Universidad Nacional Autónoma de México

Centro de Ciencias Genómicas

"Papel de un ARN antisentido en la regulación de la proteína iniciadora de la replicación RepC, en el plásmido simbiótico de *Rhizobium etli* CFN42"

T E S I S

Que para obtener el grado de:
Doctor en Ciencias Biomédicas

Presenta:
Biól. Ramón Cervantes Rivera

Tutor: Dr. Miguel Ángel Cevallos Gaos

Cuernavaca, Morelos, México. Noviembre, 2010.

DEDICATORIA

A mi hermana ***María del Refugio Cervantes Rivera.***

AGRADECIMIENTOS

Tutor principal: **Dr. Miguel Ángel Cevallos Gaos**

Comité tutoral: Dr. David Romero Camarena
Dr. José Luis Puente García

Miembros del jurado: Dr. José Guillermo Dávila Ramos
Dra. Martha Vázquez Laslop
Dr. Juan Miranda Ríos
Dr. Fernando Navarro García

Instituto de Parasitología y Biomedicina “López-Neyra”-CSIC

Dra. Cristina Romero López
Dr. Alfredo Berzal Herranz

Medical University of South Carolina

Dr. Deepak Bastia
Dr. Bidhyut Mohanty

De igual manera, agradezco profundamente a todas las personas que de una u otra manera fueron partícipes de mis logros y frustraciones durante mi largo caminar por la ***Universidad Nacional Autónoma de México.***

RESUMEN

La replicación y segregación del plásmido simbiótico (pRetCFN42d) de *Rhizobium etli* es dependiente de la presencia del operón *repABC*, en este operón se localizan los elementos necesarios para poder llevar a cabo dichas funciones. Todos los operones de la familia *repABC* tienen los tres genes (*repA*, *repB* y *repC*), además de un gene que codifica para un RNA antisentido (ctRNA) y un sitio par (*parS*). Los productos de los genes *repA* y *repB* en conjunto con el sitio *parS* conforman el sistema de segregación, además de cumplir una función como reguladoras negativas de la transcripción del operón. El último gen del operón es *repC*, que codifica la proteína iniciadora de la replicación. El RNA antisentido (ctRNA) codificado en la región intergénica *repB-repC* es un regulador negativo postranscripcional de *repC*. En este trabajo se analizó la estructura secundaria del ctRNA y su blanco (mRNA-*inc*α). Los residuos esenciales para que se lleve a cabo la interacción inicial se localizan en el extremo 5' del ctRNA y en el asa más prominente (L1α) del mRNA- *inc*α. Con estos resultados se propone un modelo de regulación de la replicación por el ctRNA.

ABSTRACT

Replication and segregation of the *Rhizobium etli* symbiotic plasmid (pRetCFN42d) depend on the presence of a *repABC* operon, which carries all the plasmid-encoded elements required for these functions. All *repABC* operons share three protein-encoding genes (*repA*, *repB,* and *repC*), an antisense RNA (ctRNA) coding gene, and at least one centromere-like region (*parS*). The products of *repA* and *repB*, in conjunction with the *parS* region, make up the segregation system and negatively regulate operon transcription. The last gene of the operon, *repC*, encodes the initiator protein. The ctRNA is a negative posttranscriptional regulator of *repC.* In this work, we analyzed the secondary structures of the ctRNA and its target, and mapped the motifs involved in the complex formed between them. Essential residues for the effective interaction localize at the unpaired 5'-end of the antisense molecule and the loop of the target mRNA. In the light of our results, we propose a model explaining the mechanism of action of this ctRNA in the regulation of plasmid replication in *R. etli*.

ÍNDICE **Pag.**

1. INTRODUCCIÓN

a) Replicación de DNA en bacterias.

La transmisión exitosa del material genético requiere de una coordinación precisa entre el inicio de la replicación del genoma y el ciclo celular. El ciclo de duplicación del genoma en bacterias, para fines prácticos, se ha dividido esencialmente en tres fases: inicio, elongación y terminación. El inicio de la replicación, en el modelo clásico en bacterias, ocurre cuando una proteína (Rep) se une al origen de replicación (*ori*), a la vez que facilita la apertura de la doble hélice en la región del *ori*, misma en la que se inicia la síntesis del DNA *de novo*. La proteína iniciadora, además de unirse al *ori*, interacciona con otras proteínas que participan en la fase temprana de la replicación promoviendo el armado y reclutamiento del replisoma en la horquilla de replicación (Johnson *et al*., 2005; Haeusser y Levin, 2008).

Los genomas bacterianos están compuestos por al menos un cromosoma acompañado en muchas ocasiones por uno o varios plásmidos, estos últimos pueden ser circulares o lineales (Summer, 1996). Los cromosomas bacterianos suelen utilizar un mecanismo de replicación bastante conservado que se conoce como mecanismo de replicación tipo theta. En contraste, los plásmidos poseen mecanismos de replicación mucho más variados que incluyen la replicación por círculo rodante y por desplazamiento de cadena.

I. Replicación tipo theta.

Este tipo de replicación es el más ampliamente estudiado como prototipo en los cromosomas y en muchos de los plásmidos circulares de bacterias. El nombre de este tipo de replicación se deriva de la semejanza del plásmido en replicación (visto bajo el microscopio electrónico) con la letra griega theta (θ) (Figura 1) (del Solar *et al*., 1998).

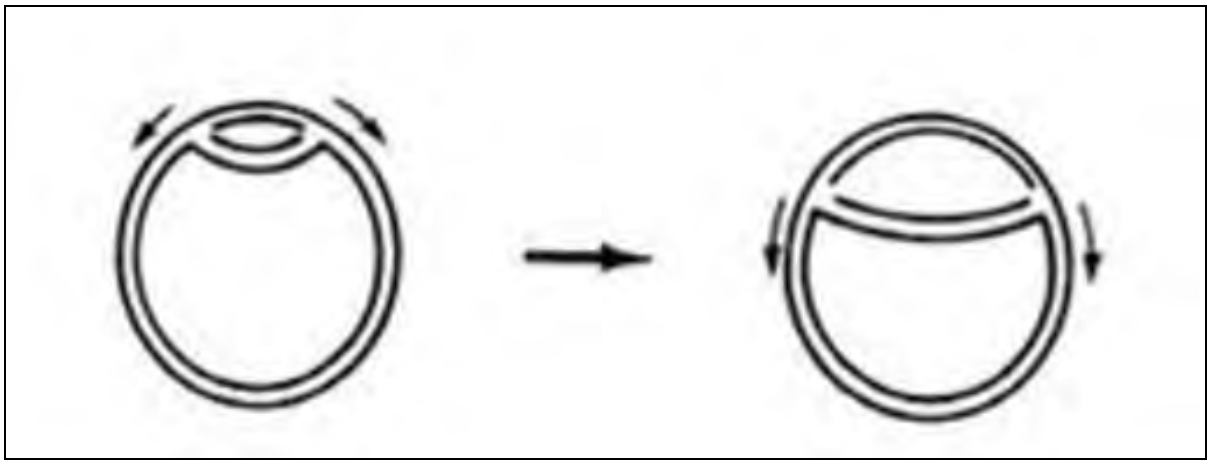


Figura 1. Esquema que muestra el mecanismo de replicación tipo theta. Tomado de del Solar *et al*., 1998.

Los replicones que usan un mecanismo tipo theta comúnmente requieren tanto de una proteína iniciadora de la replicación como de un origen de replicación. En general sólo existe un origen de replicación por cada plásmido y/o cromosoma. El origen de replicación se caracteriza por poseer en su interior o en su vecindad sitios de unión para la proteína iniciadora, sitios de metilación y además suele tener regiones ricas en adenina-timina (AT) para facilitar la apertura de la horquilla de replicación. Las horquillas de replicación en este tipo de replicones pueden desplazarse de manera unidireccional o bidireccional.

El cromosoma de *Escherichia coli* es el modelo por excelencia de replicón con replicación tipo theta bidireccional. El origen de replicación (*oriC*) del cromosoma de *E. coli* se ha delimitado en tan sólo 260 pb y contiene varios sitios de unión (cajas *dnaA*) de la proteína iniciadora de la replicación DnaA, una secuencia de reconocimiento para IHF (factor de integración del huésped) y la secuencia Fis (factor para la estimulación de la inversión) (Figura 2) (Erzberger *et al*., 2002; Davey *et al*., 2002). Existen dos clases de sitios de unión de la proteína DnaA, los sitios de baja afinidad y los de alta afinidad, además de los sitios que se conocen como cajas DnaA-ATP, sitios en los cuales se une la proteína DnaA de manera dependiente de ATP.

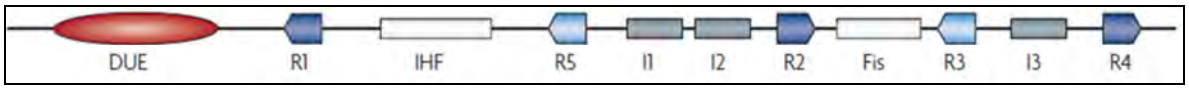


Figura 2. Origen de replicación (*oriC*) mínimo del cromosoma de *E. coli*. Se muestran cada uno de los elementos que conforman el origen. R1, R2 y R4: cajas DnaA de alta afinidad; R3 y R5, cajas DnaA de baja afinidad; IHF, factor de integración del huésped; Fis, factor de inversión simultánea y DUE, elemento de apertura del DNA; I1, I2 e I3, cajas DnaA con una secuencia diferente a la consenso. Tomado de Mott *et al.* 2007.

El inicio de la replicación en el cromosoma de *E. coli* comienza con la unión de la proteína iniciadora DnaA. La capacidad de nucleación de DnaA le permite formar grandes complejos a lo largo del origen de replicación. Una vez que la proteína DnaA satura las cajas *dnaA,* la región del *oriC* sufre una torsión facilitada por la proteína IHF en la región rica en AT, y provoca la apertura de la doble hélice, con lo que se genera el complejo abierto que expone el DNA de cadena sencilla para que se complete el ensamblado del replisoma. El siguiente paso en el inicio de la replicación es la interacción de la helicasa replicativa (DnaB) con la proteína DnaA dentro del complejo abierto del origen de replicación, la helicasa DnaB obtiene la energía para desplazase (en dirección 5'-3') de la hidrólisis de ATP. Cuando la helicasa DnaB no está asociada al DNA forma un complejo con la proteína DnaC (Seitz *et al*., 2000).

La presencia del complejo DNA-DnaB es la señal para la incorporación de la primasa DnaG (que sintetiza los pequeños prímeros iniciadores de RNA) (Frick *et al*., 2001). La síntesis de los prímeros de RNA es esencial para dar paso a la siguiente fase de la replicación, la elongación por la holoenzima DNA polimerasa III (Pol III). Una vez que se han ensamblado los replisomas en el cromosoma de *E. coli* se desplazan en ambas direcciones, la replicación en la cadena líder es continua, en tanto que en la retardada es discontinua. La replicación en la cadena retardada se lleva a cabo mediante los fragmentos de Okasaki (de 1 a 2 kb de longitud aproximadamente), cuya síntesis requiere la presencia de prímeros de RNA. Después de que ocurrió la síntesis de los fragmentos de Okasaki, los prímeros son removidos por la RNasa H (de corte específico en un heterodúplex RNA-DNA). Posteriormente, la DNA polmerasa I (Pol I) completa la cadena de DNA y finalmente la DNA ligasa une los fragmentos vecinos (Kornberg y Baker, 1992). La DNA Pol III, una polimerasa de alta fidelidad, es quien lleva a cabo el proceso de elongación del DNA *de novo*. Este atributo requiere de la participación de varias subunidades, que se ensamblan en tres sub-grupos: el núcleo, compuesto por las subunidades alfa (α), épsilon (ε) y theta (θ); la pinza deslizante (*sliding clamp*) compuesta por dos subunidades beta (β) y el cargador de la pinza (*clamp loader*) compuesto por las subunidades delta prima(δ'), gamma (γ), tau (τ), delta (δ), psi (ψ) y ji (χ). El avance de las horquillas de replicación genera una zona de DNA de cadena sencilla cuya integridad depende de las proteínas SSB (proteínas de unión a cadena sencilla de DNA), que se unen al DNA monocatenario (McHenry, 2003). La terminación de la replicación en el…

cromosoma de *E. coli* se da justo en el punto opuesto al origen de replicación, la proteína Tus se une en un estado monomérico a los sitios *ter*. La unión de las proteínas Tus funciona como un bloqueo polar del avance de las horquillas de replicación (Hill, 1992).

El mecanismo de replicación tipo theta con desplazamiento de la horquilla de manera unidireccional sólo se ha descrito en plásmidos. El plásmido más conocido que replica por un mecanismo tipo theta de manera unidireccional es el pR1 de *E. coli*. El origen de replicación (*ori*R1) de este plásmido es de sólo 188 pb (Masai *et al*., 1983). El pR1 inicia su replicación una vez que la proteína iniciadora RepA se une al origen de replicación, que se localiza corriente abajo del gene *repA* (del Solar *et al*., 1998). La unión de RepA al origen de replicación induce la torsión del *ori*, lo que provoca la apertura de la doble hélice que de algún modo se facilita por la presencia de una región rica en AT. La proteína RepA una vez unida al origen de replicación recluta a los demás miembros del replisoma y avanza la replicación de manera unidireccional (Díaz y Staudenbauer, 1982). La terminación de la replicación en este plásmido se da justo en el punto donde inicia la replicación, una vez que la horquilla ha dado la vuelta completa al plásmido.

II) Replicación tipo círculo rodante.

La replicación por círculo rodante (RCR) siempre es unidireccional, además de considerarse asimétrica puesto que la síntesis de la cadena líder con respecto a la retardada no está acoplada (del Solar *et al.*, 1993). Este tipo de replicación se ha descrito en bacterias Gram-positivas, cianobacterias y algunas especies de arqueas (Gielow *et al.*, 1991). Todos los plásmidos que replican por este mecanismo tienen un tamaño menor a 10 kb, lo cual no implica que todos los plásmidos pequeños replican por un mecanismo tipo círculo rodante.

Se pueden definir dos pasos en la replicación por el mecanismo de círculo rodante (Yasukawa *et al.*, 1991; Kleanthous *et al*, 1991): el primer paso es la unión de la proteína iniciadora al origen de cadena doble (DSO), que suele ser menor a 100 pb. En esta posición la proteína Rep induce un corte en cadena sencilla a partir del cual se inicia la síntesis de la cadena líder (del Solar *et al.*, 1998; Khan, 1997). Los orígenes de cadena sencilla (SSO) de los plásmidos que tienen un mecanismo tipo RCR contienen las dos secuencias, la de unión de la proteína Rep y la de corte. Los sitios de corte son altamente conservados en tanto que los sitios de unión de las proteínas Rep son más variables.

Se han descrito estructuras cruciformes y estructuras tallo-asa en varios de los plásmidos que tienen un mecanismo tipo RCR (Moscoso *et al*.,1995; Noirot *et al*., 1990). Estas estructuras podrían estar involucradas en el reclutamiento/utilización de la proteína Rep y por ende en el inicio de la replicación. El segundo paso es la síntesis de la cadena retardada, el inicio de la síntesis de esta cadena ocurre en los SSO, diferente al origen de cadena doble (Koepsel *et al,* 1985; Thomas *et al,.* 1990; de la Campa *et al.,* 1990)(Figura 3). La terminación en plásmidos que replican por un mecanismo tipo círculo rodante se da por una pequeña secuencia (*nic*) que se localiza justo en el sitio de corte en cadena sencilla.

Uno de los plásmidos más estudiados que replican por el mecanismo tipo círculo rodante es el pT181. En este plásmido la replicación se inicia con la unión de la proteína RepC al DSO lo que provoca la formación de una estructura que expone el sitio de corte en un asa. Hecho el corte, la proteína RepC se une de manera covalente al fosfato 5' a través de un residuo de tirosina que se localiza en el sitio activo de la proteína RepC. Posteriormente, la proteína RepC recluta a la helicasa replicativa (PcrA en el caso de las bacterias Gram-positivas), mediante una interacción proteína-proteína altamente específica. A continuación, la helicasa replicativa abre la doble hélice y las proteínas SSB se unen a la cadena desplazada para protegerla de la degradación.

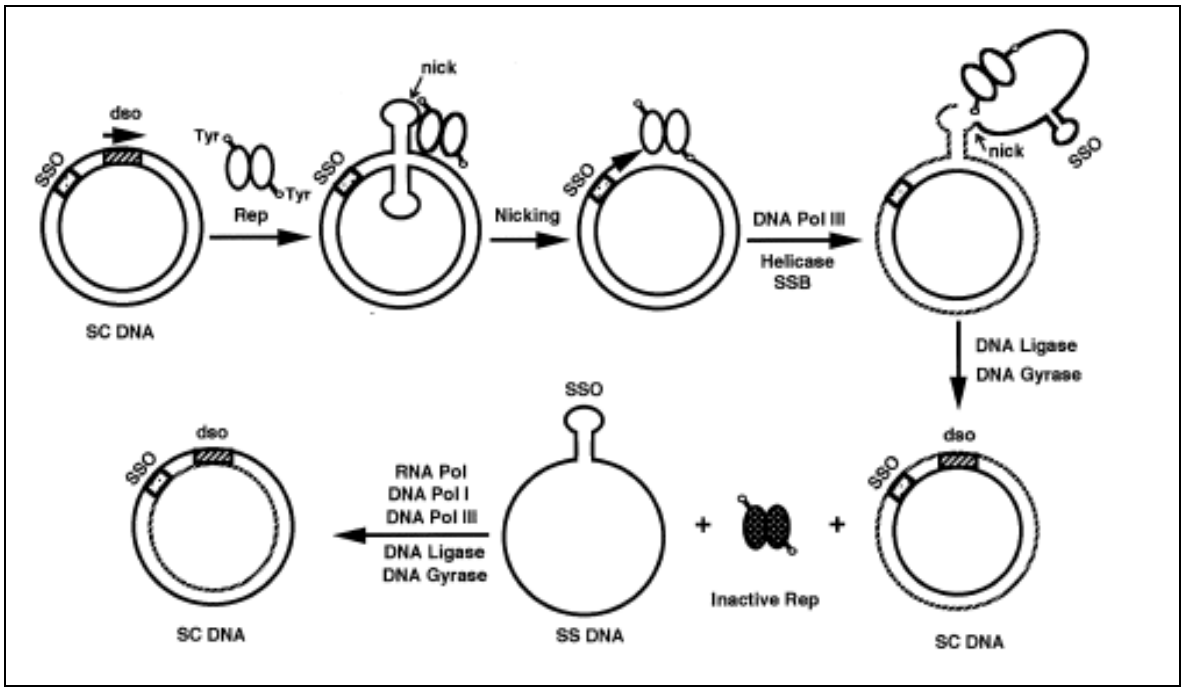


Figura 3. Modelo de replicación de un plásmido por el mecanismo de círculo rodante. Tomado de Khan *et al*., 2005.

Una vez que se ha terminado de sintetizar la cadena sencilla, la misma proteína RepC que se quedó unida al otro extremo de la cadena desplazada hace el corte y recircularíza la cadena sencilla recién replicada. La conversión del plásmido de cadena sencilla en cadena doble se inicia a partir de un prímero de RNA, y se completa gracias a la actividad de la Pol I primero, y la Pol III después (Khan, 2005).

III) Replicación por desplazamiento de cadena.

El ejemplo más estudiado de plásmido que replica por un mecanismo por desplazamiento de cadena es el plásmido promiscuo pRSF1010 (Figura 4). La replicación en este plásmido se lleva a cabo por las proteínas RepA, RepB y RepC. RepA es la helicasa replicativa que permite el avance de la horquilla de replicación, RepB es la primasa encargada de la síntesis del prímero de RNA a partir del cual inicia la síntesis del DNA *de novo*, y finalmente RepC es la proteína iniciadora de la replicación (Honda *et al*., 1991).

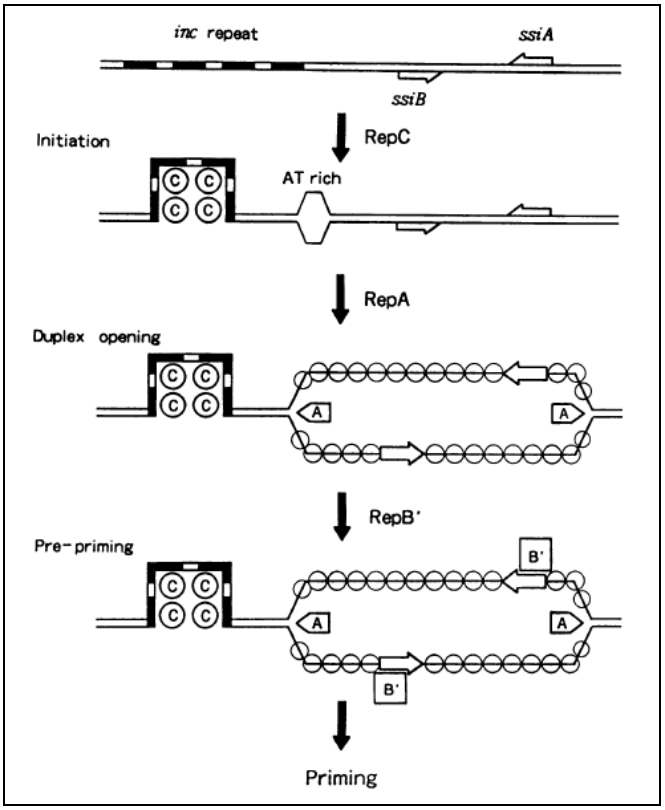


Figura 4. Representación esquemática de la replicación del plásmido pRSF1010 por el mecanismo de desplazamiento de cadena. Tomado de Honda *et al*., 1991.

El primer paso en el inicio de la replicación del pRSF1010 es la unión de RepC a los iterones (secuencias directas repetidas en tandem), que se localizan en el origen de replicación, la unión de RepC a los iterones desestabiliza la región rica en AT lo que facilita que la proteína RepA con función de helicasa se una a dicha región, adyancete a la cual se localiza el *ori*. La unión de RepA facilita la apertura de la doble hélice, lo que provoca el desplazamiento de la cadena no replicante, de tal suerte que permite la síntesis continua del DNA *de novo* de manera bidireccional. La cadena que se desplaza se protege de la degradación por la unión de las proteínas SSB. La elongación en el plásmido pRSF1010 la lleva a cabo la holoenzima DNA Pol III (del Solar *et al*., 1998).

b) Proteínas iniciadoras de la replicación.

La necesidad de una proteína iniciadora es casi omnipresente en la replicación de los diferentes plásmidos no importando mediante que sistema repliquen, como puede apreciarse en los tres diferentes modelos de replicación descritos arriba. Desgraciadamente, hay muchas proteínas iniciadoras de la replicación en las cuales no se ha realizado mucha investigación y por lo tanto no se conocen detalles de sus mecanismos de acción. No obstante, hay algunas proteínas iniciadoras de la replicación que si se tienen bien caracterizadas.

I. DnaA, proteína iniciadora de la replicación del cromosoma de *Escherichia coli*.

La capacidad de una proteína iniciadora de coordinar varios eventos de manera simultánea es esencial para el correcto ensamblado de los complejos implicados en la replicación del DNA. DnaA, la proteína iniciadora de la replicación del cromosoma de *E. coli*, tiene una organización multi-dominio (Figura 4). En años recientes se ha diseccionado y caracterizado cada uno de los dominios y sus funciones con ensayos bioquímicos y genéticos, además, se cuenta con la estructura cristalográfica de la proteína sola o formando un complejo con el origen de replicación (Mott *et al*., 2007).

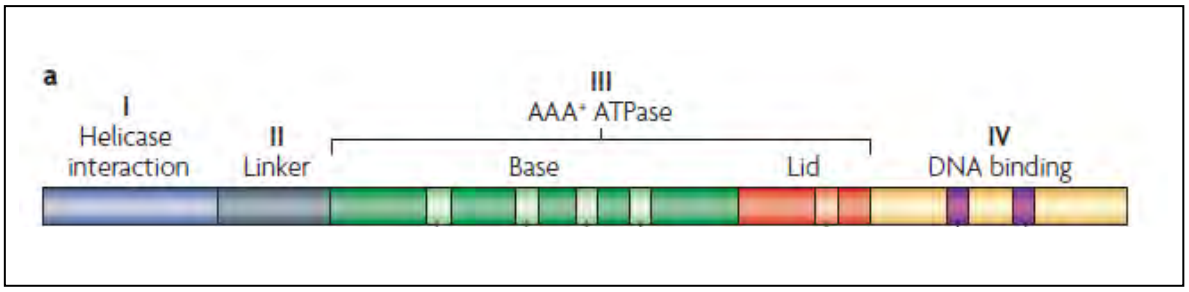


Figura 4. Esquema que muestra la estructura multidominio de la proteína DnaA iniciadora de la replicación del cromosoma de *E. coli*. Tomado de Mott *et al*., 2007.

La proteína DnaA posee cuatro dominios (I, II, III y IV): el amino terminal (dominio I) interacciona con la helicasa DnaB, y se encuentra unido al resto de la proteína por el dominio II, un enlazador (*linker*) poco conservado y con un alto grado de flexibilidad que le permite el movimiento casi independiente al dominio I del resto de la proteína. En el dominio III, residen por un lado, la actividad de ATPasa y la capacidad de oligomerizar, característica esencial en el proceso de nucleación en la región del *oriC*. Finalmente, el dominio IV se encarga de la unión específica de la proteína a las cajas *dnaA* del *oriC* del cromosoma de *E. coli* (Kaguri, 2006).

II. RepC, proteína iniciadora de la replicación del plásmido T181 de *Staphylococcus aureus*.

Desafortunadamente no todas las proteínas iniciadoras de la replicación despiertan tanto interés en la ciencia como la proteína DnaA. Sin embargo, existen otras proteínas iniciadoras, pocas, que se han estudiado con cierto grado de detalle. Una de ellas es la proteína iniciadora RepC, del plásmido pT181. Esta proteína fue la primera iniciadora de la replicación en plásmidos que replican por el mecanismo tipo círculo rodante que se describió, sin embargo, en comparación con otras iniciadoras, es poco el avance que se ha hecho con respecto al conocimiento de su estructura y sus dominios funcionales.

De manera más bien vaga se dice que RepC tiene dos dominios funcionales: uno involucrado en el reconocimiento del DSO, y otro encargado del corte en cadena sencilla a partir del cual da inicio la replicación del plásmido (Khan, 2005). Se sabe que la tirosina activa (a través de la cual se une de manera covalente a la cadena recién cortada), está en la posición 191 de RepC, además, que los aminoácidos 265 al 270 son críticos para la unión específica de la proteína al DSO (Khan, 2005).

Experimentalmente se han descrito dos diferentes estados de oligomerización de la proteína iniciadora RepC (monomérico y dimérico), en solución acuosa se encuentra como monómero, sin embargo, cuando está presente un fragmento de DNA que contiene el *ori*, forma dímeros (Chang *et al.*, 2000). No obstante, no se han identificado los dominios o los aminoácidos involucrados en este proceso. Futuros experimentos seguramente irán encaminados a la identificación de estas funciones en la proteína.

En el campo de las proteínas iniciadoras de la replicación por el mecanismo RCR queda mucho por investigar. Una exploración exhaustiva de la secuencia de RepC, podría dar un poco de luz acerca de los dominios involucrados en la interacción con otros elementos del replisoma, como la helicasa PcrA, esencial para el avance de la horquilla de replicación.

III. RepE, proteína iniciadora de la replicación del plásmido F de *Escherichia coli*.

La proteína RepE del plásmido F de *E. coli* tiene un papel preponderante en la replicación del plásmido. RepE tiene dos funciones diferentes dependiendo de su estado de oligomerización (Kline, 1985). RepE monomérica funciona como iniciadora de la replicación al unirse a los iterones del origen de replicación (*ori2*), en tanto que en su conformación dimérica reprime la transcripción del gene que la codifica al unirse a la región promotora/operadora (Figura 6) (Kawasaki *et al*., 1990). El estado monomérico o dimérico es dependiente de la actividad de la chaperona DnaK (Nakamura *et al*., 2007).

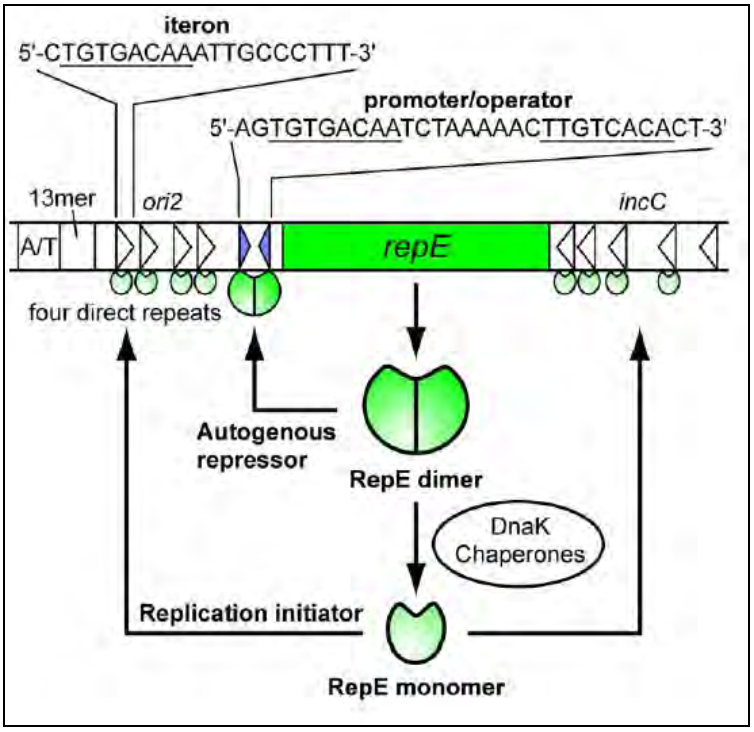


Figura 6. Esquema representativo de las funciones de RepE en la replicación del mini plásmido F de *E. coli*, Tomado de Nakamura *et al*., 2007.

RepE consiste de dos dominios; el dominio amino-terminal (NTD) y el domino carboxi-terminal (CTD), cada uno de ellos con un motivo hélice-alfa-hélice (HTH) (Brennan, 1993). El dominio NTD está involucrado en la oligomerización de la proteína y cumple funciones auxiliares en la unión al DNA. El dominio CTD es el encargado de la unión específica a los iterones del *ori2* (Nakamura *et al.*, 2007).

c) Orígenes de replicación ubicados en la región codificante de la proteína iniciadora.

El origen de replicación comúnmente se localiza en una región relativamente cercana al gene que codifica la proteína iniciadora de la replicación. No obstante, se han descrito a la fecha algunos plásmidos y bacteriófagos en los cuales el origen de replicación está dentro de la región codificante de su proteína iniciadora, este modelo de organización unimodular de la replicación aparentemente no es de amplia distribución en la naturaleza.

El pAD1 es un plásmido conjugativo de 60 kb, originalmente descrito en *Enterococcus faecalis* DS16, que codifica una citolisina que contribuye a la virulencia en modelos animales. Este plásmido codifica tres proteínas involucradas en la replicación y en la segregacíon del plásmido: RepA es la proteína iniciadora de la replicación, en tanto RepB y RepC participan en la segregación del plásmido. El origen de replicación del pAD1 se localiza dentro de la región codificante de RepA. El *oriV* mínimo es de 170 pb y contiene dos secuencias directas repetidas (DR-1 y DR-2) y una secuencia inversa repetida (IR-1) esenciales para mantener la replicación del plásmido (Francia *et al.*, 2004).

El plásmido pSX267 de *Staphylococcus xylosu* posee el origen de replicación dentro de la región codificante de la proteína Rep (RepA). Este plásmido replica por un mecanismo tipo theta cuyo origen de replicación tiene dos secuencias directas repetidas que podrían estar involucradas en la regulación de la replicación del plásmido y podría ser el sitio de unión de RepA (Gering *et al*., 1996).

El plásmido SK41 de *Staphylococcus aureus* es de gran importancia médica puesto que confiere resistencia a varios antibióticos (Berg *et al*., 1998). Los componentes esenciales para la replicación de este plásmido se delimitan en una región de 1.25 kb de longitud. El origen de replicación está dentro de la región codificante de la proteína Rep. Esta proteína se une a los cuatro iterones (cajas *rep*) de 24 pb que se localizan en la parte central del gene. Adicional a estos iterones, el pSK41 tiene dos tipos más de secuencias directas repetidas, de las cuales no se conoce su función, pero se sabe que son esenciales para la replicación eficiente del plásmido. El inicio de la replicación de este plásmido está regulada por un RNA antisentido, codificado justo corriente arriba del codón de inicio de la proteína Rep (Kwong *et al*., 2004).

El plásmido LS32 se aisló originalmente de la cepa Natto de *Bacillus subtilis*. Este plásmido replica por un mecanismo tipo tetha de manera independiente de DnaA y DNA polimerasa I (Hassan *et al.*, 1997). La replicación del pLS32 es dependiente de RepN, la proteína iniciadora de la replicación y del origen de replicación que se localiza en su región codificante. El *ori* del pLS32 tiene cinco iterones, sitios de unión de la proteína RepN. Adyacente a los iterones, este plásmido tiene una región rica en AT, característica común de los plásmidos que replican por un mecanismo tipo theta (Tanaka *et al.*, 2005).

El bacteriófago N15 es especialmente interesante; el mecanismo por el que replica tiene características tanto eucariotas como procariotas. En el estadío no replicativo del fago es una molécula lineal, poco antes de infectar la bacteria se circulariza valiéndose de sus extremos cohesivos y con el auxilio de la protelomerasa (Malinin *et al.*, 1992). La proteína iniciadora es RepA, una proteína de 1324 aa con motivos de primasa y de helicasa (Ravin *et al.*, 2000). El origen de replicación del bacteriófago N15 está dentro de la secuencia codificante de RepA (Ravin *et al.*, 2000).

La región del origen de replicación tiene tres iterones, una secuencia semejante a una caja *dnaA*, y una región rica en AT de 22 pb, todas ellas características de los plásmidos que replican por un mecanismo tipo theta. Experimentos realizados por Ravin *et al.*, 2003 y sus colaboradores han demostrado que la replicación de este bacteriófago ocurre a través de dos mecanismos: de manera bidireccional y por un mecanismo tipo theta.

La replicación del bacteriófago λ es esencialmente como el mecanismo de replicación tipo theta en plásmidos bacterianos previamente descrito. El bacteriófago λ fue el primero en el cual se describió a detalle el mecanismo tipo theta en un replicón de bacteriófago (Taylor y Wegrzyn, 1995). Este bacteriófago inicia su replicación a partir de un único origen de replicación (*oriλ*), que se ubica en la parte central del gene que codifica la proteína iniciadora de la replicación RepO. El *oriλ* tiene cuatro secuencias directas repetidas de 18 pb, que son los sitios de unión específicos de la proteína RepO, adyacente a la región en la cual se localizan los iterones, el *oriλ* tiene una región rica en AT esencial en la apertura del DNA bicatenario (Hase *et al*., 1989). Una vez que la proteína RepO se une a los iterones, el siguiente paso es el ensamblado del replisoma, compuesto esencialmente por las proteínas del huésped. Ya ensamblado el replisoma inicia la replicación bajo el esquema tipo theta, predominantemente de manera unidireccional (Hase *et al*., 1989).

d) Regulación de la replicación en plásmidos bacterianos.

La regulación de la replicación del genoma en todos los organismos vivos es esencial por dos razones: la primera, asegurarse que toda la descendencia posea una copia de la información genética, por lo que es necesario regular y coordinar muy bien la duplicación del material genético con el ciclo celular, y por otro lado, evitar una sobre replicación, ya que el gasto metabólico que esto implica, en términos de economía celular, puede conducir a las células bacterianas a una muerte por exceso de replicación de su material genético, lo que se conoce como fenómeno *runaway*. En el caso particular de los plásmidos bacterianos, se han descrito diferentes mecanismos de regulación de la replicación. Cada familia de plásmidos tiene una o más formas de regular su número de copias.

El mecanismo de regulación de la replicación en plásmidos bacterianos con la participación de iterones fue de los primeros que se describieron. El hecho de que los iterones sean elementos de secuencias directas repetidas en tándem los hace conspicuos y de inmediato hacen sospechar que podrían estar implicados en alguna función importante, además de que suelen localizarse dentro del origen de replicación, en la región promotora de la proteína Rep o en una región corriente abajo del gene que codifica la proteína Rep (Paulsson *et al*., 2006).

Se conocen al momento cuatro mecanismos por los que cuales los iterones pueden regular la replicación de plásmidos : 1) Autorepresión transcripcional que reduce la síntesis de la proteína iniciadora, 2) Dimerización de la proteína iniciadora, 3) Titulación de la proteína iniciadora, 4) Inactivación del origen de replicación por "esposado" (*handcuffing*).

I. Autorrepresión transcripcional que reduce la síntesis de la proteína iniciadora.

De las diferentes vertientes que se conocen en las que está involucrada la participación de iterones, la autorepresión transcripcional es la que más se conoce y se ha estudiado. El plásmido modelo para el estudio de este tipo de regulación es el plásmido P1 de *E. coli*. La proteína iniciadora de la replicación de este plásmido es la proteína RepA. El origen de replicación del plásmido se localiza corriente arriba del codón de inicio de la proteína RepA (Mukhopadhyay y Chattoraj, 2000). La región promotora/operadora de la proteína RepA se sobrelapa con el *ori* del plásmido, por lo tanto el arreglo en tándem a los cuales se une la proteína iniciadora se extiende hasta la región reguladora del gene *repA*. Para que se pueda iniciar la replicación del plásmido P1 es necesario que la proteína RepA sature el *ori*, lo cual implica que la región promotora/operadora de RepA quede oculta y por lo tanto inhibe su propia transcripción y de esta manera evita la sobre síntesis de la proteína Rep (Chattoraj *et al*., 1985).

II. Titulación de la proteína iniciadora.

Como ya se ha mencionado anteriormente, el inicio de la replicación en plásmidos está condicionado a que el *ori* esté saturado con la proteína Rep, es decir, que todos los sitios de unión específicos de la proteína se encuentren ocupados. Por lo tanto, la titulación de la proteína iniciadora de la replicación es una manera de prevenir la sobre replicación. Este tipo de regulación está presente en los plásmidos P1 y F de *E. coli* (Chattoraj, 2000).

III. Dimerización de la proteína iniciadora.

La dimerización de la proteína iniciadora de la replicación es un mecanismo de regulación de la replicación, que también está presente en plásmidos que tienen iterones como sistema de regulación. La mayoría de las proteínas iniciadoras de la replicación pueden interaccionar consigo mismas para formar complejos de nucleación sobre la región del *ori*. Sin embargo, cuando una proteína iniciadora de la replicación no se encuentra unida al origen de replicación, puede presentarse en dos oligoformas alternativas: como monómero y como dímero. En el caso de la proteína iniciadora de la replicación π del plásmido R6K de *E. coli*, la oligoforma capaz de unirse a los iterones del *oriγ,* y por tanto de iniciar la replicación, es el monómero. Sin embargo, la proteína en solución se encuentra predominantemente como dímero. Estudios recientes han demostrado que la…

transición de la forma inactiva en la replicación (forma dimérica) a la forma activa (forma monomérica) es dependiente de la actividad de la chaperona DnaK (Zzaman *et al*., 2004).

IV. Inactivación del origen de replicación por esposamiento (*handcuffing*).

La oligomerización de la proteína Rep puede mantener a la proteína en su estado activo como monómero, o bien, en su estado inactivo como dímero. Una vía alterna, es que la proteína dimerice una vez unida al *ori*, dando lugar al esposamiento o *handcuffing*, que provoca que dos moleculas de ADN recién replicadas se mantengan unidas por medio de la proteína iniciadora, este es otro mecanismo por el que se regula la replicación de algunos plámidos (Das y Chattoraj, 2004).

Una vez más el plásmido en el que más se ha estudiado el mecanismo de regulación por esposamiento es en el plásmido P1 de *E. coli*. Como ya se ha hecho mención, la proteína RepA se une al origen de replicación en forma monómerica, sin embargo, trabajos recientes han demostrado, de manera fehaciente, que la proteína puede dimerizar una vez que se ha unido al *ori*. De esta manera quedan unidos dos plásmidos por sus respectivos orígenes de replicación evitando que se inicie un nuevo ciclo de replicación (Chattoraj, 2000).

e) Regulación de la replicación de plásmidos bacterianos por RNAs antisentido.

Además de la regulación de la replicación a nivel de motivos de secuencia y proteínas, así como la formación de complejos DNA-proteína, en los últimos años se han descrito una gran cantidad de pequeños RNAs antisentido que tienen una participación activa en la regulación de la replicación en plásmidos bacterianos.

Los RNAs antisentido están ampliamente distribuidos en la naturaleza, tanto en los organismos eucariotas como en los procariotas. Los primeros RNAs antisentido que se descubrieron y se caracterizaron fueron los de *E. coli*, particularmente un RNA antisentido que regula la replicación del plásmido ColE1. A partir de esta primera descripción del mecanismo de acción de un RNA antisentido en el plásmido ColE1 descrito en los trabajos pioneros de Tomizawa *et al*. (1981), se abrió una puerta a la búsqueda de RNAs antisentido que participaran en funciones esenciales en bacterias. Con el paso de los años la pasión por la búsqueda y caracterización de RNAs antisentido se extendió al mundo de los eucariotas.

En los procariotas, a pesar de que la entrada en escena de los RNAs antisentido se dió con su descripción en la participación como regulador de la replicación de un plásmido, en los sucesivos 20 años se describieron RNAs, como el OxyS y el DsrA, que juegan papeles importantes en la respuesta a varios tipos de estrés (Zhang *et al*., 1998). En el campo de los plásmidos, los RNAs antisentido participan activamente en procesos como la segregación y la conjugación, sin dejar de lado los nuevos descubrimientos de RNAs antisentido que participan en la regulación de la replicación de otros plásmidos.

Los RNAs antisentido son pequeños, difusibles, no traducibles y altamente estructurados. Su función la ejercen al unirse a su RNA blanco, de esta manera controlan la expresión de su gene blanco. Existes dos tipos de RNA antisentido, los codificados en *cis*, es decir que sus genes se encuentran en la cadena opuesta a la de su blanco y los codificados en *trans*, en los que su blanco es una región diferente a la cual están codificados.

Estudios comparativos recientes sugieren que la eficacia en la regulación por un RNA antisentido está directamente relacionada con la presencia de una o varias estructuras tallo-asa ricas en Guanina-Citocina (GC), cada una de ellas con un asa de 5-6 nt de longitud y con tallos con burbujas que ayudan a estabilizar el RNA haciéndolo menos susceptible a las RNasas (Wagner *et al*., 2002). El asa por la que comúnmente inicia la interacción del RNA antisentido con su blanco frecuentemente contiene un motivo 5'-YUNR, que constituye lo que se llama un giro-U (Franch y Gerdes, 2000).

Los RNAs antisentido que regulan la replicación en plásmidos tienen varias maneras de actuar: a) Inhibición de la maduración del prímero de RNA; b) Inhibición de la traducción de la proteína iniciadora; c) Atenuación transcripcional y d) Inhibición de la formación de un pseudonudo.

I. Inhibición de la maduración del prímero de RNA.

El plásmido ColE1 tiene la peculiar característica de que no necesita una proteína iniciadora de la replicación, si no que sólo requiere un prímero de RNA (RNA II) para que este proceso inicie. El primer paso en la replicación del plásmido ColE1 es la síntesis de un pre-prímero. Durante la síntesis de dicho RNA se forma un híbrido RNA-DNA en la región del origen de replicación. El pre-prímero lo procesa la RNAsa H en el *oriV* y de esta manera se genera un extremo 3'-OH necesario para que la polimerasa pueda iniciar la elongación. La replicación está regulada por el RNA I (108 nt), que está codificado en la cadena complementaria y que al unirse con el RNA II impide que éste último tome la conformación necesaria para que pueda actuar la RNasa H en el lugar adecuado y con ello, evita la maduración del prímero de RNA (Figura 7) (Eguchi *et al*., 1991).

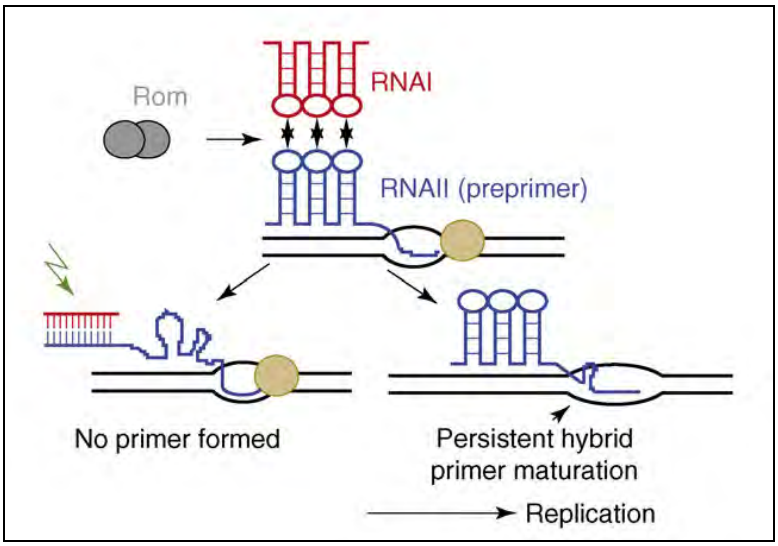


Figura 7. Representación esquemática de la regulación del inicio de la replicación del plásmido ColE1. Tomado de Brantl, 2007.

La unión del RNA I con el RNA II se inicia a partir de las asas y posteriormente se extiende a lo largo de toda la molécula. En este caso se ha demostrado que no se necesita la formación de un dúplex completo para ejercer el control sobre la replicación (Brantl *et al*., 2007).

II. Inhibición de la traducción de la proteína iniciadora.

El plásmido R1 de *E. coli*, es el modelo donde más se ha estudiado la inhibición de la traducción de la proteína iniciadora de la replicación por un RNA antisentido. El replicón básico de este plásmido tiene tres marcos abiertos de lectura: *copB*, *tapA* y *repA* que codifican las proteínas: CopB, el represor transcripcional, TAP un péptido líder de 24 aa y la proteína esencial para la replicación RepA (Figura 8) (Persson *et al*, 1988).

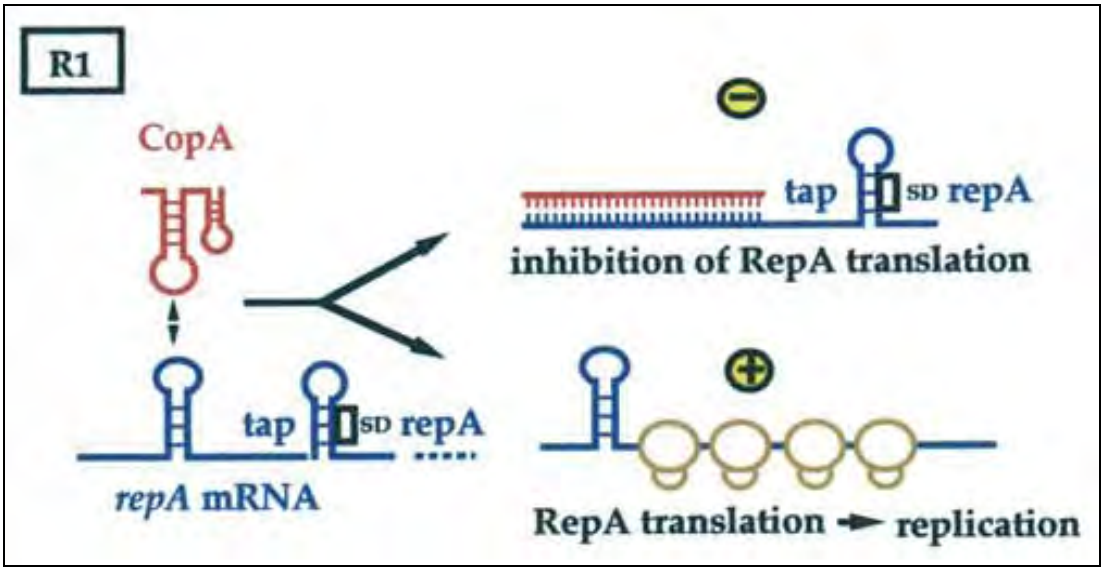


Figura 8. Modelo que muestra la regulación de la replicación del plásmido R1 de *E.coli*. por el RNA antisentido *copA*. Tomado de Brantl, 2007.

El origen de replicación *oriR* se localiza corriente abajo del gene *repA*, en tanto que el RNA antisentido CopA (de aproximadamente 90 nt), está codificado corriente arriba del codón de inicio de *repA*, pero en la cadena complementaria. La estructura secundaria de CopA está constituida por dos estructuras tallo-asa (Figura 8). El blanco de CopA es CopT, una región del mRNA de *repA*. La unión de CopA con CopT inhibe la traducción de Tap, un pequeño péptido que está acoplado traduccionalmente con *repA*. La formación del dúplex copA-copT lo convierte en un blanco de la RNasa III (Kolb *et al*., 2000).

Por ensayos de interacción *in vitro* del RNA antisentido (CopA) con su RNA blanco (CopT) se ha demostrado que la formación del complejo CopA-CopT se inicia a partir de sus asas, interacción que posteriormente se extiende a lo largo de la molécula, lo que se conoce como *kissing* extendido. La constante de formación del complejo se calcula en 3×10^6 M^{-1} s^{-1}, lo cual implica que es una reacción extremadamente rápida y eficiente. Experimentos posteriores también han demostrado que no es necesario que se forme el dúplex completo para que se inhiba la traducción de la proteína *repA*, basta con que se inicie la interacción para que el nivel de traducción de *repA* disminuya (Nordstrom, 2006).

III. Atenuación transcripcional.

El primer plásmido en el cual se describió una regulación de su replicación por atenuación transcripcional mediada por un RNA antisentido fue el pT181 de *Staphylococus*. Este plásmido replica por un mecanismo tipo círculo rodante, como ya se describió en un apartado anterior (Novick *et al*., 1989). El blanco del RNA antisentido (RNA I) en este plásmido es el mRNA de la proteína iniciadora (RepC). El RNA I del pT181 se transcribe en dirección contraria a la de *repC*, el plegamiento de este RNA da como resultado una estructura conformada por tallos-asa esenciales para que se pueda llevar a cabo la interacción del RNA I con el mRNA de *repC* (Brantl *et al*., 2000). Durante su transcripción el mRNA de *repC* puede adoptar dos conformaciones que son dependientes de la presencia o ausencia del RNA antisentido. En ausencia del RNA I, el mRNA-*repC* se pliega… formando una estructura que evita la terminación prematura de la transcripción de *repC*, en tanto que en presencia del RNA antisentido se forma un híbrido RNA I/mRNA-*repC* que provoca un cambio conformación en la estructura secundaria del mRNA-*repC* que lo convierte en un terminador *rho*-independiente, hecho que provoca la terminación temprana de la transcripción y por lo tanto, de la síntesis de RepC. (Brantl *et al*., 2000).

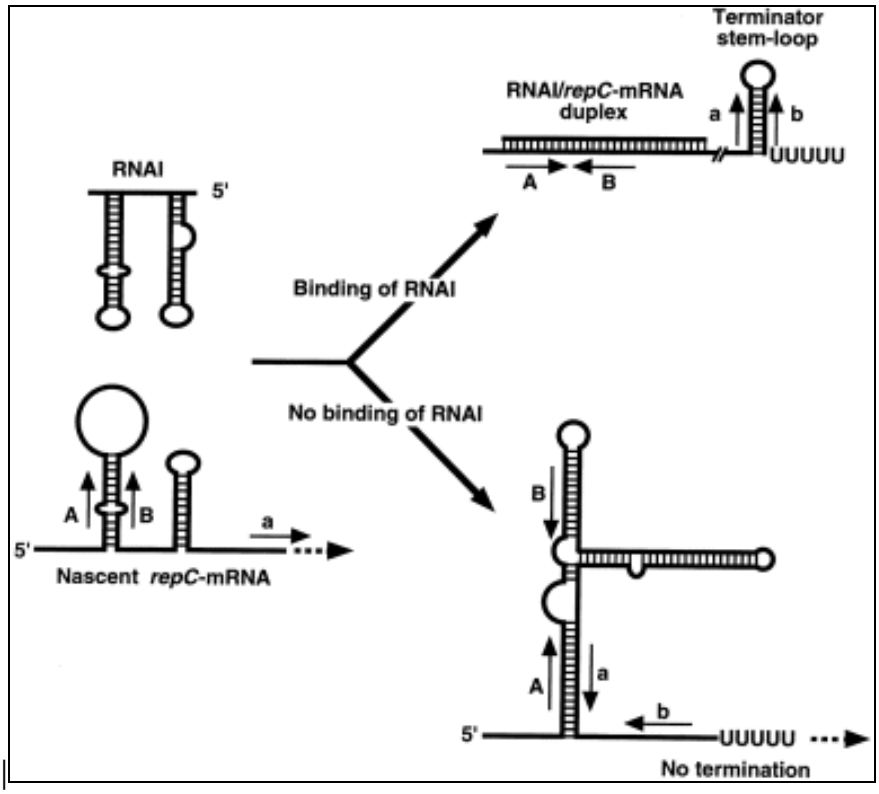


Figura 9. Regulación de la replicación del pT181 por el RNA antisentido RNA I. Tomado de Brantl, 2000.

Ensayos de unión *in vitro* del RNA antisentido con el mRNA blanco han demostrado que los tiempos en los que se lleva a cabo la formación del complejo es del orden de 10 a 20 segundos. Así también se ha demostrado que no es necesaria la formación de un dúplex perfecto para terminar de manera prematura la transcripción, la inhibición se lleva a cabo desde el primer contacto del RNA antisentido con el mRNA (Wagner and Brantl, 1998) (Figura 9).

III. Inhibición de la formación del pseudonudo.

Esta forma de regulación de la replicación se ha estudiado ampliamente en los plásmidos Incβ (pMU720) e IncIα (ColIb-P9) (Asano *et al*., 1999). Este grupo de plásmidos se caracterizan por la formación de un pseudonudo como requisito indispensable para que pueda traducirse de manera eficiente la proteína Rep (Asano *et al*., 1991).

El replicón del plásmido pColIb-P6 está compuesto por un péptido líder (RepY) de 29 aa acoplado a la proteína iniciadora de la replicación (RepZ), el origen de replicación que está corriente abajo del gene de *repZ* y un RNA antisentido (RNA I) localizado corriente arriba del codón de inicio del péptido lider RepY.

La estructura secundaria del mRNA, determinada experimentalmente por ensayos de degradación parcial *in vitro,* han mostrado la presencia de dos estructuras tallo-asa, una corriente arriba del SD de *repY* (estructura I) y la otra en la parte media del gene (estructura III) (Figura 10). La estructura III tallo-asa del mRNA ocluye el SD de *repZ*, así como la secuencia complementaria al asa de la estructura I (Asano *et al*, 1999).

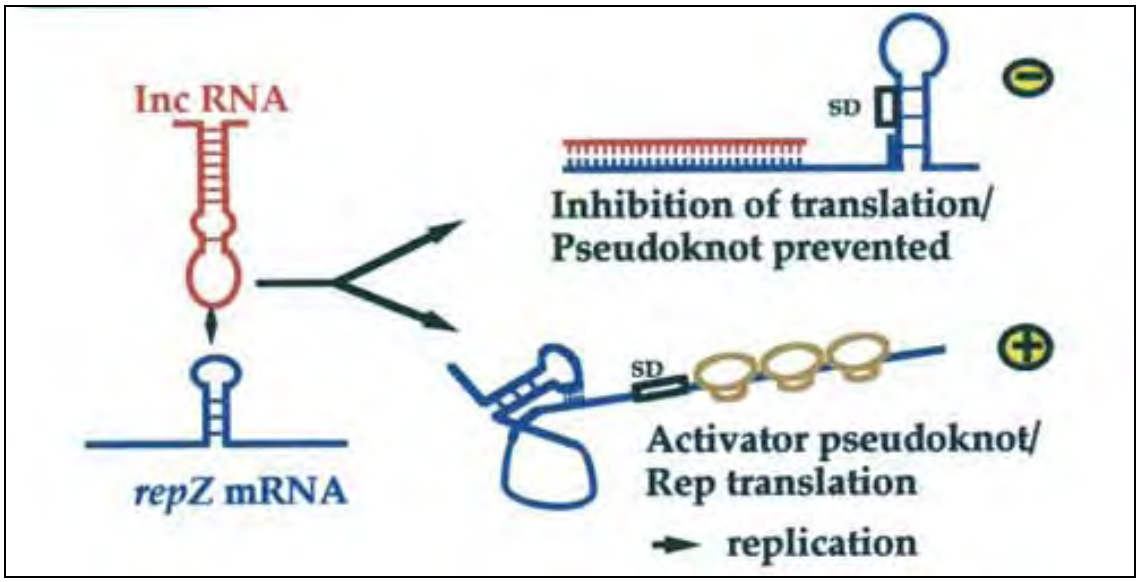


Figura 10. Esquema representativo de la regulación del plásmido ColIb-P6 por el RNA antisentido Inc RNA. Tomado de Brantl, 2000.

La formación del pseudonudo está condicionada a la ausencia del RNA I, que es un regulador negativo de la traducción de RepZ. La traducción de *repZ* es dependiente de la traducción de *repY*. El RNA I ocluye el SD de *repY*, por lo tanto evita la reducción del pequeño péptido, y de manera concomitante evita la traducción de *repZ* al estar esta acoplada con *repY*. La ausencia del RNA I permite la formación del pseudonudo y por lo tanto la traducción del péptido RepY y de la proteína iniciadora RepZ (Wagner y Brantl, 1998).

f) Partición en plásmidos bacterianos.

Además de los sistemas de regulación de la replicación, para que un linaje de bacterias pueda heredar a sus células hijas su material genético, se requiere una maquinaria especial que le permita, además de dotar a su descendencia de una copia idéntica de su genoma, mantenerlo a lo largo de las generaciones. Hablando especialmente de los plásmidos, estos se pueden perder con cierta facilidad a lo largo de los ciclos replicativos de un genoma. La evolución a dotado a los plásmidos de sistemas de segregación muy eficientes para mantener su genoma a lo largo de las generaciones.

La partición (o segregación) es el proceso por el que un organismo se asegura que su material genético se herede a cada una de sus células hijas antes de que ocurra la división celular. Los plásmidos también requieren de un sistema que asegure su adecuada segregación a las células hijas.

Los sistemas conocidos a la fecha se clasifican dentro de los siguientes grupos:

I. Sistemas "killer" o de muerte post-segregacional. Este sistema se encarga de eliminar de la población a las células que no contienen plásmido. También se conoce como mecanismo toxina-antitoxina, dicho mecanismo se basa en la acción antagónica de una toxina químicamente estable y una antitoxina menos estable.

II. Sistemas de resolución de dímeros.

Este mecanismo evita la formación de dímeros o multímeros mediante recombinación sitio específica.

III. Sistemas de partición activa.

Este sistema se refiere al movimiento dirigido de los plásmidos recién replicados hacia las células hijas (Summers, 1996). Los plásmidos bacterianos unicopia o de bajo número de copias como P1, F y R1 codifican para un sistema de partición (*par*) para lograr una segregación exitosa. El sistema *par* está compuesto por tres elementos: una región de DNA (centrómero) que actúan en *cis* y dos proteínas típicamente llamadas ParA y ParB, que actúan en *trans* y codificadas en operón. El centrómero *parC* contiene iterones. Se ha demostrado que los primeros pasos de la partición de plásmidos involucra la unión de ParB al centrómero. Subsecuentemente, múltiples moléculas de ParB se propagan alrededor, formando una gran estructura nucleoproteica llamada complejo de partición o segregosoma. ParB recluta a ParA,una ATPasa que reconoce al complejo de partición y actúa como un *switch* molecular. La hidrólisis del ATP por el complejo ParA-ParB-*parC* provee la energía para separar los productos de la replicación del plásmido, en un proceso análogo a la mitosis en células eucariotas (Ebersbach y Gerdes, 2005).

g) Incompatibilidad

Desde muy temprano en la investigación de la biología molecular de plásmidos los microbiólogos se percataron de que algo pasaba cuando intentaban introducir un plásmido en una bacteria que ya tenía otro plásmido del mismo tipo. Fue en este momento cuando se definió como incompatibilidad a la incapacidad de dos plásmidos de coexistir dentro de una célula. En un plásmido la incompatibilidad está dada por la competencia en la disposición de elementos que participan en la replicación o en la partición del plásmido (Novick, 1987).

La partición, la replicación y la incompatibilidad, son tres fenómenos estrechamente relacionados: los elementos que forman parte del complejo de partición o participan en el inicio de la replicación suelen ser fuertes factores de incompatibilidad.

Como ya se ha mencionado a lo largo de esta memoria, usualmente los sistemas de partición y la replicación en los plásmidos bacterianos suelen presentarse a manera de módulos independientes, no siendo siempre así, en el caso de los replicadores de la familia *repABC*, el módulo de replicación y partición se encuentran organizados en una misma unidad transcripcional (operón).

2. ANTECEDENTES

a) Replicadores tipo *repABC* y su anatomía.

Los replicadores de la familia *repABC* son un excelente modelo para el estudio y la comprensión de la replicación y de la partición del material genético puesto que sus módulos de replicación y de partición se encuentran codificados en una misma unidad transcripcional. Esta arquitectura los convierte en un modelo atractivo para adentrarse en la indagación de los mecanismos finos que rigen la replicación y segregación de los plásmidos. Sin embargo, no con ello se quiere decir que es un modelo sencillo de comprender, su aparente simpleza nos ha enseñado que la complejidad puede estar subyacente.

Los replicones *repABC* se encuentran ampliamente distribuidos en las α-proteobacterias, como *Rhizobium*, *Mesorhizobium*, *Sinorhisobium*, *Agrobacterium*, *Rhodobacter*, *Ruegeria* y *Paracoccus* (Nishiguchi *et al*., 1987; Tabata *et al*., 1989; Bartosik *et al*., 1998; 2001; Kaneko *et al*., 2000; Cevallos *et al*., 2002; Zhong *et al*., 2003). El caso que hoy nos ocupa es el plásmido p42d (pSym) de *R. etli* CFN42, que llamó fuertemente la atención de los rhizobiólogos al darse cuenta que en él están codificados todos los genes necesarios para llevar a cabo la interacción con plantas de frijol. Sumado a éstas nada desdeñables virtudes del p42d, tiene la virtud de pertenecer a la exclusiva familia de replicadores *repABC*.

Estructuralmente, el operón *repABC* del p42d tiene una organización simple: los dos primeros genes codifican dos proteínas (RepA y RepB) que tienen similitud con las proteínas ParA y ParB, las proteínas del sistema de segregación mejor caracterizadas (Bignell y Thomas, 2001). RepA y RepB están involucradas en la regulación postranscripcional negativa del operón *repABC* al unirse a la región promotora del operón (Ramirez-Romero *et al*., 2001; Pappas and Winans, 2003). RepB es la proteína que reconoce el sitio *par* del replicón que en conjunto con RepA (ATPasa tipo Walker) forman el sistema de partición, el sitio *par* de éste plásmido se localiza corriente abajo de *repC*, dentro de la región de incompatibilidad (*incβ*). El sitio *par* y la proteína RepA son fuertes factores de incompatibilidad (Ramirez-Romero *et al*., 2000; Soberón *et al*., 2004). De RepC se sabe muy poco, el hecho de no tener homólogos, la convierte en un enigma sin función aparente, aunque experimentos previos en nuestro laboratorio la señalan como proteína iniciadora de la replicación del p42d, además de contener el origen de replicación dentro de su secuencia codificante (datos no publicados, Cervantes-Rivera, R.). La región intergénica entre *repB* y *repC* o *incα*, es una región que ejerce una fuerte incompatibilidad. Se demostró en nuestro laboratorio que el responsable de dicho fenómeno es un RNa antisentido (ctRNA) de 59 nt, codificado en esta región (Venkova-Canova *et al*., 2004). La presencia de un RNA antisentido en la región intergénica es un fenómeno predominante en los replicadores *repABC* que mantienen la organización *repA*, *repB* y *repC*. Este pequeño RNA antisentido juega un papel crucial en la regulación de la replicación, la posición crítica en la que se localiza lo señala como el responsable más…

probable de la regulación del número de moléculas de la proteína iniciadora disponibles para que se pueda iniciar la replicación del plásmido. Para una revisión exhaustiva sobre los replicadores *repABC* consultar: Cevallos *et al*., 2008, en el **APÉNDICE 2**.

3. PLANTEAMIENTO Y JUSTIFICACIÓN

La característica más destacada de los replicadores *repABC* y en contraste con el resto de los plásmidos que hasta ahora se han descrito, es que los módulos de replicación y de segregación están organizados en una misma unidad transcripcional, lo que los convierte en un modelo único de estudio para entender los mecanismos de replicación y de segregación. En el apéndice 2 de esta tesis se encuentra una revisión detallada de los plásmidos *repABC*, publicada en el 2008 y de la cuál fui coautor. Con el fin de no parafrasear lo que en esa revisión se encuentra, considero que debe tomarse como una introducción extensa de esta tesis.

Como ya se ha mencionado a lo largo de esta memoria, se sabe que RepA y RepB están implicadas en la segregación y en la regulación negativa del operón, en tanto que RepC hemos descifrado que es la proteína iniciadora de la replicación y que dentro de su secuencia codificante se encuentra el origen de la replicación. La región intergénica *repB-repC* de 150 pb, es un fuerte determinante de incompatibilidad que hemos llamado *inc*α. Esta región es particularmente interesante ya que en ella se encuentran dos elementos indispensables para el correcto funcionamiento de la replicación: primero, ahí se encuentra codificado, pero en la cadena complementaria del mRNA del operón, un RNA antisentido (ctRNA) de tan sólo 59 nt que tiene un papel importante en la regulación de los niveles de RepC. Segundo, en esa región, el mRNA del operón forma una…

estructura tallo-asa (elemento S) que funciona como un terminador intrínseco de la transcripción. Además, al formarse dicha estructura se ocluye la secuencia Shine-Dalgarno de *repC,* lo que también pudiera impedir la traducción de RepC. Esta configuración permite que este pequeño RNA aparee con el RNA mensajero del operón *repABC* precisamente en esa región. Los dos elementos son indispensables ya que la ausencia de cualquiera de ellos conllevan a que el plásmido sea incapaz de replicar (Venkova-Canova *et al*., 2004). Como hipótesis podemos decir que el ctRNA, al aparearse con su región complementaria en el mRNA del operón, induce la formación del elemento-S y la terminación prematura del la transcripción de *repC*, el último gen del operón. Así que cualquier mutación que evite la interacción inicial entre estos dos RNAs evitará que el ctRNA ejerza su función regulatoria

4. OBJETIVOS

Objetivo general

Describir el mecanismo por el cual el ctRNA regula el inicio de la replicación del p42d de *R. etli* CFN42.

Objetivos particulares

1. Determinar la estructura secundaria del ctRNA, así como la del mRNA-*incα*.
2. Describir la dinámica de cómo ocurre la interacción entre el ctRNA y el mRNA-*incα*.
3. Elucidar la estructura del heterodúplex ctRNA/mRNA-*incα*.
4. Identificar los nucleótidos del ctRNA y del mRNA-*incα* implicados en la interacción inicial.
5. Demostrar, a través de ensayos de incompatibilidad, que las mutaciones o remociones de los nucleótidos que participan en la interacción inicial evitan que el ctRNA ejerza su función regulatoria *in vivo*.
6. Con estos datos proponer un modelo que explique la acción regulatoria del ctRNA.
7. Obtener datos experimentales que apoyen este modelo.

5. RESULTADOS

Los resultados de nuestros experimentos se encuentran relatados con detalle en el artículo publicado en la revista *Journal of Bacteriology* que publiqué como primer autor, a mediados de este año, y que se encuentra como **APÉNDICE 1** de esta tesis.

6. DISCUSIÓN

En el presente trabajo caracterizamos el mecanismo de acción del ctRNA codificado en la región intergénica *inc*α del replicón básico *repABC* del plásmido simbiótico de *R. etli* CFN42. Algunos trabajos previos del laboratorio sugirieron que la proteína RepC era la encargada de iniciar la replicación. Como parte del presente trabajo, también me enfoqué en probar que RepC es la proteína iniciadora de la replicación y que, además, el origen de replicación se localiza en la secuencia codificante de *repC*. Para demostrar la función de RepC como proteína iniciadora de la replicación, cloné la región codificadora de *repC* bajo el promotor *lac,* que es constitutivo en *R. etli*, en un vector suicida (pDOP) y la introduje a *R. etli*. Esta construcción es capaz de replicar en esta bacteria, observación que nos permitió comprobar que *repC* es suficiente y necesaria para que se lleve a cabo la replicación. De este experimento se concluye que RepC es la proteína iniciadora de la replicación y que el origen de replicación (*oriV*) está dentro de la secuencia codificante de *repC* (datos no publicados).

Como ya se ha mencionado en los apartados anteriores, en el trabajo de Venkova-Canova *et al*. (2004), se probó la presencia de un gene que codifica para un pequeño ctRNA corriente arriba del codón de inicio de *repC* y en la cadena complementaria al mRNA del operón. Ahí se comprobó que el ctRNA es el responsable del fenómeno de incompatibilidad plasmídica que posee la región intergénica *repB-repC* o región *inc*α. Además, a través de fusiones…

transcripcionales y traduccionales se describió que dicho ctRNA podría estar desempeñando una función importante como regulador negativo de la replicación por sus características intrínsecas.

Algunos autores han demostrado que los RNAs antisentido (formalmente ctRNAs) de la familia *repABC* homólogos tienen una función reguladora. Sin embargo, el mecanismo de acción no había sido descrito. Chai y Winans, 2005, propusieron un mecanismo de regulación de la replicación del pTi por un RNA antisentido a nivel de la traducción. Proponen dos conformaciones alternativas del RNA mensajero de *repC*: en ausencia del ctRNA, el mRNA de *repC* toma una conformación que permite la libre traducción de la proteína, puesto que su Shine Dalgarno (SD) se encuentra accesible al ribosoma. Por otro lado, en presencia del ctRNA, el RNA mensajero de *repC* se pliega de tal manera que el SD de *repC* queda ocluido dentro de una estructura tallo-asa (elemento S), por lo tanto la traducción de *repC* queda imposibilitada (Figura 11). Los resultados obtenidos en el caso del ctRNA del p42d muestran una regulación de la síntesis de RepC a nivel transcripcional, ya que en presencia del ctRNA la transcripción de *repC* se abate a niveles muy bajos. Contrario a lo predicho en el modelo propuesto por Chai y Winans, 2005, en el caso del p42d la replicación en ausencia del ctRNA no es posible, puesto que el ctRNA es esencial para que se puede llevar a cabo la replicación del plásmido (Venkova-Canova *et al*., 2004).

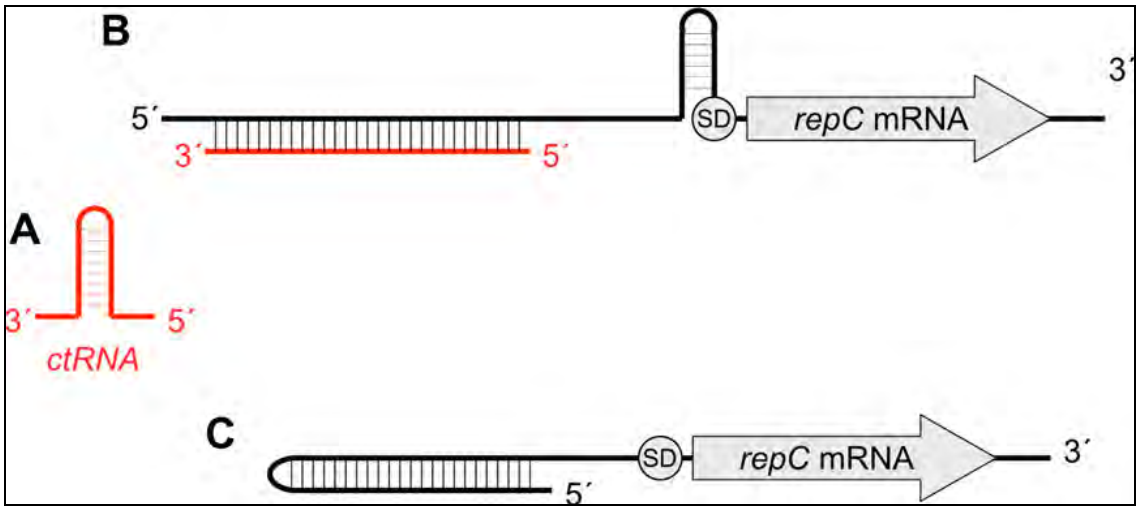


Figura 11. Modelo de regulación de la replicación propuesto por Chai y Winans, 2006. Tomado de Cevallos *et al.*, 2008.

Una de las características de los RNAs con acción en *cis* es su capacidad para formar un dúplex perfecto con su RNA blanco. Muchos de los RNAs con estructura tallo-asa que interaccionan formando un complejo suelen iniciar la formación del mismo por una interacción asa-asa (Brantl, 2006). En el caso particular del ctRNA del p42d, el mecanismo por el cual inicia la formación del dúplex es por una interacción de RNA de cadena sencilla con un asa, que podría continuarse a manera de cremallera para completar el dúplex. Como se puede apreciar en la figura 3 del artículo en el apéndice 1, la estructura secundaria del ctRNA es una típica estructura tallo-asa, con un brazo a cada lado no apareado. El mRNA-*incα* se pliega sobre si mismo dando como resultado una estructura con dos asas bien definidas y varias regiones de cadena doble. Experimentos de…

unión *in vitro* de moléculas del ctRNA con mutaciones puntuales o deleciones en su extremo 5' demostraron que la interacción primaria del ctRNA con el mRNA-*inc*α se da por el extremo 5' del ctRNA con el asa más prominente (L1α) del mRNA-*inc*α, al extenderse el dúplex provoca que haya un cambio de conformación en el resto de la molécula, lo que facilita la formación del terminador transcripcional o elemento-S.

Los RNAs con acción en *cis*, no necesitan de la participación de la chaperona Hfq para una formación eficiente del dúplex. El ctRNA del p42d no es la excepción. Experimentos de unión del ctRNA con la chaperona Hfq de *R. etli*, sobreexpresada y purificada mostraron la incapacidad del ctRNA para formar un complejo con la chaperona (resultados no publicados). Sin embargo, la formación del dúplex tiene una afinidad del orden Kd = 3.16 X 10^5 M^{-1} s^{-1}, lo que lo coloca en el mismo rango que todos los demás RNA antisentido que participan en la regulación de la replicación de otros plásmidos (Wagner *et al*., 2002).

Por otro lado, algunas de las mismas mutantes del ctRNA clonadas en un vector replicable en *R. etli* y conjugadas a la cepa que posee el plásmido parental (p42d), se observó que aquellas mutantes que son incapaces de formar el dúplex no causan incompatibilidad, en tanto que aquellas que si forman el complejo causan un fuerte fenómeno de incompatibilidad. Estos resultados que prueban la función biológica de la formación del dúplex en la regulación de la replicación del p42d.

Como ya se discutió en el artículo, la vida media de los RNAs antisentido que participan en la regulación de la replicación en otros plásmidos suele ser muy corta. En el caso del ctRNA del p42d, tiene una vida media de aproximadamente 5 minutos, más o menos en el mismo rango que el resto de RNAs antisentido con función semejante. La vida media de este tipo de reguladores suele ser corta, puesto que su síntesis es constitutiva, lo que le permite controlar fácilmente las oscilaciones en la síntesis de la proteína iniciadora.

El modelo que proponemos de regulación de la replicación por el ctRNA, concilia todos nuestros resultados experimentales. La síntesis constitutiva del ctRNA a lo largo del ciclo celular es vital para poder mantener la regulación de la replicación, aunque se ha visto que hay pequeñas variaciones en cuanto a la abundancia del ctRNA a lo largo del ciclo celular (datos no publicados). Sin embargo, la abundancia relativa del ctRNA con respecto al mensajero siempre es mayor, lo que facilita la unión a saturación del ctRNA con el mRNA-*incα*. La formación del dúplex ctRNA/mRNA-*incα* atenúa la transcripción de *repC* y por lo tanto el número de moléculas de la proteína iniciadora RepC. La regulación de la replicación por el ctRNA en términos de economía celular es un mecanismo con una alta eficiencia, dado que tiene un bajo costo energético y una alta especificidad.

La formación del complejo ctRNA/mRNA-*inc*α, además de atenuar la transcripción de *repC*, es quizá la señal de degradación por RNasas específicas. En *E. coli* se sabe que la formación del dúplex desencadena la acción de lo que se conoce como degradosoma. A la fecha se sabe que el degradosoma primario está constituido por la RNasa E, una ribonucleasa; el llamado degradosoma secundario está conformado por las ribonucleasas RNasa II y la RNasa R, también forman parte de este complejo la PNPasa, una polinucleótido fosforilasa y la RhlB, una RNA helicasa (Kaberdin *et al*., 2006), dicho complejo es el encargado de reducir a nucleótidos reutilizables los RNAs degradados. Describir las proteínas que participan en la degradación del complejo ctRNA/mRNA-*inc*α de nuestro modelo experimental es algo que aún queda por hacer.

7. PERSPECTIVAS

Como ya se mostró en el apartado de resultados (artículo que se muestra en el anexo 1), el ctRNA forma un dúplex perfecto con el mRNA-*incα in Vitro.* Nosotros proponemos que la formación del dúplex induce la terminación temprana de la transcripción de *repC* y muy posiblemente la formación de dicho dúplex es la señal para que sea blanco de la degradación por RNasas específicas.

La perspectiva más interesante es demostrar que la formación del complejo ctRNA/mRNA-*incα* es una señal para que se ensamble el degradosoma. Para probar que el complejo sigue la vía de la degradación por RNasas, proponemos los siguientes experimentos: en una columna de biotina fijar el complejo ctRNA/mRNA-*incα* y hacerle pasar un lisado total de *R. etli*, lavar y eluir las proteínas que se hayan quedado unidas al complejo. Posteriormente, identificar las proteínas por espectrometría de masas. De esta manera identificar las proteínas que interactúan con el complejo. Un paso más fino sería, ya una vez que se tengan identificadas las proteínas que forman parte del degradosoma, producir anticuerpos contra una o dos de ellas y por coinmunoprecipitación capturar todas las proteínas que participan de la vía de degradación.

8. REFERENCIAS

Asano, K., Hama, C., Inoue, S., Moriwaki, H. and Mizobuchi, K. 1999. The plasmid ColIb-P9 antisense Inc RNA controls expression of the RepZ replication protein and its positive regulator *repY* with different mechanisms. J. Biol. Chem. 274:17924-33.

Asano, K., Kato, A., Moriwaki, H., Hama, C., Shiba, K., Mizobuchi, K.A.1991. Positive and negative regulations of plasmid ColIb-P9-repZ gene expresión at the translational level. J. Biol. Chem. 266:3774-3781.

Bartosik, D., Baj, J., Wlodarczyk, M. 1998. Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. Microbiology. 144:3149-3157.

Bartosik, D., Szymanik, M., Wysocka, E. 2001. Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. J. Bacteriol. 21:6234-6243.

Berg, T., Firth, N., Apisiridej, S., Hettiaratchi, A., Leelaporn, A. and Skurray, R.A. 1998. Complete nucleotide sequence of pSK41: evolution of staphylococcal conjugative multiresistance plasmids. J. Bacteriol.180:4350-9.

Bignell, C. and Thomas, C. M. 2001. The bacterial ParA-ParB partitioning proteins. J. Biotechnol. 91:1-34.

Brantl, S. 2007. Regulatory mechanisms employed by *cis*-encoded antisense RNAs. Curr. Opin. Microbiol. 10:102-9.

Brantl, S. and Wagner, E.G. 2000. Antisense RNA-mediated transcriptional attenuation: an *in vitro* study of plasmid pT181. Mol. Microbiol. 35:1469-82.

Brennan, R.G. 1993. The winged-helix DNA-binding motif: another helix-turn-helix takeoff. Cell. 74:773-6.

Cevallos, M.A., Porta, H., Izquierdo, J., Tun-Garrido, C.,Garcia-de-los-Santos, A., Davila, G., Brom, S. 2002. *Rhizobium etli* CFN42 contains at least three plasmids of the *repABC* family: a structural and evolutionary analysis. Plasmid. 48:104-116.

Chai, y., and Winans, S.C. 2005. A small antisense RNA dowregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid, Mol. Microbiol. 56:1574-1585.

Chang, T.L., Kramer, M.G., Ansari, R.A., Khan, S.A. 2000. Role of individual monomers of a dimeric initiator protein in the initiation and termination of plasmid rolling circle replication. J. Biol. Chem. 275:13529-34.

Chattoraj, D.K. 2000. Control of plasmid DNA replication by iterons: no longer paradoxical. Mol. Microbiol. 7:467-76.

Chattoraj, D.K, Snyder, K.M. and Abeles, A.L. 1985. P1 plasmid replication: multiple functions of RepA protein at the origin. Proc. Natl. Acad. Sci. 82:2588-92.

Das, N. and Chattoraj, D.K. 2004. Origin pairing ('handcuffing') and unpairing in the control of P1 plasmid replication. Mol. Microbiol. 54:836-49.

Davey, M.J., Jeruzalmi, D., Kuriyan, J., and O'Donnell, M. 2002. Motors and switches: AAA+ machines within the replisome. Nat. Rev. Mol. Cell. Biol. 3:826-835.

Díaz, R. and Staudenbauer, W.L. 1982. Origin and direction of mini-R1 plasmid DNA replication in cell extracts of *Escherichia coli*. J. Bacteriol. 150:1077-1084.

de la Campa, A.G., del Solar, G., Espinosa, M. 1990. Initiation of replication of plasmid pLS1. The initiator protein RepB acts on two distant DNA regions. J. Mol. Biol. 213:247-62.

del Solar, G., Giraldo, R., Ruiz-Echevarría, M.J., Espinosa, M, Díaz-Orejas, R. 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. 62:434-64.

del Solar, G., Moscoso, M. and Espinosa, M. 1993. Mol. microbiol. 8:789-796.

Ebersbach, G. and Gerdes, G. 2005. Plasmid segregation mechanism. Annu. Rev. Genet. 39:453-479.

Eguchi, Y. and Tomizawa, J. 1991. Complexes formed by complementary RNA stem-loops. Their formations, structures and interaction with ColE1 Rom protein. J. Mol. Biol. 220:831-42.

Erzberger, J.P., Pirruccello, M.M. and Berger, J.M. 2002. The structure of bacterial DnaA: implications for general mechanisms underlying DNA replication initiation. EMBO. J. 21:4763-4773.

Franch, T. and Gerdes, K. 2000. U-turns and regulatory RNAs. Curr. Opin. Microbiol. 3:159-64.

Francia, M.V., Fujimoto, S., Tille, P., Weaver, K.E. and Clewell, D.B. 2004. Replication of *Enterococcus faecalis* pheromone-responding plasmad pAD1: location of the minimal replicon and *oriV* site and RepA involvement in initiation of replication. J. Bacteriol. 186:5003-16.

Frick, D.N. and Richardson, C.C. 2001. DNA primases. Annu. Rev. Biochem. 70:39-80.

Gering, M., Götz, F. and Brückner, R. 1996. Sequence and analysis of the replication region of the *Staphylococcus xylosus* plasmid pSX267. Gene. 82:117-22.

Gielow A, Diederich L, Messer W. 1991. Expression of the replication protein Arp of phasyl shows dual regulation by an antisense promoter. EMBO. J. 10:3061-6.

Haeusser, D.P. and Levin, P.A. 2008. Curr. opin. microbiol. 11:94-99.

Hase,T., Nakai, M. and Masamune, Y. 1989. Transcription of a region downstream from lambda *ori* is required for replication of plasmids derived from coliphage lambda. Mol. Gen. Genet. 216:120-5.

Hassan, A.K., Moriya, S., Ogura, M., Tanaka, T., Kawamura, F. and Ogasawara, N. 1997. Suppression of initiation defects of chromosome replication in *Bacillus subtilis dnaA* and *oriC*-deleted mutants by integration of a plasmid replicon into the chromosomes. J. Bacteriol. 179:2494-502.

Hill, T.M. 1992. Arrest of bacterial DNA replication. Annu. Rev. Microbiol. 46:603-633.

Honda, Y., Sakai, H., Hiasa, H., Tanaka, K., Kamono, T. and Bagdasarian, M. 1991. Functional division and reconstruction of a plasmid replication origin: Molecular dissection of the *oriV* of the broad-host-range plasmad RFS1010. Proc. Natl. Acad. Sci. 88:179-183.

Johnson, A. and O'Donnell. 2005. Cellular DNA replicases: Components and dynamics at the replication fork. Annu. Rev. Biochem. 74:283-315.

Kaberdin, A.R. and Blasi, U. 2006. Translation and the fate of bacterial mRNAs. FEMS Microbiol. Rev. 30:967-979.

Kaguni, J.M. 2006. DnaA: controlling the initiation of bacterial DNA replication and more. Annu. Rev. Microbiol. 60:351-75.

Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., Tabata, S. 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA. Res. 7:331-338.

Kawasaki, Y., Wada, C. and Yura, T. 1990. Roles of *Escherichia coli* heat shock proteins DnaK, DnaJ and GrpE in mini-F plasmid replication. Mol. Gen. Genet.. 220:277-82.

Khan, S.A. 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. 61:442-455.

Khan, S.A. 2005. Plasmid Rolling-circles replication: highlights of two decades of research. Plasmid. 53:126-136.

Kleanthous, H., Clayton, C.L. and Tabaqchali, S. 1991. Characterization of a plasmid from *Helicobacter pylori* encoding a replication protein common to plasmids in gram-positive bacteria. Mol. Microbiol 5:2377-2389.

Kline, B.C. (1985). A review of mini-F plasmid maintenance. Plasmid. 14:1-16.

Koepsel, R.R., Murray, R.W., Rosenblum, W.D., Khan, S.A. 1985. Proc. Nail. Acad. Sci. 82:6845-6849.

Kolb, F.A., Engdahl, H.M., Slagter-Jäger, J.G., Ehresmann, B., Ehresmann, C., Westhof, E., Wagner, E.G. and Romby, P. 2000. Progression of a loop-loop complex to a four-way junction is crucial for the activity of a regulatory antisense RNA. EMBO. J. 19:5905-15.

Kornberg, A., and T. Baker. 1992. DNA replication. W. H. Freeman & Co. New York, N.Y.

Kwong, S.M., Skurray, R.A. and Firth, N. 2004. *Staphylococcus aureus* multiresistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense regulation. Mol. Microbiology. 51:497-509.

Malinin, A.Y., Vostrov, A.A., Rybchin, V.N. and Svarchevsky, A.A. 1992. Structure of linear plasmid N15 end. Mol. Gen. Mikro-biol. Virus. 5:19-22.

Masai, H., Kaziro, Y. and Arai, K. 1983. Definition of oriR, the minimum DNA segment essencial for initiation of R1 plasmid replication in vitro. Proc. Natl. Acad, Sci. 80:6814-6818.

McHenry, C.S. 2003. Chromosomal replicases as asymmetric dimers: studies of subunit arrangement and functional consequences. Mol. Microbiol. 49:1157-1165.

Moscoso, M., del Solar, G. and Espinosa, M. 1995. *In vitro* recognition of the replication origin of pLS1 and of plasmids of the pLS1 family by the RepB initiator protein. J. Bacteriol. 177:7041-9.

Mott, M.L. and Berger, J.M. 2007. DNA replication initiation: mechanisms and regulation in bacteria. Nat. Rev. Microbiol. 5:343-54.

Mukhopadhyay, S. and Chattoraj, D. K. 2000. Replication-induced transcription of an autorepressed gene: the replication initiator gene of plasmid P1. Proc. Natl. Acad. Sci. 97:7142-7.

Nakamura, A., Wada, C. and Miki, K. 2007. Structural basis for regulation of bifunctional roles in replication initiator protein. Proc. Natl. Acad. Sci. 104:18484-9.

Nishiguchi, R., Takanami, M. and Oka, A. 1987. Characterization and sequence determination of the hairy root inducing plasmid pRiA4b. Mol. Gen. Genet. 206:1-8.

Noirot, P., Bargonetti, J., Novick, R.P. 1990. Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. Proc. Natl. Acad. Sci. 87:8560-4.

Nordström, K. 2006. Plasmid R1-replication and its control. Plasmid. 55:1-26.

Novick, R.P. 1987. Plasmid incompatibility. Microbiol. Rev. 51:381-95.

Novick, R.P., Iordanescu, S., Projan, S.J., Kornblum, J. and Edelman, I. 1989. pT181. plasmid replication is regulated by countertranscript-driven transcriptional attenuator. Cell. 59:395-404.

Pappas, K.M., Winans, S.C. 2003. The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. Mol. Microbiol. 49:441-455.

Paulsson, J. and Chattoraj, D.K. 2006. Origin inactivation in bacterial DNA replication control. Mol. Microbiol. 61:9-15.

Persson, C., Wagner, E.G. and Nordström, K. 1988. Control of replication of plasmid R1: kinetics of *in vitro* interaction between the antisense RNA, CopA, and its target, CopT. EMBO. J. 7:3279-88.

Ramirez-Romero, M.A., Soberon, N., Perez-Oseguera, A., Tellez-Sosa, J. and Cevallos, M.A. 2000. Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. J. Bacteriol. 182:3117-3124.

Ramirez-Romero, M.A., Tellez-Sosa, J., Barrios, H., Perez-Oseguera, A., Rosas, V. and Cevallos, M.A. 2001. RepA negatively autoregulates the transcription of the *repABC* operon of the *Rhizobium etli* symbiotic plasmid basic replicon. Mol. Microbiol. 42:195-204.

Ravin, V., Ravin, N., Casjens, S., Ford, M.E., Hatfull, G.F, and Hendrix, R.W. 2000. Genomic sequence and analysis of the atypical temperate bacteriophage N15. J. Mol. Biol. 299:53-73.

Seitz, H., Weigel, C. and Messer, W. 2000. The interaction domains of the DnaA and DnaB replication proteins of *Escherichia coli*. Mol. Microbiol. 37:1270-1279.

Soberón, N., Venkova-Canova, T., Ramirez-Romero, M.A., Tellez-Sosa, J. and Cevallos, M.A. 2004. Incompatibility and the partitioning site of the *repABC* basic replicon of the symbiotic plasmid from *Rhizobium etli*. Plasmid 51:203-216.

Summers, D. 1996. The biology of plasmid. Blackwell science.

Tabata, S., Hooykaas, P.J., Oka, A. 1989. Sequence determination and characterization of the replicator region in the tumor-inducing plasmad pTiB6S3. J. Bacteriol. 171:1665-1672.

Tanaka, T., Ishida, H. and Maehara, T. 2005. Characterization of the replication region of plasmid pLS32 from the Natto strain of *Bacillus subtilis*. J. Bacteriol. 187:4315-26.

Taylor, K. and Wegrzyn, G. 1995. Replication of coliphage lambda DNA. FEMS. Microbiol. Rev. 17:109-19.

Thomas, C.D., Balson, D.F. and Shaw, W.V. 1990. *In vitro* studies of the initiation of staphylococcal plasmid replication. Specificity of RepD for its origin (*oriD*) and characterization of the Rep-*ori* tyrosyl ester intermediate. J. Biol. Chem. 265:5519-30.

Venkova-Canova, T., Soberón, N.E., Ramírez-Romero, M.A. and Cevallos, M.A. 2004. Two discrete elements are required for the replication of a *repABC* plasmid: an antisense RNA and a stem-loop structure. Mol. Microbiol. 54:1431-1444.

Wagner, E.G., Altuvia, S., Romby, P. 2002. Antisense RNAs in bacteria and their genetic elements. Adv. Genet. 46:361-98.

Wagner, E.G. and Brantl, S. 1998. Kissing and RNA stability in antisense control of plasmid replication. Trends Biochem. Sci. 23:451-4.

Yasukawa, H., Hase, T., Sakai, A. and Masamune, Y. 1991. Proc. Natl. Acad. Sci. 88:10282-10286.

Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R. and Storz, G. 1998. The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein. EMBO. J. 17:6061-8.

Zhong, Z., Caspi, R., Helinski, D., Knauf, V., Sykes, S., O'Byrne, C., Shea, T.P., Wilkinson, J.E., DeLoughery, C. and Toukdarian, A. 2003. Nucleotide sequence based characterizations of two cryptic plasmids from the marine bacterium ruegeria isolate PR1b. Plasmid 49:233-252.

Zzaman, S., Reddy, J.M. and Bastia, D. 2004. The DnaK-DnaJ-GrpE chaperone system activates inert wild type pi initiator protein of R6K into a form active in replication initiation. J. Biol. Chem. 279:50886-94.

JOURNAL OF BACTERIOLOGY, July 2010, p. 3268–3278 Vol. 192, No. 13
0021-9193/10/$12.00 doi:10.1128/JB.00118-10

Analysis of the Mechanism of Action of the Antisense RNA That Controls the Replication of the *repABC* Plasmid p42d▽†

Ramón Cervantes-Rivera,[1] Cristina Romero-López,[2] Alfredo Berzal-Herranz,[2] and Miguel A. Cevallos[1]*

Programa de Genómica Evolutiva, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apartado Postal 565-A, Cuernavaca, Morelos, México,[1] and Instituto de Parasitología y Biomedicina López-Neyra, CSIC, Parque Tecnológico de Ciencias de la Salud, Avda. del Conocimiento s/n, Armilla, 18100 Granada, Spain[2]

Received 2 February 2010/Accepted 22 April 2010

Replication and segregation of the *Rhizobium etli* symbiotic plasmid (pRetCFN42d) depend on the presence of a *repABC* operon, which carries all the plasmid-encoded elements required for these functions. All *repABC* operons share three protein-encoding genes (*repA*, *repB*, and *repC*), an antisense RNA (ctRNA) coding gene, and at least one centromere-like region (*parS*). The products of *repA* and *repB*, in conjunction with the *parS* region, make up the segregation system, and they negatively regulate operon transcription. The last gene of the operon, *repC*, encodes the initiator protein. The ctRNA is a negative posttranscriptional regulator of *repC*. In this work, we analyzed the secondary structures of the ctRNA and its target and mapped the motifs involved in the complex formed between them. Essential residues for the effective interaction localize at the unpaired 5′ end of the antisense molecule and the loop of the target mRNA. In light of our results, we propose a model explaining the mechanism of action of this ctRNA in the regulation of plasmid replication in *R. etli*.

Small noncoding RNAs are widespread in nature and play essential roles in regulating gene expression, mostly at a posttranscriptional level. Prokaryotic regulatory small RNAs fall into two classes: *cis*-encoded RNAs and *trans*-encoded RNAs. The first, also known as countertranscript RNAs (ctRNAs) (17), are frequently found in accessory genetic elements, including plasmids, phages, and transposons (5). They are encoded by the complementary strands of their target genes, and for this reason, the ctRNAs and the target mRNAs can form extended-pairing hybrids (3). In contrast, *trans*-encoded RNAs and their targets are encoded at separate chromosomal loci, with their complementary regions being limited to short stretches. Moreover, *trans*-encoded RNAs frequently require the action of the RNA chaperone protein Hfq (6) to improve their stability or to facilitate RNA-RNA complex formation.

The participation and mechanisms of action of ctRNAs in the control of plasmid replication are well established: ctRNAs prevent plasmid replication by inhibiting the production of plasmid-encoded molecules (RNAs or proteins) involved in the initiation of plasmid replication. The plasmid replication rate is inversely proportional to the intracellular concentrations of the corresponding ctRNAs. ctRNAs are constitutively expressed and diffusible, usually have short half-lives, and interact rapidly with their targets, thereby effectively correcting fluctuations in plasmid copy number (4). However, the sequences of ctRNAs, the structural basis of the interaction between ctRNAs and their targets, and the relationship between ctRNAs and plasmid replication control have been analyzed for only a small number of plasmid systems (5). Here we describe the molecular details of the mechanism of action of a ctRNA in the replication control of a plasmid belonging to the *repABC* family.

repABC plasmids, commonly found in alphaproteobacteria, are characterized by the presence of a *repABC* operon, which carries all the elements required for plasmid replication and segregation (8, 23). In general, *repABC* operons consist of three protein-encoding genes, an antisense RNA (ctRNA) gene, and at least one centromere-like region (*parS*). The products of the first two genes, *repA* and *repB*, in conjunction with the *parS* region, compose the segregation system, and *repC*, the last gene of the operon, encodes the initiator protein (2, 25). The ctRNA gene is located in the large *repB-repC* intergenic region and is encoded on the strand complementary to that encoding the operon mRNA (Fig. 1a) (9, 18, 30). The symbiotic plasmid, formally pRetCFN42d (p42d), of *Rhizobium etli* CFN42 carries a *repABC* operon whose expression depends on a single promoter (11, 26). The transcription of the *repABC* operon is negatively autoregulated by RepA and RepB (26, 22), and in all of them the ctRNA (55 to 59 nucleotides [nt]) is a negative posttranscriptional regulator of *repC*. These molecules also act as strong incompatibility factors, since it was shown that genes encoding ctRNAs introduced in *trans* displace their cognate plasmids (30, 18, 9).

The tumor-inducing plasmid (pTiR10) from *Agrobacterium tumefaciens* belongs to the *repABC* family. Chai and Winans (9) described a transcriptional/translational attenuation theoretical model explaining how the ctRNA (*repE*) of pTiR10 acts; this model is easily applicable to other *repABC* operons. In the model, the *repABC* mRNA exhibits two alternative secondary structures, depending on whether or not the ctRNA is paired with the *repABC* mRNA. In the absence of the ctRNA, the section corresponding to the *repB-repC* intergenic region of the

* Corresponding author. Mailing address: Programa de Genómica Evolutiva, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apartado Postal 565-A, Cuernavaca, Morelos, México. Phone: 7773 11 46 63. Fax: 7773 17 55 81. E-mail: mac@ccg.unam.mx.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 30 April 2010.

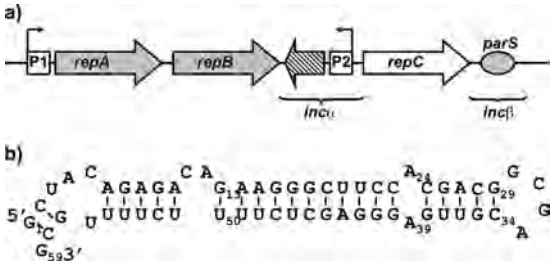


FIG. 1. (a) Schematic representation of the *Rhizobium etli* p42d *repABC* operon. Gray arrows represent genes encoding proteins involved in plasmid segregation and in the negative regulation of the operon. The white arrow shows the position of the gene encoding RepC, the initiator protein. The dashed arrow indicates the position of the gene encoding the small antisense RNA (ctRNA). Squares mark the positions of the promoters: P1, the operon promoter, and P2, the ctRNA gene promoter. The oval shows the position of *parS*. Brackets enclose two regions involved in plasmid incompatibility. (b) ctRNA folding prediction obtained with the Mfold software (33).

repABC mRNA folds into a large stem-loop structure, but the predicted *repC* Shine-Dalgarno (SD) sequence and the *repC* initiation codon remain single stranded, allowing *repC* translation. When the ctRNA interacts with *repABC* mRNA, the *repC* leader sequence folds into a stem-loop structure that resembles an intrinsic terminator; however, this structure also occludes the SD sequence and the *repC* initiation codon, thereby blocking *repC* translation. These observations suggest that *repC* levels could be modulated by a transcriptional/translational attenuation mechanism.

In this work, we analyzed the *in vitro* secondary structures of the following: (i) the ctRNA encoded in the *repB-repC* intergenic region of the *R. etli* CFN42 symbiotic plasmid p42d, (ii) the *repABC* mRNA (the target) in the *repB-repC* intergenic region (mIncα) of the same plasmid; and (iii) the complex formed between the ctRNA and mIncα, using a combination of RNA probing assays and an *in silico* RNA secondary structure prediction algorithm (Mfold) (33). Also, we calculated the kinetic parameters of hybrid formation between mIncα, the ctRNA, and various ctRNA mutants and the half-life of the ctRNA. Further, we conducted an *in vivo* incompatibility analysis of mutant ctRNAs. Our findings indicate that the interaction between the ctRNA and *repABC* mRNA is very fast and requires the nonpaired 5′ end of the ctRNA and a loop in the target mRNA. This system operates by a novel mechanism in the plasmid replication system, contrasting with the loop-loop kissing previously reported (31). Nevertheless, our observations are generally consistent with the model proposed by Chai and Winans (9).

MATERIALS AND METHODS

DNA templates and RNA synthesis. Oligodeoxyribonucleotides were synthesized in a 3400 DNA synthesizer (Applied Biosystems, Massachusetts); the DNA sequences of the oligonucleotides used in this work are listed in Table 1. The wild-type ctRNA (ctRNAwt) was generated by annealing and extension of the oligonucleotides ctRNA-T7-L and ctRNA-T7-U, essentially as described previously (1). Similarly, the ctRNA variant templates ctRNA-CGCU, ctRNA-mut2-5, ctRNA-mut2-5-CGCU, ctRNA-Δ(2-11), and ctRNA-Δ(1-8) were created by annealing and extension of the primers ctRNAasaCGCU, ctRNAmut2-5, ctRNAmut2yloop, ctRNAdel11, and ctRNAdel8, respectively, with ctRNA-T7-U. Other ctRNA templates—ctRNA-Δ(2-11)(52-59) and ctRNA-Δ(52-59)—were synthesized by annealing and extension of the primers ctRNAdel10-5′,8-3′ and ctRNAdel8-3′ with the primer Complement 3′. The template for transcribing the *repBC* intergenic region was obtained by PCR using the primers RNAT7-Bam-Alfa-U and MluAlfa-L. In competition binding assays, an oligonucleotide complementary to the 5′ end of the ctRNA (ODN5′) and another oligonucleotide complementary to the 3′ end of the same molecule (ODN3′) were used.

RNAs were synthesized by *in vitro* transcription and purified as previously described (1). The transcription reactions were carried out in volumes of 100 μl, with 3 μg of template DNA in a reaction buffer containing 40 mM Tris-HCl (pH 7.9), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 20 mM dithiothreitol (DTT), 0.01% Triton X-100, 1 mM nucleoside triphosphates (NTPs), 0.5 U of RNase inhibitor (Ambion Inc.), and 0.5 U of T7 RNA polymerase purified as previously described (20). Reaction mixtures were incubated at 37°C for 2 h, and the template DNA was then eliminated with 3 U RQ1 DNase (Promega). Samples were cleaned in Sephadex G-25 (GE) columns and then purified from 10% denaturing polyacrylamide gels.

RNA-RNA interaction assays. (i) Affinity assays. Binding reactions were carried out with 5 nM ^{32}P-labeled ctRNAwt or its variants with various amounts of unlabeled mRNA-*inc*α (mIncα) (0.625, 1.25, 2.5, 3.74, 5.0, 12.5, 25.0, 50.0, or 125.0 nM) in TMN binding buffer (1×) [20 mM M Tris acetate (OAc) (pH 7.5), 10 mM $Mg(OAc)_2$, 100 mM NaCl]. The RNAs were incubated separately at 65°C for 10 min, followed by incubation at 37°C for 10 min to allow the attainment of the native conformation. RNAs were mixed in the proportions described above and then incubated at 37°C for 10 min. The reactions were stopped by diluting the samples with an equal volume of denaturing application buffer: 94% formamide, 17 mM Na_2EDTA, 0.025% Xylene cyanol, and 0.025% bromophenol blue (24). Hybrids were analyzed in 10% polyacrylamide–7 M urea gels run at 12 W for 90 min at room temperature. Gels were quantified using the Image Quant 5.2 software program (GE Healthcare).

(ii) Time courses. Binding reactions were carried out in TMN buffer at 37°C with 5 nM ^{32}P-labeled ctRNA and 50 nM unlabeled mIncα. Samples were withdrawn at the times indicated, and the reactions were stopped by the addition of one volume of application buffer. Samples were analyzed as described above.

(iii) Competition assays. Labeled ctRNAwt, mIncα, and the primers ODN5′and ODN3′ were first denatured and renatured in TMN 1× buffer as described above. Subsequently, 5 nM labeled ctRNAwt was mixed with 760 pM of ODN5′ or ODN3′ or both in TMN (1×) buffer, and incubated at 37°C for 10 min. Finally, 50 nM mIncα was added, and the binding reaction mixes were incubated for 10 min at 37°C. Reactions were stopped by addition of one volume of application buffer. Samples were analyzed in 10% polyacrylamide–7 M urea gels, and images of dried gels were obtained and quantified as described above.

Secondary structure probing assays. Fifty femtomoles of labeled ctRNAwt, its variants, or their target (mIncα) were used in the RNA probing assays. Partial digestion reactions were performed with 0.1 U of RNase T1 (Industrial Research, New Zealand) in TMN (1×) buffer containing 2 μg tRNA (Ambion Inc.) and incubated at 37°C for 2 min. Chemical partial hydrolysis was performed at 37°C for 10 min, using 30 mM lead acetate (Merck) in TMN (1×) buffer supplemented with 1 μg of tRNA. Partial digestion reactions were carried out with 0.1 U of RNase V1 (Industrial Research) in TMN (1×) buffer containing 2 μg tRNA at 4°C for 30 min. The secondary structures of the ctRNAwt/mIncα complex were analyzed by mixing 50 fmol of labeled RNA with 5 pmol of its nonlabeled counterpart in TMN (1×) buffer at 37°C for 10 min. Enzymatic and chemical hydrolysis were performed as described above. Chemical hydrolysis was stopped by adding EDTA to a final concentration of 10 mM. Samples were extracted with phenol-chloroform and precipitated with ethanol. Products were separated on high-resolution denaturing polyacrylamide gels (12%). Images were obtained as described above.

Bacterial strains and growth conditions. *Escherichia coli* DH5α (12) and S17-1 (27) were grown at 37°C in Luria-Bertani medium, and *Sinorhizobium meliloti* 1021 and CFNX101, a *Rhizobium etli recA*::ΩSpc^r derivative containing p42d (19), were grown at 30°C in PY medium supplemented with 10 mM $CaCl_2$ (21). Nalidixic acid (20 μg/ml) and kanamycin (30 μg/ml) were added as appropriate.

Site-directed mutagenesis and plasmid incompatibility. To determine the incompatibility phenotype of the ctRNA variants studied *in vitro*, we introduced the same changes into a recombinant plasmid containing the whole *repB-repC* intergenic region (*inc*α) (Table 2) by site-directed mutagenesis, using the QuikChange II kit according to the manufacturer's instructions (Stratagene). Additionally, we made three constructs: one of them expressing a chimeric ctRNA gene consisting of the left arm from p42d ctRNA and the stem-loop and right arm from the ctRNA from *Sinorhizobium meliloti* 1021 pSymA (ctRNADqS). The second construct expresses a ctRNA possessing a left arm with a nonrelated sequence of 15 nucleotides (5′-GCCAGACAGAGAGGG), but the rest of the molecule is identical to p42d ctRNA (ctRNA5N). Finally, we

TABLE 1. Oligonucleotides used in this work

Oligonucleotide	Sequence[a]
ctRNA-T7-L	5′-TAA TAC GAC TCA CAT AGC TAC AGA GAC AGAAGG CTT CCA CGCAGG CGA CGT TGA GGG AGCTCTTTT CTT TTGCG
ctRNA-T7-U	5′-CGC AAA AGA AAA GAG CTC CCT
ctRNAasaCGCU	5′-**TAA TAC GAC TCA CTA TA**G CTA CAG AGA CAGAAG GGC TTC CAC GAC GCG CTC GTT GAG GGA GCTCTT TTC TTT TGC G
ctRNAmut2-5	5′-**TAA TAC GAC TCA CTA TA**G AAT GAG AGA CAG AAG GGC TTC CAC GAC GGC GAC GTT GAG GGA GCT CTT TTC TTT TGC G
ctRNAmut2yloop	5′-**TAA TAC GAC TCA CTA TA**G AAT GAG AGA CAG AAG GGC TTC CAC GAC GCGCTC GTT GAG GGA GCT CTT TTC TTT TGC G
ctRNAdel11	5′-**TAA TAC GAC TCA CTA TA**G AGA AGG GCT TCC ACG ACG GCG ACG TTG AGG GAG CTC TTT TCT TTT GCG
ctRNAdel8	5′-**TAA TAC GAC TCA CTA TA**G ACA GAA GGG CTT CCA CGA CGG CGA CGT TGA GGG AGC TCT TTT CTT TTG CG
ctRNAdel10-5′,8-3′	5′-**TAA TAC GAC TCA CTA TA**G AGA AGG GCT TCC ACG ACG GCG ACG TTG AGG GAG CTC TTT TC
ctRNAdel8-3′	5′-**TAA TAC GAC TCA CTA TA**G CTA CAG AGA CAG AAG GGC TTC CAC GAC GGC GAC GTT GAG GGA GCT CTT TTC
Complement 3′	5′-GAA AAG AGC TCC CTC AAC GT
RNAT7Bam-Alfa-U	5′-GGA TCC TAA TAC GAC TCA CTA TAG CCC GCA AAA GAA AAG AGC TCC CTC
MluAlfa-L	5′-ACG CGT TGT CTC TCA CCT TTC AGC
ODN5′	5′-GTC TCT GTA GC
ODN3′	5′-CGC AAA AG
tRNA-Met	5′-GCC TTC AGG TTA TGA GCC TGA CG
S-SymA-U	5′-CCG CGG CCC GCA AAA GAA AAA GGC
K-SymA-L	5′-GGT ACC TCA CGA CAC CCC CCG CCC
AlfaDHind-U	5′-AAG CTT CCC GCA AAA GAA AAG AGC TCC
AlfaD-L	5′-TGT CTC TCA CCT TTC AGC AGG CAA
ctRNADqS-U1	5′-GGA AGC CTT CCT CTC TGT CTC TGT AGC AAG AAC AGA ATC GCA TT
ctRNADqS-U2	5′-GGA TCC GCC CGC AAA AGA AAA AGG CCC CCA ACG AAC AAG CCG TGG AAG CCT TCC TCT CTG
ctRNA5n-U1	5′-GAA GCC CTT CTC TCT GTC TGG CAA GAA CAG AAT CGC ATT TCC
ctRNA5n-U2	5′-GGA TCC CCG GCA AAA GAA AAG AGC TCC CTC AAC GTC GCC GTC GTG GAA GCC CTT CTC TCT GTC TGG
ctRNAart-U1	5′-ACC AGG ATT CTG TCT CTG TAG CAA GAA CAG AAT CGC ATT TCC
ctRNAart-U2	5′-GGA TCC CCC GCA AAA GAA AAT CTT GGT TCA CCT AGT ATT AGG TCA ACC AAG ATT CTG TCT CTG
RepC-HL	5′-TTC GAG GGG AAG ACA ATC AAT TGT
RepB-HL	5′-CGC CTA CGG CCG CCA CTC CCT GTT TCC GTT G

[a] Underlined and bold sequences correspond to the T7 promoter.

also constructed a mutant ctRNA with its left and right arms identical in sequence to the arms present in the p42d ctRNA but with a stem-loop structure that is completely different in sequence although identical in size and structure (ctRNAart) (see Fig. S1 in the supplemental material).

All of these mutants were constructed using PCR technology. The mutant ctRNADqS was obtained in two successive PCRs: in the first amplification reaction, a plasmid carrying the ctRNA from pSymA was used as a template and the oligonucleotides ctRNADqS-U1 and K-SymA-L were used as primers. In the second PCR, the product of the first PCR was used as a template and ctRNADqS-U2 and K-SymA-L as primers. Mutant ctRNA5N was also obtained by two successive PCRs: primers ctRNA5n-U1 and AlfaD-L were used to raise the first amplification product, using a plasmid carrying the ctRNA gene from

TABLE 2. Plasmids used in this work

Plasmid	Relevant characteristics	Reference
pBBR1MCS2	Km^r cloning vector replicable in *Rhizobium*	Kovach et al., 1995 (16)
pCR-ctRNAwt 42d	pBBR1MCS2 carrying a ctRNA wild type from p42d *R. etli*	This work
pCR-ctRNAwt SymA	pBBR1MCS2 carrying a ctRNA wild type from pSymA *S. meliloti* 1021	This work
pCR-ctRNA-CGCU	pBBR1MCS2 carrying a ctRNA with modified loop (from 5′-GCGA to 5′-CGCU)	This work
pCR-ctRNAmut2-5	pBBR1MCS2 carrying a ctRNA with nucleotides 2-5 changed in the left arm from CUAC to AAUG	This work
pCR-ctRNAmut2-5 CGCU	pBBR1MCS2 carrying a ctRNA with both mutations: modified loop (from 5′-GCGA to 5′-CGCU) and nucleotides 2-5 changed in the left arm from CUAC to AAUG	This work
pCR-ctRNA-Δ(1-8)	pBBR1MCS2 carrying a ctRNA with lack eight of the first nucleotide of the left arm	This work
pCR-ctRNADqS	pBBR1MCS2 carrying a ctRNA with stem-loop and 3′ end from *inc*α region of *S. meliloti* pSymA and 5′ end (13 nt) from *inc*α region of *R. etli* p42d	This work
pCR-ctRNA5N	pBBR1MCS2 carrying a ctRNA with stem-loop and 3′ end from *inc*α region of *R. etli* p42d and 5′ end with nonrelated nucleotides	This work
pCR-ctRNAart	pBBR1MCS2 carrying a ctRNA with stem-loop with nonrelated nucleotides, 3′ end and 5′ end from ctRNA of p42d	This work

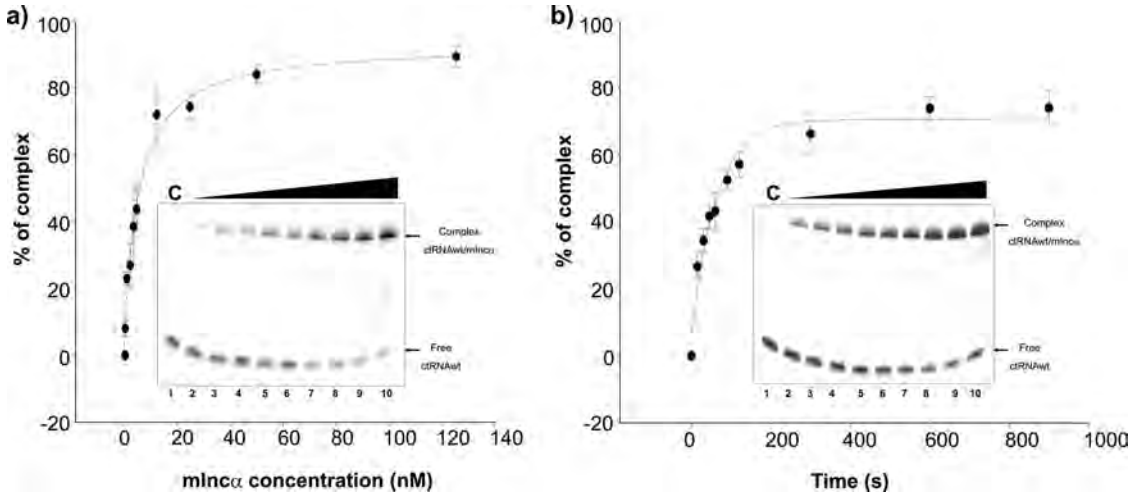


FIG. 2. ctRNA binding kinetics. Binding assays were performed as described in Material and Methods. (a) A graphic representation of the relative complex formation between the ctRNA and mIncα as s function of the mIncα concentration. (b) A graphic representation of the relative complex formation between the ctRNA and mIncα as a function of the time of incubation.

p42d as a template. In the second PCR, the oligonucleotides ctRNA5n-U2 and AlfaD-L were used as primers and the amplification product of the first PCR was used as a template. The mutant pCR-ctRNAart was obtained in a similar way: In the first PCR, the oligonucleotides ctRNAart-U1 and AlfaD-L were used as primers, and a plasmid with the ctRNA gene from p42d was used as a template. In the second PCR, the product of the first PCR was used as a template, and the oligonucleotides ctRNAart-U2 and AlfaD-L were used as primers. As controls, p42d ctRNA was obtained by PCR using the primers AlfaDHind-U and AlfaD-L, and a plasmid carrying the p42d ctRNA was used as a template. The ctRNA from pSymA was obtained by PCR using the primers S-SymA-U and K-SymA-L, and a plasmid carrying the pSymA ctRNA was used as a template. Amplification products were inserted into pBBR1MCS2 (Km^r), a vector that replicates in *Rhizobium* (16). All constructs were confirmed by DNA sequencing, performed at the DNA sequence facility of the Biotechnology Institute (IBT-UNAM).

The pBBR1MCS2 derivatives were introduced into *R. etli* CFNX101 using *E. coli* S17-1 as a donor strain. Strains were grown in the appropriate antibiotic-free liquid medium to stationary phase, mixed in a donor-recipient ratio of 1:1 on PY plates, and incubated overnight at 30°C. Colonies were resuspended in PY medium, and serial dilutions were plated on PY agar containing nalidixic acid and kanamycin. The plasmid profiles of at least four transconjugants per conjugation were determined by the in-gel lysis procedure described by Wheatcroft et al. (32) and analyzed applying the considerations described by Ramírez-Romero et al. (25).

Northern blot analysis. An overnight culture of *Rhizobium etli* CFNX101 was diluted in PY to an optical density at 620 nm (OD_{620}) of 0.2. This culture was further grown to an OD_{620} of 0.7 on a shaker, supplemented with nalidixic acid at a final concentration of 20 μg/ml, and sampled 2 ml. Sodium azide was added a final concentration of 100 mM and immediately transferred to new tubes containing a chilled solution of 0.2 M EDTA in glycerol (2:1). Total RNA was isolated from the samples with Trizol as recommended by the manufacturer (Invitrogen) and resolved on a 1% agarose gel with formaldehyde at a final concentration of 1% by electrophoresis. Gels were blotted onto Nylon membranes (Hybond-N^+; GE) using a semidry transfer cell (Bio-Rad). Membranes were probed with a ^{32}P-labeled oligonucleotide complementary to the *repB* coding sequence (RepB-HL) and an oligonucleotide complementary to the *repC* coding sequence (RepC-HL); the autoradiographic images were obtained with a phosphorimager (Storm 820; GE Healthcare) and quantified using Image Quant 5.2 software (GE Healthcare). Samples were normalized with $tRNA^{Met}$ (primer tRNA-Met).

Determination of ctRNA half-life. An overnight culture of *Rhizobium etli* CFNX101 was diluted in PY (21) to an OD_{620} of 0.2. This culture was further grown to an OD_{620} of 0.7 on a shaker, supplemented with rifampin (dissolved in dimethyl sulfoxide [DMSO]) at a final concentration of 50 μg/ml, and sampled (25 ml) at various intervals (0, 1, 5, 10, 15, 30, and 60 min). Sodium azide was added to each sample to a final concentration of 100 mM, and the samples were immediately transferred to new tubes containing a chilled solution of 0.2 M EDTA in glycerol (2:1). Total RNA was isolated from the samples with Trizol as recommended by the manufacturers (Invitrogen) and resolved on 10% polyacrylamide–7 M urea gels by electrophoresis. Gels were blotted onto Nylon membranes (Hybond-N^+; GE) using a semidry transfer cell (Bio-Rad). Membranes were probed with a ^{32}P-labeled oligonucleotide complementary to the ctRNA (Complement 3′); the autoradiographic images were obtained with a phosphorimager (Storm 820; GE Healthcare) and quantified using Image Quant 5.2 software (GE Healthcare). Samples were normalized with $tRNA^{Met}$ (primer tRNA-Met). Images were obtained and quantified as described above. *R. etli* cells subjected to a treatment with DMSO (1:1,000) alone served as a control for ctRNA analysis.

RESULTS

Binding assays of ctRNA with *repABC incα* mRNA (mIncα). The 155-bp *repB-repC* intergenic region, or *incα* region, of *R. etli* p42d contains an antisense RNA (ctRNA) gene. The promoter of this gene, including its −10 and −35 hexameric elements, was mapped experimentally (30). The size, 59 nt, of the ctRNA was calculated by Venkova-Canova et al., using *in vitro* and *in vivo* techniques (30). An *in silico* analysis of its secondary structure predicted a single stem-loop-structured domain (Fig. 1b). Genetic analysis suggested that this structure is essential for the termination of its own transcription (30).

The ctRNA/mIncα duplex affinity was determined by gel shift analysis. Various concentrations of unlabeled mIncα were tested with a fixed concentration of labeled ctRNA to determine the dissociation constant (K_d) of the ctRNA/mIncα complex. This analysis gave a K_d value of 5.17 nM (Fig. 2a and Table 3). To determine the binding rate constant (k_a) of the ctRNA/mIncα complex, a constant amount of mIncα was incubated with labeled ctRNA, and the time course of the binding between these molecules was followed (Fig. 2b and Table 3). The rate constant was calculated to be $\cong 3.16 \times 10^5\ M^{-1}\ s^{-1}$ at 37°C.

***In vitro* secondary structure probing of the ctRNA, mIncα region, and ctRNA/mIncα complex.** The ctRNA is encoded in the complementary strand to the *repB-repC* intergenic region. This fact allows the assumption of an interaction involving both

TABLE 3. Binding affinity constants and K_d pairing rate constants (k_a) of the ctRNA and its derivatives[a]

ctRNA or derivative	Binding affinity parameter[b]		Pairing rate constant (k_a)[c]	
	B_{max} ± SD (%)	K_d ± SD (nM)	a ± SD	k_a (M^{-1} s^{-1})
ctRNAwt	93.09 ± 2.62	5.17 ± 0.51	64.162 ± 4.770	3.160×10^5
ctRNA-CGCU	103.96 ± 3.28	10.29 ± 1.03	72.709 ± 3.642	2.200×10^5
ctRNA-Δ(52-59)	96.09 ± 1.17	2.60 ± 1.29	84.065 ± 4.184	18.300×10^5
ctRNA-mut2-5	59.92 ± 15.33	161.41 ± 64.11	18.157 ± 3.916	0.260×10^5
ctRNA-mut2-5-CGCU	41.02 ± 9.49	163.22 ± 58.41	5.109 ± 0.447	0.780×10^5
ctRNA-D(2-11)	13.40 ± 0.88	244.25 ± 22.45	ND	ND
ctRNA-D(1-8)	46.39 ± 7.53	118.90 ± 32.73	11.395 ± 0.928	0.680×10^5
ctRNA-D(2-11)(52-59)	ND[d]	ND	ND	ND

[a] Values are the means of results from three independent assays ± SD.
[b] Binding affinity parameters were calculated according to the equation $y = (B_{max} \cdot x)/(K_d + x)$, where y is the percentage of complexed RNA, B_{max} is the amplitude of the reaction, x is the concentration of the target RNA, and K_d is the dissociation constant.
[c] Binding rate constants were determined from time courses assays. Data were fitted to the equation $y = y0 + a \cdot (1 - e^{-bt})$, where y is the percentage of complexed RNA, a represents the amplitude of the binding reaction, b is the reaction rate constant, and bt is the time.
[d] ND, not determined.

domains. RNase and lead probing assays were undertaken to further study this hypothesis.

For studying the ctRNA structure and the residues involved in the association with its target, it was ^{32}P 5′ end labeled and partially digested with RNase T1, RNase V1, or lead acetate. RNase T1 preferentially cuts at 5′ of guanosines in single-stranded regions; lead cleaves 3′ to any nucleotide in non-paired RNA regions, and the RNase V1 preferentially digests double-stranded or stacked RNA regions. The resulting degradation pattern was used for secondary structure prediction with MFold software (Fig. 1b). The analysis showed that the ctRNA is composed of a 14-nt-long helix (Hct) and capped by an apical loop containing a GNRA motif (Lct) (14). Two single-stranded regions flank this stem-loop: the left arm, of 13 nt (E1ct), and a U-rich right arm of 9 nt (E2ct) (Fig. 3). In summary, it contains the essential features of an intrinsic transcriptional terminator. Partial digestions of the 5′-labeled ctRNA in the presence of the mIncα RNA showed a clear resistance to cleavage by both RNase T1 and lead, suggesting the involvement of the whole antisense molecule in the interaction with its target.

The secondary structure of the mIncα RNA was subsequently analyzed to identify the nucleotides involved in the interaction with the ctRNA. Target RNA was 5′ end labeled and subjected to treatment with nucleases and lead, as described above. The resulting degradation pattern was employed for secondary structure prediction with MFold software (Fig. 4). These studies indicated that mIncα folds into a complex structure that contains two major paired regions, H1α and H2α: H1α is a 20-bp-long helical region interrupted by four internal loops; H2α presents two helixes, connected by an internal loop and closed by a six-nucleotide apical loop (L1α). H1α and H2α are joined by a 20-nt-long linker that can be folded in a short stem-loop domain (L2α). The loop sequence (5′-GUAGCA) of this structure is complementary to the first 5 nt of the ctRNA left arm. The 3′ end of mIncα is a single-stranded region of 20 nt containing the SD sequence, the *repC* initiation codon remains unpaired, and the region complementary to the ctRNA is completely paired (Fig. 4b). Partial digestions of the 5′-end-labeled mIncα in the presence of the ctRNA revealed a patent conformational change (Fig. 5). This new folding includes two stem-loop domains separated by a 18-nt-long single-stranded region (E2c) (Fig. 5b and c). The first stem-loop motif contains 5-paired nucleotide (H1c), a bulge of two unpaired nucleotides (A_{71} and A_{86}), and a loop (L1c) of 8 nt. The second stem-loop motif or S-element contains 17 paired nucleotides (H2c), a 5-nt loop (L2c), one bulged nucleotide (G_{111}), and a bubble of two unpaired nucleotides (A_{120} and C_{136}); the putative *repC* SD sequence is occluded within the S element.

Relative abundances of *repB* and *repC* transcripts. ctRNA binding to its target RNA induces in the latter the formation of a stem-loop structure (the S element), which can be interpreted as an intrinsic terminator or, taking into account that the S element occludes the *repC* Shine-Dalgarno sequence, as a *repC* translational attenuator. To discriminate between these two hypotheses, the relative abundances of the *repB* and *repC* transcripts—up- and downstream from the S-element—were calculated through a Northern blot analysis, using as probes two oligonucleotides, one of them (RepB-HL) complementary to the *repB* 3′ end and the other (REpC-HL) complementary to the *repC* 5′ end. If the relative abundances of *repB* and *repC* are the same, it would indicate that the S element works as a translational attenuator. On the other hand, the *repB* transcript being more abundant than that of *repC* would indicate that the S element works as an intrinsic transcriptional terminator. As shown in Fig. 6, *repB* is approximately 17-fold more abundant that *repC*, indicating that the S element works as an intrinsic terminator.

Binding assays of ctRNA mutant derivatives with mIncα. Most *cis*-encoded antisense RNAs have a complex secondary structure containing one or several stem-loop elements. Antisense RNAs initiate the interaction with their cognate RNA through complementary apical loops. However, in some cases, the initial interaction between an antisense RNA and its target involves a single-stranded region and a stem-loop motif: examples include the toxin-antitoxin stabilization system (*hok*/*sok*) of plasmid R1 and RNA-IN/RNA-OUT of insertion sequence IS*10* (28, 15).

The ctRNA secondary structure described here possesses three candidate regions for initiation of the interaction with its mIncα target site: the left arm, the right arm, and the apical loop. To determine which regions are required to initiate binding to mIncα, we synthesized a collection of truncated versions

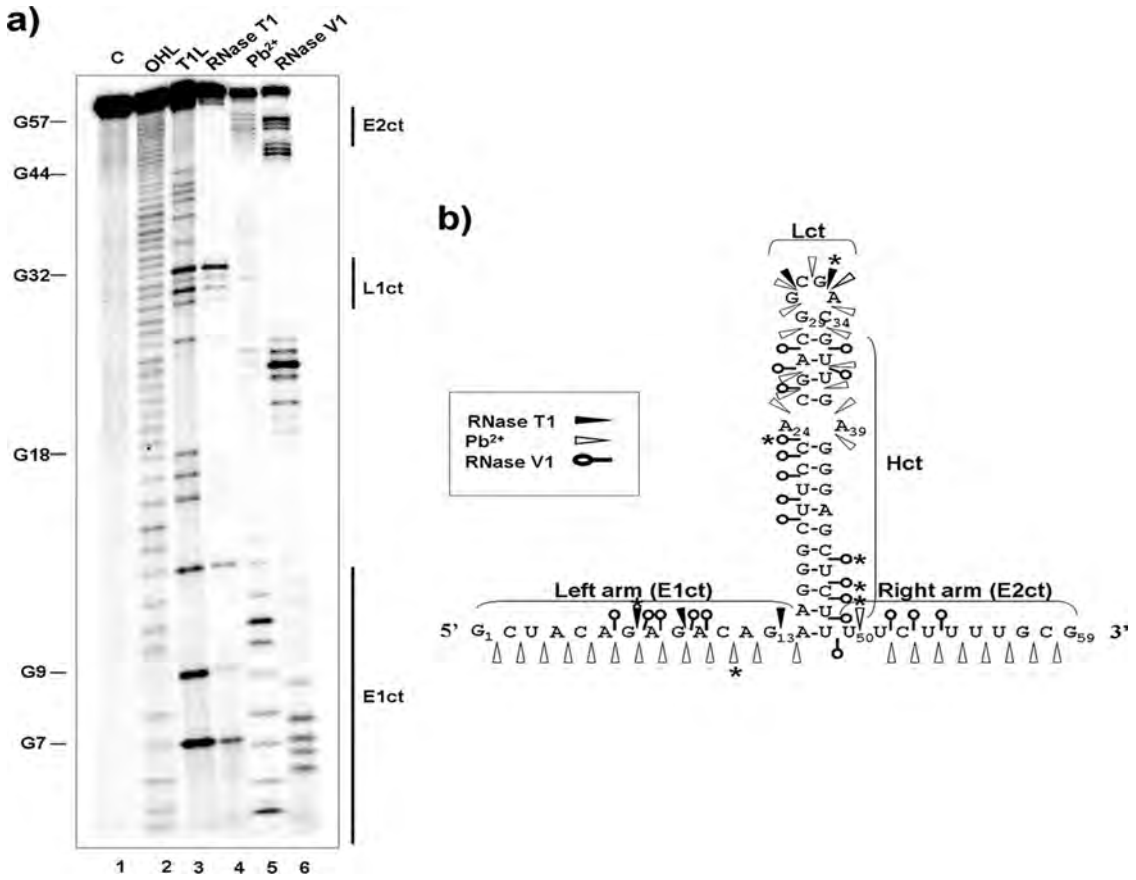


FIG. 3. ctRNA secondary structure. (a) Autoradiogram of a polyacrylamide gel used to resolve 5′-labeled ctRNA after treatment with RNase T1, lead acetate (Pb^{2+}), and RNase V1. Lane 1 (C), undigested probe; lane 2 (OHL), alkaline ladder; lane 3 (T1L), RNase T1 partial digestion of denatured ctRNA used as a ladder; lane 4, RNase T1 partial digestion of ctRNA; lane 5, Pb^{2+} partial degradation of ctRNA; lane 6, RNase V1 partial digestion of ctRNA. Vertical bars indicate loop and arms of the ctRNA. (b) Secondary structure model for ctRNA. Relevant loops and linear and helix regions are marked with the letters L, E, and H, followed by a number and the letters "ct." Major cuts are indicated by asterisks.

of the ctRNA, as well as versions carrying point mutations. The collection of constructs tested included the following: (i) a deletion mutant lacking nucleotides 2 to 11 in the left arm [ctRNA-Δ(2-11)], (ii) a deletion mutant lacking nucleotides 1 to 8 in the left arm [ctRNA-Δ(1-8)], (iii) a ctRNA derivative in which nucleotides 2 to 5 in the left arm were changed from CUAC to AAUG [ctRNAmut2-5], (iv) a ctRNA in which the loop was changed from 5′-GGCGAC to 5′-GCGCUC [ctRNA-CGCU]; (v) a deletion mutant without nucleotides 52 to 59 in the right arm [ctRNA-Δ(52-59)]; (vi) a deletion mutant lacking nucleotides 2 to 11 and 52 to 59 [ctRNA-Δ(1-11)(52-59)]; and (vii) a mutant in which nucleotides 2 to 5 and those located in the loop were changed as described above (ctRNA-mut2-5CGCU). Structure probing assays indicated that all RNA variants conserved the central stem-loop motif of the wild-type ctRNA (data not shown). We then studied the interaction kinetics of these derivatives for the target mRNA. The thermodynamic and kinetic parameters of the interaction between these derivatives and their targets are shown in Fig. 7 and Table 3. These results indicate that the terminal nucleotides of the ctRNA left arm are crucial for the interaction with the target mRNA. Only those ctRNA variants carrying nucleotide sequence changes or truncations affecting the left arm severely affect the complex formation. Little effect is observed when the apical loop sequence is altered, indicating that it is not required for the interaction with the target. An analysis of the sequence and secondary structure of the target RNA shows that the complementary region corresponding to the first few nucleotides of the ctRNA left arm lies within the loop L1α, suggesting that this loop is probably involved in the first steps of the ctRNA-target interaction (Fig. 4b).

An oligonucleotide complementary to the left arm of the ctRNA strongly inhibits duplex formation between ctRNA and its target site. To analyze the role of the left arm of the ctRNA in duplex formation, two deoxyribooligonucleotides were synthesized, one of them complementary to the left arm of the ctRNA (ODN5′) and the other to the right arm (ODN3′), and were treated individually or together as competitors for ctRNA/mIncα hybrid formation (Fig. 8). Under the conditions used in this experiment, around 50% of the labeled ctRNA was able to bind its target. A similar percentage of ctRNA/mIncα hybrids was observed when the reaction was done in the presence of ODN3′. In contrast, the yield of ctRNA/mIncα duplex was substantially reduced in the presence of ODN5′. Binding

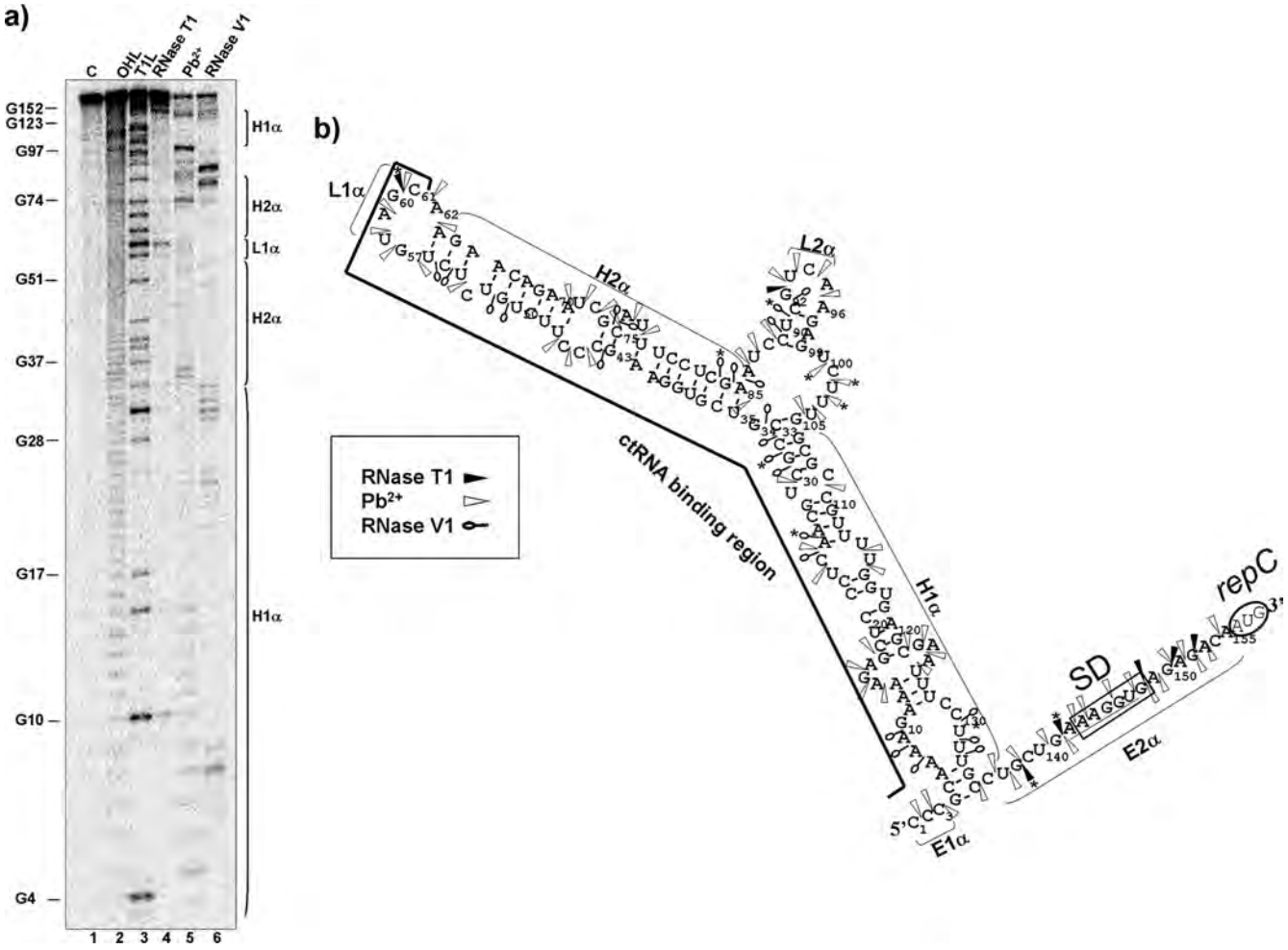


FIG. 4. Secondary structure of the target mIncα. (a) Autoradiogram of a polyacrylamide gel used to resolve 5′-labeled mIncα after treatment with RNase T1, lead acetate (Pb^{2+}), and RNase V1. Lane 1 (C), undigested probe; lane 2 (OHL), alkaline ladder; lane 3 (T1L), RNase T1 partial digestion of denatured mIncα used as a ladder; lane 4, RNase T1 partial digestion of mIncα; lane 5, Pb^{2+} partial degradation of mIncα; lane 6, RNase V1 partial digestion of mIncα. Vertical bars indicate loop and helixes of the mIncα. (b) mIncα secondary structure consistent with cleavage patterns. Black arrowheads indicate RNase T1 sites, white arrowheads indicate Pb^{2+} sites, and open circles indicate RNase V1 cleavages. Major cuts are indicated by asterisks (*). Relevant loop and helix regions are marked with the letters L and H, followed by a number and the Greek letter, "α." The mIncα region complementary to the ctRNA is marked with a black line. "SD" indicates the position of *repC* Shine-Dalgarno sequence. The *repC* initiation codon (AUG) is encircled.

assays in the presence of both ODN5′ and ODN3′ resulted in percentages of ctRNA/mIncα complex very similar to that in the presence of ODN5′ alone (Fig. 8). Therefore, these results suggest that complex formation strongly requires the ctRNA left arm and that the right arm does not participate significantly during the first steps of the ctRNA/mIncα interaction.

Mutations in the left arm of the ctRNA but not those located in the loop change its incompatibility properties. We used a plasmid incompatibility assay to confirm our *in vitro* observations that the left arm of ctRNA is essential for the interaction with mIncα. If this also applies *in vivo*, mutations or deletions in the left arm of the ctRNA gene might abolish incompatibility with p42d, the parental plasmid. In contrast, the incompatibility properties of ctRNA might be unaffected by mutations in regions not involved in the first steps of the RNA-RNA interaction.

To test this hypothesis, we constructed plasmids with the *repBC* intergenic sequence carrying mutations in the ctRNA gene to produce ctRNA variants and employed them in our studies (Table 2; see also Fig. S1 in the supplemental material). These constructs were introduced into CFNX101, an *R. etli recA* derivative containing p42d, and the plasmid profiles of the transconjugants were evaluated. The mutant with the modified loop, pCR-ctRNA-CGCU, was as incompatible as the wild-type ctRNA gene. In contrast, the constructs pCR-ctRNAmut2-5, pCR-ctRNAmut2-5CGCU, pCR-ctRNA-Δ(1-8), and pCR-ctRNA5N, carrying ctRNA genes with modified left arms, were compatible. In addition, plasmid pCR-ctRNAart, expressing the ctRNA gene with the left and right arms identical in sequence to the arms present in the wild-type ctRNA but containing a stem-loop of the same size with a different sequence, was found to be compatible. However, pCR-ctRNADqS, a plasmid with a chimeric gene containing the p42d ctRNA left arm and the stem-loop and right arm from

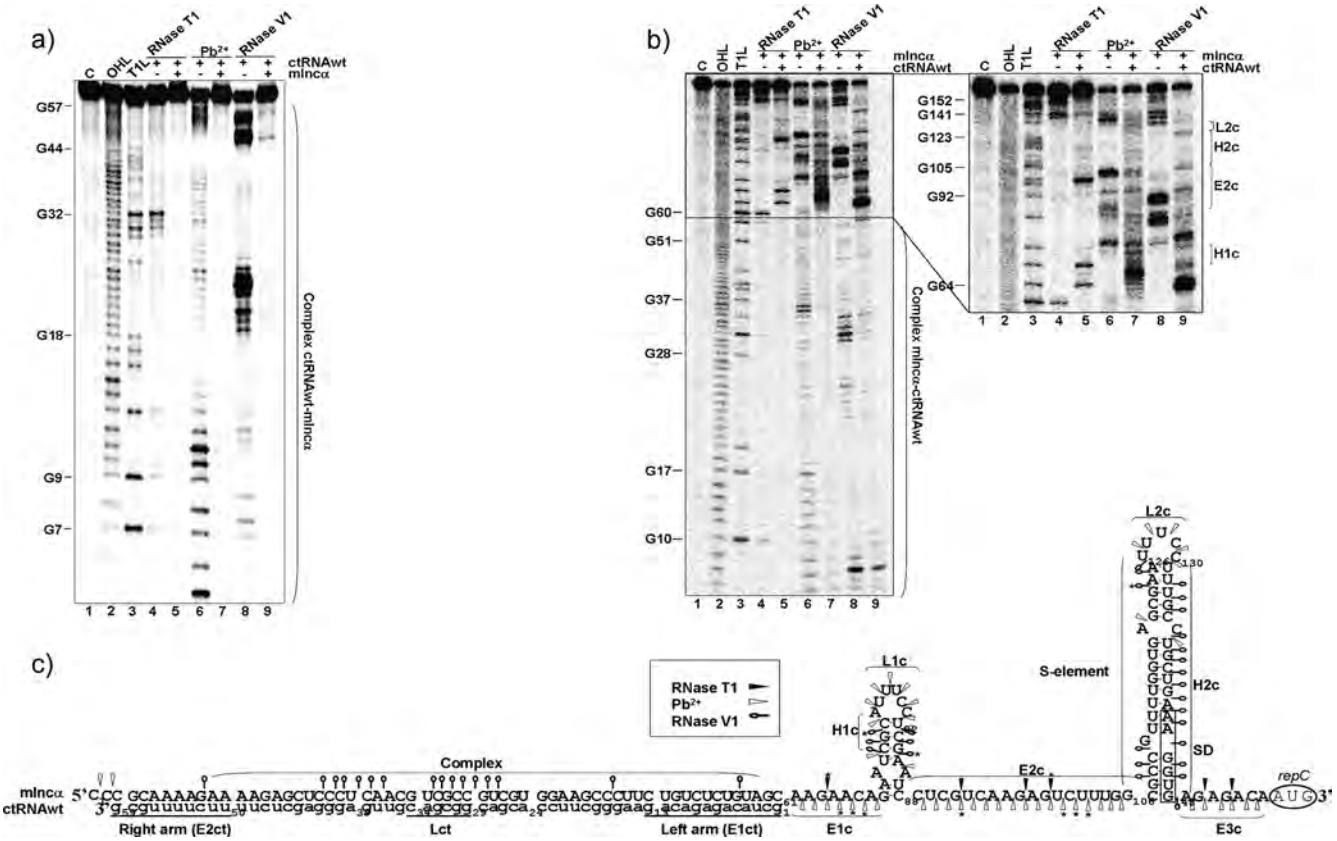


FIG. 5. ctRNA/mIncα complex secondary structure. (a) Autoradiogram of a polyacrylamide gel of 5′-labeled ctRNA in the presence or absence of the target RNA (mIncα) treated with RNase T1, lead acetate (Pb^{2+}), and RNase V1. Lane 1 (C), undigested probe; lane 2 (OHL), alkaline ladder; lane 3, (T1L), RNase T1 partial digestion of denatured ctRNA used as a ladder; lane 4, RNase T1 partial digestion of ctRNA; lane 5, RNase T1 partial digestion of ctRNA in the presence of a 10× excess of target mIncα; lane 6, Pb^{2+} partial degradation of ctRNA; lane 7, Pb^{2+} partial degradation of ctRNA in the presence of a 10× excess of target mIncα; lane 8, RNase V1 partial digestion of ctRNA; lane 9, RNase V1 partial digestion of ctRNA in the presence of a 10× excess of target mIncα. (b) Autoradiograms of polyacrylamide gels of 5′-labeled mIncα in the presence or absence of the target ctRNA treated with RNase T1, Pb^{2+}, and RNase V1. Lane 1 (C), undigested probe; lane 2 (OHL), alkaline ladder; lane 3, (T1L), RNase T1 partial digestion of denatured mIncα used as a ladder; lane 4, RNase T1 partial digestion of mIncα; lane 5, RNase T1 partial digestion of mIncα in the presence of a 10× excess of target ctRNA; lane 6, Pb^{2+} partial degradation of mIncα; lane 7, Pb^{2+} partial degradation of mIncα in the presence of a 10× excess of target ctRNA; lane 8, RNase V1 partial digestion of mIncα; lane 9, RNase V1 partial digestion of mIncα in the presence of a 10× excess of target ctRNA. (c) ctRNA/mIncα complex secondary structure model consistent with cleavage patterns. Uppercase letters indicate sequence corresponding to mIncα, and lowercase letters indicate the ctRNA sequence. Black arrowheads indicate RNase T1 sites, white arrowheads indicate the Pb^{2+} sites, and open circles indicate RNase V1 cleavage. Relevant loops, helixes, and linear regions are marked with letters L, H, and E, followed by a number and a second letter, c. "SD" indicates the position of *repC* Shine-Dalgarno sequence. Major cuts are indicated by asterisks (*).

pSymA ctRNA, was incompatible. These results show that *in vivo*, the left arm is essential but not sufficient for the ctRNA-target interactions and the involvement of the stem-loop domain is also needed to exert incompatibility.

Half-life of the ctRNA. *cis*-encoded small antisense RNAs involved in the regulation of plasmid replication are generally short-lived, consistent with efficient regulation of plasmid copy number. To determine the half-life of the ctRNA, *R. etli* CFNX101 (with p42d) was cultivated in rich medium to an OD_{620} of 0.7. Cell culture was stopped by the addition of rifampin. *R. etli* CFNX101, treated similarly but without the addition of rifampin, served as a control. Samples were then taken at different time points, and the total RNA was isolated from the samples. The abundance of the ctRNA was quantified by Northern blot analysis using a ctRNA probe and a $tRNA^{Met}$ probe as a loading control. The half-life of the ctRNA was calculated to be around 5 min (Fig. 9).

DISCUSSION

A common feature of prokaryotic antisense RNAs is the presence of well-defined secondary structures that act as key structure elements, which play essential roles in the rapid and efficient interaction with their targets. The secondary structure of the ctRNA encoded in the *repABC* operon of plasmid p42d consists of a single stem-loop flanked by two unpaired regions; the loop in this molecule lacks the U-turn (5′-YUNR) motif frequently found in other *cis*-encoded antisense RNAs (10, 13) but contains a GNRA motif usually used as a stabilizing motif in RNA-RNA interactions. We have previously shown that the

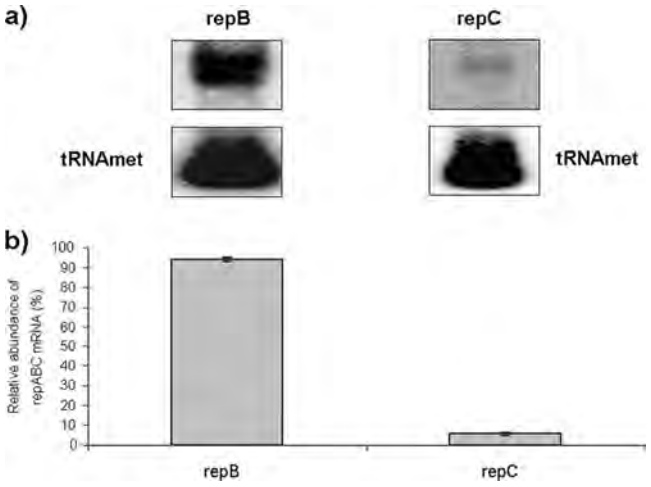


FIG. 6. Relative abundances of *repB* and *repC* transcripts. (a) Autoradiogram of a Northern blot of *Rhizobium etli* CFNX101 total RNA hybridized against two oligonucleotides: one complementary to the *repB* 3′ end and the other complementary to the *repC* 5′ end. The RNA load was normalized, hybridizing the same samples against an oligonucleotide complementary to $tRNA^{Met}$. (b) Graphic representation of the normalized hybridization signals.

stem-loop structure and the 3′ U-rich tail of the ctRNA are essential for transcriptional termination (30). On the other hand, the target RNA (mIncα) forms a large stem-loop structure interrupted by several bulges and internal loops, but the region containing the *repC* Shine-Dalgarno sequence remains unpaired. The interaction of the ctRNA with the target RNA (mIncα) induces a refolding of the sequence placed downstream of the anchoring site, resulting in the formation of two new stem-loop structures in the unpaired region of the target RNA; one of them, the S element, behaves as an intrinsic transcriptional terminator.

The ability of ctRNAs to correct fluctuations in plasmid copy number appropriately depends on the efficiency of binding to their targets and on their short half-lives. The ctRNA present in the p42d *repABC* operon fulfills these requirements: its half-life was calculated to be around 5 min, which is in the same range (1 to 5 min) as those of the ctRNAs from other plasmids (4, 7), and its pairing rate constant was $3.1 \times 10^5\ M^{-1}\ s^{-1}$, again in the same range as those of other sense-antisense interactions (i.e., RNA-OUT/RNA-IN and RNAI/RNAII) (15, 29).

Several lines of evidence indicate that the left arm of the ctRNA is responsible for efficient interaction with the target. First, ctRNA derivatives carrying changes or deletions in the first few nucleotides of the left arm show substantially impaired binding abilities. In contrast, mutations in other regions of the ctRNA do not affect binding properties. Second, oligonucleotides blocking the left arm of the ctRNA severely affect its capacity to bind the target RNA. In contrast, oligonucleotides blocking the right arm of the ctRNA do not affect the ability of the ctRNA to bind its target. Third, plasmid incompatibility tests also indicated that the left-arm ctRNA is essential to exert this phenotype: the constructs expressing ctRNAwt or a ctRNA derivative with an extensive change in the loop were equally effective at displacing p42d, whereas the constructs expressing a ctRNA with mutations in the left arm or carrying deletions in this region were compatible with p42d. However, a construct expressing a ctRNA with both arms identical to the arms of p42d but harboring a stem-loop structure similar in size but not in sequence to that present in p42d ctRNA was compatible. These results suggest that some nucleotides of the stem-loop are also required to exert incompatibility against p42d.

As mentioned earlier, the pSymA ctRNA is compatible with p42d and shares an overall 70% of sequence identity with p42d ctRNA. However, sequence identity between ctRNA stem-loops is only 63%. A plasmid expressing a pSymA ctRNA but with a left-arm sequence identical to the same region of p42d ctRNA is capable of displacing p42d, supporting the notion that besides the interaction between the ctRNA left arm and mIncα, some additional pairing between these two molecules is required to exert incompatibility.

A reexamination of the target RNA structure revealed that the loop (L1α) is complementary to the first nucleotides of the ctRNA. It is therefore possible that under *in vitro* conditions, the interaction begins at this location and propagates, with a

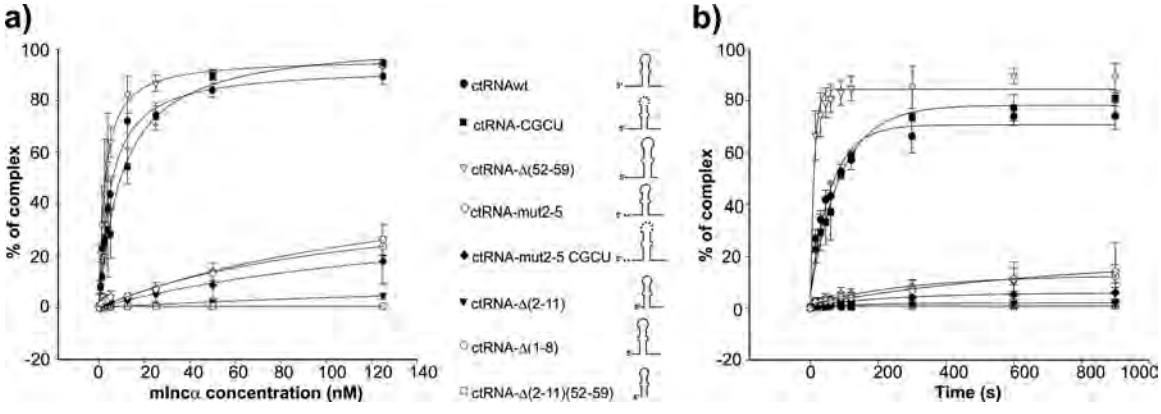


FIG. 7. Binding kinetics of the ctRNA and its truncated or mutant derivatives for the target RNA (mIncα). The binding assays were performed as described in Materials and Methods. (a) A graphic representation of the relative complex formation between the ctRNA or its derivatives and mIncα, as a function of the mIncα concentration. (b) A graphic representation of the relative complex formation between the ctRNA or its derivatives and mIncα as a function of the incubation time.

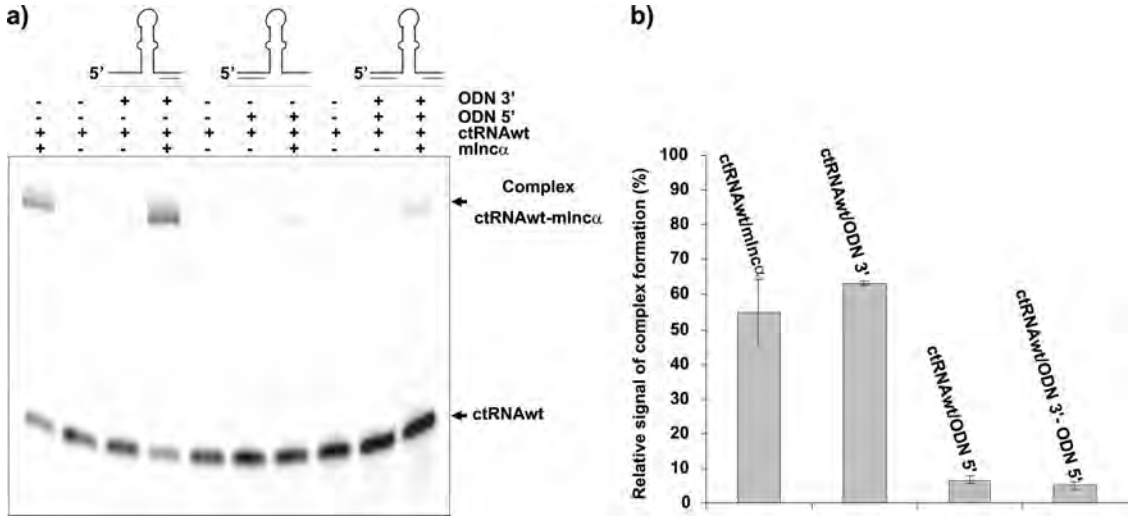


FIG. 8. Inhibition of ctRNA/mIncα complex formation by specific oligonucleotides. (a) Electrophoretic mobility shift assay (EMSA) of ^{32}P-labeled ctRNA and its target (mIncα) in the presence of oligonucleotides complementary to the ctRNA left arm (ODN5′), complementary to the ctRNA right arm (ODN3′), or both, as described in Materials and Methods. (b) Graphic representation of the relative complex formation obtained from three independent EMSA assays.

zipper-like mechanism, until the formation of the duplex is complete. However, our *in vivo* experiments suggest that extended pairing is not a requirement to induce the mIncα structural change needed to occlude the *repC* Shine-Dalgarno sequence. Similar observations indicating that full complex formation between a ctRNA and its target is not required to exert their regulatory functions have also been made for other plasmid systems (4).

As far as we know, this is the first time this type of RNA-RNA mechanism (apical-loop/single-stranded region) in a replication control system has been reported. However, this mechanism has been found previously in the transposition regulation system of IS*10* (15) and in the postsegregational killing system *hok*/*sok* of the plasmid R1 (28).

Our findings are consistent with following model of regulation: the target RNA, the mIncα encoded between *repB* and *repC*, adopts either of two radically different conformations, one in the absence of the ctRNA and the other in its presence. In the absence of the ctRNA, the target RNA folds in such way that the Shine-Dalgarno sequence of *repC* is free and ready for translation; when the critical concentration of RepC is reached in the cell, plasmid replication begins. In contrast, when the ctRNA pairs with the target, it induces the formation of two stem-loop structures; one of them operates as an intrinsic terminator, aborting *repC* transcription and, in consequence, plasmid replication.

An *in silico* analysis of the *repB-repC* intergenic region of other *repABC* plasmids showed that the ctRNA gene is in the

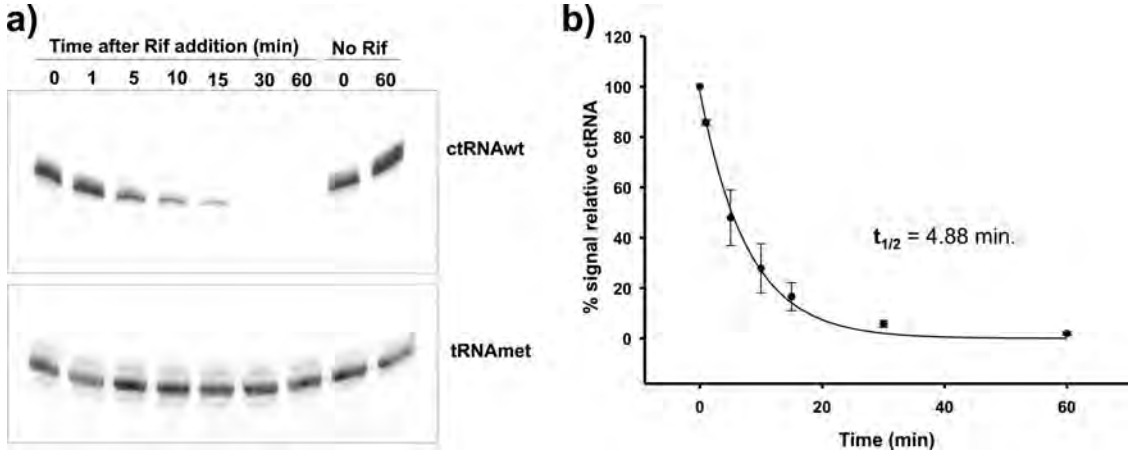


FIG. 9. ctRNA half-life determination. (a) Northern blots of total RNA from samples taken at the indicated times (min) after rifampin addition and controls without rifampin, hybridized with an oligonucleotide complementary to the ctRNA and with an oligonucleotide complementary to tRNA-Met, as described in Materials and Methods. (b) Graphic representation of ctRNA decay observed in three independent experiments.

same relative position in most of them (8). Moreover, in all cases, the ctRNA has a similar stem-loop structure (30, 18, 9). These observations suggest that all ctRNAs in *repABC* plasmids have a similar mechanism of action.

ACKNOWLEDGMENTS

This work was supported by the Consejo Nacional de Ciencia y Tecnología (México) (grant number 46738-Q to M.A.C.), the Ministerio de Educación y Ciencia (Spain) (grant number BFU2006-02508 to A.B.-H.), and the Programa de Apoyo a Proyectos de Investigación e Inovación Tecnológica (México) (grant number IN20580 to M.A.C.). R.C.-R. was supported during the Ph.D. program (Programa de Doctorado en Ciencias Biomédicas-Universidad Nacional Autónoma de México) by scholarships from the Consejo Nacional de Ciencia y Tecnología and Dirección General de Estudios de Posgrado (México).

We are greatly indebted to Ángeles Pérez-Oseguera, América Rivera-Urbalejo, Francisco Pedraza-López, and Vicente Augustin-Vacas for their technical support.

REFERENCES

1. **Barroso-delJesus, A., M. Tabler, and A. Berzal-Herranz.** 1999. Comparative kinetic analysis of structural variants of the hairpin ribozyme reveals further potential to optimize its catalytic performance. Antisense Nucleic Acid Drug Dev. **9:**433–440.
2. **Bartosik, D., J. Baj, and M. Wlodarczyk.** 1998. Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. Microbiology **144:**3149–3157.
3. **Brantl, S.** 2002. Antisense RNAs in plasmids: control of replication and maintenance. Plasmid **48:**165–173. doi:10.1016/S0147-619X(02)00108-7.
4. **Brantl, S.** 2004. Plasmid replication control, p. 47–59. *In* E. Funnell and G. J. Phillips (ed.), Plasmid biology. ASM Press, Washington, DC.
5. **Brantl, S.** 2007. Regulatory mechanisms employed by *cis*-encoded antisense RNAs. Curr. Opin. Microbiol. **10:**102–109. doi:10.1016/j.mib.2007.03.012.
6. **Brennan, R. G., and T. M. Link.** 2007. Hfq structure, function and ligand binding. Curr. Opin. Microbiol. **10:**125–133. doi:10.1016/j.mib.2007.03.015.
7. **Brenner, M., and J. Tomizawa.** 1991. Quantitation of ColE1-encoded replication elements. Proc. Natl. Acad. Sci. U. S. A. **88:**405–409.
8. **Cevallos, M. A., R. Cervantes-Rivera, and R. M. Gutiérrez-Rios.** 2008. The *repABC* plasmid family. Plasmid **60:**19–37. doi:10.1016/j.plasmid.2008.03.001.
9. **Chai, Y., and S. C. Winans.** 2005. A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. Mol. Microbiol. **56:**1574–1585. doi:10.1111/j.1365–2958.2005.04636.x.
10. **Franch, T., and K. Gerdes.** 2000. U-turns and regulatory RNAs. Curr. Opin. Microbiol. **3:**159–164. doi:10.1016/S1369-5274(00)00069-2.
11. **González, V., P. Bustos, M. A. Ramírez-Romero, A. Medrano, H. Salgado, I. Hernández-González, C. Hernández-Celis, V. Quintero, G. Moreno-Hagelsieb, L. Girard, O. Rodríguez, M. Flores, M. A. Cevallos, J. Collado-Vides, D. Romero, and G. Dávila.** 2003. The mosaic structure of the symbiotic plasmid of *Rhizobium etli* and its relations with other symbiotic genomic compartments. Genome Biol. **4:**R36. doi:10.1186/gb-2003-4-6-r36.
12. **Hanahan, D.** 1983. Studies of transformation of *E. coli* with plasmids. J. Mol. Biol. **166:**557–560.
13. **Heidrich, N., and S. Brantl.** 2003. Antisense-RNA mediated transcriptional attenuation: importance of a U-turn loop structure in the target RNA of plasmid pIP501 for efficient inhibition by the antisense RNA. J. Mol. Biol. **333:**917–929. doi:10.1099/mic.0.2006/002329-0.
14. **Jucker, F. M., and A. Pardi.** 1995. GNRA tetraloops make a U-turn. RNA **1:**219–222.
15. **Kittle, J. D., R. W. Simons, J. Lee, and N. Kleckner.** 1989. Insertion sequence IS10 anti-sense pairing initiates by an interaction between the 5′ end of the target RNA and a loop in the anti-sense RNA. J. Mol. Biol. **210:**561–572.
16. **Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop, and K. M. Peterson.** 1995. Four new derivatives of the broad-host range cloning vector pBBR1MCS, carrying different antibiotic resistance cassettes. Gene **166:**175–176. doi:10.1016/0378-1119(95)00584-1.
17. **Kumar, C. C., and R. P. Novick.** 1985. Plasmid pT181 replication is regulated by two countertranscripts. Proc. Natl. Acad. Sci. U. S. A. **82:**638–642.
18. **MacLellan, S. R., L. A. Smallbone, C. D. Sibley, and T. M. Finan.** 2005. The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. Mol. Microbiol. **55:**611–623. doi:10.1111/j.1365-2958.2004.04412.x.
19. **Martínez-Salazar, J., D. Romero, M. L. Girard, and G. Dávila.** 1991. Molecular cloning and characterization of the *recA* gene of *Rhizobium phaseoli* and construction of *recA* mutants. J. Bacteriol. **173:**3035–3040.
20. **Nacheva, G. A., and A. Berzal-Herranz.** 2003. Preventing nondesired RNA-primed RNA extensión catalyzed by T7 RNA polymerase. Eur. J. Biochem. **270:**1458–1465. doi:10.1046/j.1432-1033.2003.03510.x.
21. **Noel, K. D., A. Sánchez, L. Fernández, J. Leemans, and M. A. Cevallos.** 1984. *Rhizobium phaseoli* symbiotic mutants with transposon Tn5 insertions. J. Bacteriol. **158:**48–155.
22. **Pappas, K. M., and S. C. Winans.** 2003. The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. Mol. Microbiol. **49:**441–455. doi:10.1046/j.1365-2958.2003.03560.x.
23. **Pappas, K. M.** 2008. Cell-cell signaling and the Agrobacterium tumefaciens Ti plasmid copy number fluctuations. Plasmid **60:**89–107. doi:10.1016/j.plasmid.2008.05.003.
24. **Persson, C., E. G. Wagner, and K. Nordström.** 1988. Control of replication of plasmid R1: kinetics of *in vitro* interaction between the antisense RNA, CopA, and its target, CopT. EMBO J. **7:**3279–3288.
25. **Ramirez-Romero, M. A., N. Soberon, A. Perez-Oseguera, J. Tellez-Sosa, and M. A. Cevallos.** 2000. Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. J. Bacteriol. **182:**3117–3124.
26. **Ramirez-Romero, M. A., J. Tellez-Sosa, H. Barrios, A. Perez-Oseguera, V. Rosas, and M. A. Cevallos.** 2001. RepA negatively autoregulates the transcription of the *repABC* operon of the *Rhizobium etli* symbiotic plasmid basic replicon. Mol. Microbiol. **42:**195–204. doi:10.1111/j.1365–2958.2001.02621.x.
27. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host-range mobilization system for *in vivo* genetic engineering transposon mutagenesis in Gram negative bacteria. Biol. Technol. **1:**784–791.
28. **Thisted, T., N. S. Sørensen, E. G. Wagner, and K. Gerdes.** 1994. Mechanism of postsegregational killing: Sok antisense RNA interacts with Hok mRNA via its 5′-end single-stranded leader and competes with the 3′-end of Hok mRNA for binding to the mok translational initiation region. EMBO J. **13:**1960–1968.
29. **Tomizawa, J., and T. Itoh.** 1981. Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. Proc. Natl. Acad. Sci. U. S. A. **78:**6096–6100.
30. **Venkova-Canova, T., N. E. Soberón, M. A. Ramírez-Romero, and M. A. Cevallos.** 2004. Two discrete elements are required for the replication of a *repABC* plasmid: an antisense RNA and a stem-loop structure. Mol. Microbiol. **54:**1431–1444. 10.1111/j.1365-2958.2004.04366.x.
31. **Wagner, E. G., and S. Brantl.** 1998. Kissing and RNA stability in antisense control of plasmid replication. Trends Biochem. Sci. **23:**451–454. doi: 10.1016/S0968-0004(98)01322-X.
32. **Wheatcroft, R., G. D. McRae, and R. W. Miller.** 1990. Changes in *the Rhizobium meliloti* genome and the ability to detect supercoiled plasmids during bacteroid development. Mol. Plant Microbe Interact. **3:**9–17.
33. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. **9:**133–148.

Plasmid 60 (2008) 19–37

Contents lists available at ScienceDirect

Plasmid

journal homepage: www.elsevier.com/locate/yplas

Review

The *repABC* plasmid family

Miguel A. Cevallos [a,*], Ramón Cervantes-Rivera [a], Rosa María Gutiérrez-Ríos [b]

[a] *Programa de Genómica evolutiva, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Apartado Postal 565-A, Cuernavaca, Morelos, Mexico*
[b] *Departamento de Microbiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Apartado Postal 510-3, Cuernavaca, Morelos, Mexico*

ARTICLE INFO

Article history:
Received 7 January 2008
Revised 4 March 2008
Available online 22 April 2008

Communicated by Dhruba K. Chattoraj

Keywords:
Plasmid replication
Plasmid segregation
Plasmid evolution
α-Proteobacteria

ABSTRACT

repABC plasmids are widely distributed among α-proteobacteria. They are especially common in Rhizobiales. Some strains of this bacterial order can contain multiple *repABC* replicons indicating that this plasmid family includes several incompatibility groups. The replication and stable maintenance of these replicons depend on the presence of a *repABC* operon. The *repABC* operons sequenced to date share some general characteristics. All of them contain at least three protein-encoding genes: *repA*, *repB* and *repC*. The first two genes encode proteins involved in plasmid segregation, whereas *repC* encodes a protein crucial for replication. The origin of replication maps within the *repC* gene. In contrast, the centromere-like sequence (*parS*) can be located at various positions in the operon. In this review we will summarize current knowledge about this plasmid family, with special emphasis on their structural diversity and their complex genetic regulation. Finally, we will examine some ideas about their evolutionary origin and trends.

1. Introduction

Bacteria belonging to *Agrobacterium* or *Rhizobium* genera can establish close relationships with plants. The consequences of such interactions are very diverse: most *Agrobacterium* species are plant–pathogens, whereas members of the genus *Rhizobium* are nitrogen-fixing legume symbionts. In the late 1960s and early 1970s it became clear that *Agrobacterium* and *Rhizobium* possess large plasmids of low copy-number and that most of the genes involved in plant–microbe interaction were carried on such plasmids (Higashi, 1967; Escobar and Dandekar, 2003).

The first molecular description of the replication and partitioning of one of these plasmids was a paper by Nishiguchi and co-workers in 1987; they reported the locus responsible for the replication and stable maintenance of pRiA4b, a plasmid in *Agrobacterium rhizogenes* A4. This locus consists of an operon of three genes: *repA* is the first gene to be transcribed and *repC* the last. Similar operons have subsequently been found in many plasmids and secondary chromosomes in at least 19 α-proteobacteria genera (Table 1). Interestingly, some bacterial strains contain several plasmids, each carrying one *repABC* operon. Moreover, replicons harboring two functional *repABC* operons have also been found. The most evident cases of this peculiarity are in *Rhizobium etli* CFN42 and *Rhizobium leguminosarum* 3841, each of which has six plasmids all belonging to the *repABC* family. Two replicons in *R. etli* CFN42 and one in *R. leguminosarum* 3841 contain two *repABC* operons. This suggests, different *repABC* plasmids belong to different incompatibility groups (Young et al., 2006; Gonzalez et al., 2006).

The *repABC* operons have several characteristics in common: the three genes are always in the same relative order: *repA* is upstream from *repB*, with *repC* as the downstream gene of the operon. The large intergenic sequence between *repB* and *repC* genes contains a gene encoding a small antisense RNA. RepA and RepB have sequence similarities with proteins involved in active segregation of plasmids and chromosomes, and RepC is a replication initiator protein (Bartosik et al., 1998; Ramirez-Romero et al., 2001; Pappas and Winans, 2003b). The small RNA, as we will show, plays an essential role in the

* Corresponding author. Fax: +52 777 317 55 81.
E-mail address: mac@ccg.unam.mx (M.A. Cevallos).

0147-619X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.plasmid.2008.03.001

Table 1
List of bacterial species containing replicons with *repABC* operons

Species name/strain	Replicon	*n*	Size	Reference	GenBank
Acidiphilium cryptum JF-5	pACRY01	1	203.5 Kb	Unpublished	NC_009467
Agrobacterium rhizogenes A4	pRiA4b	1	nd	Nishiguchi et al. (1987)	X04833
Agrobacterium rhizogenes K599	pRi2659	1	185.4 Kb	Mankin et al. (2007)	EU186381
Agrobacterium rhizogenes MAFF03-01724	pRi1724	1	217.5 Kb	Moriguchi et al. (2001)	NC_002575
Agrobacterium tumefaciens C58	ChrL[*1]	1	2.07 Mb	Goodner et al. (2001)	NC_003305
	pATC58	1	542.7 Kb		NC_003064
	pTiC58	1	214.2 Kb		NC_003065
Agrobacterium tumefaciens	pTiB6S3	1	nd	Tabata et al. (1989)	M24529
Agrobacterium tumefaciens R10	pTiR10	1	194.1 Kb	Zhu et al. (2000)	AF242881
Agrobacterium tumefaciens MAFF301001	pTi-SAKURA	1	206.4 Kb	Suzuki et al. (1998)	NC_002147
Bradyrhizobium sp. BTAi1	pBBta01	1	228.8 Kb	Giraud et al. (2007)	NC_009475
Brucella abortus bv 1 str 9–941	ChrII	1	1.16 Mb	Halling et al. (2005)	NC_006933
Brucella melitensis bv abortus 2308	ChrII	1	1.15 Mb	Chain et al., 2005	NC_007624
Brucella melitensis 16 M	ChrII	1	1.17 Mb	Del Vecchio et al. (2002)	NC_003318
Brucella suis 1330	ChrII	1	1.2 Mb	Paulsen et al. (2002)	NC_004311
Brucella ovis ATCC 25840	ChrII	1	1.16 MB	Unpublished	NC_009504
Mesorhizobium sp. BNC1	p1	2	343.9 Kb	Unpublished	NC_008242
	p2	1	131.2 Kb		NC_008243
	p3	1	47.5 Kb		NC_008244
Mesorhizobium loti MAFF303099	pMLa	1	351.9 Kb	Kaneko et al. (2000)	NC_002679
	pMLb	1	208.3 Kb		NC_002682
Nitrobacter hamburgensis X14	p1	1	294.8 Kb	Unpublished	NC_007959
	p2	1	188.3 Kb		NC_007961
	p3	1	121.4 Kb		NC_007961
Oceanicola batsensis HTCC2597 (Unfinished genome)	Obat HTCC2597a	1	nd	Unpublished	NZ_AAMO01000009
	Obat HTCC2597b	1	nd		NZ_AAMO01000007
Ochrobactrum anthropi ATCC 49188	ChrII	1	1.89 Mb	Unpublished	NC_009668
	pOANT01	1	170.3 Kb		NC_009669
	pOANT02	1	101.4 Kb		NC_009670
	pOANT03	2	93.5 Kb		NC_009671
Oligotropha carboxidovorans	pHCG3	1	133 Kb	Fuhrmann et al. (2003)	NC_005873
Paracoccus versutus UW1	pTAV320	1	nd	Bartosik et al. (1998)	U60522
Rhizobium etli CFN42	p42a	2	194.2 Kb	Gonzalez et al. (2006)	NC_007762
	p42b	1	184.3 Kb		NC_007763
	p42c	1	250.9 Kb		NC_007764
	p42d	1	371.2 Kb	González et al. (2003)	NC_004041
	p42e	1	505.3 Kb	González et al. (2006)	NC_007765
	p42f	2	642.5 Kb		NC_007766
Rhizobium leguminosarum bv viciae 3841	pRL7	2	151.5 Kb	Young et al. (2006)	NC_008382
	pRL8	1	147.4 Kb		NC_008383
	pRL9	1	352.7 Kb		NC_008379
	pRL10	1	488.1 Kb		NC_008381
	pRL11	1	684.2 Kb		NC_008384
	pRL12	1	870 Kb		NC_008378
Rhizobium sp. NGR234	pNGR234a	1	536.1 Kb	Freiberg et al. (1997)	NC_000914
Rhodobacter sphaeroides 2.4.1.	pB	1	114.1 Kb	Unpublished	NC_007488
	pD	1	100.8 Kb		NC_007490
Roseovarius nubinhibens ISM (Unfinished genome)	RnuISM	1	nd	Unpublished	NZ_AALY01000005
Roseovarius sp. 217 (Unfinished genome)	Ros217a	1	nd	Unpublished	NZ_AAMV01000007
	Ros217b	1	nd		NZ_AAMV01000021
	Ros217c	1	nd		NZ_AAMV01000023
Roseovarius sp. HTCC2601 (Unfinished genome)	Ros HTCC2601a	1	nd	Unpublished	NZ_AATQ01000004
	Ros HTCC2601b	1	nd		NZ_AATQ01000009
	Ros HTCC2601c	1	nd		NZ_AATQ01000027
Roseovarius sp. TM1035	RosTM1035	1	nd	Unpublished	NZ_ABCL01000010
Ruegeria sp. PR1b (Unfinished genome)	pSD25	1	148.6 Kb	Zhong et al. (2003)	NC_004574
Sagittula stellata E-37 (Unfinished genome)	SsteE37	1	nd	Unpublished	NZ_AAYA01000007
Sinorhizobium meliloti SM11	pSymA	1	nd	Stiens	Pers. comm.
	pSymB		nd		
	pSmeSM11a	1	144.1 Kb	Stiens et al. (2006)	DQ145546
	pSmeSM11b	2	181.2 Kb	Stiens et al. (2007)	EF066650
Sinorhizobium meliloti 1021	pSymA	1	1.35 Mb	Barnett et al. (2001)	NC_003037
	pSymB	1	168.3 Kb	Finan et al. (2001)	NC_003078
Sinorhizobium meliloti MBA9	pMBA9a	1	nd	Watson and Heys (2006)	AY914873
Sinorhizobium medicae WSM419	pSMED01	1	1.5 Mb	Unpublished	NC_009620
	pSMED02	1	1.24 Mb		NC_009621
	pSMED03	1	219.3 Kb		NC_009622
Stappia aggregatal AM 12614 (Unfinished genome)	Sta IAM12614a	1	nd	Unpublished	NZ_AAUW01000020
	Sta IAM12614b	1	nd		NZ_AAUW01000021

Table 1 (*continued*)

Species name/strain	Replicon	*n*	Size	Reference	GenBank
Sulfitobacter sp. NAS-14.1 (Unfinished genome)	Sulfit NAS141a	1	nd	Unpublished	NZ_AALZ01000014
	Sulfit NAS141b	1	nd		NZ_AALZ01000012
	Sulfit NAS141c	1	nd		NZ_AALZ01000013
	Sulfit NAS141d	1	nd		NZ_AALZ01000015
	Sulfit NAS141e	1	nd		NZ_AALZ01000011
Sulfitobacter sp. EE-36 (Unfinished genome)	SulfitEE36	1	nd	Unpublished	NZ_AALV01000012
Xanthobacter autotrophicus Py2	pXAUT01	1	316.1 Kb	Unpublished	NC_009717

This list only includes strains for which the DNA sequences of *repABC* operons have been deposited in GenBank. Column marked with "*n*" indicates the number of *repABC* operons present in each replicon, or in case of unfinished genomes, the number of *repABC* operons in that particular contig. "nd" indicates that the replicon size has not been determined.

control of plasmid replication (Venkova-Canova et al., 2004; Chai and Winans, 2005a; MacLellan et al., 2005). In addition to sequence diversity among *repABC* operons, there are other differences. Some possess elements not found in any other *repABC* operons or in only a few members of the family. These differences can be broadly classified into three categories: (i) differences involving transcriptional regulatory elements, (ii) differences related to the number and position of *par*-sites (centromere-like sequences, formally *parS*), and (iii) the presence of peptide-encoding minigenes within the *repABC* operon (Fig. 1).

In this paper, only replicons containing a *repABC* operon will be considered to belong to the *repABC* plasmid family. This review addresses only *repABC* replication-partitioning systems and not the gene content of the members of this plasmid family. We will discuss current knowledge about the elements constituting the replication-partitioning system of this plasmid family. We will consider, specifically, their complex genetic regulation, and then the origin and evolution of this family of plasmids.

2. The partitioning system of *repABC* replicons

Plasmid-partitioning systems are polymer-based DNA segregation machines that consist of two proteins and a centromere-like sequence or *parS*. One of these proteins is an ATPase able to polymerize into filaments, and the second protein binds both *parS* and the ATPase, acting as an adaptor between the plasmids and the filaments that are responsible for the segregation process (Ebersbach et al., 2006; Garner et al., 2007; Lim et al., 2005; Møller-Jensen et al., 2002).

Plasmid-partitioning protein genes are usually organized into a bi-cistronic operon in which the upstream gene always encodes the ATPase (ParA), and the downstream gene, the *parS*-binding protein (ParB). Plasmid-partitioning systems are classified into two types, depending on their ATPase: Type I systems contain a Walker-type ATPase (MinD/ParA superfamily), and Type II an ATPase belonging to the actin/Hsp70 superfamily (Bork et al., 1992; Gerdes et al., 2000; Koonin, 1993). Type I partitioning systems are further subdivided in two categories, Ia and Ib, according to the phylogenetic relationships between the ATPases, their size, the presence or absence of an amino-terminal region with abilities to bind DNA, and the position of the centromere-like sequences.

Type Ia loci encode large *parS*-binding proteins and large ATPases, containing a DNA-binding domain in their N-terminal, and their *parS* loci are located downstream from the gene encoding the *par*-sites-binding protein. These ATPases also regulate the transcription of their own operons, sometimes aided by the *parS*-binding protein (Gerdes et al., 2000; Hirano et al., 1998; Jensen et al., 1994). In this partitioning-system class, the *par*-site is close to and downstream from the partitioning operon. In contrast, proteins encoded in Type Ib systems are usually small, and the

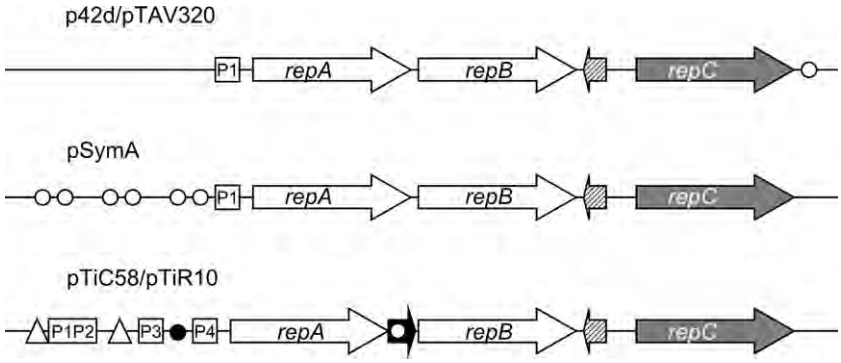


Fig. 1. Diversity of the genetic organization of replication/partitioning regions of the *repABC* replicons. White arrows represent *repA* and *repB* genes, encoding proteins involved in partitioning, and in the negative transcriptional regulation of the operon. White circles show the position of *parS*-sites. Black circles indicate *vir* boxes. Hatched arrows indicate the position of the antisense RNA gene. Gray arrows show the positions of *repC* genes, encoding the proteins limiting for replication. The black arrow represents the *repD* gene. Square boxes indicate the promoter positions, and triangles, the position of *tra* boxes.

par-sites are located in the operon promoter region (Gerdes et al., 2000).

repABC replicons have very low copy-numbers and therefore require an efficient mechanism of segregation to ensure their stable propagation. RepA and RepB proteins, and an intriguing *par*-site(s) provide this mechanism, and this has been demonstrated for *Agrobacterium tumefaciens* pTiB6S3, *Paracoccus versutus* pTAV320, *Rhizobium etli* p42d, and *Sinorhizobium meliloti* 1021 pSymA. Genetic and molecular analyses of these plasmids showed that insertions, frame-shift mutations, or deletions in *repA* or *repB* substantially decrease plasmid stability (Tabata et al., 1989; Bartosik et al, 1998; Ramirez-Romero et al., 2000; MacLellan et al., 2006). The size of the protein components and sequence similarities suggest that the segregation systems of *repABC* replicons are Type Ia. However, their *parS* loci are not close to and downstream from the gene encoding the *parS*-binding protein, as is typical for other members of this subtype.

The *parS* loci of pTAV320, p42d, *S. meliloti* pSymA, and the *A. tumefaciens* pTiC58 and pTiR10 have been identified. They consist in one or more copies of a 16-pb palindromic consensus sequence (**GTT**NN**C**N**GC**N**G**NN**AAC**) and fulfill three requisites: first, they are essential for plasmid stability; second, they are RepB-binding sites; and finally, they are strongly incompatible with their respective parental plasmid when provided *in trans* (Bartosik et al., 1998; Soberón et al., 2004; MacLellan et al., 2005; Chai and Winans, 2005b).

A direct relationship between RepB binding, incompatibility, and plasmid stability has been demonstrated for *S. meliloti* 1021 pSymA: point mutations introduced into the *par*-sites of pSymA that eliminate plasmid incompatibility also cause both diminished affinity for RepB and reduced plasmid stability (MacLellan et al., 2005).

The number of *par*-site elements present in the *parS* locus of a *repABC* replicon varies widely. Some plasmids, including p42d, contain only one *par*-site consensus element. Others, for example pTAV320 and pTiR10, contain two of these elements, and plasmid pSymA has six such elements (Bartosik et al., 2001; Chai and Winans, 2005b; MacLellan et al., 2005; Soberón et al., 2004). However, only one *par*-site consensus sequence is needed for adequate plasmid partitioning (Chai and Winans, 2005b; MacLellan et al., 2005).

The diversity of *par*-sites is not restricted to the number of *par*-site elements in a particular *parS* locus. The position of *parS* within the *repABC* operon differs considerably from one replicon to the next. In some cases (p42d, pTAV320), the *parS* locus is located downstream from *repC*, in others (pTiC58, pTiR10) these sequences are situated within the *repA–repB* intergenic region, and the rest (*e.g.* pSymA) carry *par*-sites upstream from *repA* (Chai and Winans, 2005b; MacLellan et al., 2005; Soberón et al., 2004).

To continue the identification and mapping of potential *par*-sites initiated by Chai and Winans (2005b), we conducted an in *silico* search for 16 bp palindromes containing the *par*-site consensus sequence in all *repABC* operons listed in Table 1. The search included the region from 400 bp upstream from *repA* to 200 bp downstream from *repC* in each operon. Most of the putative *par*-sites were identified in the positions already described for other *repABC* operons. In some *repABC* operons, the potential *par*-sites were found within coding regions, although the functional significance of these sequences needs to be tested. However, in several *repABC* operons we were unable to find a *par*-site consensus sequence, suggesting that these sites are more divergent than suspected, and consequently that many escape detection.

The positioning of *parS* has some interesting consequences: the *parS* loci of pTiC58 and pTiR10 contain two *par*-site sequences organized as inverted repeats. These sequences are situated within a small gene (*repD*) of only 78 codons mapping in the *repA–repB* integenic region. This minigene is part of the *repABC* operon and is transcribed and translated at significant levels. However, it seems to be irrelevant for plasmid segregation: (1) a minigene of the same size as *repD* is present only in the *repA–repB* intergenic region of other pTi plasmids, and even there their sequence is not conserved; (2) a plasmid carrying a mutation that disrupts RepD translation is as stable as the parental plasmid; (3) moving *parS* upstream from *repA* has only small effects on stability and plasmid copy-number; (4) other plasmids, for example p42a and p42b from *R. etli* and pRL8JI from *R. leguminosarum,* also contain a consensus *par*-site sequence within the *repA–repB* intergenic region, but no *repD* equivalent gene is present in these replicons.

The presence of *parS* in the *repA–repB* intergenic region does not affect plasmid partitioning, but it has a role in RepC expression. Indeed, RepB binds *par*-sites, assisted by RepA, forming a nucleoprotein complex that interferes with the progress of RNA polymerase. This interference attenuates *repC* transcription with a concomitant reduction of RepC levels (Chai and Winans, 2005b). Obviously, this regulatory mechanism is only possible in plasmids harboring a *par*-site between or, theoretically, within the *repA* and *repB* genes.

There is evidence that *repABC* plamids require elements in additional to *repA*, *repB*, and *par* for efficient segregation. Bartosik and co-workers (2001) showed that a plasmid vector carrying the *parS* locus from pTAV320 gained stability if the cognate proteins RepA and RepB were provided in *trans*. However, this stability was only half of that of the parental plasmid, pTAV320, which is 100% stable. One possible reason why this system was not fully stable is that the relative amounts of RepA, RepB, and *par* were not the same as those in the parental plasmid, because the vectors carrying them did not all have the same plasmid copy-number. Similar observations have been reported for the replication and partitioning system of p42d: a construct carrying the *repABC* operon of this plasmid was 100% stable in the absence of selective pressure. However, when the *parS* locus was moved 3.5 kb away from its original position, the stability was halved. A similar construct lacking the *parS* locus altogether was completely lost within a few generations (Soberón et al., 2004). Chai and Winans (2005b) moved the *parS* locus of pTiR10 from between the *repA* and *repB* genes to upstream from the *repA* promoter region and downstream from *repC* but close to the operon: the effects on plasmid stability were minor. These various results suggest that proper segregation of *repABC* replicons requires not only the presence of RepA, RepB, and *parS* in the ade-

quate proportions, but also that the *parS* locus be near or within the *repABC* operon for full activity. Further experimentation is needed to clarify the details of this issue.

Work in recent years with new visualization techniques, specially fluorescence *in situ* hybridization (FISH) and protein fusion with the green fluorescent protein and its derivatives, our conception of the organization of the bacterial cell has changed: it is now clear that the bacterial cell components, DNA and proteins, have particular subcellular positions during the cell cycle. The dynamic changes of some of the cellular machinery involved in cell division and genome segregation have been described in exquisite detail (Dye and Shapiro, 2007). The segregation of *repABC* replicons has been analyzed with these novel tools. *S. meliloti* 1021 and *A. tumefaciens* C58 have multipartite genomes: the former has one chromosome and two megaplasmids of the *repABC* family, and the latter contains one circular chromosome, one linear chromosome, and two large plasmids; the linear chromosome and plasmids are *repABC* replicons. Kahng and Shapiro (2003) analyzed the localization of replication regions of these bacteria and they found that all *repABC* operons, and the origin region of the circular chromosomes all reside at the poles of the cell. Moreover, compatible *repABC* operons, coexisting in the same cell share the cell pole, but their positions do not overlap. These observations can be interpreted in two ways: *repABC* replicons are targeted to non-overlapping sites, or alternatively, the segregation machinery of one *repABC* plasmid slightly displaces the machinery of the other by some sort of steric impediment. Recent investigations suggest that plasmid segregation systems do not depend on external host factors, and are consistent with the second interpretation (Ebersbach et al., 2006; Garner et al., 2007). Kahng and Shapiro (2003) also showed that the cell poles of *S. meliloti* and *A. tumefaciens* are not default regions for replication origins, because a broad-host range plasmid, like RK2, localizes at mid- and quarter-cell positions as in *Escherichia coli*. This suggests that *repABC* replicons have a specific replisome at the cell poles.

3. RepC and the origin of replication

RepC encoded by the downstream gene of the *repABC* operon is the limiting factor for replication. Frame-shift mutations, deletions, and insertions in this gene abolish the capacity of plasmids harboring them to replicate (Tabata et al., 1989; Ramirez-Romero et al., 2000).

The RepC sequence does not have homologs with another assigned function. Most *repC* genes are present in *repABC* operons, and in all cases that have been tested these operons confer plasmid maintenance on non-replicating vectors into which they are inserted (Bartosik et al., 1998; Tabata et al., 1989; Nishiguchi et al., 1987; Turner and Young, 1995; Ramirez-Romero et al., 1997; Li and Farrand, 2000; Maclellan et al., 2005; Cevallos et al., 2002). *repC* genes are also found in other genetic contexts: frequently, *repC* genes are associated with a ctRNA gene similar in sequence to those in the *repB–repC* intergenic sequence of *repABC* operons, despite the absence of nearby *repA* and *repB* genes. Nevertheless, regions with these characteristics drive replication on their own (Bartosik et al., 1997; Burgos et al., 1996; Izquierdo et al., 2005; Stiens et al., 2006). At least in *R. leguminosarum* pRL7JI, a *repC* gene has been found in close association with *repA* but not with *repB* or *ctRNA* genes, and they are also able to sustain replication in *Rhizobium* (Pérez-Segura and MAC, unpublished results). A few replicons contain a sequence similar to *repC* but without other elements usually present in *repABC* operons; there is no available information about whether or not they retain replication properties.

We propose here that all replicons containing a *repC* gene able to sustain replication must be considered to be members of the *repC superfamily*. We also propose that this superfamily includes at least three closely related but different families: the *repABC* family, constituted of all replicons containing a *repABC* operon; the *repC* family of replicons containing *ctRNA–repC genes* but without closely associated *repA* or *repB* genes (*i.e.*, pRmeGR4a from *S. meliloti* GR4; Izquierdo et al., 2005); and the *repAC* family which contain a replication system consisting only of a *repA* and a tightly linked *repC* gene (such as the *repAC* genes of pRL7JI described above; Young et al., 2006).

The first evidence that *repC* is the minimal region that is essential and sufficient for replication was obtained by inserting the *repC* gene of *P. versutus* pTAV320 into a suicide vector in two orientations: under the control of a Plac promoter and in the opposite direction. *P. versutus* transconjugants were only obtained with construct in which *repC* was under Plac promoter control (Bartosik et al., 1998). However, the replication properties of this construct were not analyzed.

Recently, a similar construct was obtained using the p42d *repC* gene, and it was able to replicate with a high copy-number when introduced into a *R. etli* strain cured of the parental plasmid. Similar results have been obtained with the *repC* genes from pSymA and pATC58 (Cervantes and MAC unpublished results). These observations also indicate that the *oriV* of *repABC* plasmids map within the RepC coding sequence, and that overexpression of RepC does not cause a runaway phenotype.

Several plasmids contain their origins of replication within the gene encoding their initiatior proteins. They include *Enterococcus faecalis* pAD1, *Sthaphylococcus xylosus* pSX267, *Staphylococcus aureus* pSK41, *Bacillus subtilis* pLS32, and the N15 prophage linear plasmid. The origins of replication of all of these plasmids contain iterons (tandem direct repeats) essential for their function (Francia et al., 2004; Gering et al., 1996; Kwong et al., 2004; Tanaka et al., 2005; Ravin et al., 2003). However, DNA sequence analysis failed to identify iterons in *repC* genes, although there is an A+T-rich region of approximately 160 bp close to the central part of all *repC* genes. A+T-rich regions are frequently found in bacterial origins of replication and are thought to be a duplex unwinding element that facilitates the formation of the replication bubble. Silent mutations that partially eliminate the A+T-rich region in the p42d *repC* gene suppress the replication abilities of the plasmid carrying them implicating this region in the replication of *repABC* plasmids (Cervantes, Bastia, Mohanty, and MAC, manuscript in preparation).

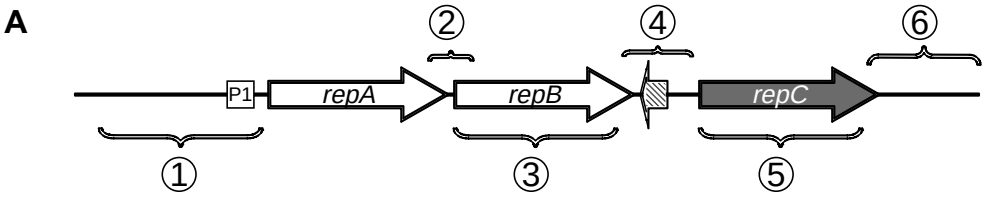
The origin of replication and the mechanism of replication of p42d were identified by 2D gel analysis of replication intermediaries (Huberman et al., 1987). The origin of replication of p42d is within the *repC* gene and close to both its 3′ end and the the A+T-rich region. This plasmid uses a theta mechanism of replication (Cervantes, Bastia, Mohanty, and MAC, manuscript in preparation). However, numerous molecular details, including where RepC binds and how this protein is able to induce replication, remain to be described.

4. Regulation

In low copy-number plasmids, genes were involved in replication and in partitioning map at different *loci* and are controlled by distinct regulatory mechanisms. One feature that makes *repABC* plasmids unique is that all elements required for replication and partitioning are parts of a single transcriptional unit: the *repABC* operon.

The *repABC* operons possess a very intricate regulatory network that ensures production of appropriate amounts of the initiator protein and of partitioning products to have a stable replication and to segregate the newly replicated plasmids into the daughter cells. This network includes both transcriptional and post-transcriptional regulatory mechanisms; some are common to all *repABC* plasmids whereas others are only found in a few members of the family.

The DNA sequences of 78 *repABC* operons of 19 α-proteobacteria genera, from three orders—Rhizobiales, Rhodo-



B

Replicon		sequence	reg.	Replicon		sequence	reg.
Arhi_pRiA4b		**GT**GGTCA**GC**TGACC**AC**	6	Ret_p42e		**GTT**GGCC**GC**GGCC**AAC**	1
Atu_pTiSAKURA	(a,b)	**GTT**GGGA**GC**TCCC**AAC**	2	Ret_p42e		**GTT**GGCC**GC**GGCC**AAC**	3
Atu_pTiR10	(a,b)	**GTT**GGGA**GC**TCCC**AAC**	2	Rleg_pRL7		**GTT**GGTC**GC**GACC**AAC**	6
Atu_pTiC58	(a,b)	**GTT**GGGA**GC**TCCC**AAC**	2	Rleg_pRL8		**GTT**GCCA**GC**CGGA**AAC**	2
Atu_pTiB6S3	(a,b)	**GTT**GGGA**GC**TCCC**AAC**	2	Rleg_pRL10		A**TT**GTCA**GC**TGAC**AA**T	1
Atu_ChrII		**GTT**AGCC**GC**GGC**AAC**	1	Rleg_pRL10		A**TT**GTCA**GC**TGAC**AA**T	4
Bra_pBBta01		**GTT**GTCA**GC**TGAC**AAC**	1	Rleg_pRL11		**GTT**GGCC**GC**GGCC**AAC**	1
Bra_pBBta01		**GTT**GTCA**GC**TGAC**AAC**	2	Rsph_pB	(a,b)	**GTT**GACA**GC**TGTC**AAC**	6
MesBNC_pl1		**GTT**AGCC**GC**GGCT**AAC**	1	Rsph_plD	(a,b)	**GTT**CACA**GC**TGTG**AAC**	6
MesBNC_pl1		**GTT**AGCC**GC**GGCT**AAC**	2	RnuISM	(a,b)	**GTT**CACA**GC**TGTG**AAC**	6
MesBNC_pl1		**GTT**GACT**GC**AGTC**AAC**	3	Ros217a		**GTT**CACA**GC**TGTG**AAC**	6
MesBNC_pl2		**GTT**CTCA**GC**TGAG**AAC**	1	Ros217c		**GTT**TCCG**GC**CGGA**AAC**	5
MesBNC_pl2		**GTT**CTCA**GC**TGAG**AAC**	3	RosHTCC2601c		**GTT**GACA**GC**TGTC**AAC**	6
MesBNC_pl3		**GTT**GTCC**GC**GGAC**AAC**	1	RosHTCC2601c		**GTT**GACA**GC**TGTC**AAC**	4
MesBNC_pl3		**GTT**GTCC**GC**GGAC**AAC**	3	RosTM1035		C**TT**GAAC**GC**GTTC**AA**G	1
Mlot_pMLa		**GTT**GGCC**GC**GGCC**AAC**	3	Rue_pSD25		C**TT**GAAC**GC**GTTC**AA**G	1
Mlot_pMLa	(a,b)	**GTT**GGCC**GC**GGCC**AAC**	1	Rue_pSD25		C**TT**GAAC**GC**GTTC**AA**G	3
NhamX14_pl1		**GTT**GACAC**C**GGTC**AAC**	1	Rue_pSD25		**GT**CCACC**GC**GGTGG**AC**	6
Oant_pOANT01	(a,b)	**GTT**CCCC**GC**GGGG**AAC**	1	Rue_pSD25		**GT**CCACC**GC**GGTGG**AC**	6
Oant_pOANT03		**GTT**GTCC**GC**GGAC**AAC**	1	Smel_pSymA	(a,b,c)	**GTT**CCCC**GC**GGGG**AAC**	1
Oant_pOANT03		**GTT**GTCC**GC**GGAC**AAC**	3	Smel_pSymA		**GTT**CCCC**GC**GGGG**AAC**	1
Pver_pTAV320		CA**T**AACA**GC**TGTT**AA**G	6	Smel_pSymA		**GTT**CCCC**GC**GGGG**AAC**	1
Ret_p42a1		**GTT**GGGA**GC**TCCC**AAC**	2	Smel_pSymA		**GTT**CCCC**GC**GGGG**TAC**	1
Ret_p42a1		**GTT**CTTT**GC**AAAG**AAC**	3	Smel_pSymA		**GTT**CCCC**GC**GGGG**AAC**	1
Ret_p42a2		**GTT**GGGA**GC**TCCC**AAC**	1	Smel_pSymA		**GTT**CCCC**GC**GGGG**AA**T	1
Ret_p42b	(a,b)	**GTT**TACT**GC**GGTA**AAC**	2	Smel_pSymB		**GTT**GTCC**GC**GGAC**AAC**	2
Ret_p42c		**GTT**GTCA**GC**TGAC**AAT**	1	Smed_pSMED01		**GTT**GTCC**GC**GGAC**AAC**	1
Ret_p42c		**GTT**GTCA**GC**TGAC**AAT**	4	SulfitNAS141b		A**TT**GTCA**GC**TGAC**AA**T	6
Ret_p42d		**GT**GGTCA**GC**TGACC**AC**	6	SulfitEE36		A**TT**GTCA**GC**TGAC**AA**T	4

Consensus palindromic sequence: **GTT**NNNN**GC**NNNN**AAC**

Fig. 2. Relative position of *parS*-sites in *repABC* operons. (A) White arrows represent *repA* and *repB* genes, and the gray arrow the *repC* gene. The hatched arrow indicates the position of the antisense RNA gene. The square box shows the promoter region. The brackets show the regions containing possible *parS*-sites and the plasmids that possess a *parS*-site in these specific regions. (B) List of the *repABC* operons indicating the position and sequence of their putative *parS*-sites. Letters in bold-type indicates conserved positions. The first letters of the name indicate the bacterial species followed by the replicon name. Arhi, *Agrobacterium rhizogenes*; Atu, *Agrobacterium tumefaciens*; Bra, *Bradyrhizobium* sp.; Mes, *Mesorhizobium* sp.; Mlot, *Mesorhizobium loti*; Nham, *Nitrobacter hamburgensis*; Oant, *Ochrobactrum anthropi*; Pver, *Paracoccus versutus*; Ret, *Rhizobium etli*; Rleg, *Rhizobium leguminosarum*; Rsph, *Rhodobacter sphaeroides*; Rnu, *Roseovarius nubinhibens*; Ros, *Roseovarius* sp.; Rue, *Ruegeria* sp.; Smel, *Sinorhizobium meliloti*; Smed, *Sinorhizobium medicae*; Sulfit, *Sulfitobacter* sp. Letters (a, b) or (a, b, c) after plasmid name indicate that this *parS*-site sequence is present in that region two or three times, respectively. If one plasmid is repeated two or more times in the plasmid list indicates that that plasmid contains two or more *parS*-site sequences in different regions of the operon. Plasmids not listed here do not contain a canonical *par*-site.

bacteriales, and Rhodospirillales—have been deposited in GenBank. Most of these sequences have been obtained by recent genome projects, and consequently it has been possible to identify the genes flanking each of the *repABC* operons: 10 are divergently transcribed from a *traI–trb* operon, and the others do not have neighboring genes in common (Fig. 2). The *repABC* operons for which the regulation has been studied in detail include some of those linked and not linked to *traI–trb* operons. The *traI–trb*-linked operons, for example the *repABC* operon of *Agrobacterium* plasmid pTiC58, have a very sophisticated regulation; their plasmid copy-number may be influenced both positively and negatively by external factors including quorum sensing and host-released factors. In contrast, those not linked to a *traI–trb* operon, like the *R. etli* p42d *repABC* operon, seem to have a much simpler regulation.

4.1. Plasmid copy-number control

Agrobacterium tumefaciens is the etiological agent of crown gall disease and induces tumors in dicotyledonous plants. The tumors produce large quantities of opines, a bacterial nutrient that can be only used by *Agrobacterium*. A crucial step in the *Agrobacterium*–plant interaction is the transfer from the bacteria to the plant nucleus of a DNA fragment–the T-DNA–that carries genes involved in tumor induction and in opines synthesis. The T-DNA is found in large conjugative *repABC* plasmids (pTi), that also carry genes coding for opine transporters and opine catabolic enzymes (Zhu et al., 2000). The replication and segregation properties of two Ti plasmids have been studied in detail: pTiC58 and pTiR10.

Agrobacterium Ti plasmids modulate their plasmid copy-number and their conjugative transfer in response to *Agrobacterium* population density and to the presence of chemical signals generated by the host plant. pTi *repABC* operons, involved in plasmid replication and segregation, are transcribed in the opposite direction to *traI–trb* genes. The intergenic region between these two gene sets contains promoters and DNA sequences involved in their transcriptional regulation (Pappas and Winans, 2003a; Li and Farrand, 2000). The pTiR10 *repABC–traI–trb* region is currently the best characterized. The pTiR10 *repABC* operon contains four promoters: P1, P2, P3, and P4, with P4 being closest to and P1 the furthest from *repA* (Fig. 1). All of them resemble sigma-70 consensus sequences suggesting that they are transcribed by the major *Agrobacterium* sigma factor SigA (Pappas and Winans, 2003b). The *repABC* P1 to P3 promoter activities and the transcription of the *traI–trb* operon are influenced positively by TraR, a LuxR homolog, and a central player in the *Agrobacterium* quorum-sensing regulatory system. TraR dimers exert their effects in the presence of *N*-3-oxooctanoyl-L-homoserine lactone (3-*O*-C_8-AHL), an autoinducer synthesized by TraI, and this involves their binding to *tra* boxes II and III. The *tra* box II is immediately upstream from the P1 −35 element and the *tra* box III is upstream from the P3 −35 element (Pappas and Winans, 2003b). The exogenous addition of 3-*O*-C_8-AHL activates P1, P2, and P3, and produces a concomitant increase in the plasmid copy-number, probably raising the intracellular concentration of RepC, the initiator protein. This increase in the plasmid copy-number was associated with an increase in tumorigenesis (Pappas and Winans, 2003a).

In contrast to the behavior of pTiR10, two other *Agrobacterium* R10 *repABC* replicons which are not linked to *traI–trb* operons—the linear chromosome and pAtR10—do not change their copy-number in presence of the autoinducer (Pappas and Winans, 2003a). The same has been described for the *Agrobacterium* strain C58: in presence of the autoinducer the copy-number of pTiC58 increases but that of pAtC58, a second *repABC* plasmid in this strain, does not (Li and Farrand, 2000; Lee et al., 2006).

Transcription of the TraR gene is regulated by a complex cascade of events: expression of this gene is repressed by an opine-responsive regulator (AccR in the case of nopaline-type Ti plasmids and OccR in octopine-type plasmids). In the absence of opines, AccR (or OccR) represses expression of TraR, but the inhibition of *traR* transcription is relieved in the presence of the corresponding opine (Piper et al., 1999; von Bodman et al., 1992). In mutant Ti derivatives lacking *accR* (pTiC58Δ*accR*), the conjugative system is expressed constitutively (Tra^c) and the copy-number is moderately elevated. Deletion mutants of *traM,* a gene coding an AccR antiactivator, are also Tra^c and have high plasmid copy-numbers (Li and Farrand, 2000).

Other gene regulators contribute to controlling pTi plasmid-copy-number: the transmembrane sensor kinase VirA, and the response regulator VirG constitute a two-component system that coordinates Ti *vir* gene expression, and modulates the transcriptional activity of the *repABC* operon both positively and negatively. In acidic extracellular media, VirA can detect various factors, including phenolic compounds and some monosaccharides, released by the host plant (Zhu et al., 2000).

Phospho-VirG, the active form of the protein, binds to DNA sequences called *vir* boxes, present in the upstream promoter regions of the genes controlled by this protein. Addition of acetosyringone, a phenolic compound that is widely used as inducer of this system, provokes an increase in plasmid copy-number of pTiC58 and pTiR10; surprisingly, it also increases the copy-number of pAtC58, a cryptic plasmid also present in *Agrobacterium* strain C58. This effect is abolished in mutants lacking *virA* or *virG*. Molecular and genetic analysis of pTiR10 identified a *vir* box in between the P3 and P4 promoters. Addition of acetosyringone enhances only P4 promoter activity (Cho and Winans, 2005); P4 is able to sustain plasmid replication in the absence of the autoinducer and of factors released by the host (Pappas and Winans, 2003a).

Salicylic acid (SA), a plant metabolite crucial for the plant response to pathogens, reduces the ability of *Agrobacterium* to infect by repressing the transcription of all *vir* genes, *tra* genes and of the *repABC* operon (Yuan et al., 2007; Anand et al., 2008). It has been recently suggested that this effect is mediated by inhibiting the transcription of *virA* and *virG* genes. Unexpectedly, SA activates the expression of the *attKLM* operon, which encodes enzymes participating in the degradation of 3-*O*-C_8-AHL (Yuan et al., 2007). In contrast, Anand and co-workers (2008) propose that the observed *vir*, *tra,* and *repABC* down-regulation is a consequence of a direct competition between SA

and the phenolic inducers. This controversy needs to be resolved soon.

Indole-3-acetyc acid (IAA), a plant hormone involved in a wide variety of functions, can shut down *virA* and *virG* expression, and probably also has a role in modulating Ti plasmid copy-number (Liu and Nester, 2006).

These observations suggest that in free life *Agrobacterium* maintains Ti plasmids at one copy per chromosome using P4 promoter activity only. However, on the surface of the plant tumor, *Agrobacterium* increases its population density: consequently, the bacteria detect their own population's quorum-sensing signals and the factors released by the host plant, and as a result the Ti plasmid conjugation system is induced and copy-number is increased. The increase in the gene dose due to the change in pTi plasmid copy-number enhances the tumorigenic capacities of *Agrobacterium*, providing opines more abundantly for population maintenance. Once the tumor is established, the SA produced by the plant as part of its response to bacterial infection will shut down *vir* genes and will limit the spread of the infection. Thus, the life cycle of *Agrobacterium* can be satisfactorily explained by gene regulation (Li and Farrand, 2000; Cho and Winans, 2005).

Rhizobium and its close relatives (collectively known as *Rhizobia*) establish symbiotic associations with legume plants. Rhizobia infect the roots of these plants and induce in them, nitrogen-fixing nodules. The relationship begins when a *Rhizobia* species perceives chemical signals, usually a flavonoid, excreted by the roots of its legume host; they then initiate the synthesis of specific lipochito-oligosaccharides, named Nod factors. These factors induce the formation of the nodule meristem on the legume root and allow the invasion of root cells by *Rhizobia* (Geurts et al., 2005; Long, 2001). Most of the genes involved in these interactions are encoded in large *repABC* plasmids.

Two *Rhizobium* symbiotic plasmids, pRL1JI and pNGR234a, share various properties with Ti plasmids: their *repABC* operons and *traI–trb* are adjacent and transcribed in opposite directions. The conjugative and replication properties of pRL1JI respond to quorum sensing, and pNGR234a responds to factors released by the host.

A transcriptional analysis of the *Rhizobium* sp. NGR234 symbiotic plasmid found that the addition of daidzein, a flavonoid that induces *nod* genes in this strain, also induces the transcription of the *repABC* operon. The induction mechanism is obscure: the *repABC* promoter region does not contain nod boxes, sites at which NodD, the transcriptional activator responding to flavonoids, binds to DNA (Perret et al., 1999).

The transfer and replication genes of pRL1JI are transcriptionally co-regulated by a cascade of three different quorum-sensing regulators—TraR, BisR, and CinR—and two different autoinducers including 3-*O*-C_8-AHL synthesized by TraI and *N*-(3-hydroxy-7-*cis*-tetradecenoyl) homoserine lactone (3-OH-$C_{1:14}$-HSL) made by CinI. TraR and BisR are encoded by plasmid-borne genes that lie downstream from the *traI–trb* operon, on the same DNA strand. As in pTi plasmids, the pRL1JI *repABC* operon is transcribed in opposite direction to the *traI–trb* operon. However, *cinI* and *cinR* are chromosomal and linked. TraR activates the transcription of *repABC* and *traI–trb* operons as a population density response to the 3-*O*-C_8-AHL synthesized by TraI, in a similar way as described above for the *Agrobacterium* pTi plasmids. Expression of TraR depends on BisR and 3-OH-$C_{1:14}$-HSL; however, BisR also represses *cinI*. Thus, a strain carrying pRLJ1 synthesizes basal amounts of 3-OH-$C_{1:14}$-HSL insufficient to trigger conjugation or enhance plasmid copy-number. In contrast, strains lacking pRLJ1 make considerable amounts of 3-OH-$C_{1:14}$-HSL because the chromosomal *cinI* is repressed by BisR. When a strain containing pRL1JI (a donor strain) encounters a strain without pRL1JI (a recipient strain) they form a match: the recipient strain makes sufficient quantities of 3-OH-$C_{1:14}$-HSL to induce TraR through BisR. In the presence of 3-*O*-C_8-AHL, TraR induces an increase in plasmid copy-number in the donor and the conjugative transfer of the plasmid. The pRLJ1 *repABC* operon contains two promoters and one of them is constitutive, so the plasmid's replication does not depend only on quorum sensing (McAnulla et al., 2007).

Not all symbiotic plasmids contain a *traI–trb* operon or are regulated by quorum sensing: the *R. etli* symbiotic plasmid (p42d) contains a *traI* pesudogene upstream from the *repABC* operon and on the complementary strand, but lacks the *trb* genes. The intergenic region between the *traI* pseudogene and the *repABC* operon is 388 bp long, which is similar in length to the analogous regions in Ti plasmids, but without a trace of *tra* boxes. Molecular and a genetic analysis showed that this *repABC* operon contains only one promoter located 52 bp upstream from the *repA* initiation codon, and activated by SigA, the *Rhizobium* housekeeping sigma factor (Ramirez-Romero et al., 2006) (Fig. 1). This promoter can maintain stable replication without requiring any other sequence located upstream from its −35 element (Ramirez-Romero et al., 2001). These findings suggest that the transcription of the p42d r*epABC* promoter is not influenced by external signals. Probably, only *repABC* plasmids that contain a *repABC* operon transcribed divergently from a *traI–trb* operon are able to detect quorum-sensing signals and increase their plasmid copy-numbers as a consequence.

Detailed analyses of p42d and pTiR10 revealed that RepA and RepB are negative regulators of *repABC* transcription, as expected for plasmids containing a type Ia partitioning system. RepA binds to an operator region a few bases downstream from the *repABC* operon, stimulated by RepB. RepA, a weak ATPase, binds the operator region in the presence of ATP or ADP. The core operator region consists of an imperfect palindrome sequence, and deletion of this region results in constitutive expression (Pappas and Winans, 2003b; Ramirez-Romero et al., 2001).

The *parS* locus also plays an important role in the negative transcriptional regulation of *repABC* operons: constructs containing the operator region, the *parS* locus, and *repA* and *repB* genes repress their own transcription more strongly than similar constructs lacking the *parS* locus. This finding indicates that *repABC* operons can repress their expression through loops interlinking a RepA–RepB complex with the operator and *parS* of the same plasmid or between two copies of the plasmid (Chai and Winans, 2005b; Pérez-Oseguera and MAC unpublished results).

4.2. The promoter regions of repABC operons

One approach to analyzing elements common to different *repABC* operons is to identify conserved DNA sequences in the first 400 bp upstream from the *repA* initiation codon in each. To do this, we used the MEME/MAST system (http://meme.sdsc.edu/meme/), which detects highly conserved regions without utilizing a DNA multialigment tool (Bailey and Elkan, 1994; Bailey and Gribskov, 1998). We first divided *repABC* operons into two groups: those linked and those unlinked to *traI–trb* operons.

Of the 77 *repABC* sequences available in GenBank, only 10 are linked to a *traI–trb* operon: six belongs to *Agrobacterium* tumor- and-root nodule-inducing plasmids; two more belongs to *Rhizobium* plasmids, and one to a *Nitrobacter hamburgensis* plasmid. MEME/MAST analysis identified five different motifs that are present in the 5′ regions of these operons: a stretch of 31 bp that includes the operator region of pTiR10; a segment of 43 bp containing the −35 hexameric element of the pTiR10 *repABC* promoter 4; another 43 bp segment harboring *tra* box III and the complete P3 promoter of the pTiR10 *repABC* operon; a 50-bp stretch including the *tra* box II; and a 48-bp stretch including the Shine–Dalgarno sequence and the first codons of the *traI* gene (see Fig. 3A). Only pTi plasmids contain all five motifs and other plasmids contain only one or two of these motifs but in all cases including the P4 promoter described in pTiR10 *repABC* operon.

The 67 *repABC* operons with no association with *traI–trb* operon were similarly analyzed. Thirty-nine operons did not show significant similarities, but the others share at least one of two motifs: a region of 56 bp with sequences highly similar to the p42d *repABC* promoter region; and a motif harboring *parS* sequences similar to those described for pSymA (Figs. 1 and 3B). Surprisingly, when the complete collection of *repABC* promoter regions were analyzed as one group, it became clear that some of the *repABC* promoter regions not linked to the *traI–trb* operon contain some elements typical of *repABC* operons associated with *traI–trb* operons, for example *tra* boxes. These observations suggest that these sequences share a common ancestor.

4.3. Postranscriptional regulation

A universal characteristic of *repABC* operons is the presence of a large and conserved intergenic region between the *repB* and *repC* genes, which if introduced in *trans* is incompatible with its parental plasmid (Tabata et al., 1989; Li and Farrand, 2000; Ramirez-Romero et al., 2000). Most of these regions are between 150 and 250 bp long, but a few members of the family have extremely large intergenic regions (663 bp in *Oligotropha carboxidovorans* pHCG3 and 611 bp in one of the *repABC* operons in *R. etli* p42a). The 5′ end of the intergenic region, close to *repB*, tends to be more conserved than the 3′ end. Genetic analyses of p42d, pSymA, pSymB, and pTiR10 identified a gene coding a 55–59 nt small non-translated RNA (ctRNA) in the most conserved part of the *repB–repC* intergenic region, and in the opposite orientation to the *repABC* operon (Venkova-Canova et al., 2004; MacLellan et al., 2005; Chai and Winans, 2005a). Sequences highly similar to the ctRNA coding region are present in almost all *repABC* operons. Promoters of the genes encoding these small antisense RNAs are conserved in sequence and show similarities with the *E. coli* σ70 consensus promoter (Venkova-Canova et al.,

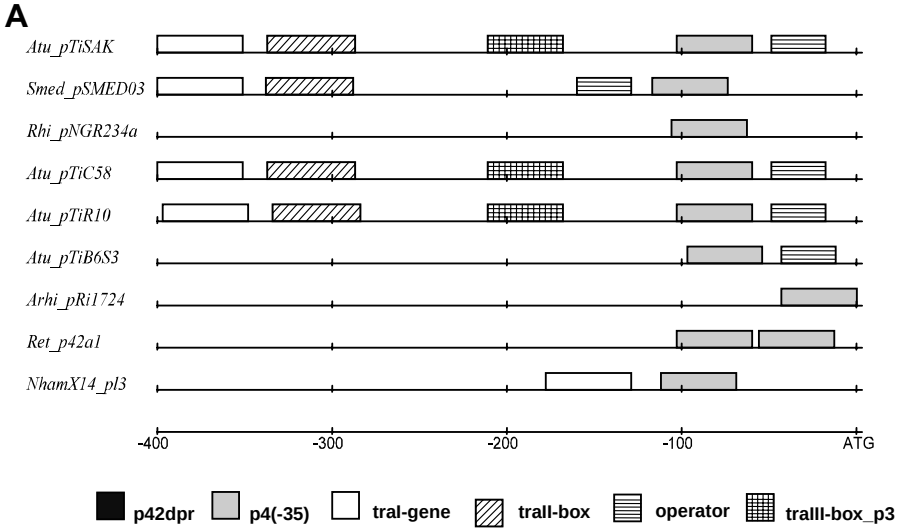


Fig. 3. Conserved motifs in *repABC* promoter regions. Schematic representation of the 400 bp upstream from the *repA* initiation codon of sequences sharing at least one of the following motifs identified by the MEME/MAST algorithm: black boxes show the positions of a 56-bp motif containing sequences similar to the p42d *repABC* operon; gray boxes indicate the position of a 43-bp motif containing sequences similar to the −35 hexameric element of the P4 pTiR10 promoter; white boxes indicate the beginning of the *traI* gene; diagonally hatched boxes indicate the position of a 50-bp motif containing a *tra* boxII; horizontally hatched boxes show the localization of a 31-bp motif including sequences similar to the operator region of pTiR10 *repA* gene; cross-hatched boxes indicate the positions of a 43-bp motif containing *tra* III boxes. (A) Promoter regions of *repABC* operons linked to *traI–trb* operons. (B) Promoter regions of *repABC* operons not linked to *traI–trb* operons.

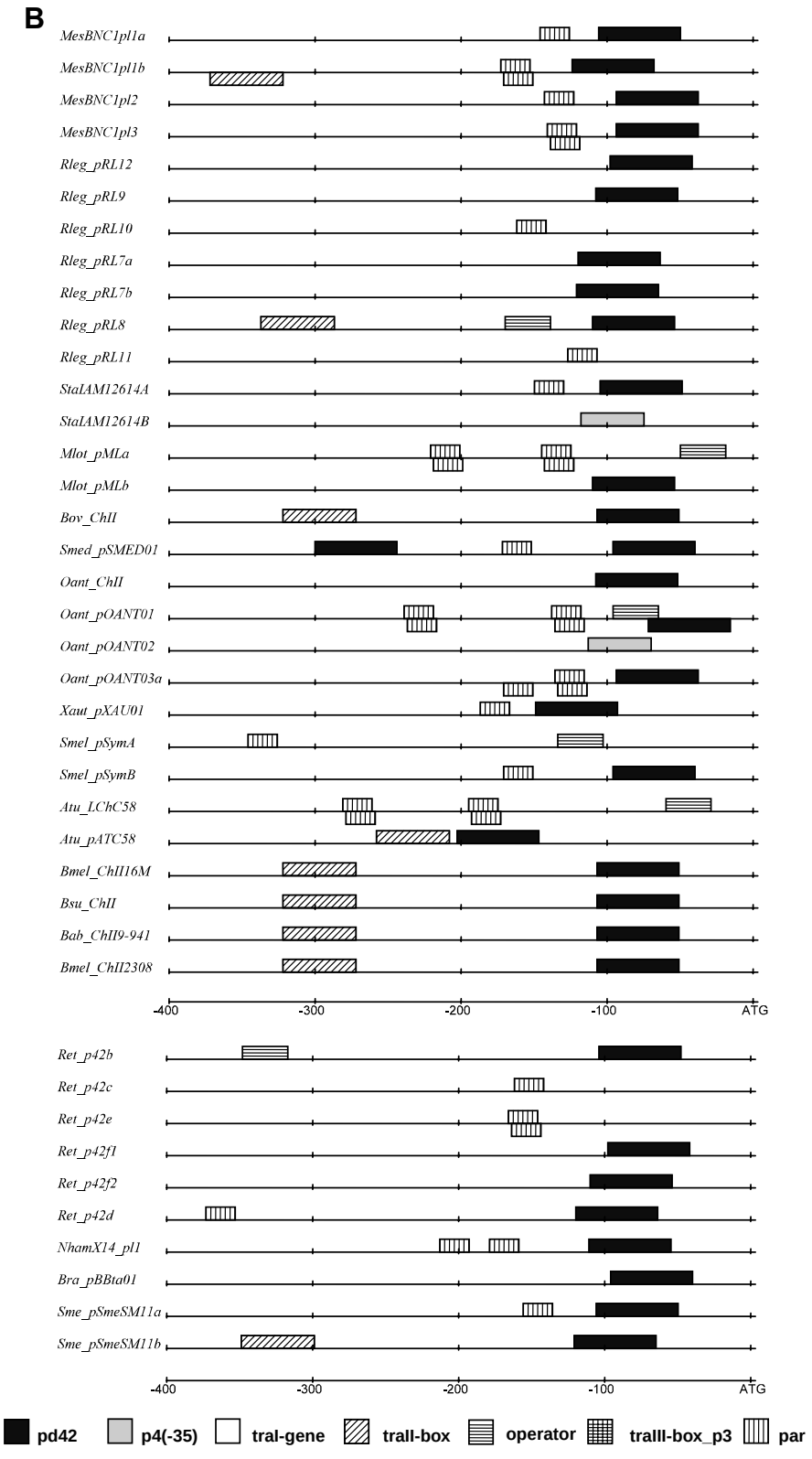


Fig. 3 (*continued*)

2004; MacLellan et al., 2005; Chai and Winans, 2005a). In fifty of the *repABC* operons belonging to the Rhizobiales order, the −10 and −35 hexameric elements and also in the spacer region of the ctRNA promoter sequence are conserved. The consensus ctRNA promoter is TTGACAGT GATT CGTGGAAATGTGATTCT (−10 and −35 hexameric elements are underlined).

The predicted secondary structure of the ctRNAs is a single stem-loop with a U-rich tail at the 3′ end, resembling transcription intrinsic terminators. The sequences and lengths of the stem-loops vary, and so do the positions of internal bulges and bubbles that may be important for RNA stability and melting with the target RNA (Brantl, 2007). Only some of these RNAs contain a 5′ YURN motif in their loops that in other plasmid systems facilitates RNA loop-loop interactions (Weaver, 2007). The ability of these RNA structures to terminate transcription has been confirmed experimentally (Venkova-Canova et al., 2004; MacLellan et al., 2005; Chai and Winans, 2005a).

Intergenic *repB–repC* regions containing mutations that abolish ctRNA gene promoter activity are unable to exert incompatibility against their cognate plasmids, when provided *in trans*. This indicates that the ctRNA is a *trans* incompatibility factor, the only element mediating incompatibility that maps to this region (Venkova-Canova et al., 2004; MacLellan et al., 2005; Chai and Winans, 2005a).

The ctRNA gene is upstream from *repC*, the gene encoding the initiator protein, and the ctRNA itself is a *trans* incompatibility factor. Therefore, ctRNA may modulate RepC levels and thereby plasmid copy-number. Indeed, there is evidence that the antisense RNA downregulates *repC* expression. Recently, Chai and Winans (2005a) described an attractive model that mechanistically explains some of the observations. They proposed that the *repABC* mRNA has two alternative secondary structures in which the ctRNA is or is not paired with its complementary section on the *repABC* mRNA. In the absence of the ctRNA, the section of the *repABC* mRNA including the *repB–repC* intergenic region folds in a large stem-loop structure with only the *repC* initiation codon and the predicted *repC* SD sequence being single stranded; in this conformation translation of *repC* is allowed. In presence of the ctRNA, the *repABC* mRNA sequence upstream from and close to the *repC* initiation codon forms a stem-loop structure occluding the Shine–Dalgarno of *repC*, and thereby suppressing *repC* translation. S1 nuclease protection assays were used to assess this model by quantifying the *repABC* mRNA. In presence of the ctRNA, there was substantially less of the transcript corresponding to the segment downstream from the putative structure occluding the SD *repC* sequence than of the upstream segment; this was not the case it its absence. Therefore, this occluding structure is formed in presence of the ctRNA, and it is a transcription attenuator (Fig. 4) (Chai and Winans, 2005a). In agreement with this proposal, pTiR10 constructs carrying suboptimal ctRNA promoters replicate with a moderately increased plasmid copy-number. Nevertheless, mini *repABC* replicons derived from pTiR10 or p42d are unable to express the ctRNA, or carrying truncated versions of the ctRNA gene plus *repC*, are unable to replicate in their native hosts (Venkova-Canova et al., 2004; Chai and Winans, 2005a). It has been proposed that in absence of ctRNA control, RepC is overexpressed resulting in a lethal runaway phenotype. However, as described above, constructs containing *repC*, with an *E. coli* consensus SD sequence under the control of a constitutive promoter, replicate with a higher copy-number without killing the cell. Possibly, the ctRNA plays a role in the replication initiation complex. To complicate the picture further, work with translational fusions suggests that p42d encodes a small peptide, immediately upstream from *repC*, that is translationally coupled to *repC*. Moreover, mutations that destroy the initiation codon of this putative peptide abolish replication. Peptides in the same relative position but with different sequences are present in other *repABC* plasmids (Venkova-Canova et al., 2004, Venkova-Canova unpublished results).

To understand better how ctRNA controls plasmid replication, it would be valuable to determine the structures of the ctRNA, of the mRNA encoded in the *repB–repC* intergenic region, and of the ctRNA–mRNA hybrid. It would also be useful to determine the half-lives of ctRNA, and of the mRNA in the presence and absence of the ctRNA. The role, if any, of peptides translationally coupled to RepC also needs to be assessed.

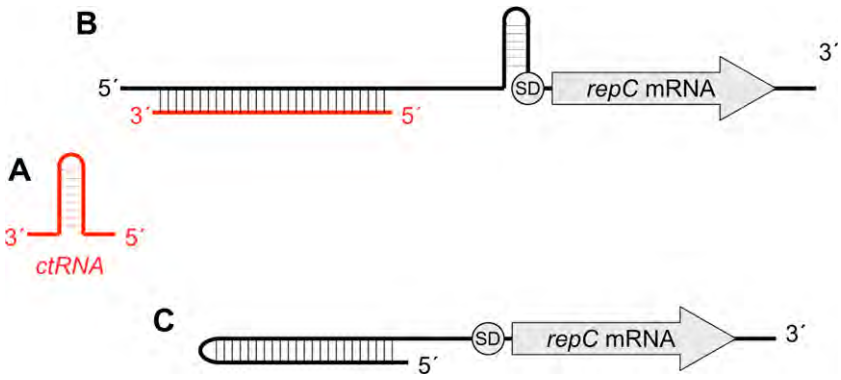


Fig. 4. Model proposed by Cho and Winans, 2005 describing the mechanism of action of the ctRNA in the control of replication of *repABC* plasmids. (A) Representation of the secondary structure of the ctRNA. (B) Schematic representation of the ctRNA–repABC mRNA heteroduplex. Note that the *repC* Shine–Dalgarno sequence is occluded in a stem-loop structure, attenuating *repC* translation. (C) In the absence of the ctRNA, the *repABC* mRNA adopts an alternative folding, in which the *repC* Shine–Dalgarno sequence is fully accessible to ribosomes.

5. Genes influencing plasmid incompatibility and plasmid stability encoded elsewhere than the *repABC* operon

p42d and pTi-SAKURA contain genes that influence plasmid stability and incompatibility properties but are not encoded in the *repABC* operon although they do map nearby. The influence of these genes is determined by their close proximity to the *repABC* operons. Gene order and gene content are not conserved among *repABC* replicons, so the effect of these genes is probably restricted to these two plasmids. Quintero and co-workers (2002) showed that a construct carrying the p42d basic replicon can replicate in a *R. etli* derivative lacking the parental plasmid. However, when this construct is introduced into a *R. etli* strain carrying the parental plasmid, the incoming plasmid forms a stable co-integrate with p42d rather than displacing it. The incoming plasmid is able to exert incompatibility against the parental plasmid in two conditions: in a *R. etli recA* derivative, and when the incoming plasmid contains, in addition to the p42d basic replicon, the action site (*ris*) of a site-specific recombinase (RinQ) in the same relative orientation as in the parental plasmid. RinQ and its action site are downstream from *repC* and are not incompatibility determinants by their own. They resolve co-integrates through the duplicated *ris*-site product of the co-integrate formation and promote the displacement of one of the resulting plasmids.

Recently, Yamamoto and coauthors (2007) described the second example: they found two genes, *tiorf24* and *tiorf25*, downstream from *repC* in pTi-SAKURA, that enhance plasmid stability and increase the incompatibility properties of constructs containing only the pTi-SAKURA *repABC* operon. These genes are organized as a single transcription unit, and they enhance plasmid stability if provided in *cis*. *tiorf24* is an orphan gene and *tiorf25* shows significant similarities with *stbB* of *Pseudomonas syringae pv tomato*, a gene which is part of a gene cluster involved in plasmid stability. These observations indicate that neighboring genes can influence the replication and partitioning properties of a *repABC* operon. Gene associations can change during plasmid evolution and some of them are likely to modulate the performance of the *repABC* operons. Presumably, only gene associations that enhance the stability of a *repABC* replicon will persist in evolution.

6. Incompatibility and origins of *repABC* plasmids

Plasmids unable to coexist in the same cell line are considered to be incompatible, and this phenomenon is generally the consequence of some degree of interference between their replication and/or partitioning mechanisms (Novick, 1987). Incompatibility restricts the ability of a bacterial cell to acquire similar plasmids, but favors the presence of a wider selection of plasmids. Plasmid incompatibility plays makes a large contribution to shaping bacterial evolution.

An intriguing characteristic of some of the sequenced α-proteobacteria is the abundance of *repABC* replicons (Table 1). In some species, all the plasmids present belong to the *repABC* plasmid family. The largest number of *repABC* replicons known to be present in a single strain is in *R. etli* CFN42 and *R. leguminosarum* 3841, two members of this bacterial family, each with six replicons. Moreover, some of these plasmids contain two *repABC* operons. This is very common in the bacterial order Rhizobiales, suggesting that *repABC* plasmids may have originated from this bacterial order. The frequent association between *traI–trb* and *repABC* operons suggests an ancestral relationship, supported by the observation that some *repABC* operons not linked to *traI–trb* retain motifs commonly found in *repABC* operons linked to *traI–trb*.

The presence of several *repABC* replicons in the same strain implies that this plasmid family includes several incompatibility groups; this is of significance concerning the origins and evolution of *repABC* replicons, and in particular, the appearance of new incompatibility groups within the same plasmid family. Two different models have been proposed. The first involves the formation of a bireplicon containing two replication/partition regions; this bireplicon could be the product of gene duplication, of co-integration between an incoming plasmid and a resident plasmid of the same incompatibility group, or of a failure of plasmid dimer resolution. Once a bireplicon is formed, one copy of the replication/partition region can accumulate mutations, and even lose its functionality without jeopardizing maintenance of the plasmid. Alternatively, such mutations could produce new variants of replication genes with new incompatibility properties; a new incompatibility group would then be generated when the bireplicon resolves into two plasmids, and both are able to persist without interference with the other (modified from Sykora, 1992). The second model is based on horizontal gene transfer: two strains containing the same plasmid evolve independently, allowing divergent evolution of plasmid determinants involved in replication and/or partition. This leads to the generation of different incompatibility groups. Then, a plasmid of one of these strains is mobilized (horizontal transfer) to another strain already containing a plasmid with the same common ancestor, and due to the divergent evolution, both plasmids can then coexist without interference (Cevallos et al., 2002).

A method to discriminate between the two models is to construct and analyze a phylogenetic tree. If plasmids of the same plasmid family present in one strain are derived according to the first model, they will share a common clade. If plasmid evolution conformed to the second model, plasmids of the same family present in the same strain will be in different clusters of the phylogenetic tree (Cevallos et al., 2002).

Three phylogenetic trees (Fig. 5A–C) were constructed using the RepA, RepB, and RepC protein sequences. The patterns of these trees are consistent with an intricate evolutionary story that involves three forces shaping the *repABC* plasmid evolution: horizontal transfer, gene duplication and divergence, and intra-operon recombination. The three phylogenetic trees constructed here are different; nevertheless, they share some similarities. The sequences included in our analysis represent *repABC* operons of three bacterial orders: rhizobiales, rhodospirillales, and rhodobacterales, and they are completely

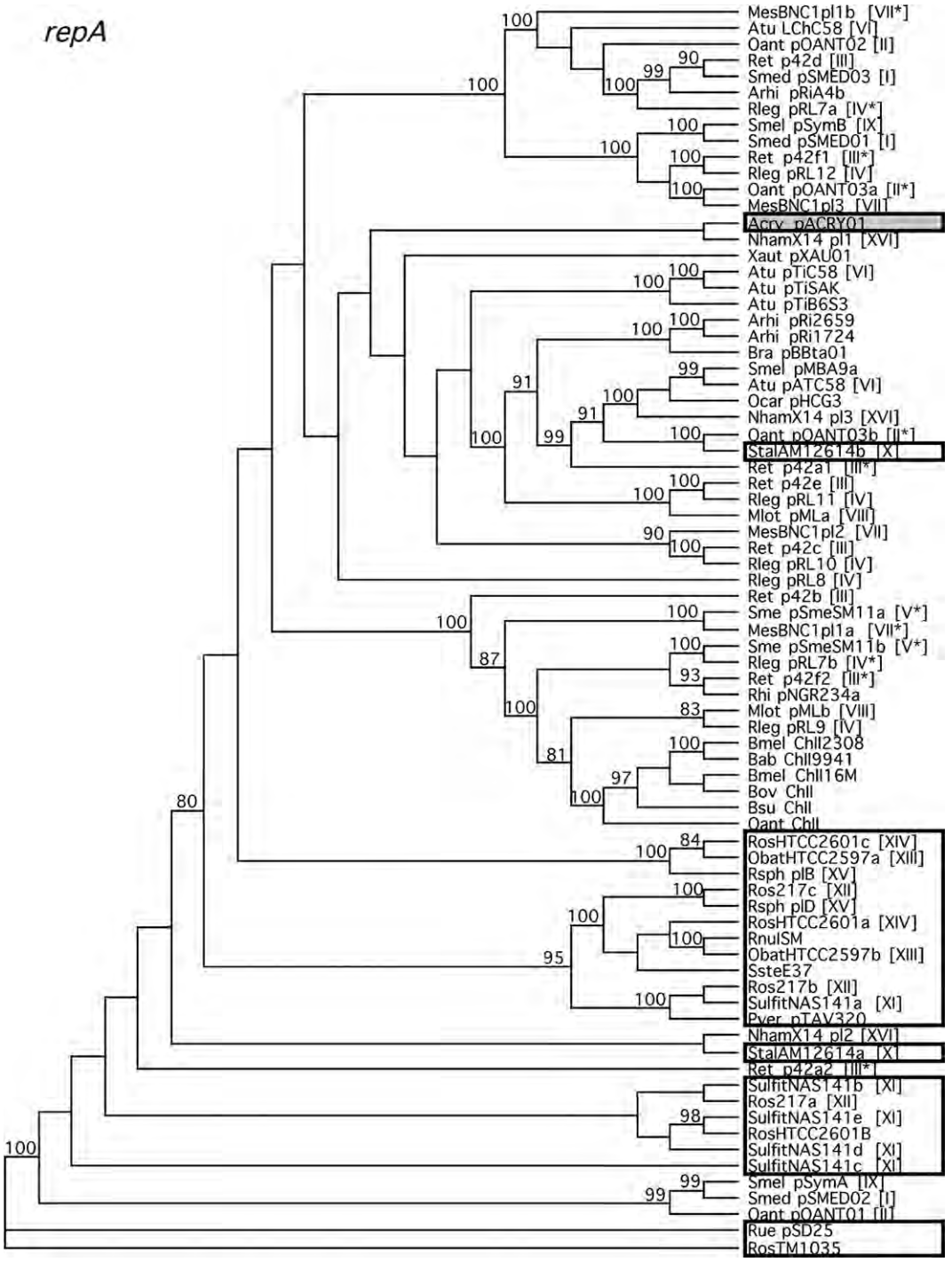


Fig. 5. Unrooted neighbor-joining phylogenetic trees constructed with: (A) RepA, (B) RepB, and (C) RepC protein sequences. Numbers in tree nodes indicate bootstrap values higher than 79% (1000 replicas). The *repABC* operons sharing the same Roman numeral are found in the same bacterial strain. The *repABC* operons marked with asterisks are carried on the same replicon. The first letters of the name indicate the bacterial species followed by the replicon name. Smed, *Sinorhizobium medicae*; Oant, *Ochrobactrum anthropi*; Arhi, *Agrobacterium rhizogenes*; Ret, *Rhizobium etli*; Rleg, *Rhizobium leguminosarum*; Atu, *Agrobacterium tumefaciens*; Smel, *Sinorhizobium meliloti*; Mes, *Mesorhizobium* sp.; Mlot, *Mesorhizobium loti*; Bov, *Brucella ovis*; Bab, *Brucella abortus*; Bsu, *Brucella suis*, Bmel, *Brucella melitensis*; Sta, *Sttapia aggregata*; Sulfit, *Sulfitobacter* sp.; Ros, *Roseovarius* sp.; Rue, *Ruegeria* sp.; Acry, *Acidiphilum cryptum*, Ocar, *Oligotropha carboxidovorans*; Obat, *Oceanicola batensis*; Rnu, *Roseovarius nubinhibens*; Sste, *Sagittula stellata*, Rsph, *Rhodobacter sphaeroides*; Pver, *Paracoccus versutus*, Nham, *Nitrobacter hamburgensis*; Bra, *Bradyrhizobium* sp.; Xaunt, *Xantobacter autotrophicus*, Replicons boxed in gray belong to the bacterial order Rhodospirillales, those boxed in white belong to the Rhodobacterales, and the other belong to the bacterial order Rhizobiales.

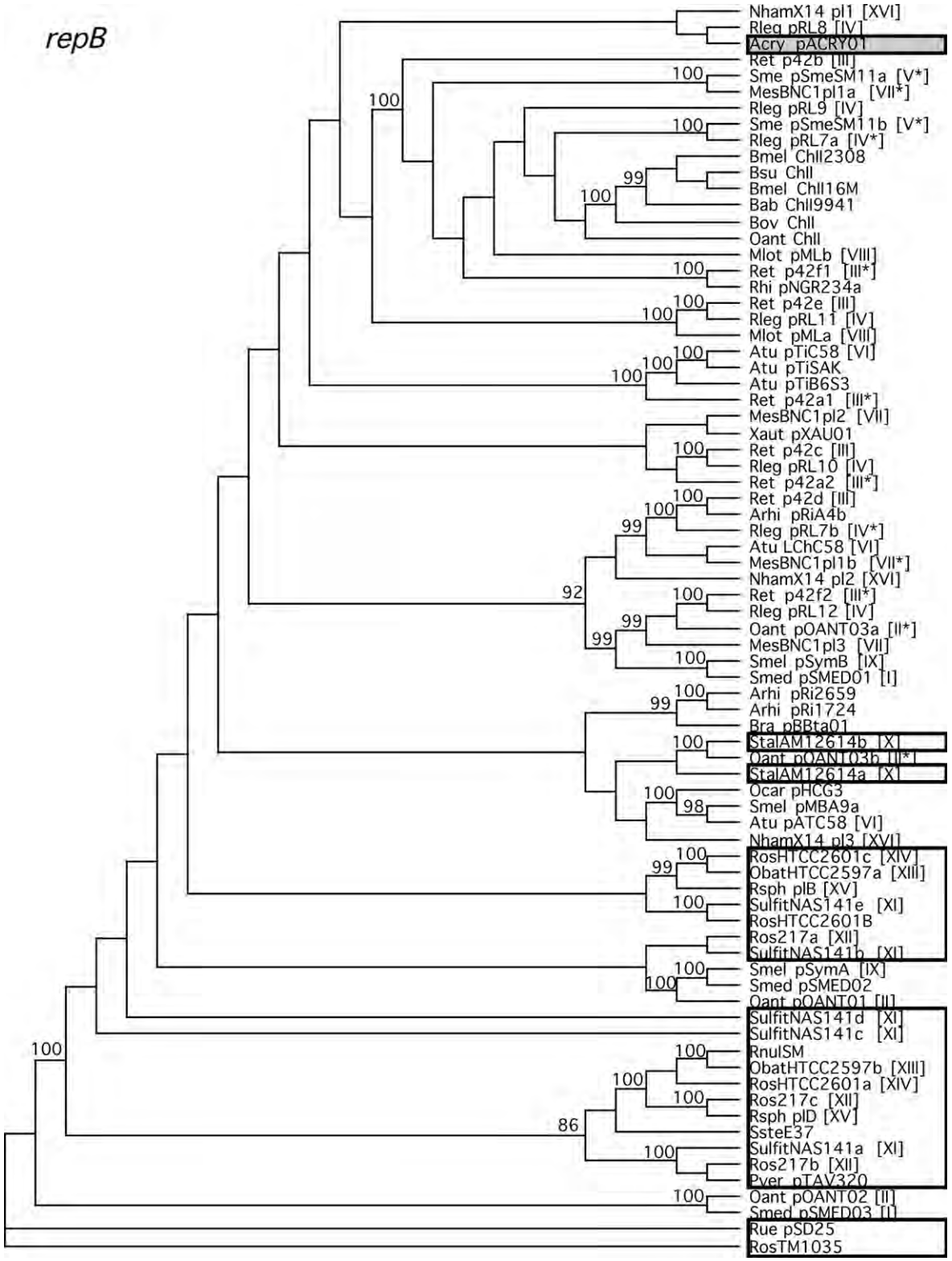


Fig. 5 (*continued*)

scattered throughout all trees, indicating that bacterial strains of these orders freely interchange genetic material. It is very common to find proteins, especially in RepA and RepB trees, present in the same strain and even encoded

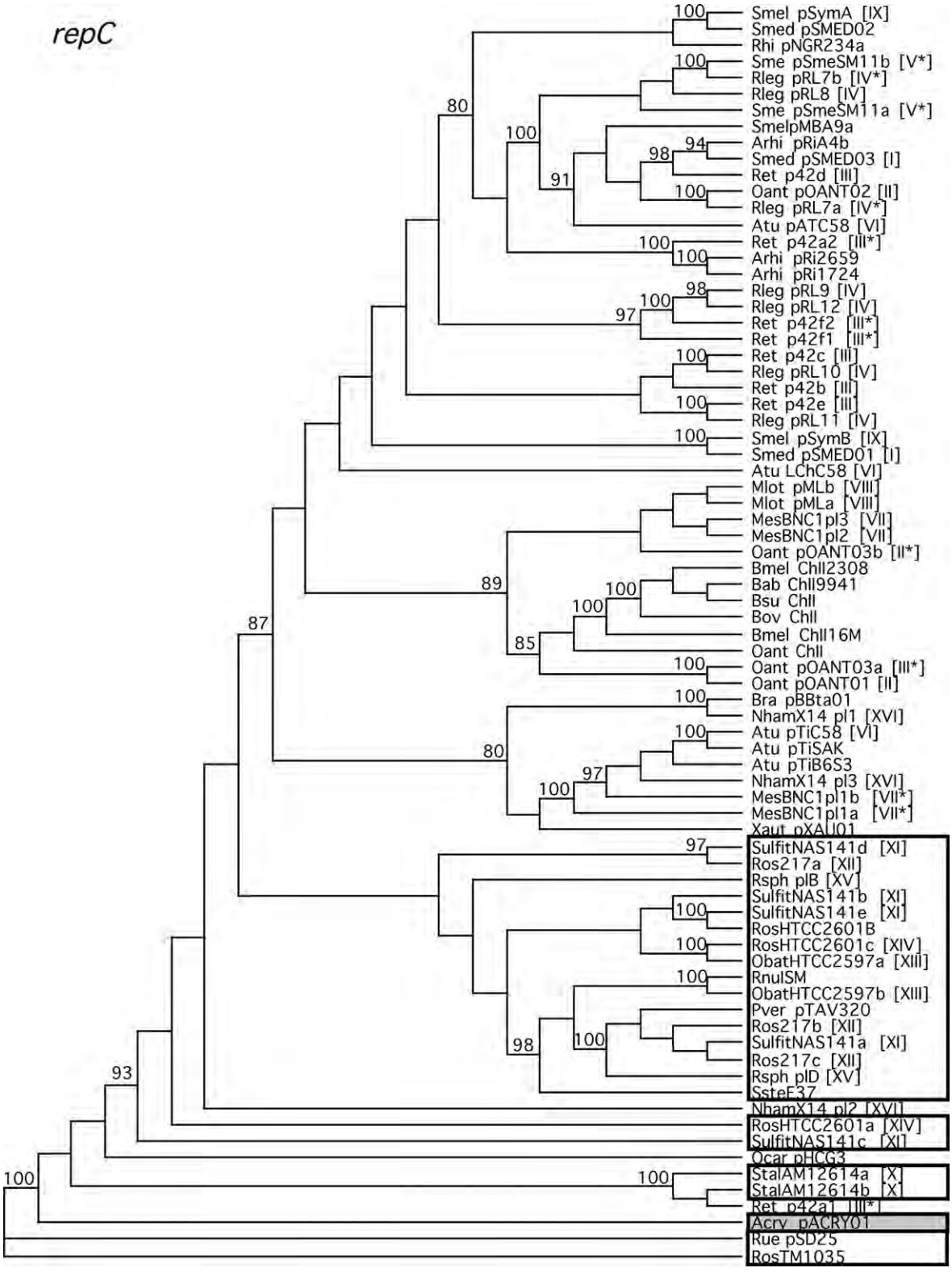


Fig. 5 (*continued*)

by the same plasmid, but which are located in different clades. Proteins of *repABC* operons located in the same plasmid are not phylogenetically close related. Finally, there are also numerous examples of clades containing

proteins of organisms that are not closely related. All these observations indicate that horizontal gene transfer has made a major contribution to *repABC* plasmid evolution, and in general, these data support the idea that new plasmid incompatibility groups arise by independent evolution and gene transfer.

Gene duplication and divergence has a role in the evolution of *repABC* replication-partitioning systems, and its influence is evident in the RepC tree. We identified four independent cases of recent *repC* gene duplications. Plasmids pRL9 and pRL12 of *R. leguminosarum* 3841 each contain a *repABC* operon, and their *repC* genes share 97% identity. RepC proteins of *M. loti* MAFF303099 plasmids pLMa and pLMb have 78% of identity. Similarly, RepC proteins of *Mesorhizobium* sp. BNC1 plasmids pl2 and pl3 share 69% amino acid identity. Finally, *Ochrobactrum anthropi* ATCC49188 plasmids pOANT03 and pOANT01 have RepC proteins with 74% identity.

Intra-operon recombination is another mechanism that has contributed to *repABC* evolution. RepC proteins of *R. leguminosarum* pLR9 and pRL12 are almost identical (see above), however their RepA and RepB proteins are very divergent, but very closely related to RepA and RepB proteins of other organisms. This pattern is also observed for pLMa and pLMb, for pOANT03 and pOANT01, and for pl2 and pl3 (Castillo-Ramírez and MAC, manuscript in preparation).

7. Role of operon fusions in *repABC* evolution

Plasmids of the *repABC* family consist of two linked genetic elements that in other plasmids of low copy-number usually map at two distinct loci: a partitioning locus and a replication locus, each one possessing its own regulatory elements. The origin of *repABC* plasmids was probably an event involving the fusion of these two different genetic *loci*, one containing a *repAB* operon, and other including an antisense RNA gene and *repC*. This notion is supported by the observation that the *repAB* genes and *repC* are independent elements in several α-proteobacteria genomes. For example, many *Sinorhizobium* plasmids contain replication regions consisting of a *repC* gene and an antisense RNA gene encoded upstream from *repC*, transcribed in the opposite direction, and its product being complementary to the leader region of the *repC* mRNA. These plasmids are now classified into the *repC* family, which is related to the *repABC* plasmids but differ in that they do not contain a partitioning system in the vicinity of the replication genes. The functionality of these regions has been demonstrated (Izquierdo et al., 2005). Plasmids belonging to this family are widely distributed in native populations of *S. meliloti*, *S. fredi*, and *R. tropici* (Burgos et al., 1996; Villadas et al., 1995; Stiens et al., 2006). They have also been found in *P. versutus*, a bacterium that is not closely related (Bartosik et al., 1997).

Genetic loci containing *repAB* genes alone (*Rhodobacter sphaeroides* 2.4.1 plasmid C, GenBank NC_007489; *Bradyrhizobium* sp. BTAi1 plasmid pBTai1, Cytryn et al., 2008), or in association with other putative initiator protein genes (*Gluconobacter oxidans* 621H pGOX1 or *Ruegueria* sp. PR1b pSD20) have been identified (Prust et al., 2005; Zhong et al., 2003).

It appears likely that in the ancestral (pleisomorphic) condition, a negatively autoregulated *repAB* operon and its *par*-site located downstream from the *repB* integrated a *ctRNA–repC* cassette. The insertion was between the *repAB* genes and *parS* and its cognate transcription-terminator, but oriented in such a way that the *repAB* and *repC* genes were transcribed in the same direction. Later, the *repC* promoter lost its function by accumulation of deleterious mutations, such that *repC* transcription became dependent on the *repAB* promoter (Izquierdo et al., 2005). However, not all *repABC* plasmids have a *parS* locus downstream from *repC*: some contain their *par*-site upstream from *repA*, whereas in others *parS* maps between the *repA* and *repB* genes (see above). We propose that these *repABC* structures are derived new characters (apomorphic). The generation of these new structures required, first, the duplication of the *parS* locus, then the translocation of one copy to a new localization. Finally, the *parS* locus was lost by accumulation of deleterious mutations.

The *repABC* plasmids contain at least three elements that exert incompatibility: RepA–RepB, the antisense RNA encoded in the *repB–repC* intergenic region, and *parS* (Ramirez-Romero et al., 2000; Bartosik et al., 2001; Venkova-Canova et al., 2004; MacLellan et al., 2004; Chai and Winans, 2005a; Chai and Winans, 2005b; MacLellan et al., 2006). RepA–RepB and the *parS* locus is associated with the partitioning gene cassette. The antisense RNA gene is linked to de *ctRNA–repC* gene locus. If *repABC* plasmids had originated from fusion events involving *repAB–parS* and *ctRNA–repC*, we predict that the spectrum of plasmids eliminated by incompatibility, by a specific *repABC* operon, would be the sum of plasmids displaced by *repAB–parS* elements and those eliminated by the *ctRNA*. We also predict that it would be common to find *repABC* plasmids incompatible with two different plasmids carried by other isolates of the same strain (dual incompatibility).

The genomes of some α-proteobacteria, for example *Gluconobacter oxydans* 621H and *Roseobacter desnitrificans* OCh114, contain *repC* genes not associated with either ctRNA or *repAB* genes. It is therefore likely that gene excision has had a role in the generation of new plasmid families from *repABC* operons. It would be of great interest to determine whether these genes are able to sustain replication. Also, a *repA–repC* gene association with the ability to sustain replication has been described in the *R. leguminosarum* plasmid pRL7 (Pérez-Segura and MAC, unpublished results).

We have only scratched the surface of *repABC* biology. So far, we have identified the elements that constitute a basic *repABC* replicon. We also have a general idea of the roles of each of these elements in the replication and segregation of *repABC* replicons. However, much experimental work remains to be done if we are to understand *repABC* replication and segregation machineries. We need to identify the chromosomal elements associated with these functions, and the number and nature of their macromolecular interactions. It is essential to establish how these machineries are assembled and dissembled and to determine their relative positions within the bacterial cells, during the cell

cycle. The new genomic, proteomic, and visualization techniques will undoubtedly prove to be the tools used to resolve these questions in the near future.

Acknowledgments

We wish to thank Nancy Mena Ocampo and Antonio Loza for their technical support. Also we want to thank Michael Stiens for sharing the *S. meliloti* SM11 DNA sequences of plasmids pSymA and pSymB before publication. This work was supported by PAPIIT Grants IN205203 and CONACYT Grant 46738-Q.

References

Anand, A., Uppalapati, S.R., Ryu, C.M., Allen, S.N., Kang, L., Tang, Y., Mysore, K.S., 2008. Salicylic acid and systemic acquired resistance play a role in attenuating crown gall disease caused by *Agrobacterium tumefaciens*. Plant Physiol. 146, 703–715.

Bailey, T.L., Elkan, C., 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In: Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, California, pp. 28–36.

Bailey, T.L., Gribskov, M., 1998. Combining evidence using *p*-values: application to sequence homology searches. Bioinformatics 14, 48–54.

Bartosik, D., Wlodarczyk, M., Thomas, C.M., 1997. Complete nucleotide sequence of the replicator region of *Paracoccus* (*Thiobacillus*) *versutus* pTAV1 plasmid and its correlation to several plasmids of *Agrobacterium* and *Rhizobium* species. Plasmid 38, 53–59.

Bartosik, D., Baj, J., Wlodarczyk, M., 1998. Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. Microbiology 144, 3149–3157.

Bartosik, D., Szymanik, M., Wysocka, E., 2001. Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. J. Bacteriol. 21, 6234–6243.

Barnett, M.J., Fisher, R.F., Jones, T., Komp, C., Abola, A.P., Barloy-Hubler, F., Bowser, L., Capela, D., Galibert, F., Gouzy, J., Gurjal, M., Hong, A., Guisar, L., Hyman, R.W., Kahn, D., Kahn, M.L., Kalman, S., Keating, D.H., Palm, C., Peck, M.C., Surzycki, R., Wells, D.H., Yeh, K.-C., Davis, R.W., Federspiel, N.A., Long, S.R., 2001. Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. Proc. Natl. Acad. Sci. USA 98, 9883–9888.

Bork, P., Sander, C., Valencia, A., 1992. An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. Proc. Natl. Acad. Sci. USA 89, 7290–7294.

Brantl, S., 2007. Regulatory mechanisms employed by cis-encoded antisense RNAs. Curr. Opin. Microbiol. 10, 102–109.

Burgos, P.A., Velásquez, E., Toro, N., 1996. Identification and distribution of plasmid-type A replicator region in rhizobia. Mol. Plant Microbe Interact. 9, 843–849.

Cevallos, M.A., Porta, H., Izquierdo, J., Tun-Garrido, C., Garcia-de-los-Santos, A., Davila, G., Brom, S., 2002. *Rhizobium etli* CFN42 contains at least three plasmids of the *repABC* family: a structural and evolutionary analysis. Plasmid 48, 104–116.

Chain, P.S., Comerci, D.J., Tolmasky, M.E., Larimer, F.W., Malfatti, S.A., Vergez, L.M., Aguero, F., Land, M.L., Ugalde, R.A., Garcia, E., 2005. Whole-genome analyses of speciation events in pathogenic brucellae. Infect. Immun. 73, 8353–8361.

Chai, Y., Winans, S.C., 2005a. A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. Mol. Microbiol. 56, 1574–1585.

Chai, Y., Winans, S.C., 2005b. RepB protein of an *Agrobacterium tumefaciens*Ti plasmid binds to two adjacent sites between repA and repB for plasmid partitioning and autorepression. Mol. Microbiol. 58, 1114–1129.

Cho, H., Winans, S.C., 2005. VirA and VirG activate the Ti plasmid *repABC* operon, elevating plasmid copy number in response to wound-released chemical signals. Proc. Natl. Acad. Sci. USA 102, 14843–14848.

Cytryn, E.J., Jitacksorn, S., Giraud, E., Sadowsky, M.J., 2008. Insights learned from pBTAi1, a 229-kb accessory plasmid from *Bradyrhizobium* sp. strain BTAi1 and prevalence of accessory plasmids in other *Bradyrhizobium* sp. strains. ISME J. 2, 158–170.

Del Vecchio, V.G., Kapatral, V., Redkar, R.J., Patra, G., Mujer, C., Los, T., Ivanova, N., Anderson, I., Bhattacharyya, A., Lykidis, A., Reznik, G., Jablonski, L., Larsen, N., D'Souza, M., Bernal, A., Mazur, M., Goltsman, E., Selkov, E., Elzer, P.H., Hagius, S., O'Callaghan, D., Letesson, J.J., Haselkorn, R., Kyrpides, N., Overbeek, R., 2002. The genome sequence of the facultative intracellular pathogen *Brucella melitensis*. Proc. Natl. Acad. Sci. USA 99, 443–448.

Dye, N.A., Shapiro, L., 2007. The push and pull of the bacterial cytoskeleton. Trends Cell Biol. 17, 239–245.

Ebersbach, G., Ringgaard, S., Moller-Jensen, J., Wang, Q., Sherratt, D.J., Gerdes, K., 2006. Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. Mol. Microbiol. 61, 1428–1442.

Escobar, M.A., Dandekar, A.M., 2003. *Agrobacterium tumefaciens* as an agent of disease. Trends Plant Sci. 8, 380–386.

Finan, T.M., Weidner, S., Wong, K., Buhrmester, J., Chain, P., Vorholter, F.J., Hernandez-Lucas, I., Becker, A., Cowie, A., Gouzy, J., Golding, B., Puhler, A., 2001. The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont *Sinorhizobium meliloti*. Proc. Natl. Acad. Sci. USA 98, 9889–9894.

Francia, M.V., Fujimoto, S., Tille, P., Weaver, K.E., Clewell, D.B., 2004. Replication of *Enterococcus faecalis* pheromone-responding plasmid pAD1: location of the minimal replicon and *oriV* site and RepA involvement in initiation of replication. J. Bacteriol. 186, 5003–5016.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A., Perret, X., 1997. Molecular basis of symbiosis between *Rhizobium* and legumes. Nature 387, 394–401.

Fuhrmann, S., Ferner, M., Jeffke, T., Henne, A., Gottschalk, G., Meyer, O., 2003. Complete nucleotide sequence of the circular megaplasmid pHCG3 of *Oligotropha carboxidovorans*: function in the chemolithoautotrophic utilization of CO, H_2 and CO_2. Gene 322, 67–75.

Garner, E.C., Campbell, C.S., Weibel, D.B., Mullins, R.D., 2007. Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. Science 315, 1270–1274.

Gerdes, K., Moller-Jensen, J., Jensen, R.B., 2000. Plasmid and chromosome partition: surprises from phylogeny. Mol. Microbiol. 37, 455–466.

Gering, M., Gotz, F., Bruckner, R., 1996. Sequence and analysis of the replication region of the *Staphylococcus xylosus* plasmid pSX267. Gene 182, 117–122.

Geurts, R., Fedorova, E., Bisseling, T., 2005. Nod factor signaling genes and their function in the early stages of *Rhizobium* infection. Curr. Opin. Plant Biol. 8, 346–352.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangenot, S., Segurens, B., Dossat, C., Franck, W.L., Chang, W., Saunders, E., Bruce, D., Richardson, P., Normand, P., Dreyfus, B., Pignol, D., Stacey, G., Emerich, D., Verméglio, A., Médigue, C., Sadowsky, M., 2007. Legumes symbioses: absence of nod genes in photosynthetic bradyrhizobia. Science 316, 1307–1312.

González, V., Bustos, P., Ramírez-Romero, M.A., Medrano, A., Salgado, H., Hernández- González, I., Hernández-Celis, C., Quintero, V., Moreno-Hagelsieb, G., Girard, L., Rodríguez, O., Flores, M., Cevallos, M.A., Collado-Vides, J., Romero, D., Dávila, G., 2003. The mosaic structure of the symbiotic plasmid of *Rhizobium etli* and its relations with other symbiotic genomic compartments. Genome Biol. 4, R36.

González, V., Santamaria, RI., Bustos, P., Hernandez-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S.C., Ramirez, M.A., Jimenez-Jacinto, V., Collado-Vides, J., Davila, G., 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc. Natl. Acad. Sci. USA 103, 3834–3839.

Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B.S., Cao, Y., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., Doughty, D., Scott, C., Lappas, C., Markelz, B., Flanagan, C., Crowell, C., Gurson, J., Lomo, C., Sear, C., Strub, G., Cielo, C., Slater, S., 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science 294, 2323–2328.

Halling, S.M., Peterson-Burch, B.D., Bricker, B.J., Zuerner, R.L., Qing, Z., Li, L.L., Kapur, V., Alt, D.P., Olsen, S.C., 2005. Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. J. Bacteriol. 187, 2715–2726.

Higashi, S., 1967. Transfer of clover infectivity of *R. trifolli* to *R. phaseoli* as mediated by an episomic factor. J. Gen. Appl. Microbiol. 13, 391–403.

Hirano, M., Mori, H., Onogi, T., Yamazoe, M., Niki, H., Ogura, T., Hiraga, S., 1998. Autoregulation of the partition genes of the mini-F plasmid and the intracellular localization of their products in *Escherichia coli*. Mol. Gen. Genet. 257, 392–403.

Huberman, J.A., Spotila, L.D., Nawotka, K.A., el-Assouli, S.M., Davis, L.R., 1987. The in vivo replication origin of the yeast 2 microns plasmid. Cell 51, 473–481.

Izquierdo, J., Venkova-Canova, T., Ramirez-Romero, M.A., Tellez-Sosa, J., Hernandez-Lucas, I., Sanjuan, J., Cevallos, M.A., 2005. An antisense RNA plays a central role in the replication control of a *repC* plasmid. Plasmid 54, 259–277.

Jensen, R.B., Dam, M., Gerdes, K., 1994. Partitioning of plasmid R1. The *parA* operon is autoregulated by ParR and its transcription is highly stimulated by a downstream activating element. J. Mol. Biol. 236, 1299–1309.

Kahng, L.S., Shapiro, L., 2003. Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. J. Bacteriol. 185, 3384–3391.

Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., Tabata, S., 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. DNA Res. 7, 331–338.

Koonin, E.V., 1993. A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. Nucleic Acids Res. 21, 2541–2547.

Kwong, S.M., Skurray, R.A., Firth, N., 2004. *Staphylococcus aureus* multiresistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense RNA regulation. Mol. Microbiol. 51, 497–509.

Lee, K., Park, D., Baek, C., Hwang, W., Kim, K., 2006. *repABC*-type replicator region of megaplasmid pATC58 in *Agrobacterium tumefaciens* C58. J. Microbiol. Biotechnol. 16, 118–125.

Li, P.L., Farrand, S.K., 2000. The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. J. Bacteriol. 182, 179–188.

Lim, G.E., Derman, Al., Pogliano, J., 2005. Bacterial DNA segregation by dynamic SopA polymers. Proc. Natl. Acad. Sci. USA 102, 17658–17663.

Liu, P., Nester, E.W., 2006. Indoleacetic acid, a product of transferred DNA, inhibits vir gene expression and growth of *Agrobacterium tumefaciens*C58. Proc. Natl. Acad. Sci. USA 103, 4658–4662.

Long, S.R., 2001. Genes signals in the *Rhizobium*–legume symbiosis. Plant Physiol. 125, 69–72.

Mankin, S.L., Hill, D.S., Olhoft, P.M., Toren, E., Wenck, A.R., Nea, L., Xing, L., Brown, J.A., Fu, H., Ireland, L., Jia, H., Hillebrand, H., Jones, T., Song, H.S., 2007. Disarming and sequencing of *Agrobacterium rhizogenes* strain K599 (NCPPB2659) plasmid pRi2659. In Vitro Cell. Dev. Biol., Plant 43, 521–535.

McAnulla, C., Edwards, A., Sanchez-Contreras, M., Sawers, R.G., Downie, J.A., 2007. Quorum-sensing-regulated transcriptional initiation of plasmid transfer and replication genes in *Rhizobium leguminosarum*biovar viciae. Microbiology 153, 2074–2082.

MacLellan, S.R., Smallbone, L.A., Sibley, C.D., Finan, T.M., 2005. The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. Mol. Microbiol. 55, 611–623.

MacLellan, S.R., Zaheer, R., Sartor, A.L., MacLean, A.M., Finan, T.M., 2006. Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons. Mol. Microbiol. 59, 1559–1575.

Møller-Jensen, J., Jensen, R.B., Lowe, J., Gerdes, K., 2002. Prokaryotic DNA segregation by an actin-like filament. EMBO J. 21, 3119–3127.

Moriguchi, K., Maeda, Y., Satou, M., Hardayani, N.S., Kataoka, M., Tanaka, N., Yoshida, K., 2001. The complete nucleotide sequence of a plant root-inducing Ri. plasmid indicates its chimeric structure and evolutionary relationship between tumor-inducing Ti. and symbiotic Sym. plasmids in Rhizobiaceae. J. Mol. Biol. 307, 771–784.

Nishiguchi, R., Takanami, M., Oka, A., 1987. Characterization and sequence determination of the hairy root inducing plasmid pRiA4b. Mol. Gen. Genet. 206, 1–8.

Novick, R.P., 1987. Plasmid incompatibility. Microbiol. Rev. 51, 381–395.

Pappas, K.M., Winans, S.C., 2003a. A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. Mol. Microbiol. 48, 1059–1073.

Pappas, K.M., Winans, S.C., 2003b. The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. Mol. Microbiol. 49, 441–455.

Paulsen, I.T., Seshadri, R., Nelson, K.E., Eisen, J.A., Heidelberg, J.F., Read, T.D., Dodson, R.J., Umayam, L., Brinkac, L.M., Beanan, M.J., Daugherty, S.C., Deboy, R.T., Durkin, A.S., Madupu, J.F., Kolonay, R., Nelson, W.C., Ayodeji, B., Kraul, M., Shetty, J., Malek, J., Van Aken, S.E., Riedmuller, S., Tettelin, H., Gill, S.R., White, O., Salzberg, S.L., Hoover, D.L., Lindler, L.E., Halling, S.M., Boyle, S.M., Fraser, C.M., 2002. The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. Proc. Natl. Acad. Sci. USA 99, 13148–13153.

Perret, X., Freiberg, C., Rosenthal, A., Broughton, W.J., Feilla, R., 1999. High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234. Mol. Microbiol. 32, 415–425.

Piper, K.R., Beck Von Bodman, S., Hwang, I., Farrand, S.K., 1999. Hierarchical gene regulatory systems arising from fortuitous gene associations: controlling quorum sensing by the opine regulon in *Agrobacterium*. Mol. Microbiol. 32, 1077–1089.

Prust, C., Hoffmeister, M., Liesegang, H., Wiezer, A., Fricke, W.F., Ehrenreich, A., Gottschalk, G., Deppenmeier, U., 2005. Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*. Nat. Biotechnol. 23, 195–200.

Quintero, V., Cevallos, M.A., Davila, G., 2002. A site-specific recombinase (RinQ) is required to exert incompatibility towards the symbiotic plasmid of *Rhizobium etli*. In: Mol. Microbiol.. 46, 1023–1032.

Ramirez-Romero, M.A., Bustos, P., Girard, L., Rodriguez, O., Cevallos, M.A., Davila, G., 1997. Sequence, localization and characteristics of the replicator region of the symbiotic plasmid of *Rhizobium etli*. Microbiology 143, 2825–2831.

Ramirez-Romero, M.A., Masulis, I., Cevallos, M.A., Gonzalez, V., Davila, G., 2006. The *Rhizobium etli* sigma70 SigA factor recognizes a lax consensus promoter. Nucleic Acids Res. 34, 1470–1480.

Ramirez-Romero, M.A., Soberon, N., Perez-Oseguera, A., Tellez-Sosa, J., Cevallos, M.A., 2000. Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. J. Bacteriol. 182, 3117–3124.

Ramirez-Romero, M.A., Tellez-Sosa, J., Barrios, H., Perez-Oseguera, A., Rosas, V., Cevallos, M.A., 2001. RepA negatively autoregulates the transcription of the *repABC* operon of the *Rhizobium etli* symbiotic plasmid basic replicon. Mol. Microbiol. 42, 195–204.

Ravin, N.V., Kuprianov, V.V., Gilcrease, E.B., Casjens, S.R., 2003. Bidirectional replication from an internal ori site of the linear N15 plasmid prophage. Nucleic Acids Res. 31, 6552–6560.

Soberón, N., Venkova-Canova, T., Ramirez-Romero, M.A., Tellez-Sosa, J., Cevallos, M.A., 2004. Incompatibility and the partitioning site of the *repABC* basic replicon of the symbiotic plasmid from *Rhizobium etli*. Plasmid 51, 203–216.

Stiens, M., Schneiker, S., Puhler, A., Schluter, A., 2007. Sequence analysis of the 181-kb accessory plasmid pSmeSM11b, isolated from a dominant *Sinorhizobium meliloti* strain identified during a long-term field release experiment. FEMS Microbiol. Lett. 271, 297–309.

Stiens, M., Schneiker, S., Keller, M., Kuhn, S., Puhler, A., Schluter, A., 2006. Sequence analysis of the 144-kilobase accessory plasmid pSmeSM11a, isolated from a dominant *Sinorhizobium meliloti* strain identified during a long-term field release experiment. Appl. Environ. Microbiol. 72, 3662–3672.

Suzuki, K., Ohta, N., Hattori, Y., Uraji, M., Kato, A., Yoshida, K., 1998. Novel structural difference between nopaline- and octopine-type trbJ genes: construction of genetic and physical map and sequencing of *trb*/*traI* and *rep* gene clusters of a new Ti plasmid pTi-SAKURA. Biochim. Biophys. Acta 1396, 1–7.

Sykora, P., 1992. Macroevolution of plasmids: a model for plasmid speciation. J. Theor. Biol. 159, 53–65.

Tabata, S., Hooykaas, P.J., Oka, A., 1989. Sequence determination and characterization of the replicator region in the tumor-inducing plasmid pTiB6S3. J. Bacteriol. 171, 1665–1672.

Tanaka, T., Ishida, H., Maehara, T., 2005. Characterization of the replication region of plasmid pLS32 from the Natto strain of *Bacillus subtilis*. J. Bacteriol. 187, 4315–4326.

Turner, S.L., Young, J.P., 1995. The replicator region of the *Rhizobium leguminosarum* cryptic plasmid pRL8JI. FEMS Microbiol. Lett. 133, 53–58.

Venkova-Canova, T., Soberón, N.E., Ramírez-Romero, M.A., Cevallos, M.A., 2004. Two discrete elements are required for the replication of a *repABC* plasmid: an antisense RNA and a stem-loop structure. Mol. Microbiol. 54, 1431–1444.

Villadas, P.J., Velázquez, E., Martínez-Molina, E., Toro, N., 1995. Identification of nodule-dominant *Rhizobium meliloti* strains carrying pRmeGR4b-type plasmid within indigenous soil

populations by PCR using primers derived from specific DNA sequences. FEMS Microbiol. Ecol. 17, 161–168.
von Bodman, S.B., Hayman, G.T., Farrand, S.K., 1992. Opine catabolism and conjugal transfer of the nopaline Ti plasmid pTiC58 are coordinately regulated by a single repressor. Proc. Natl. Acad. Sci. USA 89, 643–647.
Watson, R.J., Heys, R., 2006. Replication regions of *Sinorhizobium melilotip*lasmids. Plasmid 55, 87–98.
Weaver, K.E., 2007. Emerging plasmid-encoded antisense RNA regulated systems. Curr. Opin. Microbiol. 10, 110–116.
Yamamoto, S., Uraji, M., Tanaka, K., Moriguchi, K., Suzuki, K., 2007. Identification of pTi-SAKURA DNA region conferring enhancement of plasmid incompatibility and stability. Genes Genet. Syst. 82, 197–206.
Young, J.W., Crossman, L.C., Johnston, A.W.B., Thomson, N.R., Ghazoui, Z.F., Hull, K.H., Wexler, M., Curson, A.R.J., Todd, J.D., Poole, P.S., Mauchline, T.H., East, A.K., Quail, M.A., Churcher, C., Arrowsmith, C., Cherevach, A., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabbinowitsch, E., Sanders, M., Simmonds, M., Parkhill, S., Whitehead, J., 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol. 7, R34.
Yuan, Z.C., Edlind, M.P., Liu, P., Saenkham, P., Banta, L.M., Wise, A.A., Ronzone, E., Binns, A.N., Kerr, K., Nester, E.W., 2007. The plant signal salicylic acid shuts down expression of the *vir* regulon and activates quormone-quenching genes in *Agrobacterium*. Proc. Natl. Acad. Sci. USA 104, 11790–11795.
Zhong, Z., Caspi, R., Helinski, D., Knauf, V., Sykes, S., O'Byrne, C., Shea, T.P., Wilkinson, J.E., DeLoughery, C., Toukdarian, A., 2003. Nucleotide sequence based characterizations of two cryptic plasmids from the marine bacterium *Ruegeria* isolate PR1b. Plasmid 49, 233–252.
Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J., Farrand, S.K., Winans, S.C., 2000. The bases of crown gall tumorigenesis. J. Bacteriol. 182, 3885–3895.