oriT* de R388.

Para estudiar estudiar la importancia de cada módulo funcional en el proceso de la conjugación quisimos, en primer lugar, estudiar el comportamiento del origen de transferencia en ausencia de ellos. La finalidad sería eliminar los módulos, o las regiones entre ellos, que causaran un menor efecto en la conjugación para conseguir un origen de transferencia de menor tamaño que el actual. Reduciendo el número de proteínas implicadas en la conjugación, queríamos obtener un plásmido que conjugara a una frecuencia mayor que R388, y en el caso de que no fuera así, aún y todo dispondríamos de un *oriT* más fácil de manipular que el original.

Se había realizado con anterioridad un estudio detallado del efecto que tenía la eliminación de fragmentos del *oriT* (Llosa *et al.,* 1991), en el que se procedió a eliminar los diferentes fragmentos mediante la acción de la DNAsa I partiendo de un extremo del

oriT. Este método no permitía la deleción de cada módulo de manera única, sino que toda la secuencia hasta llegar hasta ese modulo era también eliminada. Nosotros, en cambio, queríamos disponer de una herramienta que nos permitiera eliminar cada elemento de estructura secundaria de manera independiente, sin modificar la secuencia que lo rodea. Para ello, tomando como molde el *oriT* salvaje de R388, diseñamos un *oriT* (al que denominamos $oriT_{R388}syn$) en el que habíamos insertado sitios de restricción flanqueando cada módulo de estructura secundaria (figura R1), introduciendo para ello el menor número posible de mutaciones simples en la secuencia original y que fue sintetizado por la compañía alemana Geneart (actualmente Life Technologies). El esquema del $oriT_{R388}syn$ diseñado se muestra en la figura R1.

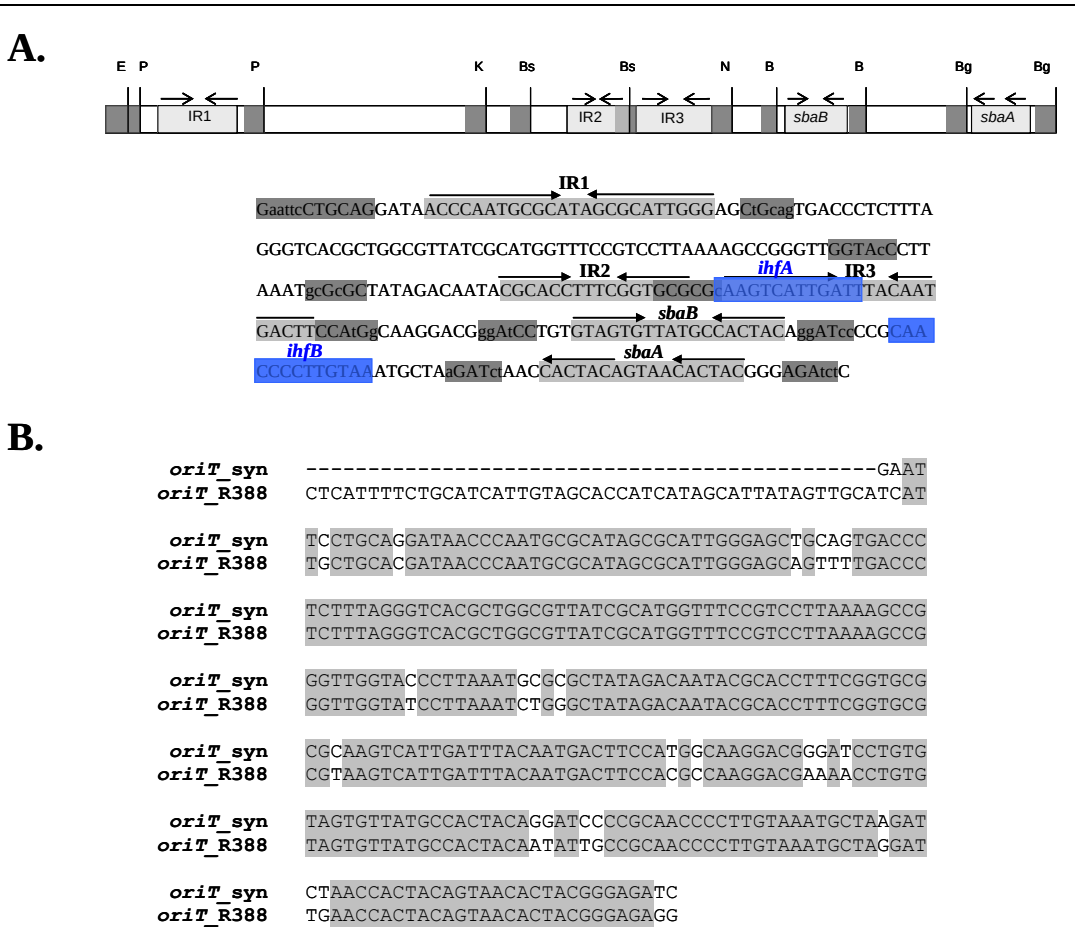


Figura R1. A. Representación esquemática y secuencia nucleotídica del origen de transferencia sintético $oriT_{R388}syn$ con los elementos de estructura secundaria representados como rectángulos gris claro, y los sitios de restricción como bloques en gris más oscuro. Las letras mostradas en la parte superior indican el sitio de restricción insertado (E; *Eco*RI, P; *Pst*I, K; *Kpn*I, Bs; *Bss*HII, N; *Nco*I, B; *Bam*HI, Bg; *Bgl*II). **B.** Alineamiento de los orígenes de transferencia nativo y sintético de R388 en el que se observan las mutaciones introducidas (en blanco).

Al encontrarse cada módulo delimitado por dos sitios de reconocimiento para una enzima de restricción, mediante una sencilla reacción de digestión seguida de una religación somos capaces de eliminar cada uno de ellos independientemente del resto de la secuencia adyacente. Es más, no solo podemos eliminar estos elementos, sino que podemos insertar en su lugar cualquier otra secuencia de interés siempre que mantengamos los sitios de restricción en los extremos. Por lo tanto, mediante la digestión, y posterior religación del *oriT*, obtenemos una serie de mutantes que carecerán del sitio elegido, lo que nos permite el estudio de su papel en el proceso de la conjugación.

Este *oriT*$_{R388}$*syn* fue clonado en el vector pSU19 dando lugar al plásmido pYTP4 y derivados (tabla M3, Materiales y Métodos) con los que estudiamos en detalle los efectos de las diferentes repeticiones en la frecuencia de conjugación. En primer lugar realizamos una serie de ensayos de movilización para ver si el *oriT*$_{R388}$*syn* era funcional, utilizando el plásmido pSU2007 como plásmido movilizador (los detalles técnicos aparecen en el apartado 2.1 de Materiales y Métodos). Como vemos en la tabla R1 la frecuencia de movilización de un plásmido conteniendo este *oriT*$_{R388}$*syn* es algo ligeramente menor (0,12 vs 0,24) que la movilización de ese mismo vector conteniendo el *oriT*$_{R388}$ nativo.

Una vez comprobado que el *oriT*$_{R388}$*syn* era funcional eliminamos los diferentes elementos de estructura secundaria. Los resultados obtenidos en estos experimentos se muestran en la tabla R1.

Plásmido	**Características**	**Frecuencia de conjugación**
oriT salvaje	*oriT* de R388	0,24 (0,1-0,55)
pYTP4	*oriTR388syn*	0,12 ($8,4x10^{-2}$-$1,73x10^{-1}$)
pYTP5	*oriTR388syn* ΔIR1	$2,11x10^{-3}$ ($1,01x10^{-3}$-$4,4x10^{-3}$)
pYTP7	*oriTR388syn* ΔIR2	0
pYTP8	*oriTR388syn* ΔIR3	$4,09x10^{-6}$ ($8,96x10^{-7}$-$1,87x10^{-5}$)
pYTP6	*oriTR388syn* que carece de toda la secuencia "upstream" hasta el IR2	$9,62x10^{-4}$ ($8,45x10^{-5}$-$1,1x10^{-2}$)
pYTP16	*oriTR388syn* con la región nic-IR2 invertida	$1x10^{-6}$($3x10^{-7}$-$1,3x10^{-5}$)

Tabla R1. Frecuencias de conjugación de los distintos mutantes de los elementos de estructura secundaria del *oriT*. Las frecuencias se calcularon dividiendo el número de transconjugantes por el número de donadores. Se muestra la media de al menos tres experimentos. Los números entre paréntesis corresponden al margen de error obtenido.

Todos los mutantes que carecían de alguno de los módulos funcionales mostraban una frecuencia de movilización menor que la del *oriT* salvaje, pero las frecuencias variaban según el módulo eliminado. Cuando se midió la movilización del mutante sin el sitio de unión de la relaxasa (pYTP7), no se pudo detectar ningún transconjugante en las múltiples veces que se repitió el experimento.

Partiendo de este mutante, construimos un clon (pYTP6) en el que introdujimos la zona que contiene el IR2 más el *nic* de manera que se insertara en una orientación contraria a la que se encuentra normalmente (figura R2). De este modo, el sitio *nic* quedaría situado en la hebra de arriba en dirección 3' respecto al sitio de unión de la relaxasa. Aunque fuimos capaces de observar algún transconjugante, la frecuencia era muy baja, indicándonos que la orientación del sitio *nic* no permite ser alterada para una correcta conjugación (tabla R1).

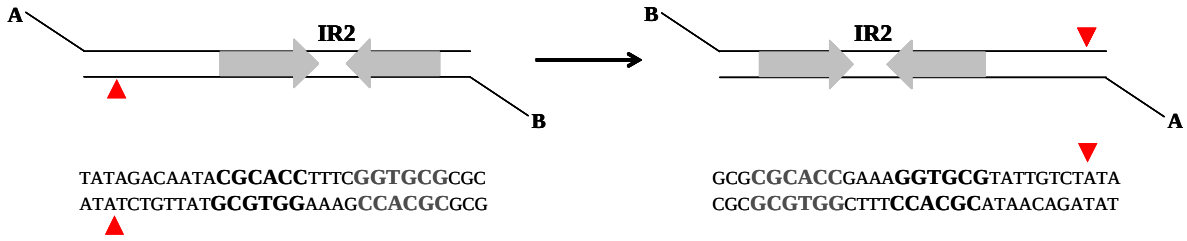


Figura R2. Representación esquemática de la obtención del clon pYTP16. El triángulo rojo representa el lugar donde se da el corte por la relaxasa, y la secuencia que conforma el brazo proximal y distal de la repetición se muestra en caracteres negros y grises respectivamente.

La eliminación de la IR3 también tiene un efecto bastante importante en la conjugación, ya que disminuye su frecuencia unas 10^5 veces. De todas formas, este dato no nos aporta la información suficiente como para afirmar que la eliminación del sitio *ihfB* es el responsable de este efecto, ya que en este caso también hemos alterado el espaciamiento entre los sitios de unión de la relaxasa y TrwA. Se probó la capacidad de movilización de dos mutantes más, uno que carece de la IR1, y otro en el cual se ha eliminado toda la secuencia que va desde el comienzo del origen de transferencia (situando éste después de StbA) hasta llegar al sitio de unión de la relaxasa.

El elemento IR1 es una repetición invertida formada por dos brazos de 9 nucleótidos separados por un espaciamiento de 3 nucleótidos. Esta repetición está situada al comienzo del *oriT*, adyacente a los iterones que conforman el sitio de unión de StbA. No se ha descrito ninguna proteína que reconozca esta repetición, así que se cree que juega un papel estructural ya que es la repetición que más tendencia muestra por formar cruciformes como se ha demostrado por *footprinting* con $KMnO_4$ (Moncalián, 2000). Esta repetición se conserva, tanto en posición como en secuencia, en los *oriT* de los IncW estudiados como pueden ser R388, R7K, pSa, pIE321 y pIE522. En los plásmidos recientemente descritos y clasificados como IncW como pMAK3, y en otros similares a IncW como pMBUI4 también está presente el IR1, aunque en pMBU14 la secuencia es diferente, y muestra además, un espaciamiento entre los dos brazos de cinco nucleótidos en vez de tres como en el resto de plásmidos (figura R3).

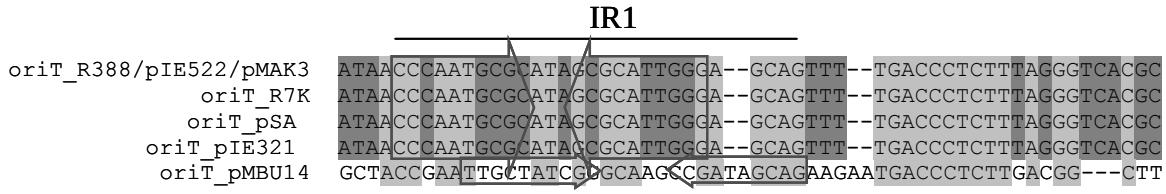


Figura R3. Alineamiento de las secuencias que abarcan el IR1 de los distintos plásmidos IncW.

La diferencia de movilización entre el mutante ΔIR1 (pYTP5) y el que carece de la secuencia que abarca desde el principio del *oriT* hasta IR2 (pYTP6) no es muy pronunciada, siendo aproximadamente tres veces menor en el pYTP6 aun siendo el fragmento eliminado mucho mayor.

2. LOS PLÁSMIDOS R388 Y R7K SE DIFERENCIAN CASI EXCLUSIVAMENTE EN EL GEN *trwA* Y LOS SITIOS DE UNIÓN DE LA PROTEÍNA EN EL *oriT*

A la hora de definir un *oriT* mínimo es interesante observar las diferencias entre *oriT*s de plásmidos cercanos evolutivamente. En ese sentido, hemos comparado los plásmidos R388 y R7K. Ambos plásmidos pertenecen ambos al grupo de incompatibilidad IncW, que engloba plásmidos de menos de 40kb que muestran un amplio rango de hospedador (Fernandez-Lopez *et al.*, 2006). Este grupo incluye tres miembros "clásicos" que son pSA (Watanabe *et al.*, 1968), R388 (Datta and Hedges, 1972) y R7K (Coetzee *et al.*, 1972), y distintos estudios han demostrado la alta similitud de secuencia que comparten (Avila and de la Cruz, 1988; Ireland, 1983; Revilla *et al.*, 2008; Valentine, 1985; Ward and Grinsted, 1982). También se engloban en este grupo otros plásmidos como pIE321 y pIE522 (Schreiter and Drennan, 2007) cuya homología respecto a R388 se comparó en un trabajo realizado por nuestro grupo de investigación. pIE321 y R7K muestran una menor homología con R388 (97% y 97,5% de identidad de nucleótidos) en comparación con pSA y piE522 (casi 100% de identidad) (Revilla *et al.*, 2008). La homología decrece cuando se comparan cuatro genes individuales, de los cuáles el que codifica para TrwA es el que muestra una mayor divergencia (tabla R2).

	Identidad de secuencia					
	R7K vs R388		**pIE321 vs R388**		**pIE321 vs R7K**	
	nt	***aa***	***nt***	***aa***	***nt***	***aa***
trwA	88.2	84.3	87.3	87.6	80.4	73.6
stbC	100	100	92.2	93	92.2	93
trwK	91.1	96.8	95.7	98.3	94.7	98.5
trwM	90.1	92.3	100	100	90.1	92.3

Tabla R2. Porcentaje de identidad nucleotídica (en negro) y aminoacídica (en gris) de los cuatro genes que más divergen en los plásmidos IncW analizados Modificado de (Revilla *et al.*, 2008).

Esto nos indicaría, que *trwA* ha sufrido una tasa de sustitución mayor que otros genes que forman el eje central genético de estos plásmidos. Curiosamente, el *oriT* de estos plásmidos también muestra diferencias, que se centran básicamente en los sitios de reconocimiento de IHF y TrwA. El sitio de reconocimiento y corte de la relaxasa, en cambio, permanece inalterado en todos ellos. El sitio de unión *ihfB* localizado entre los dos sitios de unión de TrwA en R388 también varía en R7K, no presentando en este caso la secuencia consenso de unión de IHF (material suplementario, figura S3).

Respecto a las proteínas TrwA codificadas por ambos plásmidos, muestran características similares en cuanto a peso molecular, punto isoeléctrico o predicción de estructura secundaria (tabla R3).

	$TrwA_{R388}$	**$TrwA_{R7K}$**
Número de aminoácidos	121	121
Peso molecular	13400.3	13473.3
Punto isoeléctrico	5.14	5.28
Predicción estructura 2^{aria}	RHH	RHH

Tabla R3. Comparación de las características principales de las proteínas auxiliares de R388 y R7K.

A nivel de secuencia aminoacídica, la homología entre $TrwA_{R388}$ y $TrwA_{R7K}$ es del 84%, siendo la parte N-terminal de la proteína la que difiere en mayor medida. Es en esta región N-terminal donde se encuentra el dominio RHH de unión a ADN (Moncalian and de la Cruz, 2004), y aunque hay aminoácidos como Arg10 que se mantienen en las dos proteínas, otros aminoácidos que se cree que están implicados en la unión específica al ADN como pueden ser los residuos Ser12 y Asp5 no se conservan (figura R4).

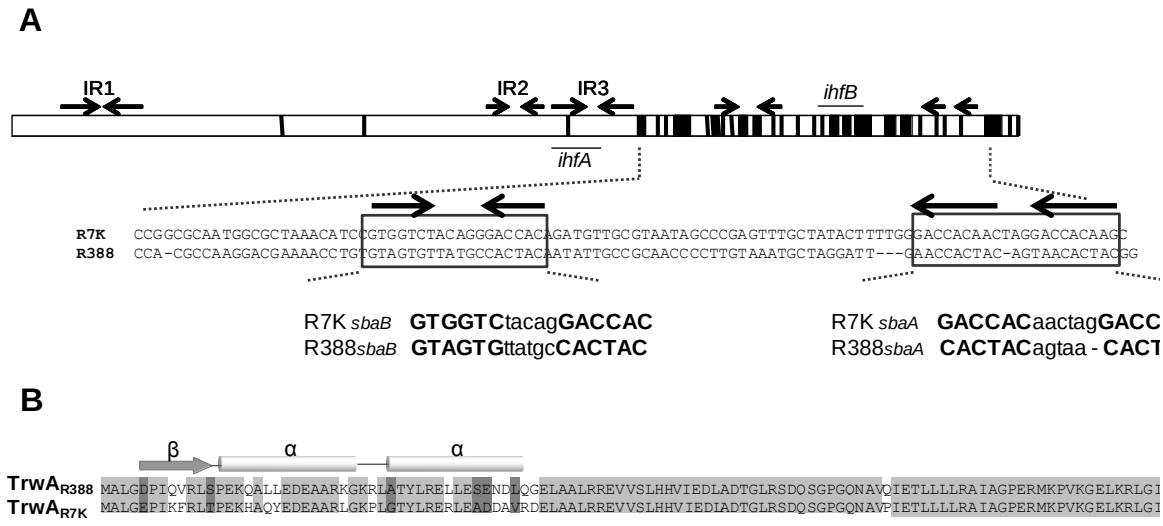


Figura R4. Comparaciones de los orígenes de transferencia y proteínas TrwA de los plásmidos R388 y R7K. **A.** Representación de la comparación de los *oriT* de R388 y R7K. En blanco se encuentran las zonas de homología, y mediante barras negras se representan las zonas de divergencia. Los sitios de unión de TrwA, de IHF y de la relaxasa (IR2), así como las demás repeticiones que aparecen en el *oriT* están señalados. En la parte inferior del esquema aparece la secuencia de los dos *oriT*, en los que se ha aumentado la zona correspondiente a los dos sitios de unión de *trwA* (*sbaA* y *sbaB*). **B.** Comparación de la secuencia primaria de la proteína TrwA codificada por cada uno de los dos plásmidos. La estructura secundaria del dominio RHH de unión a DNA se muestra en la parte superior de la secuencia.

3. CADA TrwA RECONOCE ESPECÍFICAMENTE SUS SITIOS DE UNIÓN

Se había estudiado previamente en nuestro grupo de investigación la unión de $TrwA_{R388}$ al *oriT* de R388 (Moncalian *et al.*, 1997). Se observó mediante análisis de *footprinting* y de retardo en gel que TrwA reconocía dos sitios dentro del *oriT* de R388. Estos sitios han sido nombrados como *sbaA* y *sbaB*. El elemento común es un

hexanucleótido (CACTAC), que se repite de forma directa para formar el sitio *sbaA*, y de forma invertida para formar el sitio *sbaB*. Sería lógico pensar que este hexanucleótido sería el elemento primario que reconoce la proteína en forma de tetrámero, de forma que cada uno de los dímeros que lo componen reconoce un brazo de cada repetición en cada uno de los dos sitios identificados (figura R5).

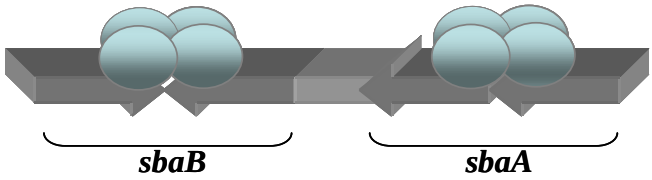


Figura R5. Ilustración que representa la imagen de cómo se produciría la unión del tetrámero de TrwA en cada sitio de reconocimiento.

Ya se ha señalado previamente que el origen de transferencia de R7K difiere principalmente del de R388 en la región que abarca los sitios de unión de IHF y de TrwA. Al observar en más detalle esta región, encontramos que la posible región de unión de TrwA en R7K también aparece una hexanucleótido repetido de forma tanto directa como invertida. Este hexanucleótido es diferente en su secuencia al encontrado en R388 siendo GACCAC en R7K y CACTAC en R388 (figura R4). Existe otra diferencia en el espaciamiento de los dos brazos que conforman la repetición directa (sitio *sbaA*); en R388 hay cinco nucleótidos separando la repetición, mientras que en R7K aparecen seis nucleótidos. En el caso de la repetición invertida, en ambos plásmidos el espaciamiento es de seis nucleótidos.

Decidimos comprobar si estos sitios encontrados, en efecto, podían ser los sitios de unión de TrwA de R7K, por lo que procedimos al clonaje del gen *trwA* de R7K en un vector de expresión y a su sobreexpresión y purificación de la proteína (Materiales y Métodos). La proteína $TrwA_{R7K}$ se purificó siguiendo el mismo protocolo que para $TrwA_{R388}$, ambas con colas de histidinas. Tras su purificación comprobamos por cromatografía de gel filtración que $TrwA_{R7K}$ eluía en la fracción correspondiente al tamaño que tendría su forma tetramérica, tal y como ocurre con $TrwA_{R388}$ (dato no mostrado).

Por hibridación de oligonucleótidos comerciales (Sigma Aldrich, St.Louis,MI), generamos los sitios ya descritos *sbaA* y *sbaB* de R388, y los posibles sitios de reconocimiento de $TrwA_{R7K}$ en R7K. Utilizando estos fragmentos, se realizaron ensayos de retardo en gel de cada sitio de unión, a los que se les añadió tanto la proteína específica como la no específica (perteneciente al otro plásmido) para comprobar en qué casos y condiciones se llevaba a cabo el reconocimiento (figura R6).

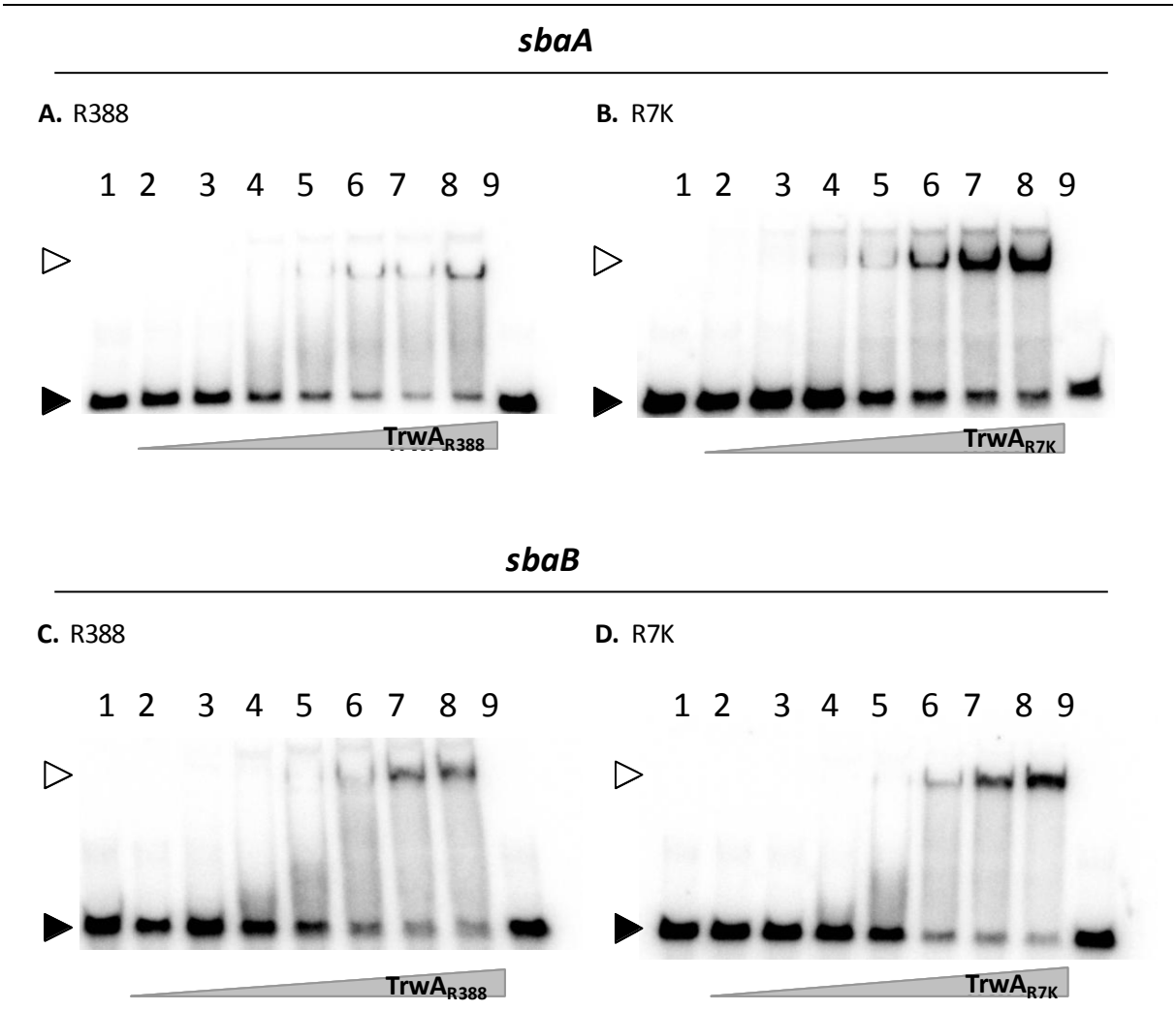


Figura R6. Ensayos de retardo en gel de TrwA de R388 y R7K con sus respectivos sitios *sbaA* y *sbaB*. En las figuras superiores se muestran los retardos realizados con el sitio *sbaA*, y en los inferiores con *sbaB*. Las figuras **A.** y **C.** corresponden a los retardos con $TrwA_{R388}$, y **B.** y **D.** a los retardos con $TrwA_{R7K}$. En todos los casos, en la calle numerada como 9 se ha añadido la proteína TrwA no específica para el plásmido. En **A.** y **B.**, los números que se muestran en la parte superior corresponden a las siguientes concentraciones de proteína: 1; 0nM, 2; 23nM, 3; 46nM, 4; 92nM, 5; 185nM, 6; 370nM, 7; 740nM, 8; 1,1µM, 9; 740nM. En C. y D., las concentraciones ensayadas son las siguientes: 1; 0nM, 2; 46nM, 3; 92nM, 4; 185nM, 5; 370nM, 6; 740nM, 7; 1,1µM, 8; 1,5µM, 9; 1,1µM.

En la parte superior de la figura R6 se muestran los experimentos realizados con los sitios *sbaA* de R388 y de R7K en presencia de cantidades crecientes de la proteína específica para cada uno de ellos (A y B respectivamente). En todos los experimentos, se muestra una calle extra (calle 9) en la que se añadió la proteína no específica para el sitio de unión ensayado, es decir, en el caso de la figura 6A., la proteína que se añadió a la mezcla que contenía el sitio $sbaA_{R388}$ fue $TrwA_{R7K}$ y en la figura 6B., la calle 9 nos muestra el resultado de la adición de $TrwA_{R388}$ al fragmento $sbaA_{R7K}$. En ambas figuras, podemos observar cómo alrededor de la calle 4 comienza a aparecer una banda que muestra un fragmento de ADN con una menor velocidad de migración, causada por la unión estable de la proteína al mismo. Según se va aumentando la cantidad de proteína añadida, la banda retardada es más intensa, ya que la mayoría del ADN pasa a estar unido a la proteína. En ninguno de los dos casos se aprecia una banda retardada cuando la proteína que se añade es la no específica (calle 9), lo que nos indica que la proteína solamente es capaz de reconocer el sitio de unión propio aún añadiendo altas concentraciones de la misma.

En las figuras R6C y R6D, se muestra un experimento similar llevado a cabo utilizando como sustrato los sitios $sbaB_{R388}$ en el primer caso, y $sbaB_{R7K}$ en el segundo. Tampoco en estos ensayos se observa un retardo en la calle correspondiente a la proteína no específica, mostrándonos que la proteína, también en este caso, solo es capaz de distinguir el sitio de reconocimiento propio. Aunque en el experimento anterior ya se observaba una banda retardada a una concentración de 92nM de TrwA, aquí no es hasta la calle 5 (370nM TrwA) que comenzamos a distinguirla, indicándonos que se necesita mayor cantidad de proteína para que se de el reconocimiento del sitio *sbaB* tanto en R388 como en R7K.

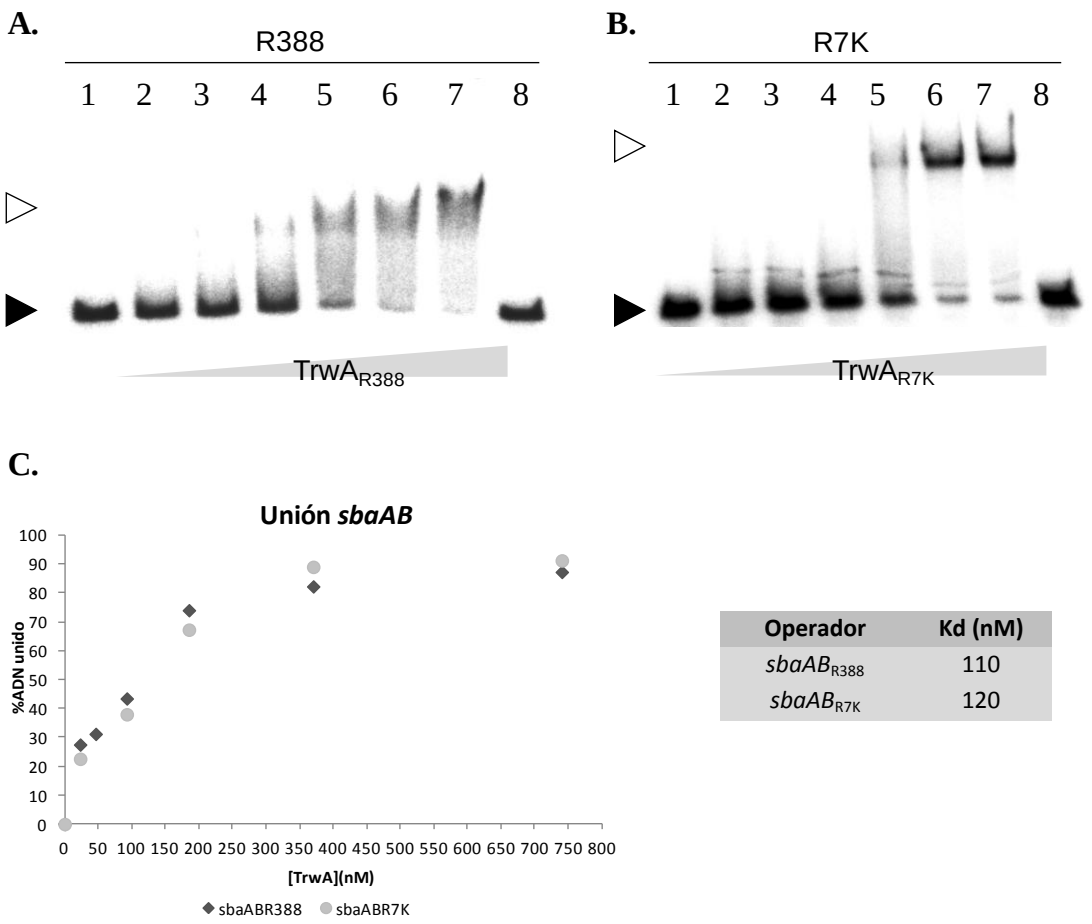
Partiendo de estos ensayos de retardo en gel mostrados en la figura R6, se calculó la constante de disociación (Kd) de cada TrwA para cada sitio de unión tal y como se describe en Materiales y Métodos (tabla R4). En ambos casos, la afinidad de TrwA es mayor por el sitio que contiene la repetición directa (*sbaA*).

Operador	**Kd (nM)**
$sbaA_{R388}$	130
$sbaA_{R7K}$	160
$sbaB_{R388}$	420
$sbaB_{R7K}$	350

Tabla R4. Constantes de disociación de los operadores de TrwA de R388 y R7K.

Una vez comprobamos la unión a los dos sitios de unión por separado, repetimos el ensayo utilizando como molde un fragmento de unas 120 pb obtenido por PCR que abarca ambos sitios de unión de R388 (figura R7A) o R7K (figura R7B).

Como ocurría al ensayar cada sitio de unión por separado, la proteína específica sólo es capaz de retardar el ADN que contiene los operadores propios.



Operador	Kd (nM)
sbaAB$_{R388}$	110
sbaAB$_{R7K}$	120

Figura R7. Retardos de los sitios *sbaAB* de R388 (**A.**) y R7K (**B.**). El triángulo blanco representa el aumento en la concentración de TrwA añadida. En la calle 8 se añadió la proteína no específica para el fragmento que estaba siendo ensayado. Las cantidades de TrwA tanto de R7K como de R388 utilizadas son las siguientes: 1; 0nM, 2; 23nM, 3; 46nM, 4; 92nM, 5; 185nM, 6; 370nM, 7; 740nM, 8; 1,1µM. **C.** Representación gráfica realizada a partir de la cuantificación de los retardos de la imagen superior, y las constantes de disociación obtenidas de ella.

Para comprobar si la especificidad mostrada *in vitro* se reproducía también en condiciones *in vivo*, diseñamos una serie de experimentos de movilización (el método se describe en detalle en la sección Materiales y Métodos). Para ello se construyeron plásmidos que contenían el *oriT* sólo y en presencia del gen *trwA* en *cis*, tanto de R388 (pSU4910 y pYTP3) como de R7K (pYTP1, pYTP2), respectivamente. Estas construcciones se introdujeron por electroporación en células a las que previamente se había introducido por conjugación el plásmido conjugativo R388 o R7K. Se midió, por lo tanto, la movilización de cada *oriT* con o sin su *trwA* en presencia de R7K y R388 como plásmidos movilizadores, para comprobar si la presencia de la *trwA* específica tenía algún efecto en la movilización.

Los plásmidos conjugativos movilizadores disponen de toda la maquinaria necesaria para una conjugación eficiente, por lo que el plásmido que contiene solo el *oriT* debería ser movilizado eficazmente en presencia del movilizador específico. Cuando el plásmido movilizador no es el correspondiente al *oriT* a conjugar, la TrwA del movilizador no podría reconocer el *oriT* (según resultados obtenidos *in vitro*), razón por la cual la frecuencia de conjugación debería verse afectada. En resultados obtenidos previamente, se mostró que en ausencia de una TrwA funcional la conjugación decaía 10^5 veces (Moncalian et al., 1997).

	Plásmido movilizador	
	R388	**R7K**
$oriT_{R388}$ (pSU4910)	0,36 (0,26-0,52)	$5,5x10^{-5}$ ($1,9x10^{-5}$-$1,6x10^{-4}$)
$oriT_{R7K}$ (pYTP1)	$1,3x10^{-5}$ ($3,3x10^{-6}$-$5,3x10^{-5}$)	0,02 (0,01-0,05)
$oriT_trwA_{R388}$ (pYTP3)	0,1 (0,05-0,25)	0,02 ($6,5x10^{-3}$-0,04)
$oriT_trwA_{R7K}$ (pYTP2)	0,07	$4x10^{-3}$ ($1,5x10^{-3}$-0,01)

Tabla R5. Frecuencias de movilización de los diferentes plásmidos que contienen el origen de transferencia seguido o no del gen *trwA* específico en *cis*. Se muestran las frecuencias obtenidas al movilizar cada construcción con R388 y R7K.

Tal y como se refleja en la tabla R5, las frecuencias de movilización de los orígenes de transferencia en presencia del plásmido específico, coinciden con aquellas obtenidas al movilizar el plásmido conjugativo (dato no mostrado) , siendo la frecuencia aproximadamente 10 veces mayor en R388 que en R7K en las condiciones en las que se realiza el experimento. La movilización cae unas 7000 veces cuando el *oriT*_R388 es movilizado por el plásmido no específico (tabla R5, pSU4910 movilizado con R7K vs R388) y unas 1500 veces en el caso de *oriT*_R7K (tabla R5, pYTP1 movilizado con R388 vs R7K). Al añadir el gen de *trwA* específico en *cis*, esperábamos que se restableciera la frecuencia original, aunque el plásmido utilizado para movilizar nuestra construcción no fuera el específico para cada uno. Los resultados obtenidos, corroboran sólo en parte esta suposición, ya que la frecuencia se restablece totalmente cuando el plásmido movilizador es R388 (tabla R5, pYTP2 con R388 vs pYTP1 con R7K), pero se restablece solo parcialmente (la frecuencia de conjugación es diez veces menor que

el salvaje) cuando lo movilizamos con R7K (tabla R5, pYTP3 con R7K vs pSU4910 con R388).

4. LA UNIÓN AL ADN Y LA MOVILIZACIÓN SE VEN AFECTADAS CUANDO ALTERAMOS LOS SITIOS DE UNIÓN DE TrwA

Una vez analizados diferentes orígenes de transferencia MOB_{F11} y habiendo observado que en todos ellos aparecen dos sitios de unión de TrwA con la misma secuencia de reconocimiento repetida de forma directa y/o invertida (figura R4), nuestro interés principal se centra en la razón de la existencia de esos dos sitios de unión de distinta conformación para TrwA.

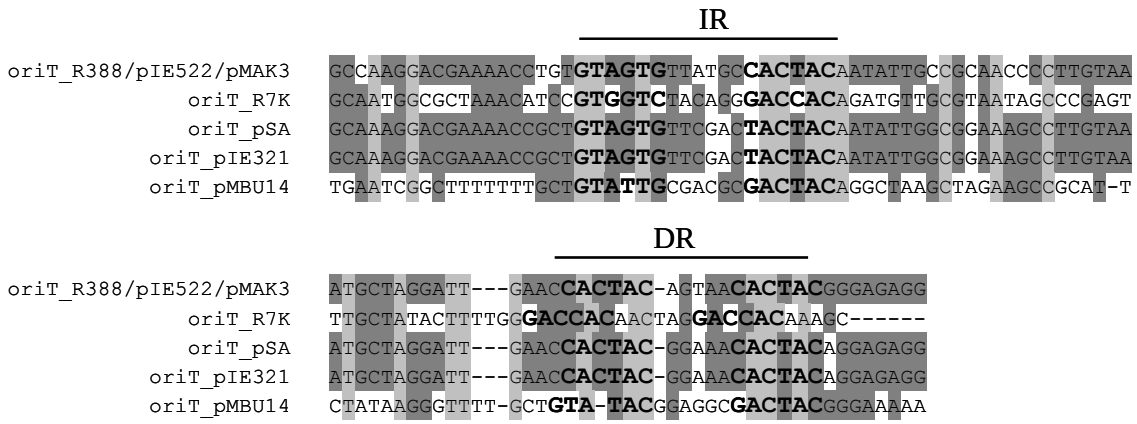


Figura R8. Alineamiento de las regiones de los sitios de reconocimiento, tanto identificados como hipotéticos, de TrwA. Se resaltan en negrita los nucleótidos que formarían los brazos de las repeticiones.

Las proteínas que, como TrwA, pertenecen a la superfamilia RHH, reconocen en su mayoría secuencias palindrómicas (Schreiter and Drennan, 2007). Solamente en el caso de la proteína omega presente en plásmidos *inc18* de bacterias Gram(+) (Weihofen *et al.*, 2006) se ha descrito que es capaz de reconocer repeticiones tanto directas como

invertidas. Los operadores que reconoce esta proteína están formados por repeticiones de 7 nucleótidos, dispuestos de manera tanto directa como invertida. Esta proteína regula la expresión de varios genes implicados en diversas funciones, y el mecanismo que ha adoptado para modular la afinidad del represor por los distintos operadores se basa en el número de las repeticiones y en su orientación, haciendo que la interacción cooperativa de la proteína con los operadores varíe según la disposición que presenten.

TrwA, como omega, muestra distinta afinidad por cada uno de los sitios que reconoce, siendo mayor la afinidad que muestra por la repetición directa que por la invertida (consultar apartado 3 de la presente sección). Por tanto, se consideró probable que esta diferencia constituyese un mecanismo para regular algunas de las funciones que la proteína lleva a cabo mediante su unión a ADN, como puede ser la represión de su propio promotor, o la estimulación del corte que realiza la relaxasa en el sitio *nic*.

Se abordó el estudio de la distribución de los sitios de unión utilizando el origen de transferencia sintético (*oriT$_{R388}$syn*) mencionado en el apartado anterior. Para ver si ambos sitios de unión de TrwA son imprescindibles para una conjugación eficiente, eliminamos cada sitio de unión por separado y medimos su frecuencia de movilización en presencia de un plásmido movilizador, siguiendo el mismo método que en las movilizaciones mencionadas con anterioridad.

Las frecuencias que obtuvimos en ambos casos mostraban una disminución importante de la capacidad de movilización, siendo ésta parecida en ausencia de uno u otro sitio (tabla R6).

Plásmido			**Frecuencia de conjugación**
	sbaB	***sbaA***	
pYTP4			0,2 (0,1-0,5)
pYTP9		—	$6x10^{-6}$ ($8x10^{-7}$, $5x10^{-5}$)
pYTP10	—		$4x10^{-6}$ ($8x10^{-7}$, $2x10^{-5}$)
pYTP11	—	—	$6x10^{-7}$ ($1x10^{-7}$, $3x10^{-6}$)

Tabla R6. Frecuencias de movilización de los mutantes que carecen de uno o ambos sitios de unión de TrwA.

El siguiente paso fue la eliminación de los dos sitios *sbaA* y *sbaB* al mismo tiempo (plásmido pYTP11). Para la obtención de este mutante, llevamos a cabo una digestión con dos enzimas distintas; *Bam*HI que tiene su diana a ambos lados del sitio *sbaB*, y *Bgl*II, que se encuentra flanqueando el sitio *sbaA* (figura R1). Así, la secuencia eliminada contenía los dos sitios de unión mas la secuencia que los separa. En este caso, la frecuencia de movilización obtenida fue aún menor que cuando se eliminaba solamente uno de los sitios, aunque no podemos achacar este efecto únicamente a la ausencia de sitios de unión de TrwA, ya que también hemos eliminado un sitio de unión de IHF además de un número añadido de nucleótidos que pueden estar implicados en una función estructural.

Para estudiar si la disposición conservada del IR y del DR era importante para la conjugación, una vez caracterizado el efecto de cada sitio de unión en la movilización, decidimos construir variantes de *oriT* que contuvieran los dos sitios de unión pero en orientaciones distintas al salvaje. Diseñamos tres construcciones diferentes valiéndonos de oligonucleótidos comerciales que llevaban los sitios de restricción adecuados para su inserción en el *oriT* sintético (figura R9). Uno de los mutantes construidos contiene dos sitios *sbaA*, a otro se le insertaron dos sitios *sbaB*, y, por último, se construyó un mutante con ambos sitios pero situados en orientaciones invertidas respecto al salvaje.

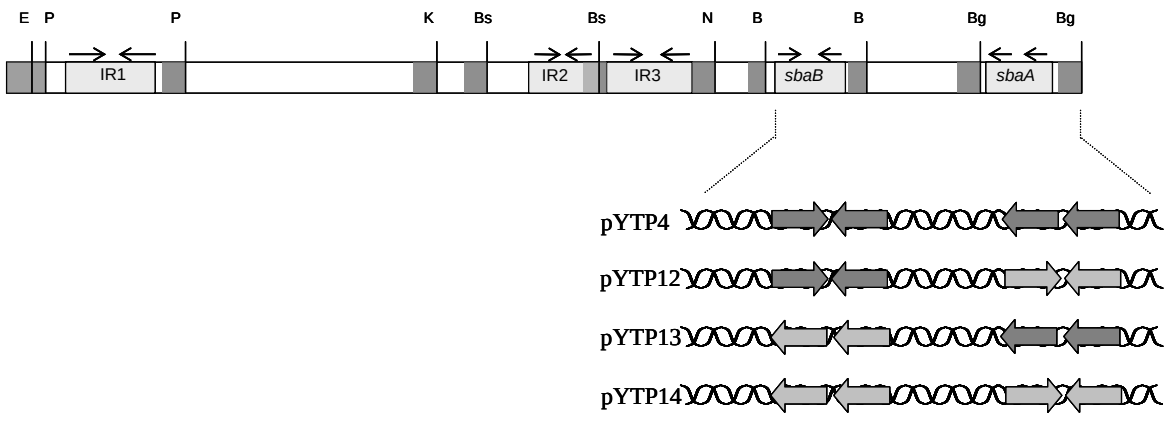


Figura R9. Representación esquemática donde se muestran los mutantes en la orientación de los sitios de unión de TrwA.

Se realizaron experimentos de movilización como en los casos anteriores, y observamos que aunque en todos los casos las frecuencias obtenidas eran menores que con el *oriT* salvaje, no disminuían tan dramáticamente como cuando uno de los sitios, o los dos, era eliminado (tabla R7).

Plásmido			**Frecuencia de conjugación**
	sbaB	***sbaA***	
pYTP4			0,22 ($8,45x10^{-2}$, 0,55)
pYTP12			$1,62x10^{-2}$ ($4x10^{-3}$, $6,53x10^{-2}$)
pYTP13			$3,61x10^{-3}$ ($1,24x10^{-3}$, $1,05x10^{-2}$)
pYTP14			$5,73x10^{-3}$ ($6,35x10^{-4}$, $5,18x10^{-2}$)

Tabla R7. Frecuencias de movilización de las construcciones que presentan diferentes distribuciones de los sitios de unión de TrwA.

Además, se realizaron ensayos de retardo en gel con fragmentos lineales de ADN de doble cadena que contenían los diferentes *oriT*s a analizar y concentraciones crecientes de TrwA (Materiales y Métodos) para conocer la afinidad que mostraba la proteína frente a los distintos sustratos. Se esperaba, de esta manera, relacionar los resultados obtenidos con los datos que previamente extraídos de los experimentos de movilización. Los resultados se muestran en la figura R10.

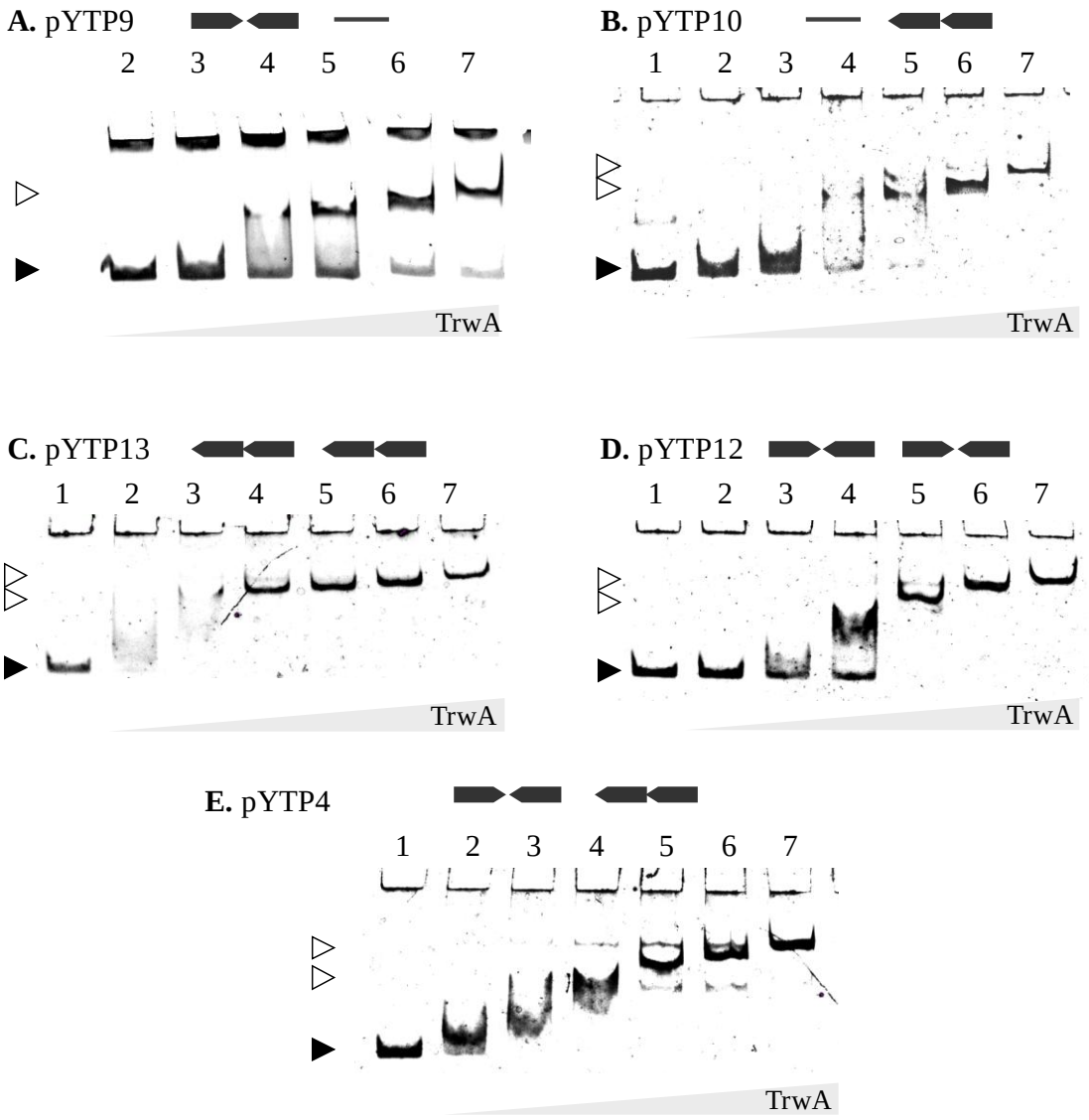


Figura R10. EMSAs de fragmentos de ADN conteniendo los *oriT* de algunas de las construcciones previamente ensayadas en experimentos de movilización. Cantidades crecientes de TrwA fueron añadidas en las siguientes concentraciones: calle 1; 0nM, calle 2; 23nM, calle 3; 46nM, calle 4; 92nM, calle 5; 185nM, calle 6; 370nM, calle 7; 740nM. El ADN no retardado se señala mediante una punta de flecha negra, y el ADN retardado a distintos niveles por una blanca. La tinción de los geles se llevó a cabo con Sybr Green.

En la figura superior podemos observar el comportamiento de los distintos mutantes de *oriT* en presencia de cantidades crecientes de TrwA. En todos los casos, aparece una banda retardada de manera gradual a medida que se aumenta la cantidad de

proteína, aunque la aparición de la misma varía entre los distintos experimentos dependiendo de la afinidad que muestra la proteína por cada fragmento.

Aun así, las diferencias que se observan al ensayar la unión de los distintos mutantes no son tan drásticas como las obtenidas en los experimentos de movilización. A continuación, procedemos a mostrar las Kd aproximadas de los distintos mutantes del origen de transferencia extraídas de estos experimentos: 20nM para el *oriTsyn* salvaje (pYTP4) y la construcción con dos sitios *sbaA* (pYTP13), 70nM para la construcción con dos sitios *sbaB* (pYTP12) y el mutante que contiene únicamente un sitio *sbaA* (pYTP10) y 140nM para el mutante al que se le ha eliminado el sitio *sbaA* (pYTP9). La afinidad que muestra la proteína por cada fragmento, sería inversamente proporcional al orden creciente utilizado para las Kd; a menor concentración de proteína, reconocería en primer lugar el *oriTsyn* y el pYTP13, seguidos del pYTP12 y pYPT10, y por último, el pYTP9.

5. EFECTO DE LAS ALTERACIONES DE LOS SITIOS DE UNIÓN DE TRWA EN LA ACTIVIDAD DEL PROMOTOR *twrABC*

TrwA actúa como represor transcripcional de su propio operón *trwABC* (Moncalian et al., 1997). Para llevar a cabo esta función es indispensable su unión a ADN, por lo que la disposición de sus sitios de unión podría verse reflejada en la actividad del promotor. De hecho ésta podría ser la razón por la que se observa un sitio de unión DR en la región promotora a pesar de ser un sitio de unión infrecuente en proteínas RHH (Schreiter and Drennan, 2007). Ya hemos comprobado en los apartados anteriores el efecto de los mutantes con diferentes combinaciones de los sitios de unión de TrwA tanto en el reconocimiento *in vitro* de la proteína como en la movilización del plásmido *in vivo*, y nos pareció interesante comprobar cómo afectan estos mutantes a la actividad del promotor.

En un trabajo previo realizado en este laboratorio se analizó la red de regulación de R388 (Fernández-Lopez, 2007). Mediante la utilización de perfiles de expresión de GFP, se detectaron y caracterizaron quince regiones promotoras y seis reguladores

transcripcionales, de los cuales todos ejercían como represores. Tres de estas proteínas reguladroas, entre las que se encuentra TrwA, reprimen únicamente sus propios promotores. El papel de TrwA en la regulación del operón *trwABC* ya había sido descrito anteriormente (Moncalian et al., 1997), y confirmado en posteriores estudios en los que se determinó cómo en presencia de TrwA la transcripción a partir de su promotor disminuye 150 veces (Fernández-López R *et al*, en prensa).

Se eligió, pues, el método descrito en el trabajo citado previamente (Fernández-López R *et al*, en prensa) para llevar a cabo nuestro análisis. Las variaciones del *oriT* se clonaron en un plásmido que contiene un gen reportero *gfpmut2*, cuya expresión queda controlada por la región promotora clonada y se midió la fluorescencia producida en presencia y ausencia del regulador, en nuestro caso, TrwA.

Para introducir los sitios de restricción que flanquean los elementos de estructura secundaria de $oriT_{R388}syn$, tuvimos que sustituir algunos nucleótidos, de tal manera que las regiones -35 y -10 del promotor *trwABC* fueron modificadas (tabla R8). Lo primero que tuvimos que comprobar, por lo tanto, era si estas mutaciones puntuales afectaban a la actividad transcripcional del promotor, para lo que clonamos $oriT_{R388}syn$ en el vector con el gen reportero (dando lugar a pYTP17) e hicimos una medición inicial. Fuimos incapaces de detectar ninguna producción de GFP en este primer experimento, de tal manera que comprobamos que las mutaciones introducidas habían abolido totalmente la expresión del promotor.

	Promotor *trwA*	
	región -35	**región -10**
$\boldsymbol{oriT_{R388}}$ **salvaje**	TTGCCG	TAGGAT
$\boldsymbol{oriT_{R388syn}}$	T**C**C**C**CG	TA**A**GAT

Tabla R8. Se recogen las secuencias de las regiones -35 y -10 del promotor de *trwA* en el *oriT* salvaje y en el sintético. Las mutaciones de este último aparecen marcadas en rojo.

Se realizó una mutagénesis dirigida por el método de Quickchange (Stratagene) para restablecer las bases originales y volvimos a medir la actividad del promotor. Esta vez, sí fuimos capaces de medir la producción de GFP, distinguiendo dos patrones claros dependiendo de si la proteína TrwA estaba presente en el medio o no (figura R11). En los casos en los que TrwA estaba presente ésta era sintetizada por el plásmido R388 que habíamos introducido en la célula que contenía el clon a estudiar.

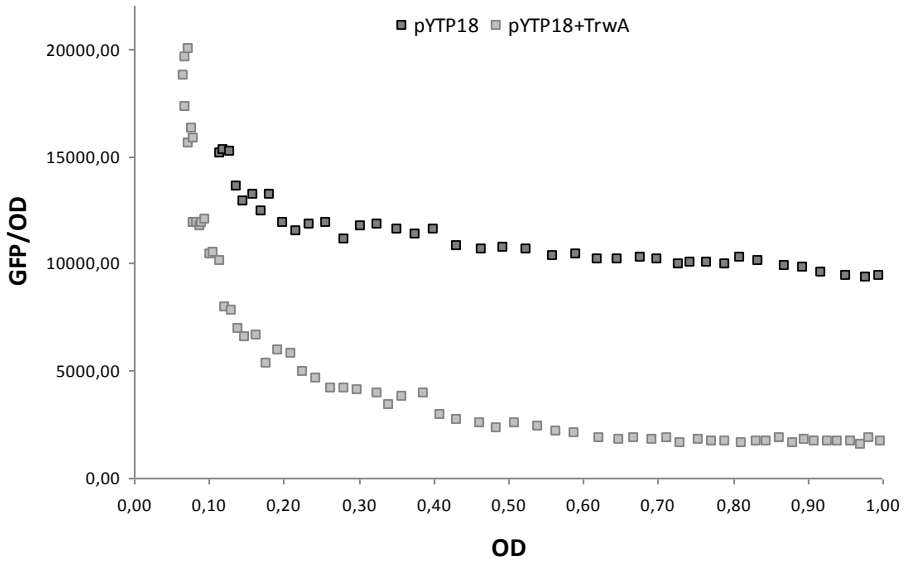


Figura R11. Actividad del promotor *trwABC* en presencia y en ausencia del regulador TrwA. Se representan en el eje Y la fluorescencia por DO y en el X la DO.

Una vez comprobado que el promotor del *oriT* sintético modificado era funcional, nos dispusimos a probar las construcciones que tenían operadores modificados de TrwA (tabla R9).

Plásmido	**Distribución de los sitios unión de TrwA**	
	sbaB	*sbaA*
pYTP18		
pYTP20		
pYTP22		

Tabla R9. Construcciones realizadas para la medición de la actividad del promotor *trwABC*.

Se eligieron las construcciones que tenían dos repeticiones directas (pYTP19) y dos repeticiones invertidas (pYTP21), y se realizó una mutagénesis mediante el método Quickchange para restablecer el sitio -35 del promotor tal y como se había hecho previamente con el *oriT*$_{R388}$*syn*, para dar lugar a los plásmidos pYTP20 y pYTP22 con el promotor restablecido.

Se realizaron las mediciones de todas las construcciones al mismo tiempo en presencia y ausencia de TrwA expresada desde R388, añadiendo como control el *oriT* original de R388 que contiene el promotor de *trwA* salvaje.

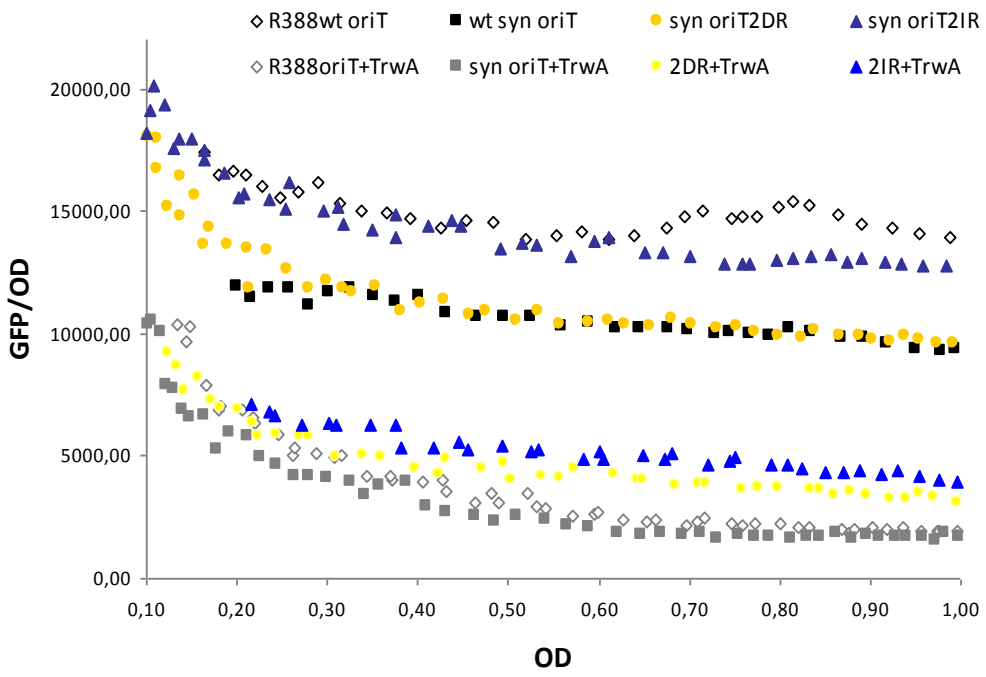


Figura R12. Actividad del promotor *trwABC* en ausencia y en presencia de TrwA en el *oriT* salvaje y en los distintos mutantes con los sitios de unión de TrwA alterados. Los valores obtenidos con cada construcción se representan con un símbolo de forma y color descritos en la parte superior de la figura, rebajándose este color unos tonos para ilustrar los datos en presencia del regulador. Se representan en el eje Y la fluorescencia por DO y en el X la DO.

En la figura R12 se muestra el resultado de las mediciones realizadas con las construcciones mencionadas anteriormente, utilizando como control el *oriT*$_{R388}$ salvaje.

En la imagen se observan dos grupos diferenciados de datos: los perfinles que muestran valores más altos de producción de GFP (posición más elevada en el eje de ordenadas) corresponden a las medidas de las construcciones en ausencia del regulador mientras que cuando R388 está presente y se produce TrwA en la célula, las curvas de expresión de GFP bajan significativamente.

Vemos en la figura R12 que en ausencia de la proteína represora TrwA, la mayor actividad la muestra el promotor del *oriT* salvaje de R388, seguido de cerca, con valores muy similares, por la construcción que contiene dos sitios *sbaB* (pYTP20). El *oriT* sintético salvaje (pYTP18) y la variante con dos sitios *sbaA* (pYTP22) muestran a su vez unos valores casi idénticos, siendo éstos menores que los obtenidos con los otros dos clones ensayados (*oriT*$_{R388}$ y pYTP20). En presencia de TrwA, la represión muestra su mayor eficiencia en el caso de *oriT*$_{R388}$, viéndose disminuida la actividad transcripcional del promotor cinco veces. El *oriT* sintético muestra una diferencia de expresión similar al *oriT* salvaje. En cambio, en las dos construcciones restantes (pYTP20 y pYTP22) la represión ejercida por TrwA es menor, siendo prácticamente la mitad que la obtenida con los otros dos plásmidos (tabla R10, figura R13).

	Ausencia de regulador	**Presencia de regulador**	**Diferencia de expresión**
***oriT*$_{R388}$**	14187 ± 779	2840 ± 1621	5
pYTP18 (IR/DR)	10479 ± 1314	2218 ± 1038	4,7
pYTP20 (IR/IR)	13581 ± 1032	5140 ± 603	2,6
pYTP22 (DR/DR)	10569 ± 799,42	4189 ± 469	2,5

Tabla R10. Variación de la actividad de los promotores de TrwA en presencia y ausencia del regulador en la célula. Se muestran los valores medios y las desviaciones estándar, así como la diferencia de expresión calculada como la división de los valores en ausencia de regulador entre los valores en presencia del mismo.

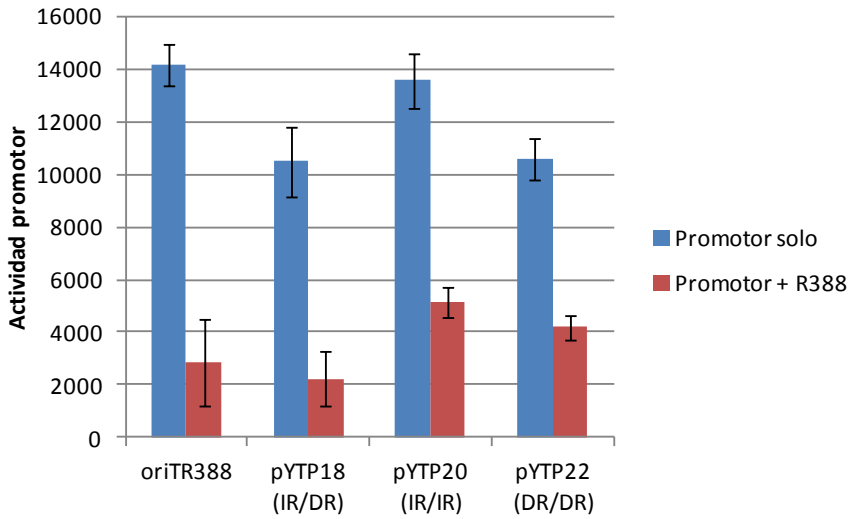


Figura R13. Interpretación de los datos del promotor representados en la tabla R10. En azul se muestran los datos obtenidos en ausencia del regulador (TrwA), y en rojo los datos en presencia del mismo.

En el experimento descrito, el represor, en este caso TrwA, se expresa controlado por su propio promotor desde R388, junto con el resto de componentes del plásmido. Para evitar cualquier interferencia que se pueda dar debido a proteínas del plásmido ajenas a TrwA así como para controlar la cantidad de regulador que se produce, utilizamos un plásmido en el que el gen *trwA* se encuentra clonado tras el promotor *araBAD* inducible por arabinosa. En este nuevo ensayo, la cantidad de represor sintetizado se controla mediante la adición al medio de una cantidad conocida de arabinosa, para lo que se utiliza una cepa de *E. coli* que expresa constitutivamente el transportador de arabinosa araE (BW27783) (Khlebnikov *et al.*, 2001).

Esto nos proporciona un método para comparar la actividad que muestran los distintos promotores a concentraciones concretas de regulador, mientras que cuando el regulador se genera directamente desde el plásmido como ocurre en el experimento anterior, sólo somos capaces de observar el efecto en presencia o ausencia del mismo. Nos permite, a su vez, comparar los tiempos de respuesta de los distintos promotores: si el promotor reacciona ante una concentración menor de regulador, significaría que éste tiene mayor facilidad para llevar a cabo la represión de esa variante en concreto, por

ejemplo, porque la capacidad de unión sea mayor que en el resto de las variantes ensayadas.

En este caso, no se puede establecer una cantidad de regulador en la que se observen diferencias significativas entre los distintos promotores. El comportamiento parece similar en todos los casos, observándose las mismas diferencias que ya se habían advertido en el caso en el que el regulador se expresaba directamente desde R388. Al no aportar información trascendente a lo que ya habíamos comprobado antes, el resultado del experimento realizado se muestra en el material suplementario (figura S1).

6. DETERMINACIÓN DE LOS CONTACTOS ESPECÍFICOS DE TRWA CON EL ADN

Aunque ya se habían definido los sitios de unión de $TrwA_{R388}$ en un trabajo previo (Moncalian et al., 1997), otro de nuestros objetivos en este trabajo de tesis doctoral era describir en más detalle cómo se producía esta unión, y conocer cuáles eran las bases implicadas en el reconocimiento de la proteína.

Para determinar cuáles son los nucleótidos específicos del operador que participan en la unión ADN-proteína, se diseñaron variaciones del sitio *sbaA* en los que se introdujeron mutaciones simétricas en ambos brazos de la repetición. Se diseñaron seis mutantes, cambiando, en cada caso, uno de los nucleótidos del hexámero que se repite en el operador, de modo que la mutación estaría presente en el brazo proximal y distal de la repetición. En todos los casos cambiamos una pirimidina por la otra pirimidina o una purina por la otra purina. La secuencia de los operadores ensayados se muestra en la tabla R11.

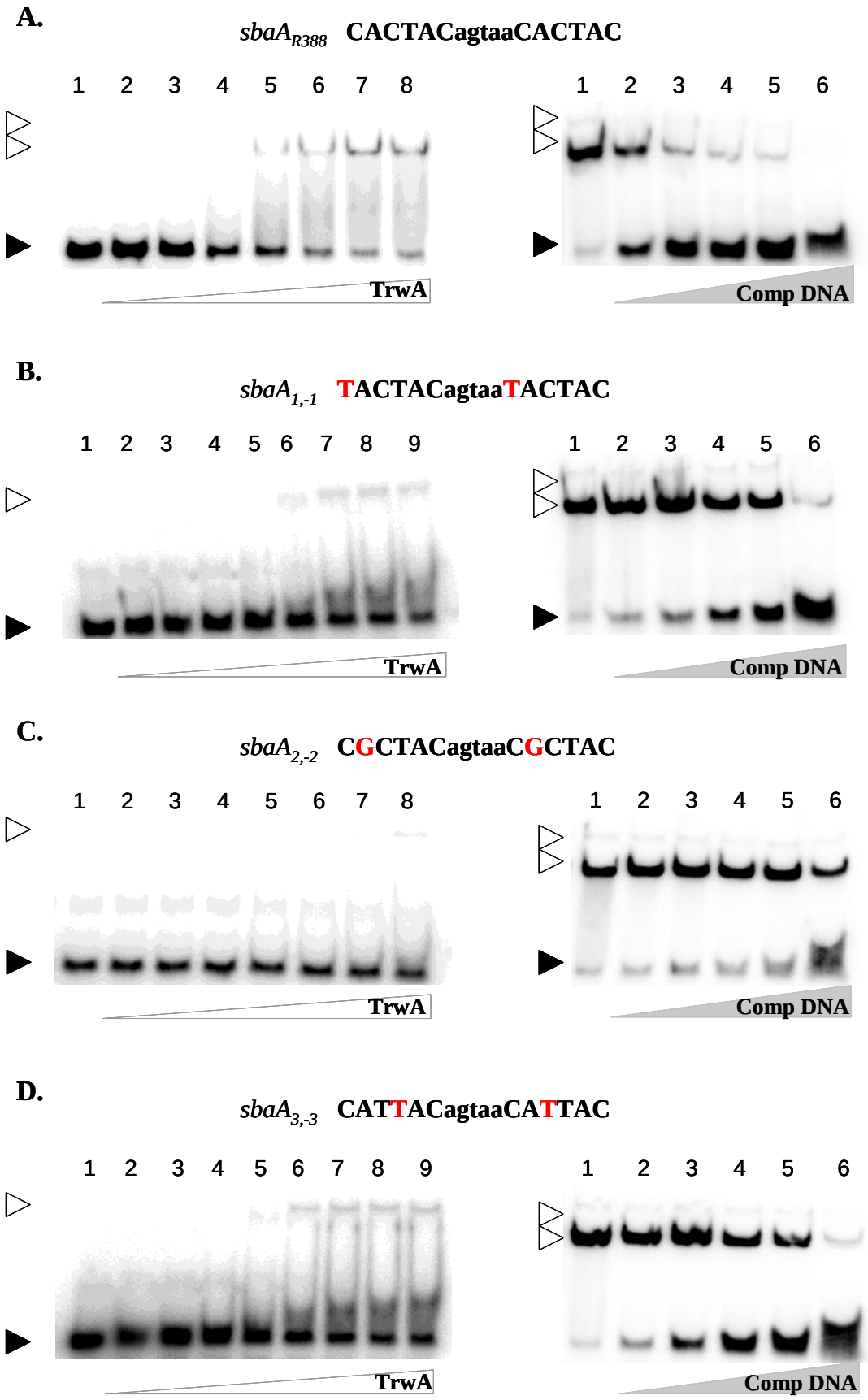
Oligonucleótido	**5’ Secuencia 3’**
sbaA$_{R388}$	CACTACagtaaCACTAC
sbaA$^{1,-1}$	**T**ACTACagtaa**T**ACTAC
sbaA$^{2,-2}$	C**G**CTACagtaaC**G**CTAC
sbaA$^{3,-3}$	CA**T**TACagtaaCA**T**TAC
sbaA$^{4,-4}$	CAC**C**ACagtaaCAC**C**AC
sbaA$^{5,-5}$	CACT**G**CagtaaCACT**G**C
sbaA$^{6,-6}$	CACTA**T**agtaaCACTA**T**

Tabla R11. Se muestra la posición de las mutaciones introducidas para construir los distintos operadores mutantes utilizados en los ensayos de unión.

La afinidad de la proteína por los distintos operadores fue medida en ensayos de retardo en gel (figura R14). Se observó que TrwA mostraba mayor afinidad por el operador salvaje, seguido por los mutantes *sbaA*$^{1,-1}$ y *sbaA*$^{3,-3}$, *sbaA*$^{4,-4}$ y *sbaA*$^{2,-2}$. La proteína no era capaz de reconocer los mutantes *sbaA*$^{5,-5}$ y *sbaA*$^{6,-6}$, aun añadiendo concentraciones muy grandes de la misma (1,6µM).

Para evitar cualquier tipo de artefacto debido a las grandes cantidades de proteína utilizadas, se realizaron ensayos de retardo en gel con el operador salvaje utilizando los operadores con mutaciones como ADN competidor. Se mezcló el operador salvaje con una concentración de proteína que previamente sabíamos que une prácticamente el 100% del ADN, y sobre esta reacción, se fueron añadiendo cantidades crecientes de los operadores mutantes.

En este caso, al competir el operador mutante con el salvaje por la unión de TrwA, mediríamos la desaparición del complejo ADN-proteína, debiendo ser la tasa de desaparición más rápida cuanto mayor fuera la afinidad del mutante competidor. Los resultados obtenidos mediante los dos tipos de retardos coincidían entre sí (figura R14).


A.
sbaA R388 CACTACagtaaCACTAC
1 2 3 4 5 6 7 8
TrwA
1 2 3 4 5 6
Comp DNA
B.
sbaA 1,-1 TACTACagtaaTACTAC
1 2 3 4 5 6 7 8 9
TrwA
1 2 3 4 5 6
Comp DNA
C.
sbaA 2,-2 CGCTACagtaaCGCTAC
1 2 3 4 5 6 7 8
TrwA
1 2 3 4 5 6
Comp DNA
D.
sbaA 3,-3 CATTACagtaaCATTAC
1 2 3 4 5 6 7 8 9
TrwA
1 2 3 4 5 6
Comp DNA

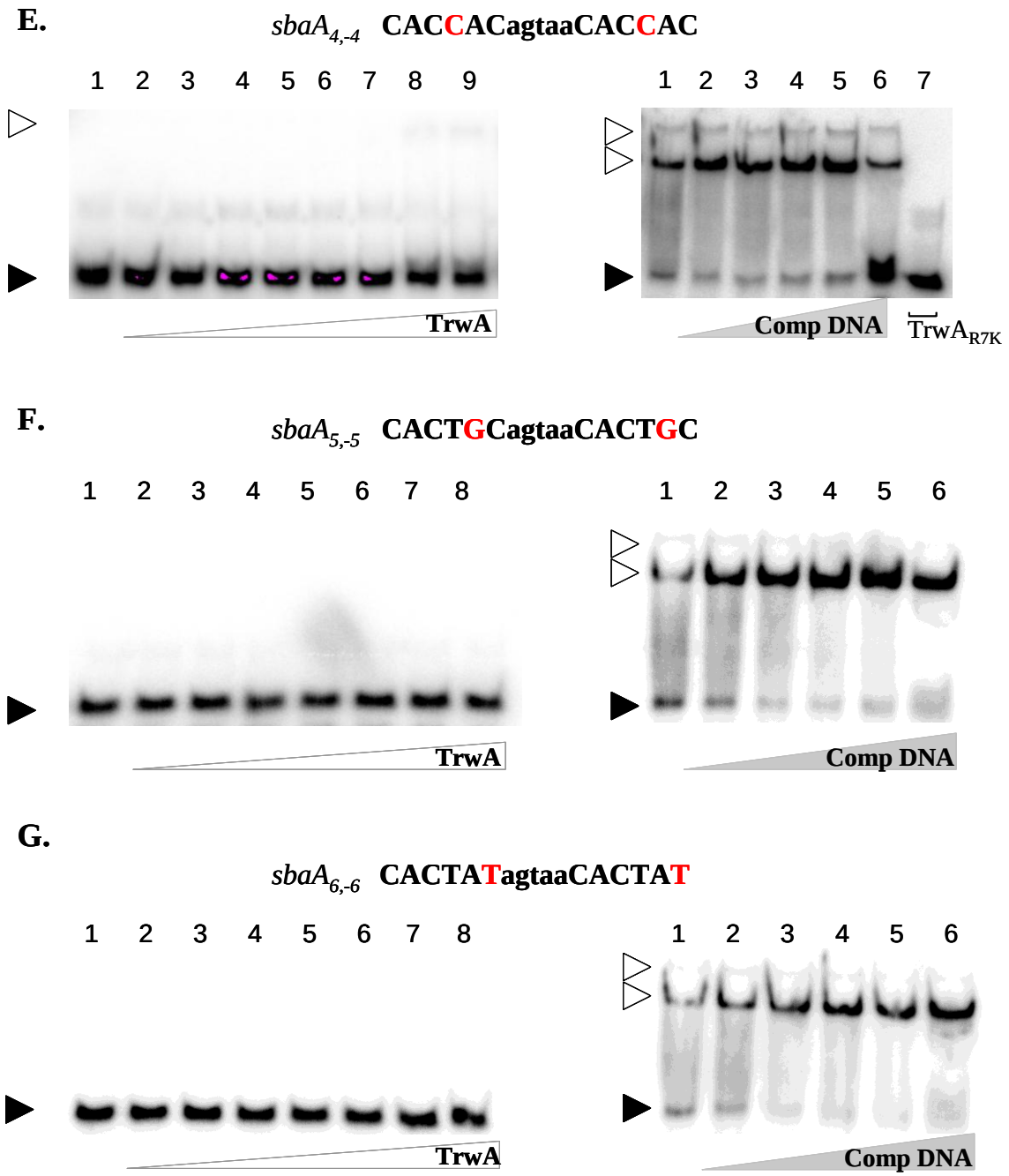


Figura R14. Se muestran los retardos de las variantes del operador *sbaA*. En las figuras de la izquierda aparecen los ensayos con cantidades crecientes de proteína, y en las figuras de la derecha aquellos realizados añadiendo cantidades crecientes de ADN competidor. En la parte superior de cada par de retardos se muestra la secuencia del ADN utilizado como molde (no se muestra la cadena complementaria). Las puntas de flecha blancas señalan los complejos de ADN-proteína, y las negras, el ADN libre. La figura A corresponde al ensayo con el operador salvaje, la figura B al mutante $sbaA^{1,-1}$, la figura C al mutante $sbaA^{2,-2}$, la figura D al $sbaA^{3,-3}$, la figura E al $sbaA^{4,-4}$, la figura F al $sbaA^{5,-5}$, y finalmente, la figura G al $sbaA^{6,-6}$. En los ensayos con cantidades crecientes de proteína se añadieron las siguientes concentraciones de TrwA: calle 1; 0, calle 2; 23nM, calle 3; 46nM, calle 4; 92nM, calle 5; 185nM, calle 6; 370nM, calle 7; 740nM, calle 8; 1100nM, calle 9; 1580nM. En los ensayos con ADN competidor las cantidades de ADN añadidas fueron las siguientes: calle 1; 100x, calle 2; 1000x, calle 3; 2500x, calle 4; 7500x, calle 5; 15000x, calle 6; 25000x. Para conseguir que la mayoría del ADN estuviera formando un complejo con la proteína se añadió 1,1µM de TrwA al operador salvaje. En el ensayo con ADN competidor de la figura E, en la calle 7, se añadió 1,5µM de $TrwA_{R7K}$ para comprobar si esta proteína reconocía a ese operador que divergía únicamente en una pb del suyo.

En el caso del mutante *sbaA*$^{4,-4}$, el cambio de ese nucleótido (CAC**C**AC vs CAC**T**AC) hace que la repetición del hexámero varíe únicamente en un nucleótido con la encontrada en el operador correspondiente en R7K (**G**ACCAC). Para ver si $TrwA_{R7K}$ reconocía este operador, se añadió una concentración en exceso de proteína pero no se observó ningún tipo de desplazamiento en la banda correspondiente al ADN. Aunque los dos hexámeros que conforman la repetición tienen casi la misma secuencia que en R7K, ya habíamos comentado previamente que el espaciamiento entre ambos es distinto en R388 y en R7K, siendo de un nucleótido más en el último. No sólo cambia el número de nucleótidos espaciadores, sino que la secuencia que presentan también varía. Por tanto, sería plausible suponer que estos son los motivos por los cuáles la proteína no es capaz de reconocer este operador. Tampoco resulta descartable que este primer nucleótido (G) del operador sea indispensable para que se produzca la unión en R7K, aunque en R388 su mutación no provoque el cese de reconocimiento por parte de la proteína.

Según estos resultados, los residuos más importantes implicados en la unión serían los situados en las posiciones 5 y 6 del hexanucleótido, ya que su modificación provoca que la proteína no sea capaz de reconocer el operador. Aunque el resto de las mutaciones no provocan un efecto tan drástico, todas ejercen un efecto en la unión, que se ve disminuida. Entre ellos destaca el provocado por la sustitución de la timina de la posición 4 del hexanucleótido, por lo que parece jugar un papel esencial en el reconocimiento por la proteína. La secuencia de estos tres nucleótidos se conserva en R388, R7K y en pMBU14, en *Anaeromyxobacter* se conserva únicamente la posición 4 y en *Psycrobacter* no coincide ninguno de estos nucleótidos.

7. LA ESPECIFICIDAD DE TrwA NO DEPENDE SOLO DE LOS AMINOÁCIDOS POLARES DE SU HOJA BETA

Ya se ha comentado en un apartado anterior que las proteínas accesorias TrwA de R388 y R7K difieren casi exclusivamente en la secuencia de los aminoácidos polares de la hebra beta del dominio de unión a ADN (figura R15). Estos aminoácidos son los implicados en el reconocimiento específico del ADN en las proteínas englobadas en la

superfamilia RHH, que incluye a varios factores de transcripción de procariotas (Schreiter and Drennan, 2007).

Se consideró la posibilidad de modificar la especificidad por su operador de cualquiera de las TrwA de los dos plásmidos que estudiamos (R388 y R7K) cambiando los aminoácidos polares implicados en la unión del ADN por aquellos situados en la misma posición en su homónima. Ya se han realizado estudios de este tipo en proteínas RHH y encontramos al menos dos casos documentados en los que se consiguió cambiar la especificidad de una proteína por otra mediante la mutación de un pequeño número de aminoácidos localizados en la hebra beta del dominio de unión a ADN (Raumann *et al.*, 1995; Wong *et al.*).

Para localizar los aminoácidos que forman parte de la hoja beta implicados en la unión a ADN, utilizamos un servidor de predicción de estructura secundaria (Jpred 3) (Cole *et al.*, 2008), y tras hacer un alineamiento de las proteínas, en un primer intento elegimos mutar dos de ellos. La proteína sobre la que se eligió hacer las mutaciones fue TrwA de R7K, y los cambios realizados fueron K8Q y T12S.

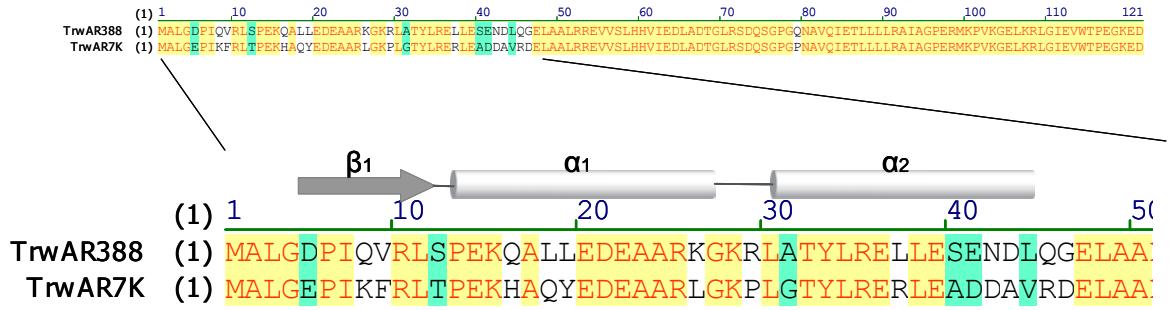


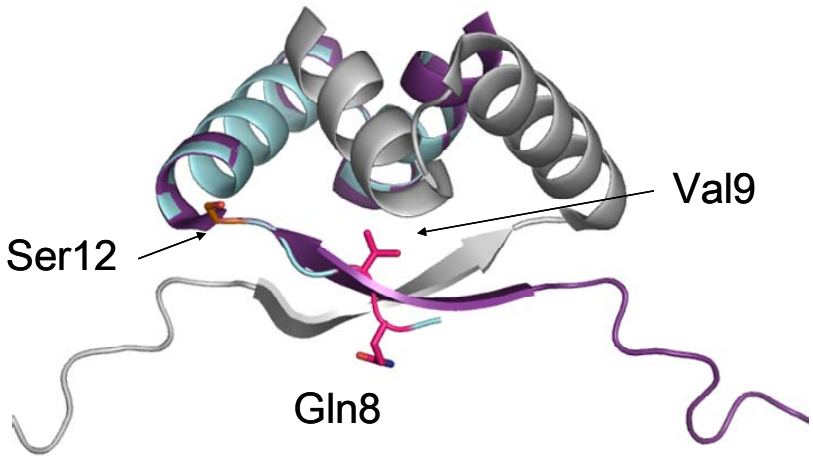
Figura R15. Alineamiento de las proteínas TrwA de R388 y R7K. En amarillo se resaltan los aminoácidos que son comunes para ambas, y en verde los que aún siendo distintos, comparten la misma carga. Los que se muestran en blanco son los que no comparten secuencia ni homología. En la imagen inferior se ha hecho un enfoque sobre el dominio de unión a ADN.

La proteína con la doble mutación K8Q, T12S se purificó siguiendo el protocolo para la TrwA salvaje. Tras la purificación se comprobó por cromatografía de exclusión molecular que TrwAK8Q, T12S también era un tetrámero en solución. Se realizó un retardo en gel utilizando como ADN molde los sitios *sbaAB* tanto de R388 como de R7K, pero aunque se utilizó una concentración en exceso de la proteína, no se observó unión a ninguno de los dos operadores.

En el tiempo en el que se planteó este experimento y se obtuvo este primer resultado, se publicó la estructura resuelta por RMN de la zona de unión a ADN de la proteína accesoria del plásmido R64, NikA (Yoshida *et al.*, 2008). Al lanzar la secuencia de TrwA a servidores que predicen la estructura terciaria que adopta una proteína comparándola con otras proteínas de estructuras ya conocidas (Phyre 2.0), la proteína resuelta con la que más similitud mostraba era precisamente NikA.

En esta predicción, la serina 12 que creíamos implicada en la unión, quedaba fuera de la hebra beta (figura R16), por lo que es posible que no establezca contactos directos con el ADN. Se decidió entonces, mutar la fenilalanina situada en la posición 9 en R7K por la valina de R388, ya que aún no siendo un aminoácido polar que participe en el reconocimiento específico, sí que podría, debido a su tamaño, provocar una variación en la conformación del dominio de unión que estuviera impidiendo la interacción. En la figura R20 se muestra la posición de estos residuos en el modelo estructural de TrwA obtenido en el servidor Phyre.

A



B

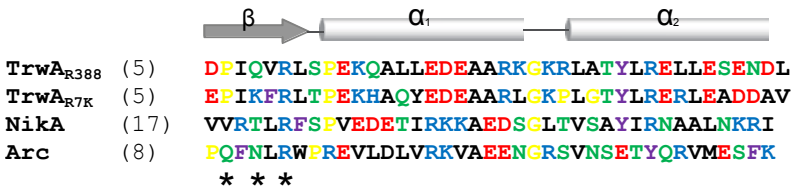


Figura R16. Se muestra la superposición de los extremos amino terminales de NikA y TrwA (A) y el alineamiento del dominio de unión a ADN de estas proteínas junto con Arc y TrwA de R7K (B). (A) En la imagen modelada por Phyre2 los dos monómeros de NikA están coloreados en morado y gris, y en azul claro se muestra TrwA. Se destacan los dos aminoácidos que se mutaron en un principio (S12 y Q8) y la posición de V9 (B) Los aminoácidos del alineamiento se han coloreado según sus propiedades: en amarillo prolina o glicina, en blanco los apolares, en verde los polares neutros, en morado los aromáticos, en azul los de carga positiva y en rojo aquellos con carga negativa. Se señalan con un asterisco los aminoácidos polares implicados en la unión específica al ADN en Arc.

Se purificó la proteína con la triple mutación (TrwAK8Q, F9V, T12S) y mediante un experimento de retardo, se comprobó que ésta tampoco era capaz de reconocer ni el operador propio ni el de R388.

Se decidió mutar el único aminoácido polar que restaba, E5. Sobre el doble mutante inicial, se cambió E5 por D, pero este nuevo mutante (TrwAE5D, K8Q, T12S) tampoco unía ADN.

Como última opción, se decidió mutar P9 y E5 al mismo tiempo, lo que daría lugar a una $TrwA_{R7K}$ que tendría una hebra beta idéntica a $TrwA_{R388}$. Sorprendentemente, esta nueva proteína tampoco mostraba ninguna afinidad por el operador de R388, y no era tampoco capaz de reconocer el de R7K.

Aunque los aminoácidos encargados de realizar los contactos específicos de secuencia con el ADN son los residuos polares de la hoja β, éstos no son los únicos contactos con el ADN que llevan a cabo estas proteínas. La hélice 2 del dominio RHH participa en el contacto no específico con el ADN siendo los dos últimos aminoácidos del extremo N terminal de este motivo los que contactan con el esqueleto fosfato del ADN a ambos lados del surco mayor. Esta interacción es favorable electrostáticamente porque el dipolo positivo del extremo N-terminal de la hélice está orientado directamente hacia el fosfato cargado negativamente del esqueleto de ADN. Estos contactos anclan el dominio RHH al ADN y orientan las bases polares de la hebra beta para alcanzar una óptima interacción (Schreiter and Drennan, 2007). Estos aminoácidos no coinciden en R388 y R7K, divergiendo tanto en secuencia como en polaridad. En primer lugar, en R388 encontramos una asparragina (N42) que es sustituida por un aspártico en R7K, y otro aspártico que ocupa la segunda posición en R388 (D43), es sustituido, esta vez, por una alanina en R7K. Al haberse mutado en este trabajo únicamente los aminoácidos de la hebra β, puede que no se lleven a cabo las interacciones necesarias entre éstos y los aminoácidos de la hélice 2 lo que provocaría que la proteína no sea capaz de reconocer el ADN o que, aún reconociéndolo, la unión no sea estable. Otra posible razón para que la unión al ADN se vea abolida es que, al insertar las mutaciones, se haya provocado un cambio conformacional que afecte a la estructura final de la proteína. Aunque se comprobó por cromatografía de exclusión molecular que todas las proteínas que contenían las mutaciones eluían en la fracción correspondiente a la forma tetramérica de la proteína como TrwA original, no sabemos cómo se ha llevado a cabo el plegamiento, por lo que es posible que la forma adoptada por éstas no fuera la adecuada.

8. CRISTALIZACIÓN TRWA-ADN

Para conocer fielmente cómo se produce la interacción entre TrwA y su operador y cómo ocurre la tetramerización de TrwA, uno de los objetivos que nos planteamos durante esta tesis fue la obtención de la estructura de TrwA unida al ADN.

Hasta ahora, se conocen las estructuras tridimensionales de dos de las proteínas que participan en la formación del relaxosoma de R388. La estructura del dominio relaxasa de TrwC unida a ADN fue resuelta por nuestro grupo en colaboración con el grupo del Dr. Coll hace algún tiempo (Boer *et al.*, 2006; Guasch *et al.*, 2003), y ya algunos años antes había sido publicada la estructura de IHF unida a ADN (Rice *et al.*, 1996).

Si se resolviera la estructura de TrwA junto con el ADN, podríamos modelar la estructura del relaxosoma al completo, aportándonos una visión global de las interacciones que se producen entre las proteínas y el ADN que lo forman. Conociendo los dos sitios de reconocimiento de la proteína, adquirimos oligonucleótidos comerciales que por hibridación darían lugar a cada uno de ellos. Se diseñaron varios oligonucleótidos de mayor o menor longitud, con o sin extremos protuberantes, y se realizó un rastreo de cristales a varias concentraciones de proteína.

Se obtuvieron cristales de $TrwA_{R388}$ en complejo con el sitio *sbaA* (figura R17), a una concentración de 10mg/ml de proteína, utilizando una relación de proteína:ADN de 1:1,3 aproximadamente en un medio que contenía 0,2M NH_4SO_4, MES pH=5,5, sin o con glicerol (10%) y aditivos (espermina 2nM, $MgCl_2$ 20mM), los cuáles variaban la morfología de los mismos. El protocolo para la obtención de estos cristales se detalla en Materiales y Métodos.

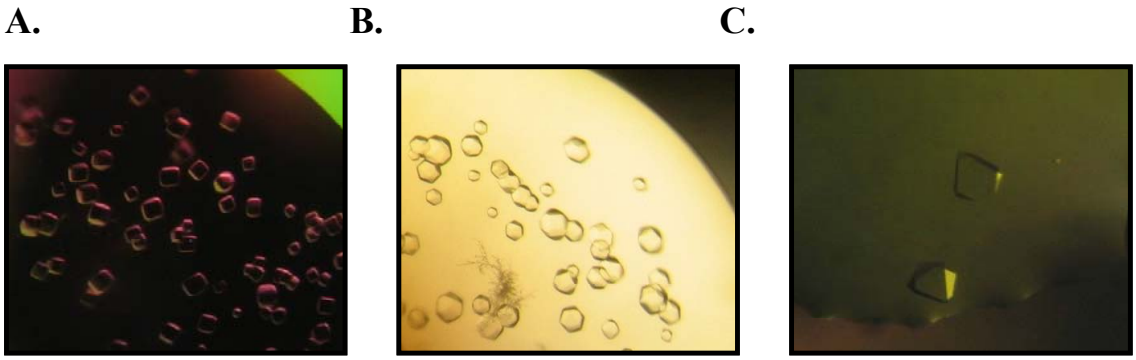


Figura R17. Cristales obtenidos con $TrwA_{R388}$ en complejo con *sbaA*. **A.** Cristales originales obtenidos sin presencia de aditivos. Al utilizar aditivos como espermina (**B**) o glicerol (**C**), la morfología de los cristales varía pasando a ser más definidas sus fases.

Los cristales mostrados en la figura R21 se consiguieron en presencia de un operador de doble cadena formado a partir de oligonucleótidos 20mer. También se obtuvo otra forma cristalina con oligonucleótidos 19mer en presencia de 0,150M de $(NH_4)_2SO_4$ y entre 20-30% de glicerol (figura R18).

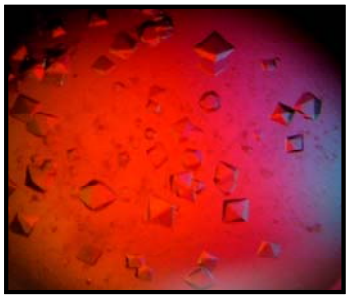


Figura R18. Cristales de $TrwA_{R388}$ en complejo con el operador *sbaA* formado por oligonucleótidos 19 mer.

En ningún caso se obtuvieron cristales de $TrwA_{R388}$ en complejo con *sbaB*. Anteriormente, nuestro grupo sí había obtenido cristales con este operador, aunque no se obtuvieron resultados satisfactorios tras su difracción. Hay ocasiones en las que se

optimiza el proceso de formación de cristales al variar la longitud de la proteína, por lo que purificamos la variante $TrwA_{2\text{-}73}$ que sigue manteniendo la capacidad de unión a ADN pero carece de los últimos 48 aminoácidos del extremo carboxilo-terminal. Con esta proteína sí se obtuvieron cristales con *sbaB* pero aparecían formando racimos y no fue posible conseguir cristales que fueran aptos para su difracción.

Con el propósito de conseguir la máxima diversidad de cristales posibles, se probó paralelamente la cristalización de $TrwA_{R7K}$ en complejo con sus sitios de unión específicos. Se consiguieron cristales en complejo con $sbaA_{R7K}$ utilizando oligonucleótidos 22mer en condiciones parecidas a las de $TrwA_{R388}$ (0,2M $(NH_4)_2SO_4$, MES pH=5,5 y PEG2K mme 5-10%). También en complejo con $sbaB_{R7K}$ se observó la aparición de cristales, utilizando unos oligonucleótidos 23mer en presencia de 0,2M $MgCl_2$, HEPES pH=7,5 y 18-20% de PEG400 (figura R19).

A. **B.**

Figura R19. Cristales obtenidos con $TrwA_{R7K}$ en complejo con *sbaA* y *sbaB*. **A.** Cristales obtenidos con *sbaA*. **B.** Cristales obtenidos en presencia de *sbaB*.

Todos estos cristales obtenidos se sometieron a difracción de rayos X en la línea BM16 del sincrotrón ESRF de Grenoble (Francia). A pesar de la buena apariencia de los cristales en ningún caso la difracción que presentaron fue mayor de 15Å, siendo imposible resolver la estructura del complejo a esa resolución. Es probable que la

proteína se encuentre en varios estados distintos dentro del empaquetamiento que adopta en el cristal y por esta razón no somos capaces de obtener un patrón de difracción eficiente.

8.1 Predicción *in silico* de la estructura de TrwA unida al ADN

Los resultados que hemos obtenido en los experimentos detallados en este trabajo (apartados 4 y 5 de la presente sección) ponen en evidencia la importancia de la interacción de TrwA con sus sitios de unión en el $oriT_{R388}$. Ante los infructuosos esfuerzos realizados para obtener la estructura tridimensional de la proteína con su operador, decidimos elaborar un modelado de la proteína para poder tener una idea aproximada de cómo se llevaría a cabo la unión al ADN.

Para ello utilizamos el programa Phyre2 (Kelley and Sternberg, 2009), que nos permitió obtener la estructura terciaria de TrwA basándose en su similitud con otras proteínas (ver Materiales y Métodos). En el apartado 7 de esta sección de resultados mostramos el modelado de la proteína partiendo de NikA de R64, única proteína accesoria de la conjugación similar a TrwA resuelta. Actualmente, nuevas proteínas RHH han sido identificadas y caracterizadas, algunas de ellas mostrando más similitud con TrwA. El servidor Phyre2 nos proporciona la homología con varias de estas proteínas RHH, de las cuales seleccionamos las tres primeras para llevar a cabo nuestro análisis (tabla R12).

Proteína	Confianza	Identidad (%)
Tipo VCA0319 (PDB:1Y9B)	90,4	20
Bp1543 (PDB:3KK4)	84,7	35
Proteína sin caracterizar de *Mycobacterium tuberculosis* (PDB:3H7D)	43,2	26

Tabla R12. Información proporcionada por el programa Phyre2 de las proteínas que se seleccionaron para el modelado de TrwA.

Las tres proteínas seleccionadas contienen un motivo RHH de unión a ADN. Es más, la homología con TrwA se centra en esta región amino terminal. Los modelos estructurales obtenidos de este modo corresponderían a la proteína libre, por lo que se modeló su unión a ADN utilizando como molde la estructura de la proteína Arc unida a su operador (PDB: 1PAR). Se llevó a cabo el modelado con cada una de ellas y observamos que la estructura era muy similar en los tres modelos (figura R20), por lo que en la presente sección se mostrará únicamente el modelado obtenido a partir de la proteína Bp1543 por ser el modelo que mejor nos permitía apreciar la posición de la hoja beta (los dos restantes se recogen en material suplementario, figura S2).

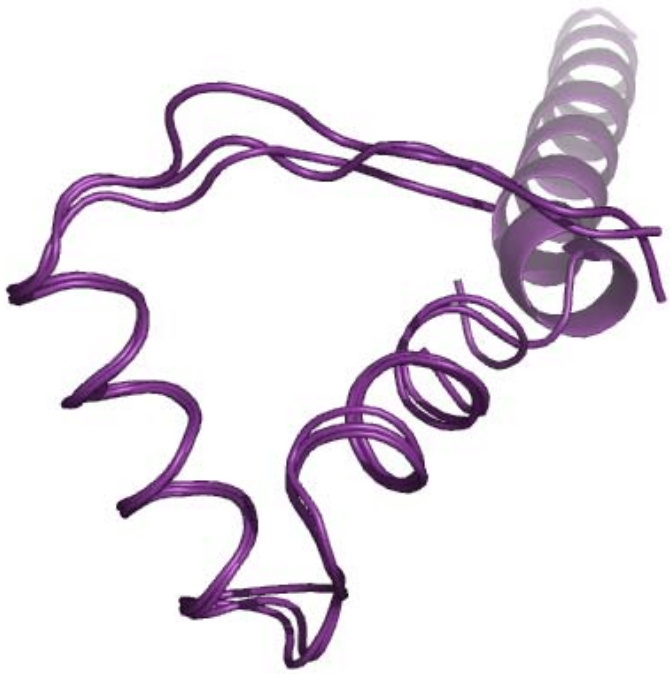


Figura R20. Superposición de las tres proteínas recogidas en la tabla R12. La superposición ha sido proporcionada por el propio servidor Phyre2, y la imagen se ha modificado mediante Pymol.

El modelado con Arc nos permite obtener una visión general del tetrámero de TrwA unido al ADN. La proteína consta del motivo RHH de unión al ADN, y en él identificamos las hojas beta formadas por el dímero de la proteína insertadas en el surco

mayor del ADN. En estas hojas beta se situarían los aminoácidos polares que llevarán a cabo contactos específicos con las bases nitrogenadas (figura R21).

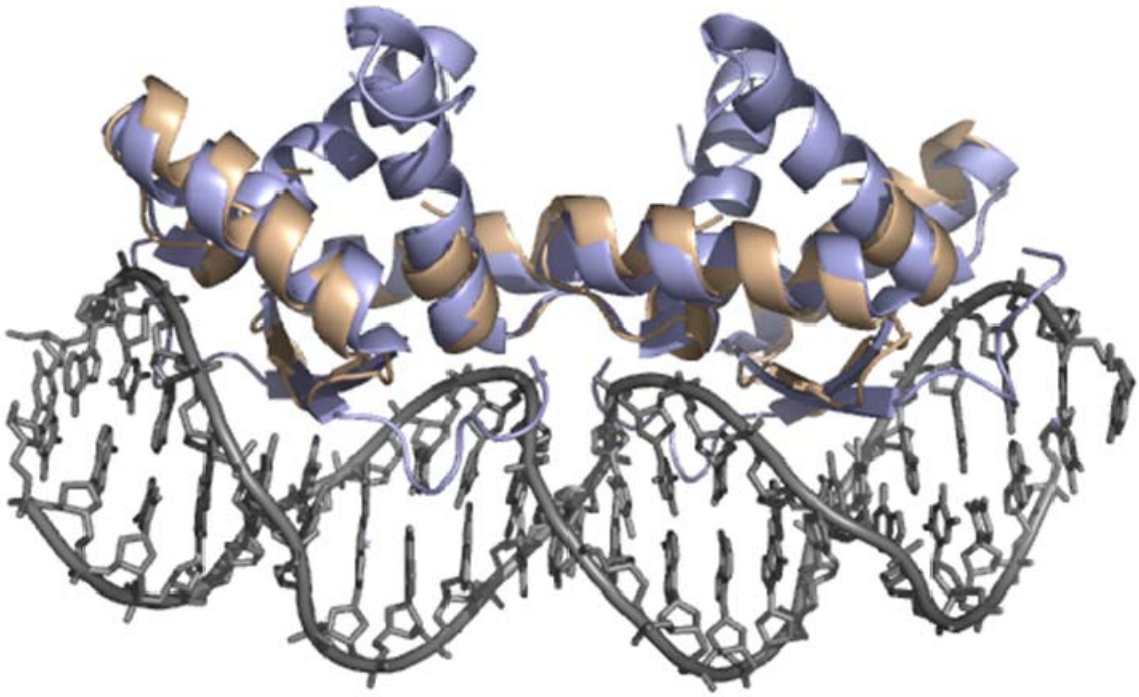


Figura R21. Modelo de TrwA unida a ADN. El molde utilizado inicialmente ha sido la proteína Bp1543, y posteriormente se ha alineado con la proteína Arc para poder visualizar el ADN. Arc aparece en azul y TrwA en color arena. El ADN se muestra en gris.

En el presente modelo, el tetrámero de TrwA coincide perfectamente con los dos dímeros de Arc que unen el ADN cooperativamente. El operador de Arc está formado por una repetición invertida, y aunque TrwA también une una repetición directa (*sbaA*), la unión no tendría por qué ser distinta, ya que debido a la simetría que muestra la proteína el hecho de que un dímero rotara para unir el brazo repetido de manera directa no modificaría la estructura general. Es necesario recordar que existen una serie de diferencias entre Arc y TrwA, siendo una de ellas que los dos brazos del IR o del DR reconocidos por TrwA están más separados que en el operador Arc.

Debemos tener en cuenta también que Arc sólo consta del dominio RHH, mientras que TrwA cuenta además con un dominio C-terminal responsable de la estabilización del tetrámero que no visualizamos en este modelo. De todas formas, es poco probable que este fragmento C-terminal de la proteína provoque alguna modificación en la unión al ADN, ya que se ha observado que los primeros 73 aa de la proteína son suficientes para lograr una unión eficiente (Moncalian and de la Cruz, 2004).

9. INTERACCIONES DE TRWA CON OTRAS PROTEÍNAS DEL RELAXOSOMA

Las interacciones que pueden llegar a darse entre las proteínas que conforman el relaxosoma han sido objeto de estudio, tanto en el sistema R388 utilizado en nuestro laboratorio, como en otros sistemas que han sido estudiados extensamente, como el de los plásmidos F (Lu *et al.*, 2008; Ragonese *et al.*, 2007; Ward *et al.*, 2002), RP4 (Schroder *et al.*, 2002; Szpirer *et al.*, 2000; Thomas and Hecht, 2007), y ColE1 (Varsaki *et al.*, 2012).

Mediante el estudio para determinar qué interacciones se dan entre estas proteínas, se trata de enlazar los dos eventos que conforman la conjugación como son el procesamiento del ADN y el transporte del mismo a la célula receptora. La sucesión de estos dos eventos implica que la relaxasa unida al ADN contacte con el poro conjugativo para que pueda ser transportada, siendo interesante conocer los contactos que pueda establecer con los distintos componentes que participan durante el proceso.

9.1. Interacción TrwA-TrwC

En el trabajo citado en la introducción (Llosa *et al.*, 2003), el análisis de las interacciones *in vitro* se llevó a cabo mediante la inmovilización de la proteína acopladora TrwB en una columna de afinidad, de la que eluiría junto con las proteínas que habían interaccionado con ella. Cuando la proteína inmovilizada en la columna era TrwA, se vio que TrwB era retenida junto con ella, pero no así TrwC.

Más recientemente, se publicó un trabajo en el que observaban una interacción directa entre TraM y el extremo C-terminal de la relaxasa TraI de F (Ragonese et al., 2007). También en ColE1 se ha observado la interacción entre la relaxasa MbeA y la proteína accesoria MbeC (Varsaki *et al.*, 2012). Basándonos en estos datos, y en ciertas evidencias experimentales que hacían creer probable una posible interacción entre TrwA y TrwC, decidimos volver a probar la interacción entre estas proteínas.

Para ello, se empleó proteína purificada TrwAhis que quedaría unida por afinidad a una columna de Niquel-agarosa y TrwC salvaje que solo coeluría junto con TrwA en el caso de que interaccionara directamente con ella, ya que carece de cola de histidinas. Se sabe por estudios previos de nuestro grupo (sin publicar) que TrwC se une a la columna de Niquel, por lo que en el tampón inicial se añadió imidazol a una concentración final de 70mM para evitar esta unión debida posiblemente a las histidinas del sitio activo de TrwC.

En un primer abordaje, se mezclaron TrwA y TrwC en proporción 1:1 en un tampón que contenía 100mM de NaCl, 50mM de TrisHCl a pH=7,2, 70mM de Imidazol y 10% de glicerol. Se incubaron las proteínas durante 10 minutos y se incorporaron a la columna que previamente había sido equilibrada con el mismo tampón. Se realizó una sucesión de lavados de un volumen de columna cada uno para eliminar el exceso de proteína no unida, y se recogieron las fracciones para su migración en un gel. Se procedió entonces a la elución de las proteínas aún unidas a la resina con un tampón con la misma composición que el anterior pero al que se le había añadido 500mM de Imidazol. Paralelamente, se realizó como control el mismo experimento pero esta vez sin TrwA, para comprobar que la cantidad de imidazol que habíamos usado en el tampón inicial era suficiente para que TrwC no se uniera inespecíficamente a la matriz de Ni-agarosa. Las fracciones recogidas de ambos experimentos se analizaron en geles desnaturalizantes de SDS-Poliacrilamida (figura R22).

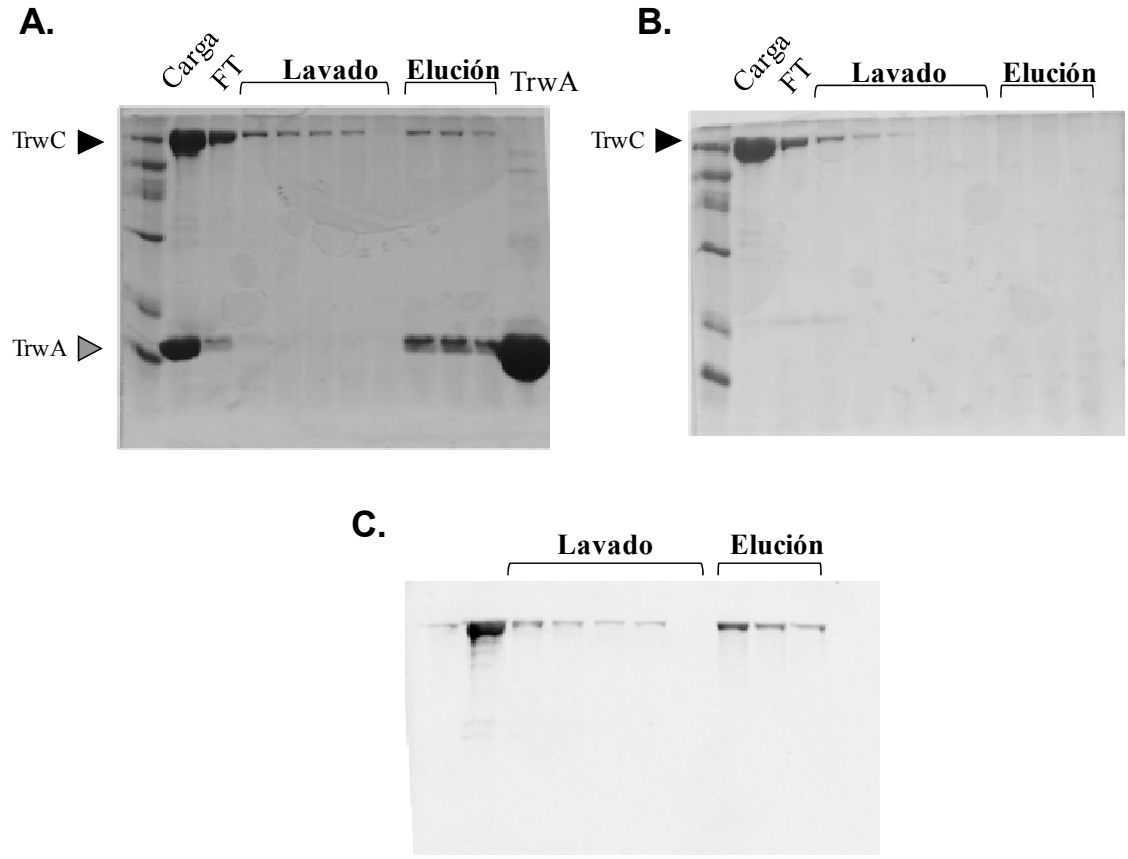


Figura R22. Interacción TrwAhis TrwC. **A.** Se observan las fracciones correspondientes a la carga en la columna, los distintos lavados y la elución con imidazol. En la última calle se muestra como control la altura a la que migra TrwA. **B.** Experimento control en el que no se añadió TrwA. **C.** Western Blot del gel A realizado con anticuerpo anti-TrwC.

El resultado obtenido parecía indicar que en presencia de TrwA, TrwC era retenida en la resina de Ni y que en ausencia de TrwA, toda la proteína se perdía en los sucesivos lavados, antes de la elución mediante imidazol.

Se decidió estudiar si era posible limitar la aparente interacción entre las dos proteínas a algún dominio en concreto, por lo que se realizó el mismo experimento, esta vez utilizando dos formas truncadas de TrwA. Una de estas proteínas fue $TrwA_{(2-73)}$, que mantiene la capacidad de unión a ADN al llevar este motivo situado en el extremo N-terminal, y la otra proteína fue $TrwA_{(36-121)}$, que contiene el dominio de tetramerización de la proteína (Moncalian and de la Cruz, 2004).

Tal y como se observa en la figura R23, TrwC coeluía con ambas formas de TrwA, aunque en el caso de $TrwA_{(2-73)}$ es más nítida la banda que corresponde a la relaxasa. En los controles negativos, TrwC eluye en los distintos lavados.

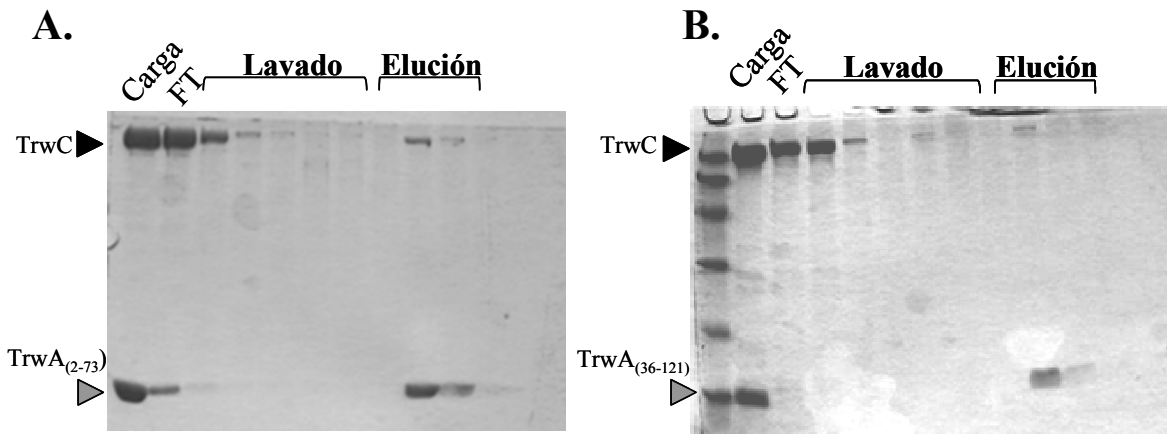


Figura R23. Interacción de los dominos N- y C-terminal de TrwA con TrwC**. A.** $TrwA_{(2-73)}$ TrwC. **B.** $TrwA_{(36-121)}$ TrwC. Ver detalles del experimento en el texto.

En la relaxasa TrwC también podemos hacer una distinción entre dominios. Podemos dividir la proteína en un dominio relaxasa que abarcaría los primeros 293 aminoácidos y un dominio helicasa que comprendería el resto de la proteína (Llosa *et al.*, 1996). Decidimos probar estas dos construcciones con TrwA entera para intentar limitar la zona de interacción entre las dos proteínas (figura R24).

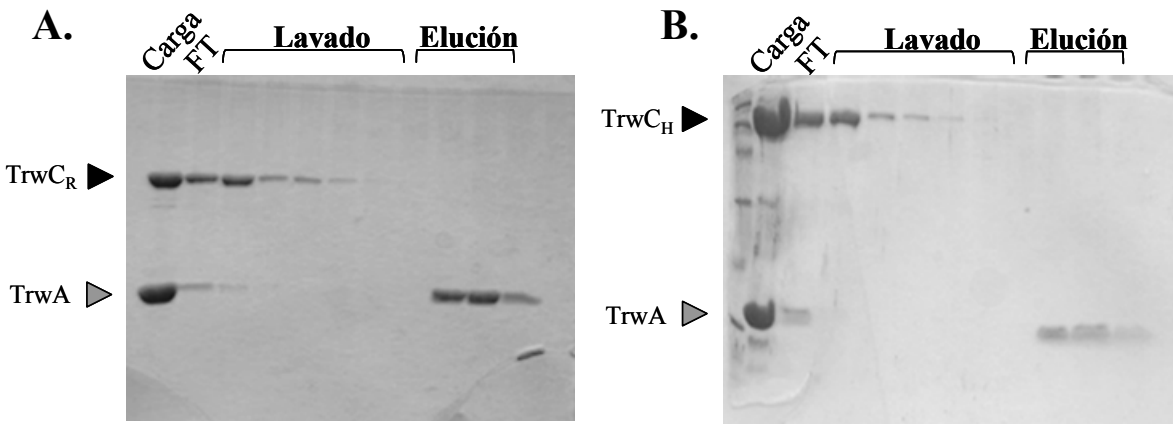


Figura R24. Interacción TrwC$_R$ y TrwC$_H$ con TrwA. **A.** Se observan las fracciones obtenidas en el experimento realizado con TrwC$_R$ y TrwA. **B.** Experimento realizado con TrwC$_H$ y TrwA.

Ningún fragmento ensayado de la relaxasa TrwC coeluía con TrwA, sino que toda la proteína había eluído previamente con los lavados, por lo que no parecía haber interacción detectable.

Antes de seguir intentando delimitar el sitio de posible interacción entre ambas proteínas, y creyendo que nos podría aportar más información, decidimos coexpresar las dos proteínas *in vivo* en C41. El gen *trwA* está clonado en un vector que le añade una cola de histidinas en el extremo C-terminal, mientras la relaxasa está clonada en un vector de expresión sin cola de histidinas. El fundamento del experimento sería el mismo que el realizado *in vitro* con las proteínas purificadas, TrwA se une por afinidad a la columna de histidinas y TrwC solo se uniría a la columna en el caso de que interaccionara con TrwA. En este caso, también se realizó un control negativo que consiste en la sobreexpresión de TrwC con el vector de expresión utilizado para TrwA vacío (pET29c) para evitar falsos positivos. Una vez sobreexpresadas y lisadas las células, se pasaron los extractos proteicos en tampón 100mM NaCl, 50mM Tris pH=7,5, 70mM imidazol por una columna de Ni y se eluyó utilizando un gradiente de un volumen de columna con un tampón que contenía 100mM NaCl, 50mM Tris pH=7,5, 0,5M Imidazol (figura R25).

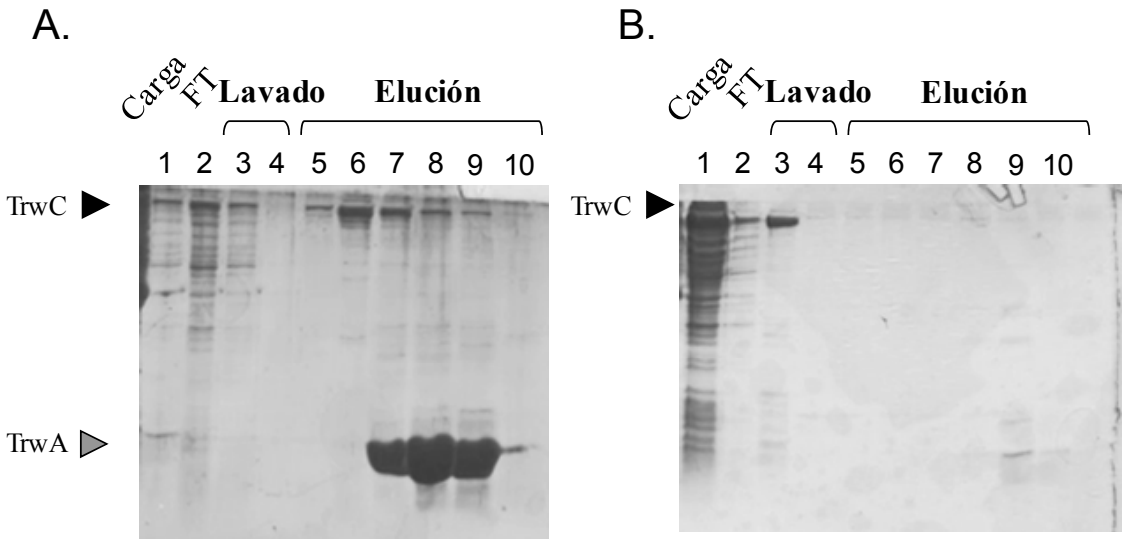


Figura R25. Interacción TrwA TrwC obtenida por la sobreexpresión conjunta de las proteínas. **A.** Se observan las fracciones obtenidas en el experimento realizado con ambas proteínas. **B.** Experimento control en el que se coexpresó TrwC con el vector vacío de TrwA.

Tal y como se observa en la figura superior R25A, aunque la elución de las dos proteínas coincide en algunas fracciones (7, 8 y 9) el patrón de elución parece independiente para cada una de ellas. En el caso de que hubiera una interacción clara, resultaría esperable ver no sólo que las proteínas coeluyen, sino que lo hacen de manera proporcional, coincidiendo los picos de elución en una concentración determinada de imidazol.

9.2. Efecto de IHF en la unión de TrwA

Las proteínas accesorias que participan en el relaxosoma pueden estar codificadas por el plásmido o por el cromosoma de la bacteria huésped. Concretamente, IHF, es una proteína codificada por el cromosoma bacteriano involucrada en varias funciones celulares que necesitan complejos proteína-ADN de mayor orden, como puede ser la replicación del ADN, la regulación transcripcional, y una variedad de sistemas de recombinación sitio específica (Friedman, 1988).

En cuanto a su papel en la conjugación bacteriana, se cree que IHF, junto con otras proteínas accesorias codificadas por el plásmido, es capaz de producir una distorsión en el ADN para generar una región con carácter de ADN de cadena sencilla en los alrededores el sitio *nic*, de tal manera que la relaxasa pueda unirse a esta región modificada y producir la reacción de corte específica (Byrd and Matson, 1997).

Se han descrito sitios de unión de IHF en orígenes de transferencia pertenecientes a plásmidos englobados en la familia de relaxasas MOB_F, como pueden ser el plásmido F o R388, entre otros (Moncalian *et al.*, 1999b; Tsai *et al.*, 1990). En plásmidos como RP4 (MOB_P), o RSF1010 (MOB_Q), IHF no participa en el procesamiento del ADN durante la conjugación. En el plásmido RSF1010, la ausencia de dependencia de proteínas codificadas por el cromosoma huésped, puede obedecer a la naturaleza de su sistema de replicación funcional en varios citoplasmas diferentes. Además, la carencia de sitios de unión para este tipo de proteínas en RSF1010, permite también que el *oriT* sea mucho más compacto (Brasch and Meyer, 1987).

Tanto en el *oriT* de F como en el de R388 se han descrito dos sitios de unión para IHF. La distribución de los sitios es similar en ambos plásmidos, el sitio de unión *ihfA* se encuentra entre el sitio *nic* y el primer sitio de unión de la proteína accesoria *sbaB*, mientras que el sitio *ihfB* está situado en la región del promotor regulado por la accesoria. Al situarse estos sitios de unión tan cerca de los sitios de unión de TrwA, quisimos comprobar si tenían un efecto en la unión de TrwA al *oriT* de R388.

Ya se había comprobado con anterioridad que tanto TrwA como IHF eran capaces de unirse *in vitro* a *oriT* de R388 por separado (Moncalian *et al.*, 1997; Moncalian *et al.*, 1999b). Así, en este caso realizamos un ensayo de retardo en gel conjunto, variando la concentración de una de las proteínas mientras que la otra se mantenía constante, para ver si al añadir las dos se observaba una diferencia en el patrón de migración del ADN. Vemos el resultado del EMSA realizado en la figura R26.

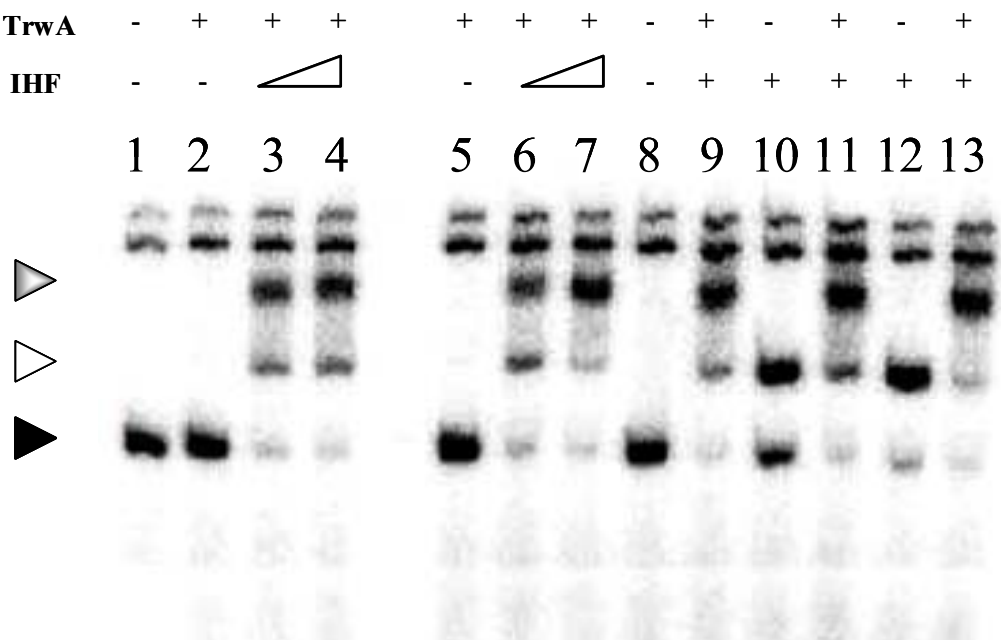


Figura R26. Ensayo de retardo conjunto de TrwA e IHF. Se representan mediante cabezas de flecha los complejos de ADN sin unir (negro), y los complejos ADN-proteína (blanco para ADN unido a IHF y gris para ADN unido a IHF y TrwA). La banda inferior marcada corresponde al fragmento que contiene el $oriT_{R388}$, mientras que las dos bandas superiores corresponden al resto del vector digerido. Calle 1: ADN en ausencia de proteínas. Calle 2: 18nM de TrwA. Calle 3: 18nM de TrwA, 2nM IHF. Calle 4: 18nM TrwA, 4nM IHF. Calle 5: 37nM TrwA. Calle 6: 37nM TrwA, 2nM IHF. Calle 7: 37nM TrwA, 4nM IHF. Calle 8: 74nM TrwA. Calle 9: 74nM TrwA, 2nM IHF. Calle 10: 2nM IHF. Calle 11: 2nM IHF, 74nM TrwA. Calle 12: 4nM IHF. Calle 13: 4nM IHF, 74nM TrwA.

El resultado obtenido indica que cuando una de las proteínas había sido añadida a la reacción, al añadir la otra, se producía un complejo nucleoprotéico de mayor tamaño que mostraba una menor movilidad en el gel. Este "super-retardo" se observa incluso cuando la concentración de la proteína añadida en primer lugar no era suficiente como para provocar un primer retardo.

En la figura R26, se observan dos complejos ADN-proteína que muestran una movilidad distinta: el de mayor velocidad de migración corresponde al complejo formado por la unión de IHF a *oriT*, mientras que y el de menor velocidad equivale al complejo TrwA-IHF-*oriT*. En este gel no tenemos constancia del complejo formado por TrwA-*oriT*, por lo que a continuación se muestra otro gel, que a pesar de tener una débil

señal radiactiva, nos da una idea de a qué altura migrarían este complejo TrwA-*oriT* y el formado por la unión de las dos proteínas TrwA e IHF y *oriT* (figura R27).

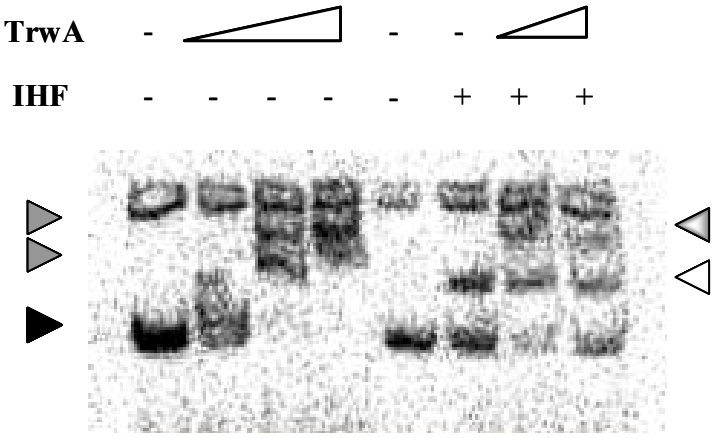


Figura R27. Ensayo de unión conjunto con TrwA e IHF. Mediante este ensayo es posible comparar las velocidades de migración correspondientes al complejo formado por TrwA-ADN y le formado por IHF-TrwA-ADN. Dos puntas de flecha grises han sido añadidas para señalar la localización del complejo TrwA-ADN y la punta de flecha gris y blanca indica el complejo TrwA-IHF-ADN. La punta de flecha blanca señalaría el complejo IHF-ADN. Las cantidades de proteína que se han añadido son las siguientes: calle 2; 92nM TrwA, calle 3; 185nM TrwA, calle 4; 370nM TrwA, calle 5; no se añadió proteína, calle 6; 2nM IHF, calle 7; 2nM IHF y 92nM TrwA, calle 8; 2nM IHF y 185nM TrwA.

La adición de cantidades crecientes de TrwA da lugar a la aparición de dos complejos ADN-proteína que muestran distinta movilidad. Estos complejos podrían corresponder a la proteína unida a uno de los operadores o a ambos. TrwA podría reconocer en primer lugar el sitio de unión *sbaA* por el que muestra más afinidad (apartado 3 de la presente sección), y al añadir una mayor concentración de proteína a la reacción uniría también el sitio de unión *sbaB* formando un complejo nucleoprotéico mayor que migraría más lentamente. Este complejo de mayor tamaño migra aparentemente a la misma altura que el supuesto complejo formado por TrwA-IHF-ADN, sin embargo, la concentración de TrwA necesaria para formar ese complejo es mucho menor cuando IHF está presente (370nM frente a 92nM).

Se ha elaborado una figura a modo explicativo para ilustrar la hipótesis de unión secuencial de TrwA que mencionamos en el párrafo anterior (figura R28)

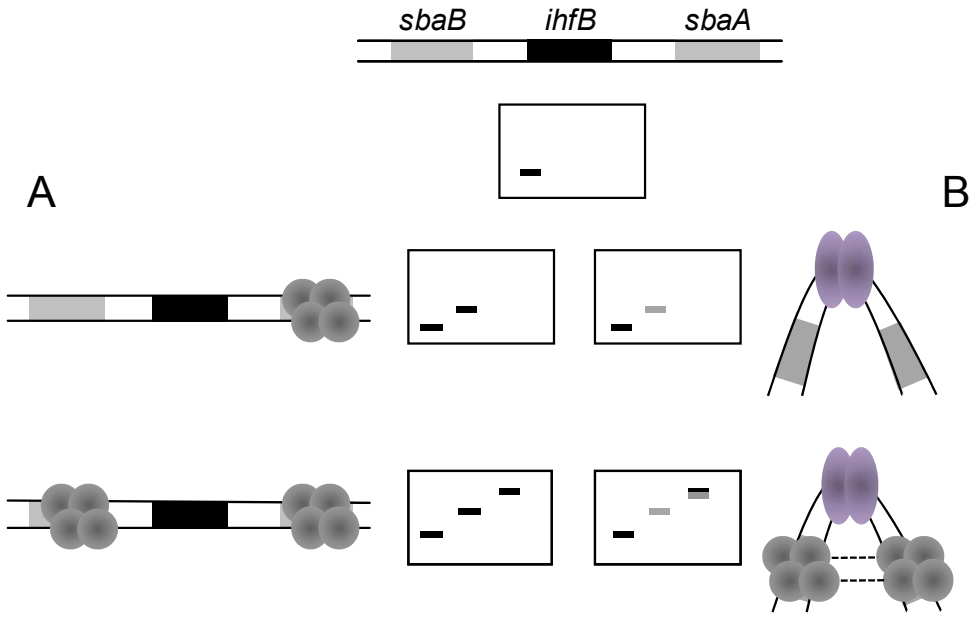


Figura R28. Representación esquemática del proceso de unión de TrwA en presencia y ausencia de IHF. TrwA se representa mediante esferas grises, e IHF mediante elipses moradas. A. En ausencia de IHF, TrwA reconocería en primer lugar el sitio *sbaA* de mayor afinidad, y al aumentar la cantidad de proteína uniría ambos sitios de unión. B. En presencia de IHF, ésta reconoce su sitio de unión y produce una torsión en el ADN que acercaría los dos sitios de unión de TrwA permitiendo así que la proteína establezca interacciones favorables entre los tetrámeros que unirán ambos sitios al mismo tiempo.

10. NUEVAS PROTEÍNAS ACCESORIAS SIMILARES A TrwA

Recientes análisis utilizando la herramienta Blast de Pubmed, nos han permitido identificar algunas proteínas que podrían tener un papel similar a $TrwA_{R388}$. Las proteínas que se describen a continuación se encuentran codificadas tanto en plásmidos como en el cromosoma y aparecen descritas como proteínas hipotéticas o que participan en movilización. Hemos analizado su localización y las proteínas adyacentes a las mismas para determinar su posible función en el proceso de la conjugación bacteriana.

En la búsqueda de homólogos por Blast hemos obtenido 11 proteínas similares a TrwA con un valor E<0.002, una cobertura mayor del 75% y una identidad superior al 26% (tabla R13).

Proteína	**Cobertura (%)**	**Valor E**	**Identidad (%)**
$TrwA_{pIE321}$	100	$1x10^{-66}$	88
$TrwA_{R7K}$	100	$2x10^{-63}$	84
Proteína de movilización (*Bordetella parapertussis*)	96	$2x10^{-44}$	64
Proteína de movilización (*Pseudomonas stutzeri* CCUG29243)	95	$2x10^{-6}$	37
Proteína hipotética C660_01185 (*Alcaligenes* sp. HPC1271)	97	$3x10^{-6}$	30
Proteína hipotética QWA_17153 (*Alcaligenes faecalis* NCIB 8687)	97	$3x10^{-6}$	31
Proteína hipotética Bpet3700 (*Bordetella petrii* DSM 12804)	88	$2x10^{-5}$	32
Proteína hipotética Gobo1_18726 (Gluconacetobacter oboediens 174Bp2)	73	$8x10^{-5}$	31
Proteína hipotética PT7_1767 (*Pusillimonas* sp. T7-7)	87	$2x10^{-4}$	32
Proteína hipotética Pcryo_2492 (Psychrobacter cryohalolentis K5)	90	$4x10^{-4}$	26
Proteína hipotética A2cp1_1112 (Anaeromyxobacter dehalogenans 2CP-1)	89	0.002	30

Tabla R13. Lista de homólogos de TrwA obtenidos por Blast. La cobertura muestra el porcentaje de la proteína que se ha utilizado para su comparación, el valor E es un parámetro que muestra el número de coincidencias que uno puede encontrarse al azar al buscar en una base de datos y la identidad representa el porcentaje de aminoácidos que coinciden en ambas proteínas.

Para establecer los posibles sitios de unión de estas proteínas tipo TrwA se llevó a cabo un alineamiento de las secuencias que identificamos como los posibles orígenes de transferencia, comparándolos con los *oriT* ya caracterizados de R388 y R7K (desglosado en las figuras R29, R30 y R31). Para ello, cabe destacar que en todos los orígenes de transferencia analizados aparece una secuencia altamente homóloga al motivo IR2 y al sitio *nic* de unión de la relaxasa TrwC. El objetivo era encontrar posibles operadores de las proteínas accesorias, para ver si mostraban una estructura

similar a aquella que encontramos en R388 y R7K (dos sitios de unión que muestran un IR cercano al sitio *nic* y un DR adyacente al inicio del gen de la proteína accesoria). A continuación se describen los sistemas analizados.

10.1 *Pseudomonas stutzeri*

La proteína implicada en movilización de *Pseudomonas stutzeri* CCUG 29243 aparece en cuarta posición (segunda obviando $TrwA_{pIE321}$ y $TrwA_{R7K}$), compartiendo un 37% de secuencia aminoacídica con $TrwA_{R388}$ y un 40% con $TrwA_{R7K}$ (figura R31). Si analizamos su posible origen de transferencia, los hipotéticos sitios de unión de la proteína se encuentran exactamente en la misma posición que en R388, y el hexanucleótido (GACTAC) que se repite para formar estos sitios coincide con aquel encontrado en el plásmido pMBUI4 (figura R29). El plásmido pMBUI4 fue identificado como plásmido similar a los MOB_F en el transcurso de esta Tesis Doctoral (comunicación personal de la Dra. Eva Top). El origen de transferencia de este plásmido diverge más respecto a R388 que el resto de los MOB_F identificados hasta la fecha.

Alineando lo que consideramos como el origen de transferencia completo de pMBUI4 y *P.stutzeri*, la identidad de secuencia es de un 63% aproximadamente. Si incluimos el *oriT* de R388 en este alineamiento, la identidad de secuencia disminuye hasta un 38%. Hay zonas en la que la identidad es mayor, como en los ya mencionados supuestos sitios de unión de TrwA y de la relaxasa. Encontramos, además, tanto en pMBUI4 como en *P. stutzeri* un sitio de unión de IHF en la misma posición que en R388 y R7K, aunque carecen del segundo sitio de unión de IHF (*ihfB*) que sí aparece en R388. Otros elementos de estructura secundaria como IR1 han sido identificados en pMBUI4 y *P. stutzeri*, los cuales coinciden en su localización con R388 pero no así en su secuencia ni en la longitud del espaciamiento entre los brazos que forman la repetición.

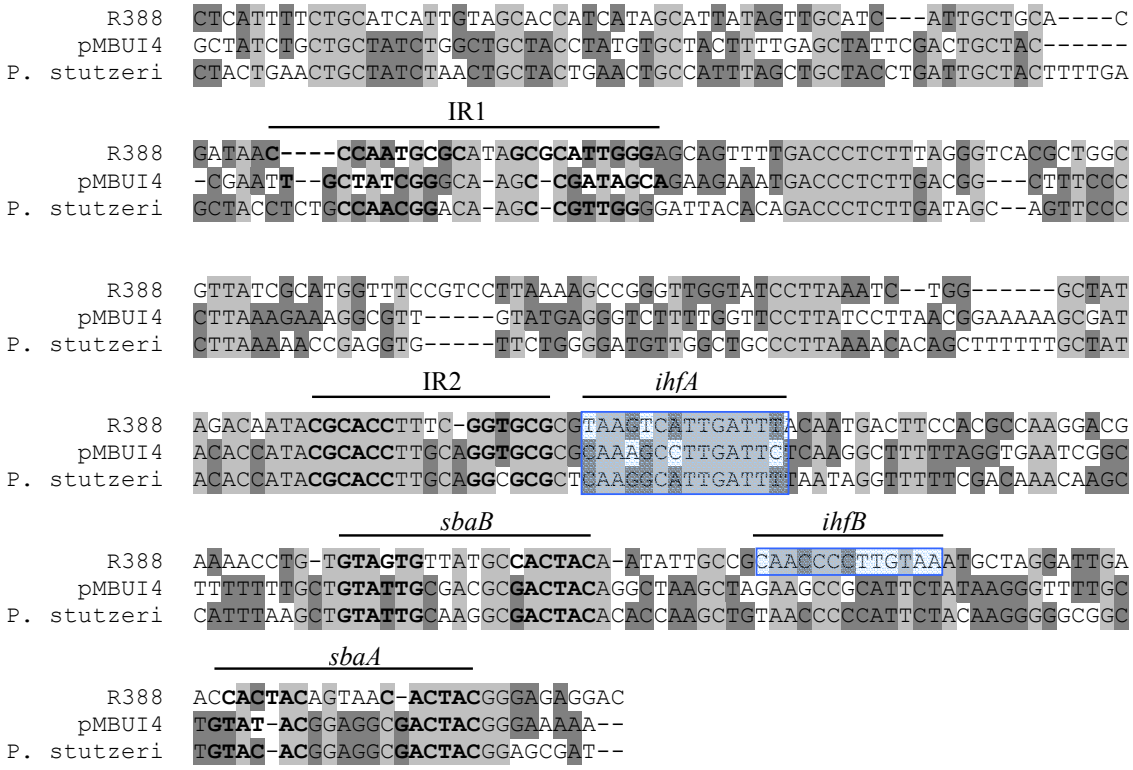


Figura R29. Alineamiento de los orígenes de transferencia de R388, pMBUI4 y *P. stutzeri*. Sombreados en gris claro aparecen los nucleótidos que se conservan en todos ellos, y en gris oscuro aquellos que se conservan solo en algunos. En negrita se realzan los nucleótidos que componen las repeticiones, y encuadrados en azul claro aparecen los sitios de unión de IHF. Los nombres de los sitios de unión de las proteínas o demás elementos de estructura secundaria se muestran en la parte superior de la secuencia.

Centrando nuestro análisis en la secuencia previa al inicio del gen que codifica para TrwA en pMBUI4, encontramos el hexanucleótido GACTAC que se repite en dos regiones. En el supuesto operador que coincide en el espacio con $sbaA_{R388}$, esta secuencia se encuentra acompañada de otro hexanucleótido de secuencia GTATAC, por lo que las diferencias en las bases 2 y 3 rompen lo que sería una repetición directa perfecta como aquella de R388. El número de pb entre los dos brazos es igual que en R7K, siendo este de 6pb en vez de 5 como en R388. En el equivalente al sitio $sbaB_{R388}$, el brazo formado por la secuencia GACTAC va acompañado por el hexanucleótido GTATTG. En este caso tampoco encontramos una repetición invertida perfecta, ya que para esto el segundo brazo debería estar compuesto por la secuencia GTAGTC en vez

de GTATTG. El espaciamiento en este caso también es de 6 pb, lo que coincide con R388 y R7K.

El supuesto sitio *sbaB* de *P. stutzeri* coincide plenamente con el de pMBUI4, aunque el hipotético sitio *sbaA* difiere en un nucleótido de uno de los brazos. El sitio *sbaA* en R388 y R7K está compuesto por una perfecta repetición directa de un hexanucleótido en el que los dos brazos de la repetición se encuentran separados por 5 pb en R388 y 6 pb en R7K. En pMBUI4 y en *P. stuzeri* se mantiene la longitud del espaciamiento de R7K, pero en estos casos la repetición directa es imperfecta, coincidiendo 4 de los nucleótidos en ambos brazos en pMBUI4 (posición 1, 4, 5 y 6) y 3 en *P. stutzeri* (posición 1,5 y 6).

Una vez observada la similitud que mostraban los posibles orígenes de transferencia de pMBU14 y *P. stutzeri* procedimos a comparar las hipotéticas proteínas TrwA de ambos. No nos sorprendió comprobar que ambas comparten un 83% de secuencia aminoacídica, siendo esta similitud aún mayor (97%) en el dominio RHH de unión a ADN (figura R31).

10.2 Bordetella parapertussis

Con mayor identidad de secuencia con TrwA, en la tercera posición en la lista, anterior a la proteína de *P. stutzeri*, se encuentra la proteína implicada en movilización del plásmido BPP5P1 aislado de *Bordetella parapertussis*. Esta proteína muestra una identidad de secuencia de 64% con $TrwA_{R388}$ y de 75% con $TrwA_{R7K}$, y cuando contrastamos su secuencia con programas de predicción de estructura secundaria los resultados revelan que pertenece a la familia de las RHH. En cuanto a los posibles sitios de unión de la proteína, hemos podido identificar dos sitios de unión que coinciden plenamente con los de R7K (figura R30). Encontramos el hexanucleótido GACCAC repetido de manera tanto directa como invertida, con un espaciamiento de 6 pb entre los brazos. Se han hallado también dos sitios de unión de IHF en las mismas posiciones que en R388 y a la misma distancia uno del otro.

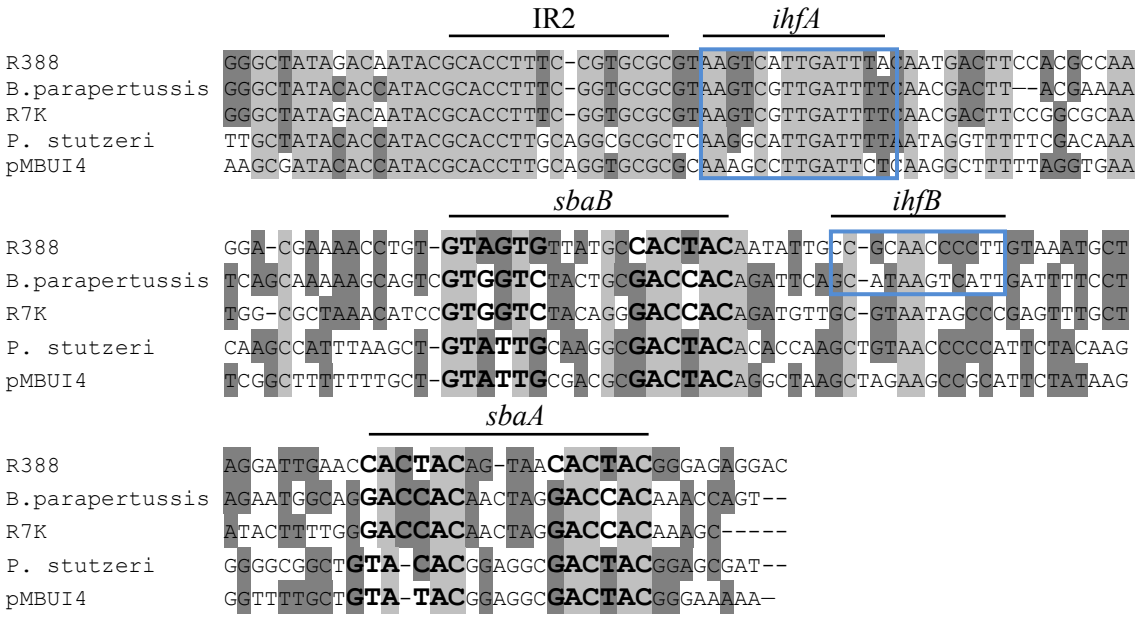


Figura R30. Alineamiento de la región que abarca desde el sitio de unión de la relaxasa hasta los sitios de unión de TrwA de R388, R7K, BPP5P1, *P. stutzeri* y pMBUI4. El código de colores es igual al utilizado en la figura R29.

En términos generales, la identidad de secuencia del *oriT* completo de R388 y BPP5P1 es del 60%, y del 65% si lo comparamos con R7K. Este plásmido tiene algunas características singulares, ya que aún siendo el más próximo a los IncW en cuanto a su TrwA, no pertenece a esta familia y tampoco presenta relaxasa ni proteína acopladora. Las proteínas del plásmido implicadas en replicación, en partición o la resolvasa no muestran homología con R388 o ningún otro plásmido perteneciente a IncW, tampoco aquellas descritas como hipotéticas muestran, aparentemente, ninguna similitud. En cambio, si contiene otras proteínas implicadas en conjugación como TrwI, TrwJ y eex (exclusión de entrada), las cuales muestran un 70%, 84% y 66% de identidad de secuencia con R388 respectivamente (69%, 85% y 61% con R7K).

10.3 Otras proteínas RHH

Siguiendo el orden que nos proporciona la herramienta Blast, a continuación aparecen cinco proteínas descritas como hipotéticas en las que no hemos sido capaces de identificar unos sitios de unión similares a TrwA. Estas proteínas que se muestran en

la tabla R6 corresponden a distintos organismos: C660_01185 (Alcaligenes sp. HPC1271), QWA_17153 (*Alcaligenes faecalis* NCIB 8687), Bpet3700 (*Bordetella petrii* DSM 12804), Gobo1_18726 (*Gluconacetobacter oboediens* 174Bp2) y PT7_1767 (Pusillimonas sp. T7-7)(figura R31). En todas las secuencias intergénicas anteriores a los genes que las codifican hemos encontrado secuencias de estructura secundaria como repeticiones y secuencias palindrómicas que pueden ser potenciales sitios de unión de proteínas, pero al no seguir el patrón de una secuencia que se repite de manera directa como invertida que venimos analizando no se describirán en detalle (se muestra un resumen en la figura R32). A continuación, aparecen otras dos proteínas hipotéticas, Pcryo_2492 (*Psychrobacter cryohalolentis* K5, plásmido 1) y A2cp1_1112 (*Anaeromyxobacter dehalogenans* 2CP-1) que muestran respectivamente una identidad de secuencia de 26% y 30% con $TrwA_{R388}$ y de 26% y 32% con $TrwA_{R7K}$. Ambas proteínas contienen un motivo RHH en su extremo amino terminal según los resultados obtenidos por el programa de predicción de estructuras secundarias Jpred.

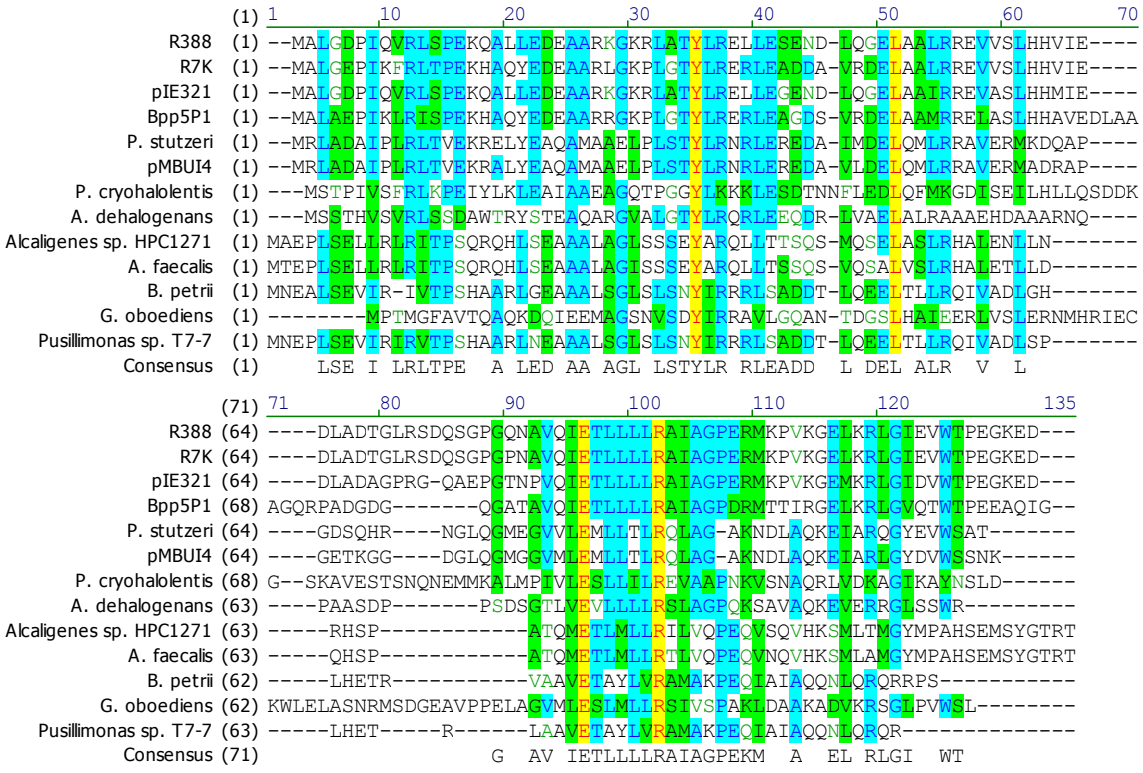


Figura R31. Alineamiento de las proteínas homólogas de TrwA proporcionadas por Blast. En sombreado amarillo aparecen los aminoácidos idénticos que se mantienen, en azul los altamente conservados y en verde los que muestran similitud. Sin sombrear, los aminoácidos e n verde indican poca similitud, y los que aparecen con caracteres negros son los no conservados.

Al analizar los posibles orígenes de transferencia, identificamos posibles sitios de unión de la proteína en ambas, aunque la proteína A2cp1_1112 muestra mayor similitud con la distribución de R388. Precisamente en el posible *oriT* de *Anaeromyxobacter dehalogenans* encontramos el hexanucleótido CTGTCA repetido de manera tanto directa como invertida. En la repetición directa, se mantiene el espaciamiento entre los dos brazos es de 5 pb como en R388, pero en la repetición invertida este espaciamiento pasa a ser de 9 pb en lugar de las 6 pb del sitio *sbaB* de R388. En *Psychrobacter cryohalolentis* se ha identificado un hexanucleótido con simetría interna que se repite de manera directa en dos posiciones. El hexanucleótido se presenta como dos tripletes espaciados por dos pb que no se conservan (TGTnnACA), y a su vez, los dos hexanucleótidos que conformarían un sitio de unión se encuentran a 3 pb de distancia en los dos sitios de unión (figura R32). Se ha identificado un posible sitio de unión de IHF entre los dos sitios de unión de la proteína. En el plásmido 1 de *P. cryohalolentis* no hay descrita ninguna proteína implicada en movilización, y al analizar las proteínas hipotéticas adyacentes a la homóloga de TrwA observamos que no muestran similitudes con proteínas de este tipo. En *A. dehalogenans*, en cambio, además de la posible proteína auxiliar de la conjugación, encontramos a continuación la proteína acopladora que muestra una identidad de secuencia de 68% con $TrwB_{R388}$ y seguidamente se sitúa la relaxasa, similar en un 63% a $TrwC_{R388}$.

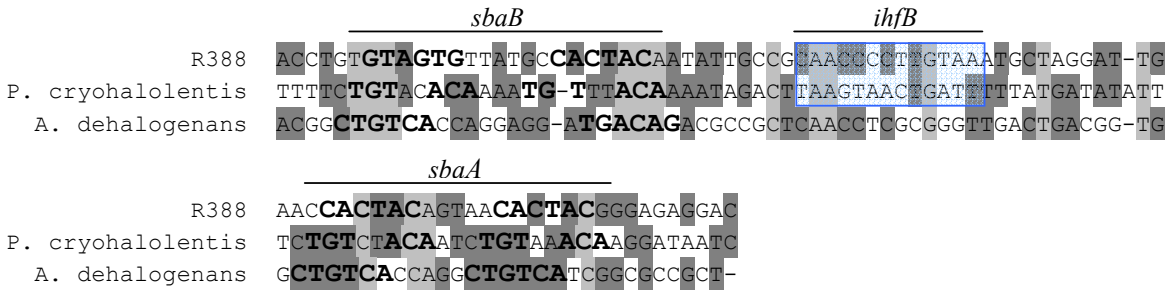


Figura R32. Alineamiento de la región de los operadores de TrwA de R388 y los posibles sitios de unión de la proteína de *P. cryohalolentis* y *A. dehalogenans*. El código de colores es el utilizado en las figuras R29 y R30.

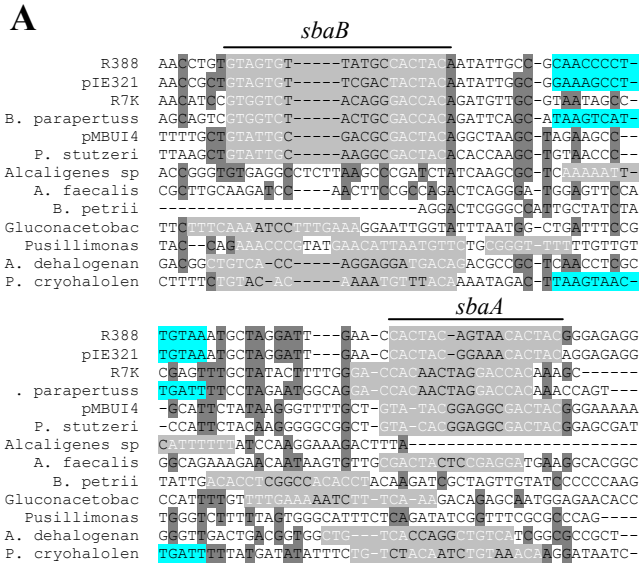
A modo de resumen, en general, en todas las secuencias intergénicas previas a las proteínas homólogas de TrwA se han podido identificar posibles sitios de unión en lugares próximos a los de R388. En algunos casos, solo hemos sido capaces de establecer un sitio de unión (*B. petrii,* Alcaligenes sp. HPC1271, *A. faecalis*), y en otros, aún encontrando dos sitios de unión, no se encuentran distribuidos a modo de repetición invertida y directa, sino que muestran la misma estructura en ambos (*Gluconacetobacter, P. cryohalolentis*). En la mayoría de los casos, los brazos que forman las repeticiones constan de seis pb, a excepción de *Gluconacetobacter* y *Pusillimonas* en los que los brazos constan de 7pb. En *Pusillimonas*, la estructura identificada es algo distinta a todas las demás. Entre los dos brazos de la repetición invertida se sitúa una secuencia palindrómica que abarca catorce de las diecinueve bases espaciadoras. Aunque en la repetición como en el palíndromo es un heptanucleótido el que se repite de manera inversa, la secuencia no coincide (figura R33).

En lo que respecta a las proteínas, en la figura R33 se muestra la secuencia de las hojas beta de unión a ADN de las mismas. Si tomamos como referencia la hebra beta de R388, observamos cómo existe una alternancia de aminoácidos hidrofóbicos y aminoácidos polares, siendo estos últimos los implicados en establecer contactos específicos con el *oriT* de R388 (Gln8, Arg10, Ser12). En casi todas las proteínas RHH homólogas de TrwA analizadas, podemos encontrar una distribución similar, aunque existen algunas excepciones. En *P. stutzeri* (y en pMBUI4, ya que comparten la misma β), el primer aminoácido polar de la hoja beta es sustituido por una prolina, aunque estas sustituciones puntuales son una característica presente en otras proteínas que no parece perjudicar la unión (Schreiter and Drennan, 2007). *Gluconacetobacter* muestra una hebra beta parcialmente distinta a lo habitual en las proteínas RHH, y en este caso resulta más difícil establecer la posición de los aa polares. En BPP5P1, sí se mantiene la típica distribución de la hebra beta, mostrando la misma secuencia que $TrwA_{R7K}$ en los dos primeros aa polares supuestamente implicados en el reconocimiento específico (Lys8, Arg10).

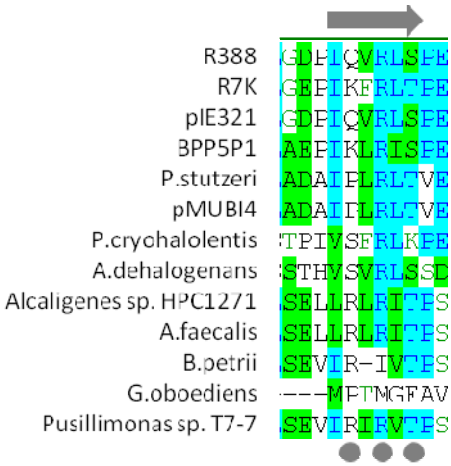
Comparando los diferentes sistemas, vemos que la homología de secuencia de las proteínas TrwA encontradas se corresponde con una homología secuencial y estructural de los correspondientes *oriT*s. Esta similitud en las proteínas auxiliares y los

oriT cabe esperar que se refleje en un mecanismo homólogo de procesamiento conjugativo del ADN.

A



B



C

Plásmido/Organismo	Repetición		Secuencia
	DR	IR	
R388	✓	✓	CACTAC
pIE321	✓	✓	CACTAC (primer nt mutado en la IR)
R7K	✓	✓	GACCAC
BPP5P1	✓	✓	GACCAC
Pseudomonas stutzeri	✓	✓	GACTAC
pMBUI4	✓	✓	GACTAC
Alcaligenes sp		✓	AAAAAT
Alcaligenes faecalis	✓		CGAnnA
Bordetella petrii	✓		ACACCT
Gluconacetobacter		✓	TTTCAAA
Pusillimonas		✓	AAACCCG | GAACATT
A. dehalogenans	✓	✓	CTGTCA
P. cryohalolentis	✓		TGTnnACA

Figura R33. Distribución y composición de los hipotéticos sitios de unión y las láminas beta de las proteínas homólogas a TrwA. **A** Se muestra un alineamiento de las zonas intergénicas previas a las proteínas donde se resaltan en gris claro los supuestos sitios de unión. La secuencia repetida de manera directa como invertida se muestra en letras blancas. En azul se han resaltado los posibles sitios de unión de IHF encontrados en posición parecida a *ihfB* de R388. **B** Alineamiento de las láminas beta de las proteínas homólogas a TrwA. La posición de la lámina beta se representa con una flecha gris y los aminoácidos polares implicados en la unión de $TrwA_{R388}$ con puntos del mismo color. Los aa sombreados en azul son los altamente conservados y en verde los que muestran similitud. Sin sombrear, los aminoácidos en verde indican poca similitud, y los que aparecen con caracteres negros son los no conservados. **C** En la tabla inferior se resalta si la distribución se da a modo de repetición directa o invertida y la secuencia que se repite en cada uno de los casos.

5.

DISCUSIÓN

5. DISCUSIÓN

Previamente a este trabajo, se habían realizado estudios muy esclarecedores sobre las características de la proteína TrwA de R388 y su unión a ADN, así como de las funciones que lleva a cabo en la conjugación (Moncalian *et al.*, 1997; Moncalian and de la Cruz, 2004) y su participación en los eventos de recombinación e integración llevadas a cabo por la relaxasa TrwC (Cesar *et al.*, 2006; Draper *et al.*, 2005; Llosa *et al.*, 1994b).

Se sabía también que TrwA pertenece a la superfamilia de reguladores transcripcionales RHH y que guarda similitudes funcionales con otras proteínas accesorias de la conjugación como puede ser TraY de F (Nelson *et al.*, 1995; Silverman and Sholl, 1996). De hecho, en el tiempo que ha durado este trabajo, se han llegado a identificar nuevas proteínas accesorias que comparten con TrwA este dominio de unión a ADN, como MbeC de ColE1 (Varsaki *et al.*, 2009) o NikA de R64 (Yoshida *et al.*, 2008). También se han incorporado a este grupo otras como TraM de los plásmidos de la familia de F, la cual se creía que pertenecía a las H-T-H hasta que la estructura de TraM de pED208 fue resuelta recientemente (Wong JJ, 2011).

En un trabajo llevado a cabo por nuestro grupo de investigación (Revilla et al., 2008), se compararon 5 plásmidos englobados en la familia IncW, mostrando que compartían una gran similitud de secuencia en su genoma y permitiendo definir el eje central común de este grupo. El único gen del *backbone* en el que se encontró una variación significativa fue *trwA*. Además, se observó que las variaciones en la secuencia de éste gen se correlacionaban con cambios en las correspondientes regiones de unión de TrwA en el *oriT*, apuntando a una coevolución que daría lugar al mantenimiento de la especificidad de la proteína.

Dentro del sistema conjugativo de R388, además de TrwA, han sido ampliamente estudiadas otras proteínas que forman parte del relaxosoma como son TrwC e IHF (Gonzalez-Perez *et al.*, 2007; Grandoso *et al.*, 2000; Llosa *et al.*, 1996; Moncalian *et al.*, 1999b), aunque no se había llegado a estudiar en profundidad las posibles interacciones que pudiesen establecerse entre ellas o la influencia que puedan ejercer unas sobre otras.

En este trabajo, hemos intentado arrojar nuevos datos sobre la especificidad que muestra TrwA por sus sitios de unión en el *oriT*, unión que es indispensable para la mayoría de las funciones que realiza TrwA durante la conjugación. Hemos estudiado, además, las posibles interacciones que pueda establecer TrwA con las demás proteínas, que junto con el *oriT*, conforman el relaxosoma de R388. De esta forma queremos caracterizar las peculiaridades de este grupo de RHH y compararlo con el resto de factores transcripcionales tipo RHH.

Se ha dividido la discusión en cinco apartados que resumen las observaciones más relevantes del presente trabajo, y que se desarrollarán a continuación.

1. Diferencias puntuales en las proteínas TrwA y sus sitios de unión en R388 y R7K son responsables de dotar de especificidad al sistema

Es sabido que la conjugación bacteriana es un mecanismo específico que permite el intercambio de genes en bacterias. Esta especificidad se encuentra coordinada a distintos niveles: la relaxasa reconoce y corta una secuencia específica dentro del *oriT* y una o más proteínas accesorias presentes en el relaxosoma se unen a sus sitios de unión en el *oriT*, permitiendo cambios estructurales en el ADN que facilitan que la relaxasa realice su función (Nelson *et al.*, 1995; Ziegelin *et al.*, 1992), y la proteína acopladora dirige el contacto específico entre las regiones de procesamiento de ADN (*dtr*) y de formación del poro conjugativo (*mpf*) (Llosa et al., 2003).

Ya se ha comentado previamente que los plásmidos R388 y R7K comparten una gran similitud de secuencia con excepción de TrwA y de sus sitios de unión en *oriT*. En este trabajo, hemos demostrado tanto por ensayos de unión *in vitro* como por experimentos de movilización *in vivo* que cada TrwA es únicamente capaz de reconocer

su propio *oriT*, y que la ausencia de este reconocimiento acarrea una disminución notable de la conjugación del plásmido (apartado 3 de Resultados). Es, pues, tentador plantear que al ser muchos otros elementos implicados en la conjugación idénticos entre estos dos plásmidos, TrwA sea la responsable de dotar de especificidad las reacciones de procesamiento de ADN que ocurren previas a la transferencia del mismo.

La frecuencia de conjugación de R7K en las condiciones en las que se lleva a cabo el experimento es 10 veces menor que la obtenida con R388. Además de las diferencias en los sitios de unión de TrwA, encontramos que el *oriT* de R7K dispone únicamente de un sitio de unión para IHF, ya que carece del sitio *ihfB* localizado entre los dos sitios de unión de TrwA en R388, lo que nos hace suponer que la conformación que adopta el relaxosoma puede variar entre los dos plásmidos. De todas formas, se observó en los experimentos de movilización que aún estando codificado el gen *trwA* en *cis* respecto al *oriT* que pretendíamos movilizar, cuando esta movilización se llevaba a cabo utilizando la maquinaria conjugativa de R7K, la frecuencia no se restablecía, sino que los valores eran 10 veces menores que los obtenidos con *oriT* solo. Esto nos puede estar indicando que la posible distinta conformación que adopte el relaxosoma de R7K durante el procesamiento de ADN no es la única responsable de esta disminución en la conjugación, sino que es más probable que sea debido a que simplemente la conjugación no sea tan eficiente como en R388, o que las condiciones en las que se lleva a cabo el experimento o el huésped elegido no sean los indicados para este plásmido que fue inicialmente aislado de *Providencia rettgeri* (Coetzee et al., 1972).

2. La disposición que adoptan los sitios de unión de TrwA en $oriT_{R388}$ es vital para asegurar el correcto funcionamiento de la proteína

En R7K y en R388 los sitios de unión de TrwA están distribuidos de la misma manera, en forma de un hexanucleótido que se repite de manera directa para formar el sitio *sbaA*, y de manera invertida para formar el sitio *sbaB*.

En TrwA tanto de R7K como de R388, también hemos observado que la proteína muestra más afinidad por el sitio *sbaA*, siendo la diferencia algo menor en R7K que en R388, quizá porque el espaciamiento entre los dos brazos de este sitio en *sbaA* tiene 6

nucleótidos en vez de 5 como R388. Creímos que esta disposición de los dos operadores de TrwA y la diferencia en la unión a cada uno de ellos, podría ser una manera de regular alguna o todas las funciones que realiza la proteína durante la conjugación, por lo que abordamos su estudio basándonos en la unión a ADN, la movilización, y el efecto en la actividad represora del promotor *trwABC*.

En primer lugar se procedió a estudiar el comportamiento del *oriT* en ausencia de cada uno de los dos sitios de unión, o cuando ambos sitios eran eliminados. Se llevaron a cabo experimentos de movilización y de unión al ADN. Los resultados obtenidos en cada uno de los experimentos apuntan hacia la necesidad de la presencia de los dos sitios de unión para un eficiente funcionamiento del $oriT_{R388}$. Las diferencias que se observan en la unión al ADN de las distintas construcciones, aparentemente no son tan drásticas como las observadas en la movilización. También en el caso de los mutantes en los que se ha eliminado uno de los sitios de unión observamos una diferencia de afinidad de la proteína, corroborando los datos de ensayos previos (apartado 3 de Resultados), (Moncalian *et al.*, 1997). En el *oriT* que carece de ambos sitios de unión no se pudo apreciar ningún retardo (resultado mostrado en material suplementario, figura S4). Este dato nos aporta un indicio de la especificidad de TrwA, la cual es incapaz de reconocer ninguna otra región del ADN dentro del *oriT* incluso cuando concentraciones muy altas de la proteína son añadidas a la reacción.

Una vez comprobada la importancia de la presencia de los dos operadores de TrwA, nos pareció interesante probar si la distribución de los sitios también era un factor tan influyente en el funcionamiento del *oriT*. Para ello, diseñamos distintos mutantes del *oriT* que contenían dos sitios *sbaA*, dos sitios *sbaB* o los dos sitios originales pero modificando sus respectivas posiciones, es decir, en este último caso sería el sitio *sbaB* el que se encontraría cerca del inicio de la proteína. Estas nuevas construcciones se probaron en ensayos de movilización y de unión al ADN y además, se llevó a cabo un análisis de la actividad del promotor, análisis que no se pudo llevar a cabo con las construcciones que carecían de uno de los sitios al resultar el promotor cercenado. En presencia de los dos sitios de unión, sin importar la disposición que adoptan, la frecuencia de movilización se restaura al menos 1000 veces al compararlo con los datos obtenidos con los mutantes que contienen un único sitio (tabla R7 Resultados). Por otro lado, los experimentos de unión a ADN in vitro avalan sólo

parcialmente los resultados obtenidos en los experimentos de movilización. El mutante con dos sitios *sbaB* muestra una afinidad tres veces menor que el *oriT* original, mientras que el mutante con dos sitios *sbaA* parece mostrar una capacidad de unión ligeramente mayor (figura R10 Resultados). Este último dato era en cierto modo previsible, ya que hemos comentado previamente que TrwA muestra mayor afinidad por la repetición directa que la invertida, siendo por lo tanto posible que reconozca mejor un *oriT* que contenga dos sitios distribuidos de esta manera. Nos llama la atención el hecho de que en los experimentos realizados con los *oriT* que contienen dos sitios de unión de TrwA, independientemente de su distribución, observamos una unión escalonada en presencia de cantidades crecientes de TrwA. Este hecho no se puede observar en la unión al *oriT* que contiene solamente uno de los sitios, por lo que hemos achacado esta característica a la unión secuencial de ambos operadores según se aumenta la cantidad de proteína. Es decir, en presencia de pequeñas cantidades de TrwA, esta se uniría al operador por el que muestra más afinidad, y según se aumentara la concentración de proteína añadida, acabaría uniéndose a los dos operadores creando un complejo nucleoprotéico mayor que mostraría una menor velocidad de migración en el gel.

Por último, decidimos medir la actividad del promotor de los distintos variantes del *oriT* para ver si éramos capaces de observar algún efecto tales como los observados en los ensayos de movilización o de unión a ADN. De manera involuntaria, hemos comprobado la importancia de la región -35 del promotor para su correcto funcionamiento, ya que a la hora de diseñar el *oriT* sintético se introdujeron en esa zona dos mutaciones para la inserción de un sitio de restricción que abolieron la producción de GFP que nos indicaba que el promotor era funcional. Una vez restablecido este módulo en todas las construcciones, procedimos a su comparación. Lo primero que nos llamó la atención fue que el promotor localizado en el *oriT* salvaje de R388 mostraba una mayor actividad que aquel en el *oriT* sintético. Achacamos esta disminución en la actividad a la mutación puntual situada en la región-10 del promotor, aunque puede también puede ser debido a una diferencia a la hora de clonar este último en el pUA66. Nos vimos obligados a utilizar el sitio de restricción *PagI* en vez del sitio *BamHI* utilizado normalmente, ya que este sitio de restricción había sido ya insertado en el $oriT_{R388S}$*syn* para delimitar un módulo funcional. Este hecho cambia ligeramente el espaciamiento que va desde el promotor hasta el inicio del gen *gfpmut2*, por lo que creemos que esto puede estar influyendo a la hora de medir la producción de GFP.

Comparando las tres variantes del *oriT* sintético que han sido clonadas de la misma manera (pYTP18, pYTP20 y pYTP22), observamos que la mayor actividad es obtenida en presencia del *oriT* con dos sitios *sbaB* (pYTP20). Por alguna razón que desconocemos, la polimerasa parece encontrar esta disposición más accesible para realizar su función, quizá debido a la conformación que adopta en ADN en presencia de las dos repeticiones invertidas (tabla R10). Esta mayor actividad transcripcional, podría traducirse en un aumento de la producción de las proteínas codificadas por el operón *trwABC*, llevando a un desequilibrio en el sistema que podría acarrear una disminución en la eficiencia de conjugación del plásmido, como el observado por la sobreexpresión de TrwA (Moncalian et al., 1997).

Aunar los resultados obtenidos con cada una de las construcciones en los distintos experimentos nos permite obtener una visión global del comportamiento de cada mutante ante las diferentes funciones en las que se ve implicado el *oriT* durante la conjugación. En algunos casos, todos los resultados parecen apuntar en un mismo sentido, como en el caso de los mutantes que carecen de uno de los sitios de unión. Los resultados obtenidos en los experimentos *in vivo* e *in vitro* con estas construcciones indican que la ausencia de uno de los sitios de unión de TrwA, o de ambos, ejerce un efecto negativo en el funcionamiento del *oriT* que podría ser debido a una unión menos eficiente de la proteína. En el caso del plásmido con dos sitios de unión *sbaB* (pYTP12/pYTP20) también parecen correlacionarse los resultados de los tres experimentos realizados. Observamos un patrón que indica una menor eficiencia en todas las funciones estudiadas, que podría explicarse, como en el caso de las construcciones con un solo sitio de unión, a una menor afinidad de la proteína por la nueva organización de los sitios de reconocimiento. No nos resulta posible extraer una conclusión similar en el caso del plásmido que contiene dos sitios *sbaA* (pYTP13/pYTP22). Aunque la movilización de este plásmido y la capacidad de represión del promotor son menores que en el *oriT* salvaje, TrwA parece unir este *oriT* mutante con la misma afinidad que el original. Por lo tanto, la disminución en la eficiencia de las funciones durante la conjugación puede deberse a que la disposición de los dios sitios de unión de TrwA en forma de repetición directa provoca una alteración en la estructura óptima del ADN para que la proteína adquiera cierta conformación, o lleve a cabo contactos específicos entre sus subunidades que se traduzcan en un correcto desarrollo de su función.

Una vez analizados todos los datos, la conclusión más clara que podemos extraer es que la presencia de los dos sitios de unión de TrwA distribuídos a modo de repetición directa e invertida es indispensable para lograr el funcionamiento más eficiente del *oriT* durante la conjugación bacteriana. La eliminación o modificación de estos sitios de unión acarrea un cambio en la estructura del *oriT* que no beneficia al funcionamiento global del plásmido, aunque características puntuales como la unión al ADN (pYTP13) o la actividad del promotor (pYTP20) puedan verse ligeramente aumentadas.

.

3. La distribución característica de los sitios de unión de TrwA de R7K y R388 también se encuentra en otros plásmidos

Cuantas más proteínas de la familia RHH han sido analizadas y clasificadas, mayor divergencia se ha encontrado en sus estrategias de unión al ADN. Las RHH "canónicas" como pueden ser el represor Arc o CopG y MetJ unen principalmente repeticiones invertidas de secuencias simétricas (las dos primeras) o repeticiones directas (MetJ), pero únicamente TrwA de R388 y el represor ω de pSM19035 habían sido identificadas como capaces de unir tanto repeticiones directas como invertidas. El hecho de que las demás proteínas no sean capaces de unir indistintamente repeticiones directas o invertidas indicaría que existe, en estas, algún elemento que rompe la simetría de los homodímeros y provoca que la proteína solo sea funcional en una orientación determinada. Puede que el problema se encuentre en la interacción entre los monómeros, quizá ésta interacción no puede darse o no es efectiva si un monómero debe rotar para reconocer un brazo del operador que se encuentra distribuido de una manera distinta. Tanto en TrwA como en omega, se estudió la afinidad por los dos tipos de operadores y se observó que ambas mostraban una ligera predilección de unión por un tipo concreto de distribución. TrwA muestra aproximadamente 3 veces mayor afinidad por la repetición directa (sitio *sbaA*) que por la repetición invertida (sitio *sbaB*) (KdDR=130nM vs KdIR=420nM) mientras que ω, a su vez, muestra unas 4 veces mayor afinidad por dos héptadas repetidas de manera invertida que de manera directa (KdDR=90nM vs KdIR=20nM).

Cada vez más proteínas se clasifican en esta familia RHH y con ello, surgen nuevas características y singularidades en lo que se refiere a su interacción con el ADN.

Así, en lo que respecta a las proteínas que participan en la conjugación, se ha observado que TraY de F también une tanto DR como IR. Estas estructuras se repiten en los dos sitios de unión de TraY, aunque estas repeticiones son en ambos casos imperfectas (Lum *et al.*, 2002). La proteína ParG (análogo funcional de ω), que junto con la ATPasa ParF se ensamblan en el centrómero *parH*, participa en el proceso de partición del plásmido TP228, también es capaz de unirse a repeticiones imperfectas del tetrámero ACTC que aparece formando repeticiones directas e invertidas (Zampini *et al.*, 2009). Se encuentran también proteínas que aún perteneciendo a este grupo RHH, no se ha identificado ninguna repetición que pudiera ser reconocida por ellas, como es el caso de la proteína MbeC involucrada en la conjugación del plásmido ColE1 (Varsaki *et al.*, 2009). Aunque la secuencia que reconoce esta proteína no se encuentre formando una repetición, se ha comprobado que se une al *oriT* del plásmido que la codifica, por lo que las funciones que desempeña durante la conjugación se estima que son iguales o parecidas a aquellas proteínas auxiliares como TrwA o TraY, entre otras. Es decir, MbeC sería la encargada, mediante su unión específica, de provocar una curvatura en el ADN que facilite la apertura de la doble hélice en esa región para que la relaxasa lleve a cabo el corte de la cadena sencilla.

En el caso de TrwA de R388 está ampliamente comprobado que la proteína reconoce y une establemente dos sitios de en el origen de transferencia de este plásmido (Moncalian *et al.*, 1997). En el presente trabajo, se ha demostrado que la presencia de ambos sitios de unión es imprescindible para el funcionamiento eficiente de la proteína, tanto en lo que respecta a la conjugación como a la regulación transcripcional que lleva a cabo como represor (apartados 4 y 5 de Resultados). El plásmido R7K que se encuentra estrechamente emparentado filogenéticamente con R388 mantiene esta distribución característica de los sitios de unión de TrwA. Se ha podido observar que en R7K TrwA también muestra una mayor afinidad por el operador que se presenta a modo de repetición directa (apartado 3 de Resultados), y aunque no se han llevado a cabo análisis de eliminación o sustitución de estos sitios de unión y podamos encontrar en el *oriT* diferencias como la carencia del sitio *ihfB*, creemos que la proteína TrwA de este plásmido muestra un comportamiento similar a su homóloga en R388. Para comprobar si la distribución de los sitios de unión de TrwA es una característica generalizada de este tipo de proteínas o sin embargo, se limita a las analizadas hasta la fecha, hemos realizado un análisis exhaustivo de las proteínas que muestran una mayor identidad de

secuencia con TrwA de R388 y R7K. Realizamos una búsqueda utilizando la herramienta BLAST de Pubmed y se obtuvieron un número de homólogos de los cuales se analizaron aquellos que mostraban una identidad superior al 26%, prestando principal atención a la zona amino terminal que abarca el dominio RHH de unión al ADN. De todas estas proteínas, algunas se encuentran descritas como proteínas implicadas en movilización, y otras aún no han sido estudiadas por lo que se muestran como hipotéticas.

Se procedió a analizar la región intergénica anterior a la proteína en cada una de ellas, para buscar posibles elementos de estructura secundaria que pudieran ayudar a identificar esta región como origen de transferencia. Algunas de estas regiones mostraban similitudes con el origen de transferencia de R388 y R7K. En concreto, en las regiones anteriores a los genes que codifican para las proteínas de movilización de *Bordetella parapertussis*, *Pseudomonas stutzeri* y *Anaeromyxobacter dehalogenans* respectivamente la tercera, cuarta y undécima proteínas con mayor homología con $TrwA_{R388}$, se encontraron posibles sitios de unión de la proteína distribuidos de la misma manera que en los plásmidos R388 y R7K ya caracterizados (apartado 10 de Resultados). En el análisis realizado se ha incluido además el plásmido pMBUI4 recientemente caracterizado como similar a IncW, ya que también se han podido identificar similitudes con R388.

Los orígenes de transferencia de estos plásmidos muestran el mismo patrón en la distribución de los sitios de unión de $TrwA_{R388}$, pero la secuencia de estos operadores es única para cada uno, con algunas excepciones. Los plásmidos pMBUI4 y *P. stutzeri* muestran prácticamente la misma secuencia en los sitios de unión de TrwA, aunque esto cobra sentido cuando al comparar las proteínas observamos que el dominio de unión a ADN difiere en un único aminoácido. R7K y el plásmido BPP5P1 de también comparten sitios de unión idénticos. En este caso, las proteínas sólo comparten un 75% de su secuencia global, aunque a nivel de la hebra beta la homología aumenta y los dos primeros aminoácidos polares son idénticos en ambos plásmidos.

La estructura de los sitios de unión a modo de repetición directa y repetición invertida parece pues, repetirse en los orígenes de transferencia de distintos organismos. Este dato, unido a los resultados presentados en este documento en el que destacamos la

necesidad de la presencia de los dos sitios de unión de TrwA para el funcionamiento eficiente de la proteína, parece indicar que la distribución de los sitios de unión de TrwA es una característica que regula y define la función de esta proteína.

Si comparamos todos los *oriT*, la mayor identidad de secuencia se centra en la región de reconocimiento y corte de la relaxasa. En los plásmidos IncW, esta región del *oriT* es prácticamente idéntica y este parecido es también extrapolable a las relaxasas. Las relaxasas de R7K y pIE321 muestran una identidad de secuencia aminoacídica de 99.8% y 98.7% respectivamente con la relaxasa de R388. La relaxasa de pMBUI4, sin embargo, es similar a $TrwC_{R388}$ solo en un 61%. Otras proteínas implicadas en conjugación también son muy similares en los distintos plásmidos MOB_F, siendo algunas como TrwD idénticas en R388, R7K y Pie321. De hecho, una de las proteínas que más difiere en estos plásmidos es TrwA (Revilla *et al.*, 2008). En el plásmido pMBUI4 las similitudes no son tan grandes como entre los otros MOB_F, pero también en este caso proporcionalmente TrwA es la proteína más divergente de todas las que participan en el aparato conjugativo. Aunque no se halla clasificado dentro de los MOB_F, en *Anaeromyxobacter dehalogenans* observamos la misma característica. La proteína homóloga de TrwA de *A. dehalogenans* es similar a la de R388 en un 30%, en contraste con el 63% y 68% de similitud que muestran las proteínas relaxasa y acopladora respectivamente. En cambio, en el plásmido BPP5P1, la diferencia no es tan drástica. En este caso, la hipotética TrwA comparte un 64% de secuencia con aquella de R388, frente al 84% de similitud de TrwJ o el 70% en el caso de TrwI.

Con todo ello, parece existir una tendencia hacia la diferenciación de TrwA de los distintos plásmidos, siendo la proteína conjugativa que se ve sometida a una mayor presión selectiva. Esto podría significar que sería precisamente TrwA la encargada de dotar de especificidad al sistema conjugativo, seleccionando mediante su unión al *oriT* el lugar donde se dará el corte que iniciará la transferencia de ADN a la célula receptora. Para lograr esta especificidad, la proteína y la región del *oriT* que abarca sus sitios de unión habrían evolucionado de manera conjunta, dando lugar a plásmidos que sólo pueden ser movilizados eficientemente en presencia de la proteína accesoria específica.

4. Posible interacción de TrwA con otras proteínas del relaxosoma de R388

Hasta la fecha se ha logrado caracterizar todas las proteínas del relaxosoma de R388, definiendo los sitios que ocupa cada una al unirse al ADN y las funciones que realiza durante el procesamiento del mismo.

Se han demostrado, también, las interacciones que establece la proteína acopladora con proteínas del relaxosoma, así como con proteínas pertenecientes al T4SS (Llosa et al., 2003). No se han podido, en cambio, aportar evidencias de ningún contacto entre proteínas pertenecientes al relaxosoma, aun sabiendo que forman un complejo estrechamente relacionado en el que las funciones que desempeña cada uno de los componentes pueden interferir en la labor del elemento adyacente.

Ya se demostró en un trabajo anterior que TrwA aumenta el corte realizado por la relaxasa (Moncalian et al., 1997). Esto se puede explicar fácilmente sin recurrir a la interacción entre ambas proteínas, ya que TrwA, así como otras proteínas accesorias podría ser capaz de curvar el ADN de tal manera que la relaxasa pudiera encontrar mayor accesibilidad al sitio *nic*. En otro trabajo realizado posteriormente, se encontró también que TrwA era necesaria para la recombinación eficiente de TrwC, y que esto ocurría incluso cuando los sitios de unión de TrwA eran eliminados del *oriT*, apuntando a una interacción directa entre ambas proteínas (Cesar et al., 2006). Esta posible interacción ya había sido descartada en un trabajo anterior (Llosa et al., 2003), pero basándonos en estos nuevos datos y en resultados publicados para el sistema F (Ragonese et al., 2007) decidimos probar de nuevo la posible interacción que pudiera darse entre la relaxasa y la proteína accesoria de nuestro sistema.

A la vista de los resultados obtenidos, resulta imposible establecer una conclusión, ya que aunque los resultados *in vitro* apuntaban a una posible interacción, la prueba realizada *in vivo* hizo tambalear esta teoría. Quizá sea posible modificar las condiciones en las que se lleva a cabo el ensayo para que la hipotética interacción sea visible y demostrable, abriendo paso a nuevos experimentos que traten de dilucidar el motivo y el mecanismo por el cuál relaxasa y accesoria puedan interaccionar.

Ya se ha comentado que el *oriT* de R388 contiene dos sitios de unión para IHF, localizados entre el sitio *nic* y el sitio *sbaB*, y entre ambos sitios de unión de TrwA. Sabemos del papel de IHF en la formación de complejos nucleoprotéicos que requieren la participación de más de una proteína, y que es capaz de inducir una curvatura en el ADN de tal manera que puede acercar dos sitios de unión de una proteína que de otra manera se encuentren demasiado espaciados para que se de una unión eficiente (Stenzel *et al.*, 1991).

Observando esta distribución en la que los sitios de unión para ambas proteínas se encuentran tan cerca unos de otros, nos pareció interesante estudiar si la capacidad de unión de alguna de las proteínas podía verse alterada por la unión previa de una de ellas, debido a algún cambio que pudiera introducir en la conformación del ADN. Los resultados obtenidos mostraban que parecía haber un componente de cooperatividad, ya que al haber una proteína presente en la reacción, la adición de la segunda producía la aparición de un supercomplejo formado por la unión de ambas al *oriT*, no importando si la cantidad de la primera era suficiente para provocar, por si misma, un retardo en el ADN.

Este resultado nos llevó a plantear un modelo, representado en la figura D.1, en el que se muestra cómo se llevaría a cabo la unión de las dos proteínas y la apariencia que tendría el complejo final formado por ambas proteínas junto con el origen de transferencia.

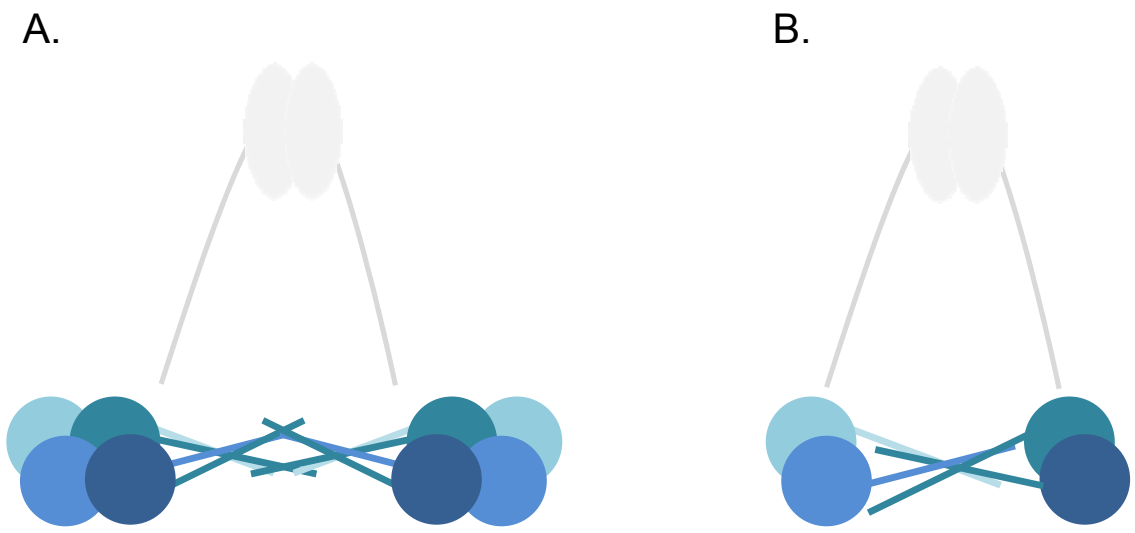


Figura D1. Propuesta de unión de TrwA en presencia de IHF. La molécula de IHF y el ADN sujeto a torsión se muestran en gris claro, y los monómeros de TrwA aparecen en distintos tonos de azul. Las líneas que asoman desde cada monómero simularían la interacción que se daría entre los distintas subunidades mediante las cremalleras de leucina contenidas en las hélices alfa.

Cuando TrwA es añadida al fragmento del *oriT* que contiene su región de unión, TrwA reconocerá y unirá en primer lugar el fragmento *sbaA* por el cual muestra mayor afinidad. Al aumentar la concentración de proteína, ésta se unirá a los dos fragmentos *sbaA y sbaB*, dando lugar a un complejo nucleoproteico con una menor velocidad de migración en gel. En presencia de IHF, al reconocer ésta su sitio de unión *ihfB* también podemos observar la aparición de una banda retardada de ADN, la cual pasa a ser de un tamaño mucho mayor cuando se añade TrwA a la mezcla. Es probable que mediante la torsión que provoca IHF en el ADN, los sitios de unión de TrwA se posicionen más cerca el uno del otro, quedando más accesibles para la proteína. En el modelo planteado se sugieren dos posibilidades de cómo ocurre esta unión: **A.** Cada tetrámero de TrwA reconoce un único de unión, reconociendo cada dímero un brazo de la repetición. Las interacciones favorables que se dan entre subunidades de las TrwA unidas a ambos sitios aumentan la cooperatividad de la misma, por lo que es necesaria una menor concentración de proteína para que el complejo ADN-proteína sea visible. En la opción **B.**, el tetrámero de TrwA estaría compuesto por dos dímeros reconociendo un brazo de cada sitio de unión cada uno, ya que al quedar ambos sitios enfrentados, se da lugar a una nueva disposición en la que podemos encontrar una repetición directa y otra

repetida formadas por la combinación de los dos sitios. Puede que esta nueva conformación resulte más accesible para la proteína, viéndose incrementada su afinidad.

5. Determinación de los contactos específicos de TrwA con el ADN

La unión de las proteínas pertenecientes a la familia RHH a la que pertenece TrwA ha sido ampliamente estudiada, y aunque la manera global de unirse al ADN de cada dominio RHH es el mismo, los detalles del contacto específico con las bases del ADN no parecen estar conservados entre los distintos miembros de esta familia. Se ha observado, por ejemplo, que proteínas como CopG, NikR y CcdA que muestran los mismos residuos aminoacídicos implicados en la unión en las posiciones 2, 4 y 6 de la hebra beta, reconocen secuencias específicas distintas de ADN. Actualmente, no existe un código que relacione directamente la secuencia aminoacídica de la hoja beta de estas proteínas con una secuencia específica de ADN, lo que hace que la predicción de secuencias de nuevos operadores para estas proteínas resulte difícil (Schreiter and Drennan, 2007).

Dado que los intentos de resolución de la estructura de TrwA junto a su operador resultaron infructuosos, decidimos abordar el estudio del contacto específico que establece esta proteína con las bases nucleotídicas de otras manera, con la intención de aportar algún indicio más sobre sus características de unión extrapolable a las RHH en general. Para ello llevamos a cabo dos estudios, un ensayo bioquímico en el que estudiamos la unión de TrwA a variables de su operador *sbaA* y una predicción *in silico* para obtener una aproximación de la estructura que adoptaría TrwA en su unión a ADN.

Del estudio mostrado en el apartado 6 de Resultados, podemos concluir que no todos los nucleótidos presentes en el operador *sbaA* de TrwA tienen la misma función en el reconocimiento por la proteína. Los nucleótidos que provocan un efecto más dramático en la unión, y por lo tanto, podríamos considerar "esenciales", se localizan en las posiciones 5 y 6 del operador, y además de en R388 y R7K, se conservan en pMBUI4 y *P. stutzeri*. De todas formas, con los datos con los que contamos hasta la fecha es imposible predecir si la importancia que adquieren en todos los plásmidos es la

misma, y menos aún si es una característica aplicable a otras proteínas pertenecientes a la familia RHH.

A día de hoy disponemos de la estructura tridimensional de varias proteínas conjugativas del plásmido R388, incluyendo proteínas del relaxosoma como la relaxasa TrwC o la proteína acopladora TrwB. Así, solo nos faltaría la estructura de la proteína accesoria para tener una visión completa de la arquitectura del relaxosoma.

Basándonos en predicciones de estructura secundaria, alineamientos con otras proteínas RHH y algunos resultados experimentales, hace ya tiempo TrwA se caracterizó como proteína perteneciente a la superfamilia RHH. Se llevó a cabo un análisis de los aminoácidos polares de la hoja beta y se estimó que los aminoácidos implicados en la unión de TrwAR388 serían Gln8, Arg10 y Ser12 (Moncalian and de la Cruz, 2004). El modelo de TrwA unida al ADN que hemos presentado en este trabajo (apartado 8.1 de Resultados), nos facilita una visión aproximada de cómo se daría el reconocimiento de ADN mediante TrwA. Ampliando la zona de las hojas beta para mostrar las cadenas laterales de los aminoácidos polares implicados en la unión de Arc y de aquéllos que se estima que son importantes en la unión de TrwA, comprobamos que algunos de los aminoácidos polares de TrwA quedarían excluidos de la interacción con el ADN (figura D2)

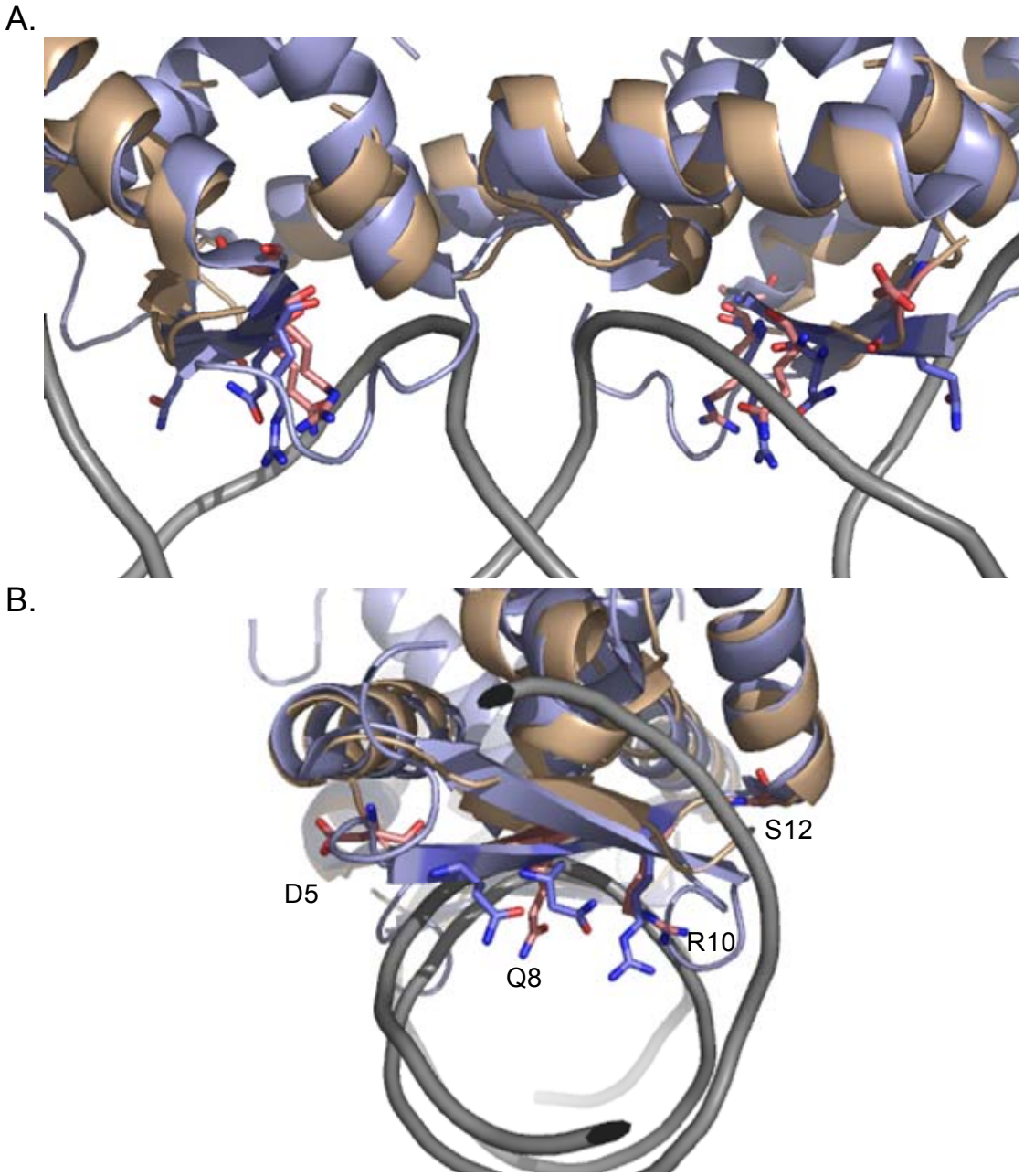


Figura D2. Vista detallada de la región de la hoja beta implicada en la unión al ADN. A. Imagen de las hojas betas introducidas en el surco mayor del ADN. Se destacan los residuos polares presentes: en azul, los de Arc, y en color salmón, los de TrwA. B. Imagen anterior girada 90º.

Los residuos Gln8 y Arg10 encajan perfectamente en el rol asignado como encargados del reconocimiento específico del ADN, ya que aparecen posicionados de manera muy similar a los aminoácidos polares de Arc. El aminoácido Ser12, en cambio, queda algo alejado de la hoja beta y parece difícil que su cadena lateral pueda interaccionar con el ADN. Se ha destacado también en este modelo el residuo Asp5, que

tampoco parece implicado en la unión al apuntar su cadena lateral hacia el exterior de la interfaz tal y como ocurre en el caso de Ser12. Este último aminoácido fue uno de los que se modificó al llevar a cabo el análisis de sustitución de los aminoácidos polares entre TrwAR7K y TrwAR388 (apartado 7 de Resultados), ya que tras constatar en el modelado con NikA que la S12 quedaba alejada de la hoja beta, creímos que Asp5 podía estar interaccionando con el ADN. Dejando de lado estos detalles, en términos generales TrwA parece encajar en la imagen canónica de las RHH, y solamente podremos contrastar los aspectos más concretos de sus interacciones específicas cuando se obtenga la estructura de la proteína.

De todos modos, debemos tener en cuenta que no podemos extraer conclusiones determinantes del modelado que hemos realizado, ya que la estructura que observamos aquí puede mostrar variaciones respecto a la original, y es probable que nos impida ver las características propias y peculiaridades de TrwA que aportarían una visión más amplia de las distintas maneras en las que se puede llevar a cabo la unión al ADN por las proteínas RHH.

6.

CONCLUSIONES

Conclusiones

Se enumeran, a continuación, las principales conclusiones de este trabajo

1. Es posible la obtención de un origen de transferencia funcional (*oriT*$_{R388}$*syn*) introduciendo mutaciones puntuales que crean sitios de restricción sin afectar a los elementos de estructura secundaria.

2. Todos los elementos de estructura secundaria del *oriT* son necesarios, en mayor o menor medida, para una conjugación eficiente. Además de su presencia, es importante que la orientación sea correcta, como en el caso de la región *nic*-IR2, que al adoptar una orientación diferente no es capaz de conjugar correctamente

3. Aunque los *oriT* de R388 y R7K muestran la misma distribución de sus elementos de estructura secundaria, existen diferencias como la ausencia del sitio *ihfB* en R7K. Esto podría conllevar una distinta estructura del relaxosoma que podría influir en la conjugación.

4. TrwA$_{R388}$ y TrwA$_{R7K}$ muestran especificidad por sus respectivos sitios de unión, siendo la afinidad mayor en ambos casos por el sitio *sbaA* formado por la repetición directa. Esta especificidad se observa también *in vivo*, ya que los plásmidos sólo son movilizados eficientemente en presencia de la proteína específica de cada uno.

5. La modificación de la distribución de los sitios de unión de TrwA influye de manera negativa en las principales funciones de la proteína, viéndose afectada la unión de TrwA al ADN, la movilización del plásmido y la represión del promotor *trwABC*.

6. No todos los nucleótidos que forman el sitio *sbaA* de unión de TrwA influyen de la misma manera en la unión de la proteína. La sustitución de aquellos situados en las posiciones 5 y 6 de los dos brazos es capaz de abolir la unión de TrwA, mientras que la sustitución del resto de los nucleótidos afecta en distinta medida la capacidad de unión. La secuencia de las posiciones 5 y 6 se conserva en R388, R7K y pMBUI4.

7. La hoja beta de TrwA no es la única responsable de dotar de especificidad a la proteína, ya que la sustitución de los aminoácidos de esta región entre R7K y R388 no provoca un cambio de especificidad, sino que repercute negativamente en la unión al operador propio.

8. Otras proteínas del relaxosoma podrían influir en la unión de TrwA. La adición de IHF en los ensayos de unión a ADN aumenta la afinidad de TrwA por el mismo y provoca la aparición de un supercomplejo nucleoprotéico en lugar de la unión secuencial observada cuando sólo TrwA está presente. Además, TrwA parece interaccionar con TrwC en condiciones concretas en ensayos *in vitro*, aunque no se ha conseguido la reproducción de la interacción TrwA-TrwC *in vivo* y tampoco se ha encontrado una explicación biológica para este fenómeno.

9. La distribución a modo de repetición directa en invertida que encontramos en R388 y R7K aparece también en otros plásmidos o cromosomas (*Bordetella parapertussis*, *Pseudomonas stutzeri*, pMBUI4 y *Aneromyxobacter dehalogenans*). Además, en todos los casos TrwA es la proteína conjugativa menos conservada, apuntando a que podría ser la encargada de dotar de especificidad al sistema conjugativo, regulando mediante su unión específica el inicio de la transferencia de ADN a la célula receptora.

7.

BIBLIOGRAFÍA

Abo, T., and Ohtsubo, E. (1993) Repression of the traM gene of plasmid R100 by its own product and integration host factor at one of the two promoters. *J Bacteriol* **175**: 4466-4474.

Abo, T., and Ohtsubo, E. (1995) Characterization of the functional sites in the oriT region involved in DNA transfer promoted by sex factor plasmid R100. *J Bacteriol* **177**: 4350-4355.

Alvarez-Martinez, C.E., and Christie, P.J. (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**: 775-808.

Andersson, J.O., Sarchfield, S.W., and Roger, A.J. (2005) Gene transfers from nanoarchaeota to an ancestor of diplomonads and parabasalids. *Mol Biol Evol* **22**: 85-90.

Arechaga, I., Pena, A., Zunzunegui, S., del Carmen Fernandez-Alonso, M., Rivas, G., and de la Cruz, F. (2008) ATPase activity and oligomeric state of TrwK, the VirB4 homologue of the plasmid R388 type IV secretion system. *J Bacteriol* **190**: 5472-5479.

Avila, P., and de la Cruz, F. (1988) Physical and genetic map of the IncW plasmid R388. *Plasmid* **20**: 155-157.

Avila, P., and de la Cruz, F. (1988) Physical and genetic map of the IncW plasmid R388. *Plasmid* **20**: 155-157.

Bartolome, B., Jubete, Y., Martinez, E., and de la Cruz, F. (1991) Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* **102**: 75-78.

Beaber, J.W., Hochhut, B., and Waldor, M.K. (2002) Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from Vibrio cholerae. *J Bacteriol* **184**: 4259-4269.

Becker, E.C., and Meyer, R.J. (2003) Relaxed specificity of the R1162 nickase: a model for evolution of a system for conjugative mobilization of plasmids. *J Bacteriol* **185**: 3538-3546.

Boer, R., Russi, S., Guasch, A., Lucas, M., Blanco, A.G., Perez-Luque, R., Coll, M., and de la Cruz, F. (2006) Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. *J Mol Biol* **358**: 857-869.

Bolland, S., Llosa, M., Avila, P., and de la Cruz, F. (1990) General organization of the conjugal transfer genes of the IncW plasmid R388 and interactions between R388 and IncN and IncP plasmids. *J Bacteriol* **172**: 5795-5802.

Bolland, S. (1991) Genes implicados en la producción del pilus conjugativo de R388.

Bradley, D.E., and Cohen, D.R. (1976) Basic characterization of W-pili. *J Gen Microbiol* **97**: 91-103.

Bradley, D.E. (1980) Morphological and serological relationships of conjugative pili. *Plasmid* **4**: 155-169.

Bradley, D.E., Taylor, D.E., and Cohen, D.R. (1980) Specification of surface mating systems among conjugative drug resistance plasmids in Escherichia coli K-12. *J Bacteriol* **143**: 1466-1470.

Brasch, M.A., and Meyer, R.J. (1987) A 38 base-pair segment of DNA is required in cis for conjugative mobilization of broad host-range plasmid R1162. *J Mol Biol* **198**: 361-369.

Byrd, D.R., and Matson, S.W. (1997) Nicking by transesterification: the reaction catalysed by a relaxase. *Mol Microbiol* **25**: 1011-1022.

Cascales, E., and Christie, P.J. (2003) The versatile bacterial type IV secretion systems. *Nat Rev Microbiol* **1**: 137-149.

Cavalli, L.L., Lederberg, J., and Lederberg, E.M. (1953) An infective factor controlling sex compatibility in Bacterium coli. *J Gen Microbiol* **8**: 89-103.

Cesar, C.E., Machon, C., de la Cruz, F., and Llosa, M. (2006) A new domain of conjugative relaxase TrwC responsible for efficient oriT-specific recombination on minimal target sequences. *Mol Microbiol* **62**: 984-996.

Cesar, C.E., and Llosa, M. (2007) TrwC-mediated site-specific recombination is controlled by host factors altering local DNA topology. *J Bacteriol* **189**: 9037-9043.

Clewell, D.B., and Helinski, D.R. (1969) Supercoiled circular DNA-protein complex in Escherichia coli: purification and induced conversion to an opern circular DNA form. *Proc Natl Acad Sci U S A* **62**: 1159-1166.

Coetzee, J.N., Datta, N., and Hedges, R.W. (1972) R factors from Proteus rettgeri. *J Gen Microbiol* **72**: 543-552.

Cole, C., Barber, J.D., and Barton, G.J. (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**: W197-201.

Couturier, M., Bex, F., Bergquist, P.L., and Maas, W.K. (1988) Identification and classification of bacterial plasmids. *Microbiol Rev* **52**: 375-395.

Christie, P.J. (2001) Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. *Mol Microbiol* **40**: 294-305.

Christie, P.J., and Cascales, E. (2005) Structural and dynamic properties of bacterial type IV secretion systems (review). *Mol Membr Biol* **22**: 51-61.

Chung, C.T., and Miller, R.H. (1988) A rapid and convenient method for the preparation and storage of competent bacterial cells. *Nucleic Acids Res* **16**: 3580.

Datta, N., and Hedges, R.W. (1972) Trimethoprim resistance conferred by W plasmids in Enterobacteriaceae. *J Gen Microbiol* **72**: 349-355.

Datta, S., Larkin, C., and Schildbach, J.F. (2003) Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure* **11**: 1369-1379.

de la Cruz, F., and Grinsted, J. (1982) Genetic and molecular characterization of Tn21, a multiple resistance transposon from R100.1. *J Bacteriol* **151**: 222-228.

de la Cruz, F., and Davies, J. (2000) Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol* **8**: 128-133.

de la Cruz, F., Frost, L.S., Meyer, R.J., and Zechner, E.L. (2010) Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* **34**: 18-40.

de la Hoz, A.B., Pratto, F., Misselwitz, R., Speck, C., Weihofen, W., Welfle, K., Saenger, W., Welfle, H., and Alonso, J.C. (2004) Recognition of DNA by omega protein from the broad-host range Streptococcus pyogenes plasmid pSM19035: analysis of binding to operator DNA with one to four heptad repeats. *Nucleic Acids Res* **32**: 3136-3147.

de Paz, H.D., Sangari, F.J., Bolland, S., Garcia-Lobo, J.M., Dehio, C., de la Cruz, F., and Llosa, M. (2005) Functional interactions between type IV secretion systems involved in DNA transfer and virulence. *Microbiology* **151**: 3505-3516.

de Paz, H.D., Larrea, D., Zunzunegui, S., Dehio, C., de la Cruz, F., and Llosa, M. (2010) Functional dissection of the conjugative coupling protein TrwB. *J Bacteriol* **192**: 2655-2669.

Dempsey, W.B., and Fee, B.E. (1990) Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. *Mol Microbiol* **4**: 1019-1028.

Di Laurenzio, L., Frost, L.S., Finlay, B.B., and Paranchych, W. (1991) Characterization of the oriT region of the IncFV plasmid pED208. *Mol Microbiol* **5**: 1779-1790.

Di Laurenzio, L., Scraba, D.G., Paranchych, W., and Frost, L.S. (1995) Studies on the binding of integration host factor (IHF) and TraM to the origin of transfer of the IncFV plasmid pED208. *Mol Gen Genet* **247**: 726-734.

Ding, Z., Atmakuri, K., and Christie, P.J. (2003) The outs and ins of bacterial type IV secretion substrates. *Trends Microbiol* **11**: 527-535.

Dower, W.J., Miller, J.F., and Ragsdale, C.W. (1988) High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res* **16**: 6127-6145.

Draper, O., Cesar, C.E., Machon, C., de la Cruz, F., and Llosa, M. (2005) Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc Natl Acad Sci U S A* **102**: 16385-16390.

Fekete, R.A., and Frost, L.S. (2002) Characterizing the DNA contacts and cooperative binding of F plasmid TraM to its cognate sites at oriT. *J Biol Chem* **277**: 16705-16711.

Fernandez-Lopez, R., Del Campo, I., Ruiz, R., Lanza, V., Vielva, L., and de la Cruz, F. Numbers on the edges: a simplified and scalable method for quantifying the gene regulation function. *Bioessays* **32**: 346-355.

Fernandez-Lopez, R., Garcillan-Barcia, M.P., Revilla, C., Lazaro, M., Vielva, L., and de la Cruz, F. (2006) Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. *FEMS Microbiol Rev* **30**: 942-966.

Fernández-Lopez, R. (2007) Evolución y Red de Regulación del Plásmido R388.

Fernández-Lopez R, d.C.I., Revilla C, de la Cruz F (2013) Low mutational robustness and transient zygotic induction in horizontal gene transfer. *PLOS Genetics (en prensa)*.

Francia, M.V., de la Cruz, F., and Garcia Lobo, J.M. (1993) Secondary-sites for integration mediated by the Tn21 integrase. *Mol Microbiol* **10**: 823-828.

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

Friedman, D.I. (1988) Integration host factor: a protein for all reasons. *Cell* **55**: 545-554.

Frost, L.S., Ippen-Ihler, K., and Skurray, R.A. (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* **58**: 162-210.

Fu, Y.H., Tsai, M.M., Luo, Y.N., and Deonier, R.C. (1991) Deletion analysis of the F plasmid oriT locus. *J Bacteriol* **173**: 1012-1020.

Furste, J.P., Pansegrau, W., Ziegelin, G., Kroger, M., and Lanka, E. (1989) Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc Natl Acad Sci U S A* **86**: 1771-1775.

Furuya, N., and Komano, T. (1995) Specific binding of the NikA protein to one arm of 17-base-pair inverted repeat sequences within the oriT region of plasmid R64. *J Bacteriol* **177**: 46-51.

Garcillan-Barcia, M.P., Francia, M.V., and de la Cruz, F. (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* **33**: 657-687.

Goldenfeld, N., and Woese, C. (2007) Biology's next revolution. *Nature* **445**: 369.

Gomis-Ruth, F.X., Moncalian, G., Perez-Luque, R., Gonzalez, A., Cabezon, E., de la Cruz, F., and Coll, M. (2001) The bacterial conjugation protein TrwB resembles ring helicases and F1-ATPase. *Nature* **409**: 637-641.

Gonzalez-Perez, B., Lucas, M., Cooke, L.A., Vyle, J.S., de la Cruz, F., and Moncalian, G. (2007) Analysis of DNA processing reactions in bacterial conjugation by using suicide oligonucleotides. *Embo J* **26**: 3847-3857.

Goosen, N., and van de Putte, P. (1995) The regulation of transcription initiation by integration host factor. *Mol Microbiol* **16**: 1-7.

Gophna, U., Charlebois, R.L., and Doolittle, W.F. (2004) Have archaeal genes contributed to bacterial virulence? *Trends Microbiol* **12**: 213-219.

Grandoso, G., Llosa, M., Zabala, J.C., and de la Cruz, F. (1994) Purification and biochemical characterization of TrwC, the helicase involved in plasmid R388 conjugal DNA transfer. *Eur J Biochem* **226**: 403-412.

Grandoso, G., Avila, P., Cayon, A., Hernando, M.A., Llosa, M., and de la Cruz, F. (2000) Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J Mol Biol* **295**: 1163-1172.

Grant, S.G., Jessee, J., Bloom, F.R., and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc Natl Acad Sci U S A* **87**: 4645-4649.

Guasch, A., Lucas, M., Moncalian, G., Cabezas, M., Perez-Luque, R., Gomis-Ruth, F.X., de la Cruz, F., and Coll, M. (2003) Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat Struct Biol* **10**: 1002-1010.

Guljamow, A., Jenke-Kodama, H., Saumweber, H., Quillardet, P., Frangeul, L., Castets, A.M., Bouchier, C., Tandeau de Marsac, N., and Dittmann, E. (2007) Horizontal gene transfer of two cytoskeletal elements from a eukaryote to a cyanobacterium. *Curr Biol* **17**: R757-759.

Guynet, C., Cuevas, A., Moncalian, G., and de la Cruz, F. The stb operon balances the requirements for vegetative stability and conjugative transfer of plasmid R388. *PLoS Genet* **7**: e1002073.

Guzman, L.M., and Espinosa, M. (1997) The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid oriT. *J Mol Biol* **266**: 688-702.

Halary, S., Leigh, J.W., Cheaib, B., Lopez, P., and Bapteste, E. (2010) Network analyses structure genetic diversity in independent genetic worlds. *Proc Natl Acad Sci U S A* **107**: 127-132.

Harley, M.J., and Schildbach, J.F. (2003) Swapping single-stranded DNA sequence specificities of relaxases from conjugative plasmids F and R100. *Proc Natl Acad Sci U S A* **100**: 11243-11248.

Hayes, W. (1953) [Observations on a transmissible agent determining sexual differentiation in Bacterium coli]. *J Gen Microbiol* **8**: 72-88.

Heinemann, J.A., and Sprague, G.F., Jr. (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340**: 205-209.

Hernando, M.A. (2000) Bioquímica de las reacciones de transferencia de cadenas de DNA catalizadas por la proteína TrwC.

Howard, M.T., Nelson, W.C., and Matson, S.W. (1995) Stepwise assembly of a relaxosome at the F plasmid origin of transfer. *J Biol Chem* **270**: 28381-28386.

Hubscher, U., Maga, G., and Spadari, S. (2002) Eukaryotic DNA polymerases. *Annu Rev Biochem* **71**: 133-163.

Ilyina, T.V., and Koonin, E.V. (1992) Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucleic Acids Res* **20**: 3279-3285.

Inamoto, S., and Ohtsubo, E. (1990) Specific binding of the TraY protein to oriT and the promoter region for the traY gene of plasmid R100. *J Biol Chem* **265**: 6461-6466.

Inamoto, S., Yoshioka, Y., and Ohtsubo, E. (1991) Site- and strand-specific nicking in vitro at oriT by the traY-traI endonuclease of plasmid R100. *J Biol Chem* **266**: 10086-10092.

Inamoto, S., Fukuda, H., Abo, T., and Ohtsubo, E. (1994) Site- and strand-specific nicking at oriT of plasmid R100 in a purified system: enhancement of the nicking activity of TraI (helicase I) with TraY and IHF. *J Biochem* **116**: 838-844.

Ireland, C.R. (1983) Detailed restriction enzyme map of crown gall-suppressive IncW plasmid pSa, showing ends of deletion causing chloramphenicol sensitivity. *J Bacteriol* **155**: 722-727.

Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M.G., and Alon, U. (2001) Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science* **292**: 2080-2083.

Karl, W., Bamberger, M., and Zechner, E.L. (2001) Transfer protein TraY of plasmid R1 stimulates TraI-catalyzed oriT cleavage in vivo. *J Bacteriol* **183**: 909-914.

Kawabata, T. (2003) MATRAS: A program for protein 3D structure comparison. *Nucleic Acids Res* **31**: 3367-3369.

Kelley, L.A., and Sternberg, M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**: 363-371.

Khlebnikov, A., Datsenko, K.A., Skaug, T., Wanner, B.L., and Keasling, J.D. (2001) Homogeneous expression of the P(BAD) promoter in Escherichia coli by constitutive expression of the low-affinity high-capacity AraE transporter. *Microbiology* **147**: 3241-3247.

Koonin, E.V., Makarova, K.S., and Aravind, L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol* **55**: 709-742.

Koraimann, G. (2003) Lytic transglycosylases in macromolecular transport systems of Gram-negative bacteria. *Cell Mol Life Sci* **60**: 2371-2388.

Krasilnikova, M.M., Smirnova, E.V., Krasilnikov, A.S., and Mirkin, S.M. (2001) A new trick for an old dog: TraY binding to a homopurine-homopyrimidine run attenuates DNA replication. *J Mol Biol* **313**: 271-282.

Kupelwieser, G., Schwab, M., Hogenauer, G., Koraimann, G., and Zechner, E.L. (1998) Transfer protein TraM stimulates TraI-catalyzed cleavage of the transfer origin of plasmid R1 in vivo. *J Mol Biol* **275**: 81-94.

Lanka, E., and Wilkins, B.M. (1995) DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* **64**: 141-169.

Lederberg, J., and Tatum, E.L. (1946) Gene recombination in Escherichia coli. *Nature* **158**: 558.

Lessl, M., and Lanka, E. (1994) Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. *Cell* **77**: 321-324.

Levin, B.R., and Bergstrom, C.T. (2000) Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes. *Proc Natl Acad Sci U S A* **97**: 6981-6985.

Loper, J.E., and Kado, C.I. (1979) Host range conferred by the virulence-specifying plasmid of Agrobacterium tumefaciens. *J Bacteriol* **139**: 591-596.

Lu, J., Wong, J.J., Edwards, R.A., Manchak, J., Frost, L.S., and Glover, J.N. (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.

Lum, P.L., and Schildbach, J.F. (1999) Specific DNA recognition by F Factor TraY involves beta-sheet residues. *J Biol Chem* **274**: 19644-19648.

Lum, P.L., Rodgers, M.E., and Schildbach, J.F. (2002) TraY DNA recognition of its two F factor binding sites. *J Mol Biol* **321**: 563-578.

Luo, Y., Gao, Q., and Deonier, R.C. (1994) Mutational and physical analysis of F plasmid traY protein binding to oriT. *Mol Microbiol* **11**: 459-469.

Llosa, M., Bolland, S., and de la Cruz, F. (1991) Structural and functional analysis of the origin of conjugal transfer of the broad-host-range IncW plasmid R388 and comparison with the related IncN plasmid R46. *Mol Gen Genet* **226**: 473-483.

Llosa, M., Bolland, S., and de la Cruz, F. (1994a) Genetic organization of the conjugal DNA processing region of the IncW plasmid R388. *J Mol Biol* **235**: 448-464.

Llosa, M., Bolland, S., Grandoso, G., and de la Cruz, F. (1994b) Conjugation-independent, site-specific recombination at the oriT of the IncW plasmid R388 mediated by TrwC. *J Bacteriol* **176**: 3210-3217.

Llosa, M., Grandoso, G., and de la Cruz, F. (1995) Nicking activity of TrwC directed against the origin of transfer of the IncW plasmid R388. *J Mol Biol* **246**: 54-62.

Llosa, M., Grandoso, G., Hernando, M.A., and de la Cruz, F. (1996) Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *J Mol Biol* **264**: 56-67.

Llosa, M., Gomis-Ruth, F.X., Coll, M., and de la Cruz Fd, F. (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* **45**: 1-8.

Llosa, M., Zunzunegui, S., and de la Cruz, F. (2003) Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc Natl Acad Sci U S A* **100**: 10465-10470.

Llosa, M., and de la Cruz, F. (2005) Bacterial conjugation: a potential tool for genomic engineering. *Res Microbiol* **156**: 1-6.

Lu, J., Wong, J.J., Edwards, R.A., Manchak, J., Frost, L.S., and Glover, J.N. (2008) Structural basis of specific TraD-TraM recognition during F plasmid-mediated bacterial conjugation. *Mol Microbiol* **70**: 89-99.

Martinez, E., and de la Cruz, F. (1988) Transposon Tn21 encodes a RecA-independent site-specific integration system. *Mol Gen Genet* **211**: 320-325.

Matilla, I., Alfonso, C., Rivas, G., Bolt, E.L., de la Cruz, F., and Cabezon, E. (2010) The conjugative DNA translocase TrwB is a structure-specific DNA-binding protein. *J Biol Chem* **285**: 17537-17544.

Mazel, D., and Davies, J. (1999) Antibiotic resistance in microbes. *Cell Mol Life Sci* **56**: 742-754.

McLeod, S.M., Burrus, V., and Waldor, M.K. (2006) Requirement for Vibrio cholerae integration host factor in conjugative DNA transfer. *J Bacteriol* **188**: 5704-5711.

Mengeritsky, G., Goldenberg, D., Mendelson, I., Giladi, H., and Oppenheim, A.B. (1993) Genetic and biochemical analysis of the integration host factor of Escherichia coli. *J Mol Biol* **231**: 646-657.

Meyer, R. (2009) The r1162 mob proteins can promote conjugative transfer from cryptic origins in the bacterial chromosome. *J Bacteriol* **191**: 1574-1580.

Miroux, B., and Walker, J.E. (1996) Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**: 289-298.

Moncalian, G., Grandoso, G., Llosa, M., and de la Cruz, F. (1997) oriT-processing and regulatory roles of TrwA protein in plasmid R388 conjugation. *J Mol Biol* **270**: 188-200.

Moncalian, G., Cabezon, E., Alkorta, I., Valle, M., Moro, F., Valpuesta, J.M., Goni, F.M., and de La Cruz, F. (1999a) Characterization of ATP and DNA binding activities of TrwB, the coupling protein essential in plasmid R388 conjugation. *J Biol Chem* **274**: 36117-36124.

Moncalian, G., Valle, M., Valpuesta, J.M., and de la Cruz, F. (1999b) IHF protein inhibits cleavage but not assembly of plasmid R388 relaxosomes. *Mol Microbiol* **31**: 1643-1652.

Moncalián, G. (2000) Arquitectura y ensamblaje del relaxosoma en el plásmido conjugativo R388. Santander: Universidad de Cantabria.

Moncalian, G., and de la Cruz, F. (2004) DNA binding properties of protein TrwA, a possible structural variant of the Arc repressor superfamily. *Biochim Biophys Acta* **1701**: 15-23.

Monzingo, A.F., Ozburn, A., Xia, S., Meyer, R.J., and Robertus, J.D. (2007) The structure of the minimal relaxase domain of MobA at 2.1 A resolution. *J Mol Biol* **366**: 165-178.

Nash, H.A. (1996) The HU and IHF proteins: accessory factors for complex protein-DNA assemblies. *Regulation of gene expression in E. coli, E.C.C. Lin, and A. Simon Lynch, eds*: 149-179.

Nedelcu, A.M., Miles, I.H., Fagir, A.M., and Karol, K. (2008) Adaptive eukaryote-to-eukaryote lateral gene transfer: stress-related genes of algal origin in the closest unicellular relatives of animals. *J Evol Biol* **21**: 1852-1860.

Nelson, W.C., Morton, B.S., Lahue, E.E., and Matson, S.W. (1993) Characterization of the Escherichia coli F factor traY gene product and its binding sites. *J Bacteriol* **175**: 2221-2228.

Nelson, W.C., Howard, M.T., Sherman, J.A., and Matson, S.W. (1995) The traY gene product and integration host factor stimulate Escherichia coli DNA helicase I-catalyzed nicking at the F plasmid oriT. *J Biol Chem* **270**: 28374-28380.

Nogueira, T., Rankin, D.J., Touchon, M., Taddei, F., Brown, S.P., and Rocha, E.P. (2009) Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. *Curr Biol* **19**: 1683-1691.

Nunez, B., and De La Cruz, F. (2001) Two atypical mobilization proteins are involved in plasmid CloDF13 relaxation. *Mol Microbiol* **39**: 1088-1099.

Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299-304.

Pansegrau, W., Ziegelin, G., and Lanka, E. (1988) The origin of conjugative IncP plasmid transfer: interaction with plasmid-encoded products and the nucleotide sequence at the relaxation site. *Biochim Biophys Acta* **951**: 365-374.

Pansegrau, W., Balzer, D., Kruft, V., Lurz, R., and Lanka, E. (1990) In vitro assembly of relaxosomes at the transfer origin of plasmid RP4. *Proc Natl Acad Sci U S A* **87**: 6555-6559.

Pansegrau, W., Lanka, E., Barth, P.T., Figurski, D.H., Guiney, D.G., Haas, D., Helinski, D.R., Schwab, H., Stanisich, V.A., and Thomas, C.M. (1994a) Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. *J Mol Biol* **239**: 623-663.

Pansegrau, W., Schroder, W., and Lanka, E. (1994b) Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. *J Biol Chem* **269**: 2782-2789.

Parker, C., Becker, E., Zhang, X., Jandle, S., and Meyer, R. (2005) Elements in the co-evolution of relaxases and their origins of transfer. *Plasmid* **53**: 113-118.

Parker, C., and Meyer, R.J. (2007) The R1162 relaxase/primase contains two, type IV transport signals that require the small plasmid protein MobB. *Mol Microbiol* **66**: 252-261.

Penfold, S.S., Simon, J., and Frost, L.S. (1996) Regulation of the expression of the traM gene of the F sex factor of Escherichia coli. *Mol Microbiol* **20**: 549-558.

Ragonese, H., Haisch, D., Villareal, E., Choi, J.H., and Matson, S.W. (2007) The F plasmid-encoded TraM protein stimulates relaxosome-mediated cleavage at oriT through an interaction with TraI. *Mol Microbiol* **63**: 1173-1184.

Raumann, B.E., Knight, K.L., and Sauer, R.T. (1995) Dramatic changes in DNA-binding specificity caused by single residue substitutions in an Arc/Mnt hybrid repressor. *Nat Struct Biol* **2**: 1115-1122.

Rest, J.S., and Mindell, D.P. (2003) Retroids in archaea: phylogeny and lateral origins. *Mol Biol Evol* **20**: 1134-1142.

Revilla, C., Garcillan-Barcia, M.P., Fernandez-Lopez, R., Thomson, N.R., Sanders, M., Cheung, M., Thomas, C.M., and de la Cruz, F. (2008) Different pathways to acquiring resistance genes illustrated by the recent evolution of IncW plasmids. *Antimicrob Agents Chemother* **52**: 1472-1480.

Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**: 1295-1306.

Rivas, S., Bolland, S., Cabezon, E., Goni, F.M., and de la Cruz, F. (1997) TrwD, a protein encoded by the IncW plasmid R388, displays an ATP hydrolase activity essential for bacterial conjugation. *J Biol Chem* **272**: 25583-25590.

Scherzinger, E., Lurz, R., Otto, S., and Dobrinski, B. (1992) In vitro cleavage of double- and single-stranded DNA by plasmid RSF1010-encoded mobilization proteins. *Nucleic Acids Res* **20**: 41-48.

Scherzinger, E., Kruft, V., and Otto, S. (1993) Purification of the large mobilization protein of plasmid RSF1010 and characterization of its site-specific DNA-cleaving/DNA-joining activity. *Eur J Biochem* **217**: 929-938.

Scholz, P., Haring, V., Wittmann-Liebold, B., Ashman, K., Bagdasarian, M., and Scherzinger, E. (1989) Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. *Gene* **75**: 271-288.

Schreiter, E.R., and Drennan, C.L. (2007) Ribbon-helix-helix transcription factors: variations on a theme. *Nat Rev Microbiol* **5**: 710-720.

Schroder, G., Krause, S., Zechner, E.L., Traxler, B., Yeo, H.J., Lurz, R., Waksman, G., and Lanka, E. (2002) TraG-like proteins of DNA transfer systems and of the Helicobacter pylori type IV secretion system: inner membrane gate for exported substrates? *J Bacteriol* **184**: 2767-2779.

Schwab, M., Gruber, H., and Hogenauer, G. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol Microbiol* **5**: 439-446.

Silverman, P.M., and Sholl, A. (1996) Effect of traY amber mutations on F-plasmid traY promoter activity in vivo. *J Bacteriol* **178**: 5787-5789.

Smillie, C., Garcillan-Barcia, M.P., Francia, M.V., Rocha, E.P., and de la Cruz, F. (2010) Mobility of plasmids. *Microbiol Mol Biol Rev* **74**: 434-452.

Stenzel, T.T., MacAllister, T., and Bastia, D. (1991) Cooperativity at a distance promoted by the combined action of two replication initiator proteins and a DNA bending protein at the replication origin of pSC101. *Genes Dev* **5**: 1453-1463.

Sugimura, S., and Crothers, D.M. (2006) Stepwise binding and bending of DNA by Escherichia coli integration host factor. *Proc Natl Acad Sci U S A* **103**: 18510-18514.

Sut, M.V., Mihajlovic, S., Lang, S., Gruber, C.J., and Zechner, E.L. (2009) Protein and DNA effectors control the TraI conjugative helicase of plasmid R1. *J Bacteriol* **191**: 6888-6899.

Szpirer, C.Y., Faelen, M., and Couturier, M. (2000) Interaction between the RP4 coupling protein TraG and the pBHR1 mobilization protein Mob. *Mol Microbiol* **37**: 1283-1292.

Taki, K., Abo, T., and Ohtsubo, E. (1998) Regulatory mechanisms in expression of the traY-I operon of sex factor plasmid R100: involvement of traJ and traY gene products. *Genes Cells* **3**: 331-345.

Tato, I., Zunzunegui, S., de la Cruz, F., and Cabezon, E. (2005) TrwB, the coupling protein involved in DNA transport during bacterial conjugation, is a DNA-dependent ATPase. *Proc Natl Acad Sci U S A* **102**: 8156-8161.

Tato, I., Matilla, I., Arechaga, I., Zunzunegui, S., de la Cruz, F., and Cabezon, E. (2007) The ATPase activity of the DNA transporter TrwB is modulated by protein TrwA: implications for a common assembly mechanism of DNA translocating motors. *J Biol Chem* **282**: 25569-25576.

Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008) Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* **11**: 472-477.

Thomas, J., and Hecht, D.W. (2007) Interaction of Bacteroides fragilis pLV22a relaxase and transfer DNA with Escherichia coli RP4-TraG coupling protein. *Mol Microbiol* **66**: 948-960.

Tsai, M.M., Fu, Y.H., and Deonier, R.C. (1990) Intrinsic bends and integration host factor binding at F plasmid oriT. *J Bacteriol* **172**: 4603-4609.

Valentine, C.R. (1985) One-kilobase direct repeats of plasmid pSa. *Plasmid* **14**: 167-170.

Varsaki, A., Moncalian, G., Garcillan-Barcia Mdel, P., Drainas, C., and de la Cruz, F. (2009) Analysis of ColE1 MbeC unveils an extended ribbon-helix-helix family of nicking accessory proteins. *J Bacteriol* **191**: 1446-1455.

Varsaki, A., Lamb, H.K., Eleftheriadou, O., Vandera, E., Thompson, P., Moncalian, G., de la Cruz, F., Hawkins, A.R., and Drainas, C. (2012) Interaction between relaxase MbeA and accessory protein MbeC of the conjugally mobilizable plasmid ColE1. *FEBS Lett* **586**: 675-679.

Vergunst, A.C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C.M., Regensburg-Tuink, T.J., and Hooykaas, P.J. (2000) VirB/D4-dependent protein translocation from Agrobacterium into plant cells. *Science* **290**: 979-982.

Wang, S., Cosstick, R., Gardner, J.F., and Gumport, R.I. (1995) The specific binding of Escherichia coli integration host factor involves both major and minor grooves of DNA. *Biochemistry* **34**: 13082-13090.

Ward, D.V., Draper, O., Zupan, J.R., and Zambryski, P.C. (2002) Peptide linkage mapping of the Agrobacterium tumefaciens vir-encoded type IV secretion system reveals protein subassemblies. *Proc Natl Acad Sci U S A* **99**: 11493-11500.

Ward, J.M., and Grinsted, J. (1982) Physical and genetic analysis of the Inc-W group plasmids R388, Sa, and R7K. *Plasmid* **7**: 239-250.

Watanabe, T., Furuse, C., and Sakaizumi, S. (1968) Transduction of various R factors by phage P1 in Escherichia coli and by phage P22 in Salmonella typhimurium. *J Bacteriol* **96**: 1791-1795.

Waters, V.L., Hirata, K.H., Pansegrau, W., Lanka, E., and Guiney, D.G. (1991) Sequence identity in the nick regions of IncP plasmid transfer origins and T-

DNA borders of Agrobacterium Ti plasmids. *Proc Natl Acad Sci U S A* **88**: 1456-1460.

Waters, V.L., and Guiney, D.G. (1993) Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. *Mol Microbiol* **9**: 1123-1130.

Watkins, R.F., and Gray, M.W. (2006) The frequency of eubacterium-to-eukaryote lateral gene transfers shows significant cross-taxa variation within amoebozoa. *J Mol Evol* **63**: 801-814.

Weihofen, W.A., Cicek, A., Pratto, F., Alonso, J.C., and Saenger, W. (2006) Structures of omega repressors bound to direct and inverted DNA repeats explain modulation of transcription. *Nucleic Acids Res* **34**: 1450-1458.

Willetts, N., and Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol Rev* **48**: 24-41.

Williams, S.L., and Schildbach, J.F. (2007) TraY and integration host factor oriT binding sites and F conjugal transfer: sequence variations, but not altered spacing, are tolerated. *J Bacteriol* **189**: 3813-3823.

Wong, J.J., Lu, J., Edwards, R.A., Frost, L.S., and Glover, J.N. Structural basis of cooperative DNA recognition by the plasmid conjugation factor, TraM. *Nucleic Acids Res* **39**: 6775-6788.

Yeo, H.J., and Waksman, G. (2004) Unveiling molecular scaffolds of the type IV secretion system. *J Bacteriol* **186**: 1919-1926.

Yoshida, H., Furuya, N., Lin, Y.J., Guntert, P., Komano, T., and Kainosho, M. (2008) Structural basis of the role of the NikA ribbon-helix-helix domain in initiating bacterial conjugation. *J Mol Biol* **384**: 690-701.

Zampini, M., Derome, A., Bailey, S.E., Barilla, D., and Hayes, F. (2009) Recruitment of the ParG segregation protein to different affinity DNA sites. *J Bacteriol* **191**: 3832-3841.

Zaslaver, A., Mayo, A.E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M.G., and Alon, U. (2004) Just-in-time transcription program in metabolic pathways. *Nat Genet* **36**: 486-491.

Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M.G., and Alon, U. (2006) A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nat Methods* **3**: 623-628.

Zhang, S., and Meyer, R. (1997) The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transfer. *Mol Microbiol* **25**: 509-516.

Ziebuhr, W., Ohlsen, K., Karch, H., Korhonen, T., and Hacker, J. (1999) Evolution of bacterial pathogenesis. *Cell Mol Life Sci* **56**: 719-728.

Ziegelin, G., Pansegrau, W., Strack, B., Balzer, D., Kroger, M., Kruft, V., and Lanka, E. (1991) Nucleotide sequence and organization of genes flanking the transfer origin of promiscuous plasmid RP4. *DNA Seq* **1**: 303-327.

Ziegelin, G., Pansegrau, W., Lurz, R., and Lanka, E. (1992) TraK protein of conjugative plasmid RP4 forms a specialized nucleoprotein complex with the transfer origin. *J Biol Chem* **267**: 17279-17286.