

UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

El Complejo CarD-CarG y su Implicación en un Nuevo Mecanismo de Regulación de la Expresión de un Sistema CRISPR-Cas

> D. Diego Bernal Bernal 2018



UNIVERSIDAD DE MURCIA

Facultad de Biología

D^a. Montserrat Elías Arnanz, Catedrática de Universidad del Área de Genética, en el Departamento de Genética y Microbiología, y D. Subramanian Padmanabhan Iyer, Profesor de Investigación del CSIC, AUTORIZAN:

La presentación de la Tesis Doctoral titulada "El complejo CarD-CarG y su implicación en un nuevo mecanismo de regulación de la expresión de un sistema CRISPR-Cas", realizada por D. Diego Bernal Bernal, bajo nuestra inmediata dirección y supervisión, y que se presenta para la obtención del grado de Doctor por la Universidad de Murcia.

En Murcia, a 17 de Septiembre de 2018

Departamento de Genética y Microbiología Campus de Espinardo 30100 Murcia T. +34 868 884 949 F. +34 868 883 963

"Solo el que sabe es libre, y más libre el que más sabe... Solo la cultura da libertad... No proclaméis la libertad para volar, sino dad alas; no la de pensar, sino dad pensamiento. La libertad que hay que dar al pueblo es la cultura"

Miguel de Unamuno

A mi familia A Noe

ÍNDICE

ÍNDICE DE MATERIAS

I. IN	I. INTRODUCCIÓN			
I.1	I.1 Expresión génica en bacterias: inicio de la transcripción v			
	su regulación3 -			
	I.1.1 La polimerasa de RNA y los factores σ en bacterias3 -			
	I.1.2 Los promotores dependientes de la familia σ^{70}			
	I.1.3 Los factores σ-ECF 12 -			
	I.1.4 Reguladores transcripcionales 17 -			
	I.1.4.1 A nivel del promotor 18 -			
	I.1.4.2 A nivel de la interacción con la RNAP 22 -			
12	Myxococcus xanthus - 24 -			
1.2	1 2 1 Características generales de las mixohacterias - 24 -			
	122 Ciclo de vida de <i>M</i> xanthus - 25 -			
	1.2.3 Respuesta a la luz en <i>M. xanthus</i> - 27 -			
	1.2.4 Factores σ en <i>M. xanthus</i>			
	I.2.5 La familia de proteínas CarD-CdnL 31 -			
	I.2.5.1 El complejo regulador CarD-CarG 31 -			
	I.2.5.1.1 El operón carD-carG 31 -			
	I.2.5.1.2 La proteína CarD 31 -			
	I.2.5.1.3 La proteína CarG 33 -			
	I.2.5.1.4 El complejo CarD-CarG 34 -			
	I.2.5.1.5 Funciones de <i>carD-carG</i> 35 -			
	I.2.5.1.5.1 Acción de <i>carD</i> y <i>carG</i> en la respuesta a la luz 35 -			
	I.2.5.1.5.2 Acción de <i>carD</i> y <i>carG</i> en el desarrollo multicelular 35 -			
	I.2.5.1.5.3 Acción de <i>carD</i> y <i>carG</i> en la fase vegetativa 36 -			
	1.2.5.2 CdnL 36 -			
1.3	Sistemas CRISPR-Cas 37 -			
	I.3.1 Función de los sistemas CRISPR-Cas 38 -			
	I.3.2 Composición de los sistemas CRISPR-Cas 38 -			
	I.3.2.1 La región CRISPR 38 -			
	I.3.2.2 Los genes <i>cas</i> 39 -			
	1.3.3 Mecanismo de acción de los sistemas CRISPR-Cas 40 -			
	1.3.4 Tipos de sistemas CRISPR-Cas			
	1.3.5 Regulacion de sistemas CRISPR-Cas 45 -			
	1.3.6 Sistemas CRISPR-Cas en M. xanthus 46 -			
1.4	Objetivos 47 -			
II. N	IATERIALES Y MÉTODOS			
II.1	Estirpes y plasmidos 51 -			
	II.1.1 Estirpes de Myxococcus xanthus			
	11.1.2 Estirpes de <i>E. coll</i> 55 -			
	II.1.3 Oligonucleotidos 60 -			
II.2	Medios de cultivo 63 -			
	II.2.1 Medios ricos para <i>M. xanthus</i> 63 -			
	II.2.2 Medios de fructificación para <i>M. xanthus</i> 64 -			
	II.2.3 Medios ricos para <i>E. coli</i> 64 -			
11.3	Condiciones de cultivo			
11.4	inducción del desarrollo multicelular de <i>M. xanthus</i>			
II.5	Tampones y soluciones 66 -			
	II.5.1 Electroforesis de DNA 66 -			

II.5.	2 Electroforesis de proteínas	- 66 -
II.5.	3 Electroforesis de RNA	- 66 -
II.5. II.5	4 Purilicación de proteínas mediante anticuernos específicos	- 67 -
II.5.	6 Determinación de la actividad β-galactosidasa	- 68 -
II.5.	7 Inmunoprecipitación de cromatina	- 68 -
II.5.	8 Northern blot	- 69 -
11	.5.8.1 Para análisis de RNAs pequeños	- 69 -
	.5.8.2 Para análisis de RNAs totales	- 70 -
II.6	Extracción de DNA	70 -
II.7	Transformación	70 -
II.7.	1 Transformación de <i>E. coli</i>	- 70 -
11.7.	2 Transformación de <i>M. xanthus</i>	- 71 -
II.8	Tratamiento enzimático del DNA	71 -
II.9	Amplificación de DNA mediante PCR	72 -
II.10	Secuenciación de DNA	72 -
II.11	Extracción, cuantificación y análisis de la calidad de RNA	72 -
II.11	1.1 Para ensayos de Northern blot	- 72 -
II.11	1.2 Para ensayos de RT-PCR o qPCR	- 73 -
II.12	Procedimiento de mutagénesis	74 -
II.12	2.1 Mutagénesis dirigida	- 74 -
II.12	2.2 Generación de estirpes de <i>M. xanthus</i> portadoras de una deleción	- 74 -
II.13	Técnicas electroforéticas	76 -
II.13	3.1 Electroforesis de DNA en gel de agarosa	- 76 -
II.13	B.2 Electroforesis de proteínas en geles de poliacrilamida-SDS	- 76 -
II.13	3.3 Electroforesis de DNA en gel de acrilamida no desnaturalizante	- 76 -
II.13 II.13	3.4 Electroforesis de RNA en gel de acrilamida desnaturalizante	- 70 -
II.14	Expresion y purificación de proteinas	- 11 -
II. 14 II. 12	1.2 Expresión y purificación de CarD	- 78 -
II.14	4.3 Expresión y purificación de CarG	- 78 -
II.14	4.4 Purificación de la polimerasa de RNA de <i>M. xanthus</i>	- 79 -
II.15	Detección de proteínas mediante anticuerpos específicos	
	(Western-blot)	79 -
II.16	Ensayo de cuantificación relativa mediante RT-PCR en tiempo real	
	(qRT-PCR)	80 -
II.17	Ensayo de RT-PCR	82 -
II.18	Ensayo de doble híbrido bacteriano (DHB)	82 -
19	Ensavo de ß-galactosidasa	83 -
II.19	$D.1$ Ensayo de β -galactosidasa en <i>M. xanthus</i>	- 83 -
II.19	9.2 Ensayo de β-galactosidasa en <i>E. coli</i>	- 84 -
II.20	Ensayo de retraso en la movilidad electroforética (EMSA)	85 -
II.21	Northern blot	85 -
II.21	1.1 Northern blot de RNAs pequeños	- 85 -

II.2	21.2 Northern blot de RNAs total 86 -	
II.22	Inmunoprecipitación de cromatina-qPCR (ChIP-qPCR) 86 -	
II.23	Chromatin Inmmunoprecipitation Sequencing (ChIP-Seq) 88 -	
II.24	Análisis bioinformático 89 -	
III. RES	SULTADOS	
III.1 III.1 III.1	El dominio N-terminal de CarD: función y estructura	
III.2	Los factores σ -ECF y su dependencia del complejo CarD-CarG 103 -	
.2 .2 .2 .2	 2.1 Los factores anti-σ CarR y DdvA: topología en la membrana e interacción con sus respectivos factores σ–ECF104 - 2.2 Dependencia de CarD-CarG de otras parejas σ-ECF/anti-σ110 - III.2.2.1 Una pareja σ-ECF/anti-σ similar a DdvS/DdvA110 - III.2.2.2 Parejas σ-ECF/anti-σ con un dominio quinasa de serina/treonina en el factor anti-σ112 - 	
III.3	Modo de acción del complejo regulador CarD-CarG 116 -	
.; .; .;	 Análisis y validación de los sitios de unión del complejo CarD-CarG 116 - Estudio de la asociación <i>in vivo</i> de CarQ y CarD al promotor P_{QRS} 120 - Análisis del efecto de CarG sobre la unión de CarD al DNA <i>in vivo</i> 122 - 	
III.4	Control multifactorial de un sistema CRISPR-Cas por la pareja	
.4 .4 .4 .4 .4	 DdvS/DdvA y el complejo regulador global CarD-CarG 124 - Promotores dependientes de DdvS y CarD-CarG en el locus CRISPR-Cas- 125 - La región líder de CRISPR4 tiene una actividad promotora despreciable 130 - La región CRISPR4 se expresa a partir del operón <i>cas</i> 132 - DdvS, CarD y CarG controlan la producción de los crRNAs 135 - La producción de los crRNAs derivados de la región CRISPR4 depende de la proteína Cas6 137 - La expresión del sistema CRISPR4 no está activada por el desarrollo 141 - 	
IV. DIS	CUSIÓN 145 -	
IV. 1	El dominio N-terminal de CarD: subdominios e interacciones 147 -	
IV. 2	Acción reguladora global del complejo CarD-CarG 150 -	
IV.3.	Factores anti-σ en <i>M. xanthus</i> : topología en la membrana y	
	asociación de dominios 151 -	
IV.4.	Sistemas CRISPR-Cas en <i>M. xanthus</i> y su regulación 158 -	
IV.5	¿Otros genes de defensa en el locus CRISPR4-Cas? 164 -	
V. CON	NCLUSIONES 167 -	
VI. BIBLIOGRAFÍA 171 -		

ÍNDICE DE FIGURAS

Figura 2	Detalle de las etapas que conducen al ensamblaje de la RNAP	4 -
Figura 2	Estructura del nucleo de la polimerasa de RNA	5 -
Figura	. Representación esquematica de los dominios conservados de los cuatro	0
Figura	Pegión promotora reconocida por la RNAD unida al factor σ ⁷⁰	- 0 - 10
Figura	Detalle de las etanas que constituyen el inicio de la transcrinción	- 10 -
Figura	Mecanismo de acción general de las pareias $\sigma_{\rm e} {\rm ECE}$	- 11 -
Figura 5	Mecanismo de activación simple	
Figura	Activación de la transcrinción nor dos reguladores transcrincionales	- 20 -
Figura	Mecanismos de represión simple	- 20 -
Figura	 Mecanismos de represión simple O Ciclo de vida de M xanthus 	- 26 -
Figura	1 Síntesis de carotenoides en M. vanthus en respuesta a la luz	- 28 -
Figura '	 2 Modelo de organización y acción de los genes implicados en la 	20 -
i igui u	resnuesta a la luz de M vanthus	- 28 -
Figura	3 Arquitectura de dominios de CarG y CarD y comparación de su	20
riguiu	dominio C- terminal con la proteína humana HMGA1a	- 32 -
Figura 2	4 Comparación de las proteínas CarD y Cdnl de M xanthus	- 37 -
Figura '	5. Estructura y composición de los sistemas CRISPR-Cas	- 39 -
Figura '	6. Mecanismo de acción de los sistemas CRISPR-Cas	- 41 -
Figura '	7. Composición y estructura tridimensional de los compleios efectores	
	de los sistemas CRISPR-Cas de tipo I v III	- 42 -
Figura ²	8. Sistemas CRISPR-Cas de <i>M. xanthus</i>	- 46 -
Figura '	9. Esquema de los cuatro loci CRISPR-cas de M. xanthus	- 47 -
Figura	0. Esquema del proceso de obtención de una estirpe portadora de una	
	deleción en <i>M. xanthus</i>	75 -
Figura 2	1. Esquema del sistema del doble híbrido bacteriano	83 -
Figura 2	2. Representación esquemática de las principales etapas	
U	de la inmunoprecipitación de cromatina	88 -
Figura 2	3. Alineamiento de las proteínas CarD y CdnL	93 -
Figura 2	4. Comparación del efecto de la eliminación de los dominios de CarD	
•	versus la eliminación de la proteína completa en <i>M. xanthus</i>	95 -
Figura 2	5. Análisis estructural de CarDNt	97 -
Figura 2	6. Análisis mutacional de la interacción entre CarDNt y βRNAP	98 -
Figura 2	7. Efecto <i>in vivo</i> de las mutaciones en residuos de CarD implicados	
	en la interacción con la RNAP	100 -
Figura 2	8. Análisis mutacional de residuos conservados en el subdominio	
	C-terminal (CarD ₆₁₋₁₇₉) de CarDNt	102 -
Figura 2	9. Efecto de la mutación R95A en la proteína CarD en M. xanthus	103 -
Figura 3	0. Predicción de la topología en la membrana de los factores anti-σ CarR	
	y DdvA	104 -
Figura 3	1. Topología en la membrana del factor anti-σ CarR	105 -
Figura 3	2. Ensayo de interacción entre los dominios del factor anti- σ CarR y el	
	factor σ-ECF CarQ	106 -
Figura 3	3. Topología del factor anti-σ DdvA	107 -
Figura 3	4. Ensayo de interacción entre los dominios del factor anti- σ DdvA y el	
	factor σ-ECF DdvS	108 -
Figura 3	5. La sobreexpresión del dominio NZASD de DdvA anula los efectos	
	de la sobreexpresión de DdvS en <i>M. xanthus</i>	109 -
Figura 3	6. Análisis de la posible oligomerización de DdvA y de DdvS	110 -
Figura 3	7. Estudio de la pareja σ -ECF/anti- σ MXAN_5506/MXAN_5507	111 -
Figura	8. Posibles parejas σ -ECF/anti- σ con un factor anti- σ atípico con un	
	dominio ZAS y quinasa de serina/treonina.	113 -
Figura 3	9. Estudio de la interacción de las posibles parejas σ -ECF/anti- σ atípicas	114 -

Figura	40.	Ensayo de posibles interacciones cruzadas entre parejas σ -ECF/anti- σ	
		con una alta homología	114 -
Figura	41.	Esquema de la estrategia utilizada para analizar la actividad del	
		promotor P1710-09 y su dependencia de CarD y CarG	115 -
Figura	42.	Estudio de la dependencia de CarD-CarG en el factor σ -ECF	
		MXAN_1709	116 -
Figura	43.	Comprobación de la expresión y funcionalidad CarD-Flag y CarG-Flag	117 -
Figura	44.	Perfiles de unión de CarD y CarG en el genoma de M. xanthus	118 -
Figura	45.	Validación mediante ChIP-qPCR de la unión de CarD a las tres	
		regiones promotoras indicadas	119 -
Figura	46.	Análisis de la dependencia de CarD y CarG mediante qRT-PCR de los	
		genes indicados	120 -
Figura	47.	Generación de la estirpe que expresa Flag-CarQ y análisis	
		de su comportamiento	121 -
Figura	48.	Análisis de la asociación in vivo de CarD a la región promotora P _{QRS}	122 -
Figura	49.	Efecto de CarG sobre la unión de CarD al DNA in vivo	123 -
Figura	50.	Distribución en el genoma de los loci activados por DdvS y dependientes	
		de CarD y CarG	125 -
Figura	51.	Esquema del sistema CRISPR4-Cas	126 -
Figura	52.	Alineamiento de las regiones -35 y -10 de los promotores dependientes de	
		DdvS e identificación del sitio de inicio de la transcripción mediante 5' RACE	126 -
Figura	53.	Medida de la actividad β -galactosidasa de los promotores	
		P _{MXAN_7291} , P _{ddvSA} , P _{MXAN_7286} y P _{MXAN_7283} en los fondos genéticos indicados	127 -
Figura	54.	Estudio de la estirpe que expresa condicionalmente flag-ddvS	128 -
Figura	55.	Analisis de la asociación <i>in vivo</i> de CarD a los promotores	400
F :		aependientes de DavS	129 -
Figura	56.	Secuencia de las regiones promotoras dependientes de DavS y	400
Figure	E7	estudio de la union de CarD-CarG a dichas regiones	130 -
Figura	5/.	Analisis de la region lider	131 -
Figura	JO .	Expresion de la region líder suvestre o mutada en un sitio neterologo	100
Figura	50	Expressión de la fusión casé://ac7	132 -
Figura	59.	Análisis de la expressión de un mDNA policistrónice a partir de	155 -
Figura	00.	Pure a expressión de un minicipa policistionico a partir de	13/
Figura	61	Análisis mediante RT_PCR de la expresión de un mRNA nolicistrónico	104 -
rigura	01.	entre los genes cas y los genes situados aguas arriba	135 -
Figura	62	Análisis mediante Northern blot del RNA total obtenido a partir de las	100 -
rigura	02.	estimes indicadas usando una sonda que hibrida desde el espaciador 52 al	
		37 de CRISPR4 -	136 -
Figura	63.	Análisis de los crRNAs producidos por las estirpes indicadas usando	100
		una sonda que hibrida en el espaciador 52 y otra en el espaciador 1 de la	
		región CRISPR4	137 -
Figura	64.	Papel de Cas6 en la formación de los crRNAs	138 -
Figura	65.	Análisis del papel de Cas6 mediante Northern blot de RNA total obtenido	
Ū		a partir de las estirpes indicadas usando una sonda que hibrida desde el	
		espaciador 52 al 37 de CRISPR4	139 -
Figura	66.	Alineamiento de las dos proteínas Cas6 de M. xanthus	140 -
Figura	67.	Alineamiento de las repeticiones de las regiones CRISPR2 y	
		CRISPR3 del sistema CRISPR2/3-Cas y de las repeticiones de la región	
		CRISPR4 del sistema CRISPR4-Cas.	140 -
Figura	68.	Formación de los crRNAs de la región CRISPR4 mediada por Cas6a	141 -
Figura	69.	Producción de cuerpos fructíferos en la estirpe silvestre y las estirpes	
		mutantes $\Delta ddvA$ y $\Delta carD$ a los tiempos indicados durante el desarrollo	
		multicelular	142 -
Figura	70.	Análisis cualitativo y cuantitativo de la expresión de la fusión	
		Líder::/acZ durante el desarrollo en los fondos genéticos indicados	143 -

Figura 71. Análisis de los crRNAs producidos durante el desarrollo por los	
sistemas CRISPR4-Cas y CRISPR2/3-Cas	144 -
Figura 72. Alineamiento de los 70 primeros aminoácidos de los factores anti- σ	
con motivo NZASD de <i>M. xanthus</i>	152 -
Figura 73. Distribución en el genoma de <i>M. xanthus</i> de las parejas σ -ECF/anti- σ	
estudiadas en este trabajo	154 -
Figura 74. Alineamiento del dominio eSTK de PkA de Mus musculus y PknB	
de <i>M. tuberculosis</i> con los dominios eSTK de los factores anti-σ estud	ados de
M. xanthus	155 -
Figura 75. Regulación del desarrollo multicelular de M. xanthus	157 -
Figura 76. Modelos de activación para eSTKs	158 -
Figura 77. Modelo de regulación del sistema CRISPR4-Cas	159 -
Figura 78. Secuencia de DdvA con sus dominios funcionales	160 -
Figura 79. Similitudes de los sistemas CRISPR4-Cas y CRISPR2/3-Cas	161 -
Figura 80. Homología de un espaciador de la región CRISPR2/3 con un fragment	to
del gen de la integrasa del fago Mx8	163 -
Figura 81. Genes de función hipotética asociados al sistema CRISPR4-Cas	165 -
Figura 82. Producción de cOA en los sistemas CRISPR-Cas de tipo III	166 -

ÍNDICE DE TABLAS

Tabla 1. Estirpes de <i>M. xanthus</i>	51 -
Tabla 2. Estirpes de <i>E. coli</i> usadas en este trabajo	- 56 -
Tabla 3. Plásmidos utilizados en este trabajo	56 -
Tabla 4. Oligonucleótidos usados en este trabajo.	60 -
Tabla 5. Antibióticos empleados en <i>M. xanthus</i>	63 -
Tabla 6. Antibióticos empleados en <i>E. coli.</i>	64 -
Tabla 7. Lista de genes activados en la estirpe Δ <i>ddvA</i> y reprimidos en	
$\Delta ddvA \ \Delta carD \ y \ \Delta ddvA \ \Delta carG$	124 -

I. INTRODUCCIÓN

I.1 Expresión génica en bacterias: inicio de la transcripción y su regulación

La supervivencia de una especie implica una constante adaptación a las condiciones ambientales en las que habita. En particular, las bacterias han desarrollado numerosas adaptaciones que les han permitido persistir desde tiempos inmemoriales en los diversos hábitats que han colonizado. Al vivir en contacto directo con el medio externo, las bacterias son más susceptibles que otros organismos a las variaciones de éste, lo que les ha llevado a desarrollar sistemas de detección, así como respuestas acordes para facilitar su supervivencia, generalmente basadas en el control de la expresión de determinados grupos de genes relacionados con la respuesta en cuestión. El proceso de expresión génica comienza con la síntesis de una molécula de RNA a partir de DNA, en un proceso denominado transcripción que es llevado a cabo por una única enzima, la polimerasa de RNA (en adelante RNAP). Este proceso engloba tres fases: iniciación, elongación y terminación, siendo el inicio de la transcripción la etapa clave en la regulación de la expresión génica. Durante esta fase, distintos factores reguladores, como los propios componentes de la RNAP, los factores que actúan en cis (promotores, operadores e intensificadores) y los elementos de acción en trans (activadores, represores y adaptadores) modulan la expresión génica. El paso regulador clave es el reconocimiento de la región promotora (secuencia específica de DNA necesaria para iniciar la transcripción sobre la que se elaborará más adelante) por la RNAP para llevar a cabo el inicio de la transcripción (Browning & Busby, 2016).

I.1.1 La polimerasa de RNA y los factores σ en bacterias

En bacterias, la RNAP está formada por un núcleo enzimático (de ~400 kDa de masa molecular) constituido por cinco subunidades ($\alpha_2\beta\beta'\omega$) y un factor σ disociable. La subunidad α participa en el ensamblaje del conjunto de la RNAP; las subunidades β y β' presentan actividad catalítica, y la subunidad ω presenta una función auxiliar (Alhadid *et al.*, 2017). El núcleo enzimático requiere la unión a un factor σ para generar una holoenzima activa, capaz de reconocer y unirse a los promotores, con el fin de provocar la apertura de la región promotora y generar un complejo abierto con el DNA para la síntesis de RNA (Feklistov, 2013).

El núcleo enzimático de la RNAP está muy conservado evolutivamente tanto en secuencia como en estructura y función (Ebright, 2000). El ensamblaje de este núcleo enzimático tiene lugar de forma secuencial, comenzando con la dimerización de las subunidades α , sobre las que se asienta la subunidad β generando un complejo intermedio $\alpha_2\beta$, al que se une la subunidad β ' con la ayuda de la subunidad ω , que no interviene de forma directa en la transcripción (Figura 1). A este núcleo enzimático se le agrega el factor σ , que facilita el reconocimiento y la localización de la RNAP en promotores específicos (Browning & Busby, 2004; Haugen *et al.*, 2006).



Figura 1. Detalle de las etapas que conducen al ensamblaje de la RNAP. A) Pasos en el ensamblaje de la RNAP. B) Modelo illustrativo de la formación del complejo enzimático. La forma holo de la RNAP se corresponde con el complejo $\alpha^2\beta\beta'\omega\sigma$.

Estudios previos con la RNAP de *Escherichia coli*, así como con las de las bacterias termófilas *Thermus thermophilus* y *Thermus aquaticus*, y más recientemente con la RNAP de *Mycobacterium tuberculosis*, han permitido obtener datos a nivel estructural a resolución atómica, y profundizar en su modo de acción (Bae *et al.*, 2015; Basu *et al.*, 2014; Lin *et al.*, 2017; Murakami, 2013; Zhang *et al.*, 1999). Estos datos indican que la RNAP bacteriana se asemeja a las pinzas de un cangrejo, al igual que la RNAP II de levaduras (Fu *et al.*, 1999) y arqueas (Ebright, 2000), donde la extremidad pequeña se correspondería a la subunidad β , y la subunidad β' a la extremidad de mayor tamaño (Figura 2A) (Borukhov & Nudler, 2003). Ambas subunidades (β y β') generan así un amplio canal interno que alberga el centro activo en la parte posterior de la subunidad β' , caracterizada por una secuencia muy conservada NADFDGD (Asn-Ala-Asp-Phe-Asp-Gly-Asp), cuyos aspárticos coordinan un catión Mg²⁺ esencial para la actividad enzimática (Alhadid *et al.*, 2017; Zaychikov *et al.*, 1996). La subunidad β' presenta una región larga y muy conservada denominada

"hélice puente" que divide el canal interno en dos. El primero de ellos es el canal principal de unión al DNA, mientras que el segundo se corresponde con el canal secundario de entrada de nucleósidos 5'- trifosfato (NTPs) hacia el sitio activo (Figura 2B) (Murakami, 2013; Zhang *et al.*, 1999) o de pequeñas moléculas y/o proteínas que regulan la transcripción (Figura 2B) (Haugen *et al.*, 2008; Molodtsov *et al.*, 2018; Ross *et al.*, 2016). Las subunidades β y β ' de la RNAP interaccionan con las subunidades α que, pese a ser idénticas, una interacciona con la subunidad β mientras que la otra lo hace con la subunidad β ' (Zhang *et al.*, 1999). Cada una de las subunidades α contiene dos dominios independientes: el dominio N-terminal (αNTD), encargado de dimerizar y de facilitar el ensamblaje de la subunidad β y β ', y un dominio C-terminal (αCTD) que participa en la unión al DNA, de especial importancia en la regulación de determinados promotores. Ambos dominios están unidos mediante un conector flexible y desestructurado que permite llegar a las regiones promotoras (Browning & Busby, 2004).



Figura 2. Estructura del núcleo de la polimerasa de RNA. A) Estructura del núcleo de la polimerasa de RNA de *Thermus aquaticus*, tomada de Alhadid *et al.*, (2017). B) Representación esquemática del complejo de inicio de la transcripción, tomada de Cox *et al.*, (2012).

La unión del factor σ al núcleo enzimático de la RNAP para generar la holoenzima activa provoca una serie de cambios en las posiciones de todos los dominios estructurales, y el estrechamiento del canal generado por las subunidades β y β '. En la holoenzima, el factor σ se encuentra extendido, con amplias interacciones a lo largo de una de las caras de la RNAP a través de las subunidades β y β ' (Borukhov & Nudler, 2003). El área total de contacto entre el factor σ y el núcleo abarca una superficie de más de 8200 Å² (Murakami *et al.*, 2002).

Los factores σ participan en las diferentes etapas del inicio de la transcripción empezando con el reconocimiento y la unión de la holoenzima al promotor. Por tanto, los factores σ actúan como reguladores maestros de la expresión génica en bacterias (Davis *et al.*, 2017). Su número varía de unas especies a otras según la estabilidad del

ambiente, siendo mayor en organismos expuestos a constantes fluctuaciones, probablemente para favorecer una respuesta rápida y eficaz. Según su estructura y función, los factores σ se clasifican en dos grandes familias, la familia σ^{54} (también llamada RpoN) que suele estar implicada en la regulación de la expresión de genes relacionados con el metabolismo del nitrógeno y en la respuesta al estrés (Zhang & Buck, 2015), y la familia σ^{70} que engloba a la mayoría de los factores σ e incluye a los factores σ primarios o de "housekeeping" encargados de la expresión de la mayoría o de todos los genes esenciales en el crecimiento vegetativo (Feklístov *et al.*, 2014; Paget & Helmann, 2003). Durante esta tesis se ha abordado el estudio de un subgrupo de factores σ pertenecientes a la familia σ^{70} , por lo que a continuación se detallará en profundidad las principales características de dicha familia.

Esta familia está constituida por el factor σ primario y por un número variable de factores σ alternativos que pueden ser clasificados en cuatro grupos filogenéticos que comparten estructura y, con frecuencia, función (Davis *et al.* 2017; Helmann, 2002; Lonetto *et al.*, 1992).

- Grupo I: es el grupo más complejo y abundante. Comprende al factor σ⁷⁰ de *E. coli* (con un peso molecular de ~70 kDa) y sus ortólogos (denominados σ^A). Este grupo incluye a los factores σ primarios o de "*housekeeping*" (Campbell *et al.*, 2002; Murakami *et al.*, 2002).
- **Grupo II:** comprende factores σ similares al grupo I en términos de organización estructural pero no relacionados con la expresión de genes implicados en el crecimiento celular (Paget & Helmann, 2003) sino con la adaptación a estreses como la ausencia de nutrientes u otros estreses relacionados con la fase estacionaria. A este grupo pertenece el factor σ de *E. coli* σ^{s} , responsable de la expresión de genes durante la fase estacionaria (Paget, 2015).
- Grupo III: son factores que activan regulones implicados en la respuesta general a estrés, estructura flagelar, quimiotaxis y procesos del desarrollo como la formación de endosporas (Davis *et al.* 2017; Paget, 2015).
- Grupo IV: es el grupo más numeroso y heterogéneo de la familia σ⁷⁰ y engloba a los factores σ-ECF (*extracytoplasmic function*). La mayoría de los miembros de este grupo responden a las señales del medio extracitoplásmico permitiendo dar una respuesta fisiológica apropiada a la variación de las condiciones ambientales. Este grupo de factores se abordará en profundidad en el apartado I.1.3.

El alineamiento de la secuencia de los factores de la familia σ^{70} permitió establecer cuatro regiones conservadas (región 1, 2, 3, 4) que, a su vez, pueden

dividirse en subregiones en base a su función (Davis *et al.*, 2017; Gribskov & Burgess, 1986; Helmann & Chamberlin, 1988). Estas regiones fueron validadas posteriormente a partir de los estudios estructurales con el factor σ^{70} de *E. coli* (Malhotra *et al.*, 1996) y de los datos obtenidos con el factor σ^A de *Thermus aquaticus* (Campbell *et al.*, 2002). Así se definió la existencia de cuatro dominios para los factores σ primarios, unidos mediante conectores de longitud y secuencia variable (Davis *et al.*, 2017; Murakami, 2015).

- Región 1: la región amino terminal de los factores σ está dividida en dos segmentos conservados denominados 1.1 y 1.2. La subregión 1.1 solo está presente en los factores σ primarios mientras que la región 1.2 aparece también en el grupo II de los factores σ alternativos (Figura 3). Diferentes estudios permitieron determinar que la subregión 1.1, cuya superficie está cargada negativamente, bloquea la unión con el DNA interaccionando con las regiones 2 y 4, cargadas positivamente (Dombroski *et al.*, 1992; Dombroski *et al.*, 1993; Schwartz *et al.*, 2008). En cambio, la región 1.2 se une a la región discriminadora de la cadena no molde afectando a la velocidad de disociación de los complejos abiertos (Haugen *et al.*, 2006; Haugen *et al.*, 2008).
- Región 2: es la región más conservada de los factores σ y está dividida en cuatro subregiones (2.1, 2.2, 2.3 y 2.4) (Figura 3). Esta región tiene un papel fundamental en el reconocimiento, unión a la región promotora y separación de las cadenas de DNA. Ensayos de mutagénesis establecieron que la región 2.1 es necesaria para la interacción con el núcleo de la RNAP (Lesley & Burgess, 1989; Shuler et al., 1995). La región 2.2, compuesta por una hélice, interacciona con el núcleo de la RNAP y podría mediar una serie de cambios conformacionales necesarios para la discriminación de las diferentes cadenas de DNA (Callaci & Heyduk, 1998; Marr & Roberts, 1997; Sharp et al., 1999). La región conservada 2.3 está implicada en la apertura de las cadenas de DNA y presenta una serie de aminoácidos aromáticos conservados que participan en la unión a la región - 10 del promotor (descrito en el apartado siguiente) y en la estabilización del inicio de la transcripción (Helmann & Chamberlin, 1988; Helmann & DeHaseth, 1999). La región 2.4 está implicada en el reconocimiento del hexámero - 10 del promotor (Daniels et al., 1990; Kahn & Ditta, 1991; Siegele et al., 1989; Waldburger et al., 1990; Zuber *et al.*, 1989) y está muy conservada en los factores σ primarios que reconocen secuencias consenso, pero mucho menos en factores σ alternativos que presentan una región - 10 mucho más variable.



Figura 3. Representación esquemática de los dominios conservados de los cuatro grupos de factores σ . Las regiones conservadas aparecen indicadas por un mismo color. Modificada de Souza *et al.*, (2014).

- Región 3: el análisis de su secuencia indica que esta región se encuentra muy conservada en los factores σ principales, y poco o nada conservada, e incluso ausente en los factores alternativos (Figura 3). Consta de tres subregiones diferenciadas, la región 3.0, constituida por una hélice α que es perpendicular a la región 2.4 y se une a la región - 10 extendida en el surco mayor a través de los residuos de su superficie. Esta interacción es crucial para la transcripción de una serie de promotores que presentan regiones - 35 y - 10 diferente a la consenso y parece estimular la transcripción, incrementando y estabilizando la formación de complejos (Bown et al., 1999; Haugen et al., 2008; Hook-Barnard et al., 2006; Mitchell et al., 2003). La región 3.1, situada en la proximidad del primer nucleótido, presenta un dominio similar a los motivos de unión a DNA de tipo hélice-giro-hélice que parece participar en la interacción con la RNAP (Helmann & Chamberlin, 1988; Sharp et al., 1999). En cambio, la región 3.2, que se encuentra en el núcleo catalítico de la holoenzima, bloquea el canal de salida del RNA e interacciona con el extremo 5' del RNA contribuyendo al escape de la región promotora y a la disociación del factor σ (Kulbachinskiy & Mustaev, 2006; Petushkov *et al.*, 2017; Severinov et al., 1994).
- Región 4: esta región participa en el reconocimiento de la región 35 del promotor (véase siguiente apartado) y está constituida por dos subregiones separadas por un conector de longitud y secuencia variable (Figura 3). La subregión 4.1 está muy conservada en todos los factores σ mientras que la subregión 4.2 solo está conservada en factores σ primarios y menos, en factores σ alternativos. Una gran cantidad de ensayos genéticos, bioquímicos y estructurales indican que en el

reconocimiento de la región - 35 está implicada una hélice-giro-hélice de la subregión 4.2 que interacciona con ambas cadenas de DNA. La unión a la región - 35 puede favorecer que se establezcan contactos con la subunidad α de la RNAP unida al denominado elemento "UP" (véase siguiente apartado). Dicha interacción podría ser remodelada por la interacción con factores de transcripción (Haugen *et al.*, 2008; Lonetto *et al.*, 1998; Rhodius & Busby, 2000).

I.1.2 Los promotores dependientes de la familia σ^{70}

Las regiones específicas de DNA localizadas aguas arriba del inicio de la transcripción reconocidas por la holoenzima se denominan promotores. Estas secuencias pueden estar localizadas en una región intergénica situada aguas arriba del gen de interés, o en el interior de un gen adyacente, y su transcripción puede ser monocistrónica (cuando afecta a un solo gen) o policistrónica (a más de un gen) (Ross & Gourse, 2009). Los promotores contienen elementos críticos cuya modificación provoca una pérdida significativa de la eficacia del inicio de la transcripción. De forma general, los promotores constan de dos elementos críticos, separados por una distancia óptima. La naturaleza de estos elementos determina el tipo de subunidad o por la que serán reconocidos, la cantidad de RNA producido, y es probable que también la respuesta a los reguladores transcripcionales. Para una secuencia promotora dada, cambios en la temperatura, concentración de sal o de otros solutos pueden afectar a la unión y al reconocimiento de dicho promotor por la RNAP holoenzima (Ruff *et al*., 2015). Como se ha mencionado anteriormente, en las bacterias se han descrito dos grandes familias de factores σ , la familia σ^{70} y la familia σ^{54} que, entre otros aspectos, se distinguen por la arquitectura de los promotores que reconocen. Dado que en este trabajo se estudia la regulación de promotores dependientes de un subgrupo de la familia σ^{70} , a continuación se describirán los aspectos más relevantes de dichos promotores, basados principalmente en estudios realizados con *E. coli*.

El análisis de las secuencias promotoras de la familia σ^{70} revela varios elementos conservados que determinan el reconocimiento por la RNAP asociada al factor σ primario, σ^{70} (Figura 4). De entre estas regiones destaca la región -35 cuya secuencia consenso es TTGACA y la región - 10 (TATAAT) localizadas 35 y 10 pb aguas arriba, respectivamente, del inicio de la transcripción (+1), y separadas por una región de secuencia variable de unas 17 pb (Feklistov, 2013; Mitchell *et al.*, 2003; Shimada *et al.*, 2014). Se ha propuesto que las secuencias promotoras localizadas aguas arriba de la región - 10 (incluida parte de ésta) participan en los pasos iniciales

de la unión de la RNAP, así como en los cambios conformacionales que siguen a dicha unión; para algunos promotores, los elementos localizados aguas abajo de la región - 10 favorecen la estabilización del complejo abierto y determinan su vida media. Así, el elemento "UP" es una secuencia localizada aguas arriba de la región - 35, entre las posiciones - 40 a - 60, rica en pares AT, que es reconocida por el dominio C-terminal de las subunidades α de la RNAP a través del surco menor y que consta de dos módulos, distal y proximal, cada uno de los cuales interacciona con un dominio Cterminal de las subunidades α (Ross *et al.*, 2001) (Figura 4). El elemento Z, localizado entre las bases - 24 y - 18, interacciona con la subunidad β de la RNAP (residuos 40-45) (Yuzenkova et al., 2011). Se desconoce si la interacción es específica de secuencia o simplemente es específica de ciertos espaciadores (Figura 4). La región - 10 extendida (TGn), localizada aguas arriba de la región - 10, participa incrementando la actividad del complejo a través de contactos con la región $\sigma^{3.0}$ de los factores σ (Barne et al., 1997), quizás aumentando la vida media del complejo (Haugen et al., 2006) (Figura 4). La región discriminadora, localizada entre la región - 10 y el sitio de inicio de la transcripción, es una secuencia de 6-8 pb con un alto contenido en GC que está implicada en la estabilidad y en la regulación de la vida media de los complejos transcripcionales interaccionando con la región $\sigma^{1.2}$ del factor σ (Haugen et al., 2006) (Figura 4). Recientemente, se ha determinado una nueva región denominada CRE (core recognition element), localizada entre las bases - 4 a + 2, que es reconocida por la subunidad β de la RNAP cuando ya se ha formado la burbuja de transcripción y participa incrementando la vida media del complejo (Zhang et al., 2012) (Figura 4).



Figura 4. Región promotora reconocida por la RNAP unida al factor σ^{70} **.** Las bolitas representan nucleótidos de la cadena molde. X= A o T. Debajo se indica la posición respecto al sitio de inicio de la transcripción. Modificada de Ruff *et al.*, (2015).

La unión de la holoenzima a estas regiones genera un complejo de reconocimiento o complejo cerrado (RP_c) en el que el DNA permanece como doble cadena (Figura 5). Aunque la formación del RP_c es específica, la unión de la RNAP en esta etapa es reversible y por lo tanto inestable, lo que implica que in vitro puede ser revertida mediante la adición de DNA inespecífico o heparina (un polianión). Las interacciones entre la RNAP y las secuencias promotoras desencadenan una serie de cambios en ambas biomoléculas conocidos como isomerización. Esto genera un complejo abierto (RP_o) en el cuál la doble cadena de DNA es separada desde la región - 10 hasta más allá del inicio de la transcripción (aproximadamente 13 pb) gracias a la unión del factor σ a la cadena no molde (Figura 5). Así, se forma una burbuja de transcripción inicial flexible, que permite acomodar la cadena molde en el sitio activo de la polimerasa. Este complejo es estable para la síntesis de RNA y, por tanto, para la formación del complejo inicial de transcripción (RP_{ITC}) (Figura 5). Los nucleótidos entran al centro catalítico a través del canal secundario mientras que la salida del transcrito se hace a través del canal de salida. Para facilitar esta salida se requiere un cambio conformacional que provoca el desplazamiento de la región 3.2 del factor σ , así como de su extremo C-terminal. A menudo, el inicio de la transcripción no culmina en la elongación del transcrito, sino que se generan RNAs abortivos de 2-12 nucleótidos (Figura 5). Estos transcritos cortos son liberados del núcleo de la RNAP generándose de nuevo el complejo abierto e iniciando el proceso de transcripción. La síntesis con éxito de un producto de RNA de una longitud mayor permite la entrada en la fase de elongación. En esta fase se produce la liberación del factor σ mientras que el núcleo de la RNAP se desplaza en la dirección del gen promoviendo el desenrollamiento del DNA y favoreciendo la polimerización de nucleótidos al transcrito inicial (Alhadid et al., 2017; Davis et al., 2017; Gries et al., 2010; Hook-Barnard & Hinton, 2007; Marchetti et al., 2017) (Figura 5).



Figura 5. Detalle de las etapas que constituyen el inicio de la transcripción. Modificado de Alhadid *et al.*, (2017).

I.1.3 Los factores σ-ECF

Durante este trabajo se han estudiado algunos factores σ alternativos de tipo ECF, por lo que a continuación se explicarán con más detalle las principales características de este diverso grupo, así como su regulación.

La familia de factores σ -ECF fue clasificada dentro de un subgrupo de la familia σ^{70} gracias a los estudios con el factor σ^{E} de *Streptomyces coelicolor* que revelaron una gran similitud con otras proteínas descritas como reguladores positivos en diversos procesos celulares, tales como CarQ de *Myxococcus xanthus* (Lonetto *et al.*, 1994), objeto de estudio en este trabajo. Análisis filogenéticos basados en la secuencia y en el contexto genómico han permitido clasificar los factores σ -ECF en hasta 56 familias principales y 32 minoritarias, formando el grupo más numeroso y diverso de factores σ (Huang *et al.*, 2015; Jogler *et al.*, 2012; Sineva *et al.*, 2017; Staron *et al.*, 2009). El número de factores σ -ECF codificados por un genoma puede variar enormemente, desde ninguno en bacterias intracelulares obligadas como clamidias, hasta 115 en planctomicetos como *Gemmata obscuriglobus* y 118 en la mixobacteria *Plesiocystis pacifica*, el organismo con el mayor número (Jogler *et al.*, 2012). Algunos subgrupos están ampliamente extendidos en bacterias, lo que sugiere un origen ancestral, aunque la gran mayoría de estos factores σ -ECF suelen formar parte

del regulón de genes que ellos mismos controlan, cuando este regulón es pequeño. Estudios en promotores específicos sugieren que dentro de una misma familia, los factores σ -ECF a menudo reconocen varios promotores, aunque este intercambio es mucho más reducido entre diferentes familias de factores σ -ECF (Kim *et al.*, 2016; Staron *et al.*, 2009). Las bacterias tienden a presentar una gran variedad de factores σ -ECF indicando que hay una presión evolutiva para favorecer la especificidad de estos factores σ -ECF (Rhodius *et al.*, 2013). No obstante, una misma especie puede presentar dos o más factores σ -ECF relacionados con un mismo proceso, como sucede en *Bacillus subtilis*, donde los regulones activados por estos factores σ -ECF presentan un alto grado de solapamiento (Eiamphungporn & Helmann, 2008; Huang *et al.*, 1998; Mascher *et al.*, 2007). Por lo tanto, la especificidad de los factores σ -ECF está regulada por el grado de reconocimiento de los promotores, por la disponibilidad del factor σ -ECF y por el estado fisiológico celular y el del medio.

En cuanto a su estructura, los factores σ -ECF presentan normalmente un tamaño inferior a 200 aminoácidos y una arquitectura de dominios caracterizada por la presencia de solo dos de los cuatro dominios típicos de la familia σ^{70} , los dominios 2 y 4. Estos dos dominios son suficientes tanto para el reconocimiento de la secuencia promotora como para la unión al núcleo de la RNAP. La región de reconocimiento del elemento - 35 (región 4.2) presenta cierta similitud con la de otros grupos de factores σ . Sin embargo, no ocurre lo mismo con la secuencia de reconocimiento del elemento - 10 (región 2.4), lo que explica que las secuencias promotoras reconocidas por los factores σ -ECF estén parcialmente conservadas en la región - 35 pero no en la región - 10, cuyo consenso resulta ser más laxo (Lonetto *et al.*, 1994). En los últimos años se ha avanzado mucho en el conocimiento de los promotores dependientes de factores σ -ECF, estableciéndose un cierto consenso para las regiones promotoras. Así, se ha determinado una secuencia conservada 5' -AAC-3' en la región - 35 y una secuencia 5'-CGT-3' para la región - 10, aunque ésta suele ser menos frecuente (Gaballa *et al.*, 2018; Helmann, 2002; Mascher *et al.*, 2007; Staron *et al.*, 2009).

La regulación de los factores σ -ECF puede estar modulada por diferentes mecanismos a nivel de la transcripción, traducción o de recambio proteico. Sin embargo, la gran mayoría de los factores σ -ECF están regulados negativamente por unas proteínas específicas denominadas factores anti- σ que bloquean la unión a la RNAP (Figura 6). Aunque existen factores anti- σ solubles y citoplasmáticos, un gran número de ellos suelen ser proteínas transmembranales que presentan al menos dos dominios, uno de unión al factor σ y otro sensor o de señalización que reconoce una gran variedad de estímulos diferentes generados dentro o fuera de la célula (Huang *et al.*, 2015; Staron *et al.*, 2009). Los factores σ y los factores anti- σ no están conservados a nivel de secuencia. Sin embargo, a menudo se cotranscriben garantizando que se mantengan niveles estequiométricos de ambos factores (Davis *et al.*, 2017; Paget, 2015). Algunos grupos de factores anti- σ se caracterizan por presentar un motivo de unión a zinc conservado denominado ZAS (<u>z</u>inc binding <u>a</u>nti-<u>s</u>igma) en su extremo N-terminal con la secuencia consenso (HisX₃CysX₂Cys, siendo x cualquier aminoácido) (Campbell *et al.*, 2007). Estudios estructurales y funcionales revelan que al menos las tres principales clases conocidas de factores anti- σ con motivos ZAS varían en su sensibilidad redox y en su papel en la regulación de los factores σ (Sineva *et al.*, 2017).

En general, en ausencia de una señal específica extracitoplásmica, los factores anti- σ secuestran al factor σ -ECF en la membrana, impidiendo su unión al núcleo de la RNAP (Figura 6). En presencia de dicha señal, el factor anti- σ sufre un cambio conformacional o se degrada, liberando al factor σ -ECF, que se une a la RNAP e inicia la expresión de su propio operón, así como la de genes de respuesta al estímulo desencadenador. Por norma general, los factores σ -ECF están sometidos a autorregulación positiva, garantizando de este modo su efecto activador mientras el estímulo prevalece. Una vez que el estímulo cesa, el factor anti- σ asegura el fin de su acción promotora de la transcripción (Davis *et al.*, 2017; Paget, 2015). En los últimos años, la comprensión de los mecanismos de inhibición de los factores σ se ha incrementado gracias a estudios estructurales de varios complejos σ /anti- σ (Campbell *et al.*, 2008; Hastie *et al.*, 2016; Maillard *et al.*, 2014).



Figura 6. Mecanismo de acción general de las parejas σ -ECF/anti- σ .

Con frecuencia, los cambios en el entorno tienen lugar de forma abrupta, por lo que la velocidad de respuesta determina el éxito o el fracaso de la misma. Así, una respuesta que se demore en el tiempo puede ser menos efectiva que otra más limitada, pero en la que el lapso de tiempo entre la percepción del estímulo y la activación de la respuesta sea más corto. Por ello, aunque existen ejemplos de factores σ -ECF activados a nivel transcripcional, en general, los mecanismos de activación de los factores σ -ECF tienen lugar a nivel de la proteína. La presencia en la célula de moléculas de los factores σ -ECF, aún en ausencia del estímulo disparador, permite generar una respuesta rápida a dicho estímulo. A continuación, se comentarán brevemente los principales mecanismos de acción de las parejas σ -ECF/anti- σ descritos hasta la fecha.

Cascada de proteólisis del factor anti-σ anclado a membrana

Este mecanismo tiene lugar en dos de las parejas σ -ECF/anti- σ más estudiadas, σ^{E} /RseA de *E. coli* y σ^{W} /RsiW de *B. subtilis*. Un exceso de proteínas mal plegadas para la primera pareja o un estrés en la envoltura celular para la segunda, activan una cascada de proteólisis, primero a nivel extracitoplasmático y luego a nivel citoplasmático que provocan la liberación del factor σ -ECF. Por último, el factor anti- σ es degradado totalmente por la proteasa ClpXP (Chaba *et al*, 2007; Heinrich *et al.*, 2009; Hizukuri & Akiyama, 2012; Ho & Ellermeier, 2012; Li *et al.*, 2009).

✓ Cambios conformacionales del factor anti-σ

Este es el caso de las parejas $\sigma^R/RsrA$ de *Streptomyces coelicolor* y $\sigma^E/ChrR$ de *Rhodobacter sphaeroides*. La respuesta celular al estrés redox en *S. coelicolor* está bajo el control del factor σ^R , cuya pareja anti- σ es una proteína citoplasmática, RsrA. En su estado reducido, RsrA se une al zinc e interacciona con σ^R , inhibiendo su actividad transcripcional. El estrés redox induce la formación de un puente disulfuro intramolecular en el factor anti- σ RsrA, que provoca un cambio en la conformación de la proteína, liberando al factor σ^R (Kang *et al.*, 1999; Li *et al.*, 2003; Rajasekar *et al.*, 2016). De forma similar, el factor σ^E de *R. sphaeroides*, implicado en la respuesta a estrés oxidativo, es secuestrado por ChrR, que en presencia de oxígeno singlete sufre un cambio conformacional que provoca la liberación del factor σ -ECF (Anthony *et al.*, 2004; Greenwell *et al.*, 2011; Newman *et al.*, 2001).

Estudios recientes han demostrado que los factores anti- σ de membrana también pueden ser regulados por cambios conformacionales. En la bacteria *Cupriavidus metallidurans*, la unión de cobalto o níquel por parte del sensor

extracelular CnrX provoca cambios conformacionales que son transmitidos al factor anti- σ de membrana CnrY que, por último, libera al citoplasma el factor σ -ECF CnrH (Pompidor *et al.*, 2009; Trepreau *et al.*, 2011a; Trepreau *et al.*, 2011b).

✓ Cascada de interacciones proteína-proteína

En *E. coli*, la regulación del transporte de citrato férrico está mediada por la pareja Fecl/FecR. En ausencia de citrato férrico, Fecl se mantiene inactivado por su factor anti- σ FecR, situado en la membrana interna. La activación de Fecl es llevada a cabo por una cascada de interacciones proteína-proteína que se inicia cuando FecA, una proteína anclada en la membrana externa, une citrato férrico y transmite la señal a FecR/Fecl. Una vez libre, Fecl es reclutada por la RNAP para su unión al promotor del operón *fec*, lo que provoca un aumento en la expresión de los genes de captación de hierro. Este mecanismo de activación parece estar conservado para otros factores σ de tipo Fecl presentes en otras bacterias (Braun & Mahren, 2005; Brooks & Buchanan, 2008).

✓ Intercambio de pareja

Diversos estudios han permitido establecer un modelo para la respuesta general al estrés, GSR (general <u>s</u>tress <u>r</u>esponse), en α -proteobacterias. Las proteínas implicadas en este proceso son: el factor σ^{G} , el factor anti- σ NepR y el regulador de respuesta PhyR. PhyR consta de dos dominios: un dominio C-terminal encargado de recibir la señal, susceptible de fosforilación, y un dominio N-terminal que muestra similitud con el factor σ^{G} . En ausencia de fosforilación, el dominio tipo ECF de PhyR está plegado sobre el dominio C-terminal y NepR secuestra al factor σ^{G} . Cuando se produce el estrés, PhyR es activado mediante fosforilación e interacciona con NepR, liberando así al factor σ^{G} , que puede entonces asociarse con la RNAP y activar los promotores diana (Francez-Charlot *et al.*, 2015; Paget, 2015).

✓ Regulación a través del dominio C-terminal del factor σ

Existe un grupo ampliamente distribuido de factores σ -ECF, carentes de factor anti- σ , que se caracterizan por presentar una prolongación muy conservada en su extremo C-terminal. En trabajos realizados en *B. licheniformis* y *Rhodobacter sphaeroides* se observó que versiones truncadas en este dominio provocaban una hiperactividad del factor σ -ECF, mientras que la eliminación total de esta prolongación provocaba la completa pérdida de la actividad promotora. De estos datos se dedujo que esta extensión C-terminal es crucial tanto para la activación como para la inactivación de estos factores σ (Sineva *et al.*, 2017; Wecke *et al.*, 2012). Además, en
Mycobacterium tuberculosis se ha descubierto que el factor σ^J presenta un dominio Cterminal denominado SnoaL_2 (similar a la superfamilia de factores de transporte nuclear 2) que parece modular la estructura del factor σ^J y su unión al promotor o a la RNAP (Goutam *et al.*, 2017). Por otro lado, recientemente se ha identificado un subgrupo de factores σ -ECF, también sin un factor anti- σ aparente, que presentan una pequeña extensión C-terminal (CRD) rica en residuos de cisteína. Entre estos factores destaca CorE, que presenta un dominio CRD que funciona a modo de factor anti- σ percibiendo el estado redox del cobre para activar o inactivar al factor σ (Gómez-Santos *et al.*, 2011) y CorE2, activado por cadmio o zinc, que regula genes próximos que cifran una bomba de protones y una metaloenzima (Marcos-Torres *et al.*, 2016).

✓ Sistemas de 2 componentes

Además, existen otros factores σ -ECF que no poseen ni factor anti- σ ni una extensión en su extremo C-terminal, pero en cuyo entorno genético aparecen genes que codifican quinasas de serina/treonina. Estos datos sugieren que este grupo de factores σ -ECF podrían estar regulados mediante mecanismos de fosforilación de proteínas, siendo probablemente la diana de la fosforilación el propio factor σ -ECF (Jogler *et al.*, 2012). En *M. xanthus*, organismo objeto de estudio en este trabajo, existen al menos 6 posibles factores σ -ECF en cuya proximidad se localizan genes que cifran quinasas de serina/treonina (Abellón-Ruiz *et al.*, 2014).

I.1.4 Reguladores transcripcionales

Como se ha comentado anteriormente, el inicio de la transcripción está modulado por diferentes factores tales como la secuencia promotora, los factores σ o el grado de compactación del cromosoma bacteriano (Browning & Busby, 2004). Otro factor a tener en cuenta son los factores transcripcionales que modulan la acción de promotores específicos (Browning & Busby, 2016). Algunas de estas proteínas controlan un número elevado de genes, mientras que otras solo regulan uno o dos. Estos factores suelen regular la expresión de genes en respuesta a señales del medio, por lo que su actividad debe modularse mediante la unión a pequeños ligandos, modificaciones covalentes o la unión a otras proteínas reguladoras, o por su concentración en la célula (Browning & Busby, 2004). Los factores transcripcionales pueden actuar activando o reprimiendo el inicio de la transcripción a través de su unión a regiones promotoras o mediante su interacción con la RNAP.

I.1.4.1 A nivel del promotor

Cuando la unión de un factor transcripcional a un promotor incrementa la actividad del promotor dicho factor transcripcional recibe el nombre de activador. En la mayoría de los organismos procariotas, la activación transcripcional de un promotor suele depender de un solo regulador, aunque en algunos promotores depende de dos o más. Aquellos promotores regulados por un solo factor transcripcional se dice que están sujetos a una activación simple. En función del mecanismo empleado, los activadores se dividen en diferentes clases:

• **Clase I**: incluye a los activadores que se unen a una región operadora localizada aguas arriba de la región - 35 y que reclutan a la RNAP por contactos con el extremo C- terminal de la subunidad α (Figura 7A). Este tipo de activación ocurre en promotores que requieren el reclutamiento de la RNAP porque uno o más elementos del promotor no son óptimos para su reconocimiento por la RNAP. A esta clase pertenece el regulador global CRP (<u>cyclic AMP receptor protein</u>) o CAP (catabolite activator protein) en el promotor *lac* de *E. coli,* cuya actividad depende de adenosín monofosfato cíclico (AMPc). El homodímero de CRP se une al DNA en las posiciones - 61 y - 62 aguas arriba del inicio de la transcripción y contacta con el extremo C-terminal de la subunidad α estabilizando la formación del complejo abierto sin provocar un cambio conformacional (Browning & Busby, 2016; Lawson *et al.*, 2004; Lee *et al.*, 2012; Shimada *et al.*, 2011; Zheng *et al.*, 2004).



Figura 7. Mecanismos de activación simple. A) El activador se une aguas arriba de la región - 35 y contacta con α CTD de la RNAP. B) El activador se une a secuencias cercanas a la región - 35 e interacciona con el dominio σ^4 y/o α NTD. C) El activador se une a la región situada entre los elementos - 35 y - 10. Tomada de Browning & Busby, (2016).

• **Clase II**: incluye activadores que se unen a un operador situado en la región - 35 (Figura 7B). Una vez unido, determinadas regiones del activador reclutan a la RNAP mediante interacción directa con el dominio 4 del factor σ , con el dominio Nterminal de la subunidad α o con otros componentes de la RNAP. En algunos promotores, la interacción entre el activador y la RNAP es más favorable en el complejo abierto, lo que promueve el paso de complejo cerrado a abierto. En este tipo de promotores, el dominio C-terminal de la subunidad α es incapaz de unirse a los elementos UP y por ello se une a la región situada aguas arriba del activador. Estas características permiten que los activadores de clase II puedan funcionar junto a los activadores de clase I. Uno de los activadores de clase II mejor estudiados es la proteína cI del bacteriófago λ que activa su propio promotor, P_{RM}, para el mantenimiento de la lisogenia. La proteína cI se une en forma de dímero a la región - 35 e interacciona con el dominio 4 de σ^{70} . Estos contactos favorecen la formación del complejo abierto en P_{RM} (Browning & Busby, 2016; Lee *et al.*, 2012; Li *et al.*, 1997).

• **Clase III**: los activadores de esta clase se unen a un operador situado entre la región - 35 a - 10. Debido al poco espacio entre estos dos elementos, se cree que esta activación se basa en la distorsión del DNA entre las regiones - 35 y - 10 (Figura 7C). Los casos mejor conocidos corresponden a los miembros de la familia MerR, implicados en el control de la expresión de genes que determinan resistencia al mercurio o bombas de flujo cuya actividad se desencadena por xenometabolitos (Brown *et al.*, 2003; Browning & Busby, 2016; Lee *et al.*, 2012; Heldwein & Brennan, 2001).

Entre los diferentes mecanismos de activación también cabe citar un mecanismo indirecto, basado en anular la acción de un represor mediante una proteína activadora que actúa a modo de antirrepresor. Este es el caso de la proteína Ler de *E. coli*, que activa la expresión de genes de virulencia, bloqueando la acción represora de la proteína H-NS (<u>h</u>istone-like <u>n</u>ucleoid-<u>s</u>tructuring) (Bustamante *et al.*, 2001; Sperandio *et al.*, 2000) o el de las proteínas CarS-CarA de *M. xanthus* sobre las que se hablará a continuación.

En ocasiones, la activación de un promotor requiere dos o más factores transcripcionales. Se conocen cuatro mecanismos principales por los cuales puede ocurrir la activación.

 A través del reposicionamiento de un factor transcripcional previamente unido a una región promotora por un segundo regulador que reposiciona al factor transcripcional inicial para favorecer la activación de la transcripción (Figura 8A1) o modifica el plegamiento del DNA (Figura 8A2).

- 2. Mediante la interacción de ambos factores transcripcionales con las regiones promotoras o con la RNAP para activar la transcripción (Figura 8B).
- A través de la unión cooperativa de dos o más factores transcripcionales para que se produzca la activación transcripcional (Figura 8C).
- La acción represora de determinadas proteínas es anulada por otro factor transcripcional, favoreciendo la activación de la transcripción por un regulador primario (Figura 8D).



Figura 8. Activación de la transcripción por dos reguladores transcripcionales. A1) Reposicionamiento por un activador secundario. A2) Reposicionamiento por un cambio en el plegamiento del DNA. B) Mediante interacción directa al promotor o a la RNAP. C) A través de la unión cooperativa de los dos factores transcripcionales. D) Mediante la acción antirrepresora de un factor transcripcional. Tomada de Browning & Busby, (2004).

El inicio de la transcripción también puede estar regulado negativamente por proteínas denominadas represoras, que pueden actuar según distintos mecanismos:

• Por impedimento estérico, debido a que la proteína represora se une a los elementos promotores en las proximidades de las regiones - 10 y - 35 y bloquea la unión de la RNAP (Figura 9A). Uno de los represores mejor estudiados es el represor LacI de *E. coli*, que, en ausencia del inductor alolactosa, se une a la región operadora

en forma de homotetrámero siendo capaz de unirse a dos regiones operadoras. Este represor está formado por un dominio N-terminal de unión al DNA separado por un conector de un dominio regulador y de un dominio C-terminal que participa en la tetramerización. La unión a las diferentes regiones operadoras provoca el bloqueo de la transcripción y la formación de un bucle en el DNA (Müller-Hill, 1998). Otro ejemplo de proteínas represoras que bloquean la unión de la RNAP son las proteínas parálogas CarA y CarH de *M. xanthus*, que han sido estudiadas por el grupo de investigación donde se ha realizado esta tesis. Estas proteínas reprimen el promotor P_B , responsable de la expresión de los genes carotenogénicos. En oscuridad, ambas proteínas se unen a un operador que solapa con la región - 35 de su promotor diana (el promotor P_{B}) bloqueando el acceso de la RNAP y la transcripción (López-Rubio *et al.*, 2004; López-Rubio et al., 2002; Ortiz-Guerrero et al., 2011). La luz elimina la represión mediada por CarA, induciendo la expresión de CarS, un antirrepresor que estructuralmente imita al DNA operador reconocido por CarA (y CarH); al interaccionar físicamente con el dominio de unión al DNA de CarA, elimina su unión al DNA y permite la expresión génica (León et al., 2010; Navarro-Avilés et al., 2007). Aunque tanto CarA como CarH presentan su dominio N-terminal de unión al DNA fusionado, a través de una región conectora, a un dominio C-terminal de unión a B₁₂ (en su forma adenosilcobalamina o AdoB12) solo CarH depende de AdoB12 para realizar su acción reguladora (Ortiz-Guerrero et al., 2011; Pérez-Marín et al., 2008). En la oscuridad, la AdoB₁₂ favorece la tetramerización de CarH, necesaria para su unión al operador y, por tanto, para llevar a cabo la represión. En la luz, la fotólisis de la AdoB₁₂ provoca el desmantelamiento de los tetrámeros de CarH a monómeros, lo que provoca su liberación del promotor, permitiendo así la expresión génica. El descubrimiento de CarH y de su modo de acción por el grupo en el gue se ha realizado esta tesis ha llevado a la identificación de un nuevo tipo de fotorreceptores, muy extendido en bacterias, que utiliza la AdoB12 como molécula sensora de luz (Elías-Arnanz et al., 2011; Jost et al., 2015; Ortiz-Guerrero et al., 2011; Padmanabhan et al., 2017).

• Por la unión de las proteínas represoras a zonas distales del promotor provocando un cambio en la estructura del DNA que bloquea la transcripción (Figura 9B). Un ejemplo de este tipo es el represor GalR, que se une como dímero a dos operadores situados 60 pb aguas arriba y 53 pb aguas abajo del inicio de la transcripción. La unión de ambos dímeros genera un tetrámero que provoca la formación de un bucle en el DNA, en presencia de la histona HU, y el bloqueo de la transcripción (Browning & Busby, 2016; Lewis & Adhya, 2015).

• Por la unión de un anti-activador (Figura 9C), como es el caso de CytR, que se

une a una región operadora del promotor *deo*P2 de *E. coli* e interacciona con el activador (CRP) unido a ambos lados, impidiendo la interacción de CRP con la RNAP y, por tanto, la transcripción (Shin *et al.*, 2001).



Figura 9. Mecanismos de represión simple. a) Represión mediante impedimento estérico. b) Represión mediante cambio conformacional del promotor. c) Represión por inactivación de un activador. Tomada de Browning & Busby, (2016).

• En algunos casos, las proteínas represoras no impiden la unión a la RNAP sino que bloquean la formación de complejos intermedios (Ansari *et al.*, 1995; Tagami & Aiba, 1999). Además, se ha descubierto un regulador que inhibe la transcripción compitiendo con el elemento UP (Quinones *et al.*, 2006).

I.1.4.2 A nivel de la interacción con la RNAP

La simplicidad y ubicuidad de la regulación por factores σ ha restado atención a otros numerosos factores que se unen a la RNAP y regulan la transcripción. Aunque la mayoría de estos factores actúan a nivel de la elongación y la terminación de la transcripción (como los factores Nus), un pequeño número actúa a nivel del inicio de la transcripción (Browning & Busby, 2016).

Entre estos reguladores transcripcionales de unión a la RNAP destaca RbpA (<u>RNA p</u>olymerase <u>b</u>inding <u>p</u>rotein), presente solamente en actinobacterias, se une a la RNAP estabilizando los complejos cerrados. RbpA interacciona con la RNAP a través de dos de sus dominios, el dominio central que interacciona con la subunidad β de la

RNAP y el dominio C-terminal que se une al dominio 2 de los factores σ . Por lo tanto, RbpA regula positivamente promotores dependientes del factor σ mayoritario (Bortoluzzi *et al.*, 2013; Hu *et al.*, 2012; Hubin *et al.*, 2015, 2017).

Otro regulador transcripcional ampliamente estudiado es DksA, que regula promotores de síntesis de proteínas ribosomales uniéndose al canal secundario de la RNAP. DksA desempeña diversas acciones, entre las que destaca la regulación negativa de promotores ribosómicos y positiva de promotores relacionados con la biosíntesis de aminoácidos. Además, DksA puede actuar de forma sinérgica a ppGpp, un indicador del estado nutricional celular, que se une a la RNAP y disminuye el tiempo de vida media de los complejos abiertos (Lemke *et al.*, 2009; Paul *et al.*, 2005; Paul *et al.*, 2004).

De particular interés, por su relación con este trabajo, es la familia CarD CdnL TRCF (PF02559), cuyos miembros interaccionan con la RNAP a través de un mismo tipo de dominio (Cayuela *et al.*, 2003; Garcia-Moreno *et al.*, 2010; Padmanabhan et al., 2001; Srivastava et al., 2013; Stallings et al., 2009). El prototipo de la familia, la proteína CarD, es una proteína de unión al DNA objeto de estudio en esta tesis que actúa como regulador global de la transcripción formando un complejo con CarG, otra proteína reguladora que es incapaz de unirse al DNA (Elías-Arnanz et al., 2010; García-Heras et al., 2013; García-Heras et al., 2009; Padmanabhan et al., 2001; Peñalver-Mellado et al., 2006). El complejo CarD-CarG regula la expresión de un buen número de factores σ-ECF de M. xanthus (Abellón-Ruiz et al., 2014). A esta familia también pertenece el dominio RID del factor TRCF (Transcription Repair Coupling Factor) que participa en la reparación del DNA acoplado a la transcripción, o la proteína CdnL, que a diferencia de CarD y TRCF, carece de un dominio de unión al DNA (García-Moreno et al., 2010). Recientemente se ha visto que CdnL se une a la subunidad β de la RNAP y favorece la formación de complejos abiertos de transcripción estables (Bae et al., 2015; Gallego-García et al., 2017; Gallego-García et al., 2014) (Ver apartado I.2.5).

También se ha observado que algunos factores transcripcionales secuestran a la RNAP, lo que provoca una disminución del número de holoenzimas disponibles, mientras que otros, como los de algunos fagos, inhiben la RNAP bacteriana para favorecer su propia RNAP (Browning & Busby, 2016).

Además de estos tipos de regulación, existe un mecanismo adicional que implica modificaciones de bases específicas o cambios en la secuencia. Estas modificaciones pueden alterar la afinidad de un factor transcripcional o modular la unión o afinidad de la RNAP. Un ejemplo de este tipo de modificaciones químicas es la metilación del DNA que puede provocar cambios en el transcrito como se da en los operadores de los genes *pap* y *agn43* de *E. coli*, que son metilados por la <u>D</u>NA <u>a</u>denina <u>m</u>etilasa (Dam) y provoca la pérdida de unión del represor OxyR (Browning & Busby, 2016; Sánchez-Romero *et al.*, 2015; van der Woude & Henderson, 2008). Otro ejemplo a destacar sería el factor transcripcional GcrA de *Caulobacter crescentus*, que se une a la región σ_2 y dirige la unión de la RNAP- σ^{70} a secuencias promotoras consenso metiladas para promover la transcripción (Fioravanti *et al.*, 2013; Haakonsen *et al.*, 2015).

I.2 Myxococcus xanthus

I.2.1 Características generales de las mixobacterias

Las mixobacterias son bacterias Gram negativas, no patógenas, aerobias, de forma bacilar y de tamaño relativamente grande (0'7-1'2 µm de ancho y 3-12 µm de largo). Suelen vivir en el suelo, nutriéndose de material en descomposición o de otros microorganismos, aunque también podemos encontrarlas en otros hábitats como aguas dulces o ambientes marinos (Dawid, 2000; lizuka *et al.*, 1998). La mayoría de especies son típicamente mesófilas. *M. xanthus*, por ejemplo, presenta una tasa máxima de crecimiento a 34-36 °C, sufriendo daños celulares y dejando de crecer a temperaturas superiores a 40 °C, o siguiendo un crecimiento reducido a 14 °C. En cambio, las mixosporas son completamente estables a 50-60 °C (Janssen *et al.*; 1977).

Las mixobacterias forman parte de la división δ de las Proteobacterias, junto a otros dos grupos bacterianos: las bacterias reductoras de sulfato, como *Desulfovibrio*, y el género *Bdellovibrio* (Shimkets & Woese, 1992). En función de sus características morfológicas y fisiológicas (Reichenbach, 1999) y el análisis del RNA 16S (Shimkets & Woese, 1992; Sproer *et al.*, 1999) han sido clasificadas en tres subórdenes: *Sorangineae*, *Cystobacterineae* (que incluye a *M. xanthus*, como la especie más estudiada), y *Nannocystineae*.

En general, el tamaño del genoma de las mixobacterias es inusualmente grande en relación al de otros procariotas (por ejemplo, 9.14 Mb para *M. xanthus* y unas 13 Mb para *Sorangium cellulosum*) y con un alto contenido en G+C (67-71%). Esta expansión genómica podría ser debida a: (1) eventos de transferencia génica horizontal de un número elevado de genes implicados en la producción de metabolitos secundarios que probablemente son importantes para su comportamiento depredador (Yhang; & Higgs, 2014); (2) una duplicación génica masiva que afecta específicamente

a genes implicados en la señalización intercelular, la percepción sensorial de pequeñas moléculas, y el control integrado de la transcripción, que podría haber sido clave para la evolución del estilo de vida multicelular de las mixobacterias (Goldman *et al.*, 2006; Kaiser & Dworkin, 2008).

Las mixobacterias, consideradas erróneamente como hongos en sus orígenes, exhiben un comportamiento social único en el mundo procariótico que les permite generar unas complejas estructuras multicelulares, denominadas cuerpos fructíferos, en situaciones de falta de nutrientes. Este proceso requiere de una absoluta coordinación celular, lograda gracias a diversas señales intercelulares y a una compleja red reguladora que provoca la activación de genes específicos implicados en la diferenciación de las mixosporas dentro de los cuerpos fructíferos (Kaiser, 2008; Lee Kroos, 2007). La movilidad y el comportamiento depredador típico de este grupo de bacterias también depende de la comunicación entre células a través de señales. Dentro de las mixobacterias, *M. xanthus* se ha convertido en el modelo de estudio para entender la genética y los mecanismos moleculares implicados en este singular proceso.

I.2.2 Ciclo de vida de *M. xanthus*

Su ciclo de vida comprende dos fases, una fase vegetativa y otra de desarrollo multicelular que ensalzan la naturaleza social de este organismo. Ambos procesos son mediados por movimientos coordinados de un conjunto de células. *M. xanthus* presenta dos formas de movilidad independiente de flagelos, el movimiento aventurero (A) y el movimiento social (S), ambos coordinados en el espacio y en el tiempo (Muñoz-Dorado *et al.*, 2016).

El movimiento aventurero (A) impulsa el movimiento de células individuales de los bordes de la colonia para la exploración de nuevos entornos, dejando atrás senderos que pueden ser seguidos por otras células (Islam & Mignot, 2015). Para explicar este movimiento aventurero se han propuesto dos modelos. El primero de ellos propone que la fuerza motora deriva de la extrusión de mucílago, a través de estructuras embebidas en el polo posterior de la envoltura celular, que al hidratarse propulsaría la célula hacia delante (Wolgemuth *et al.*, 2002; Yu & Kaiser, 2007). El segundo modelo se base en la existencia de complejos de adhesión focal bacterianos denominados aparatos Agl-Glt anclados a la membrana que contactan con el citoesqueleto dispuesto de forma helicolidal a lo largo de la célula. Se ha sugerido que estos complejos motores se desplazan sobre el citoesqueleto en dirección contraria al movimiento de la célula, generando en su avance una fuerza igual y opuesta en esta

hélice, pero se desconoce cómo este movimiento se traduce en el movimiento celular (Faure *et al.*, 2016; Nan *et al.*, 2013; Nan & Zusman, 2016). En cambio, el movimiento social (S) está mediado por pili de tipo IV localizados en el polo celular adelantado (Islam & Mignot, 2015), fibrillas y lipolisacáridos de la pared celular (Arnold & Shimkets, 1988; Bowden & Kaplan, 1998; Hartzell *et al.*, 2008; Wall & Kaiser, 1999) que promueven el movimiento celular mediante sucesivos ciclos de extensión y retracción (Skerker & Berg 2001; Sun *et al.*, 2000). La presencia de exopolisacárido (EPS) secretado cubriendo a las células favorece la interacción de los pili de tipo IV con las células vecinas permitiendo el movimiento de grupos celulares (Mercier & Mignot, 2016).

Ambos tipos de movimiento pueden cambiar la dirección modificando la polaridad del eje celular. Estos cambios son reversibles y están controlados por tres proteínas: MgIA, MgIB y RomR. La proteína citoplásmica MgIA es una GTPasa de tipo Ras que actúa como un activador central de los sistemas de movilidad cuando se le une GTP en el polo celular adelantado. RomR es una proteína reguladora de respuesta que recluta MgIA-GTP al polo adelantado. MgIB es una proteína de tipo GAP (<u>G</u>TP hydrolysis <u>a</u>ctivating <u>p</u>rotein) que colocaliza con RomR en el polo rezagado, desde donde promueve la conversión de MgIA-GTP a MgIA-GDP, evitando su acumulación en este polo (Bulyha *et al.*, 2011). La localización de estas proteínas y, por tanto, la dirección del movimiento está regulada por Frz (Mercier & Mignot, 2016).



Figura 10. Ciclo de vida de *M. xanthus***.** A) Fase vegetativa. B) Fase de desarrollo multicelular. Tomada de Muñoz-Dorado *et al*., (2016).

A nivel nutricional, *M. xanthus* es una bacteria compleja que hidroliza proteínas y ácidos nucleicos para obtener aminoácidos esenciales y fosfato, no viéndose favorecido su crecimiento por la presencia de mono- o disacáridos en el medio. Los aminoácidos son utilizados para la glucogenogénesis, mediante la vía de Embden-Meyerhof, y como fuente de energía, al ser convertidos en acetyl-CoA para su posterior entrada en el ciclo de los ácidos tricarboxílicos (TCA) (Dworkin, 1962). En presencia de nutrientes, las células se mueven de forma coordinada, formando biofilms multicelulares conocidos como enjambres (Figura 10). Cuando los enjambres están en contacto con otros microorganismos (presa), miles de células provocan su lisis mediante la liberación de proteasas, nucleasas, lipasas, glucanasas y antibióticos, para el aprovechamiento de las presas como nueva fuente de nutrientes (Berleman & Kirby, 2009; Pérez *et al.*, 2016). Por este comportamiento han sido consideradas como microodepredadoras (Reichenbach & Höfle, 1993; Singh, 1947).

En la ausencia de nutrientes, como se ha comentado anteriormente, las células se mueven colectivamente e inician un proceso de agregación, intercambio de señales químicas extracelulares, así como de señales de contacto físico para formar los cuerpos fructíferos (Kaiser, 2004; Mauriello, 2010) (Figura 10). En los cuerpos fructíferos, las mixosporas están fuertemente unidas y, cuando las condiciones del medio vuelven a ser apropiadas, germinan y dan lugar a las formas vegetativas móviles y metabólicamente activas reanudando la fase vegetativa (Elías-Arnanz & Murillo, 1991a; 1991b). La formación de mixosporas permite la supervivencia de las células en condiciones adversas, mientras que la germinación de un elevado número de células localizadas en un mismo lugar origina un nuevo enjambre que garantiza la supervivencia del grupo.

Para ambos procesos, depredación y desarrollo, *M. xanthus* usa una mezcla heterogénea de polisacáridos, proteínas y DNA extracelular que secreta al medio, y establece complejas redes sociales, así como, una comunicación célula-célula (Muñoz-Dorado *et al.*, 2016).

I.2.3 Respuesta a la luz en *M. xanthus*

Las bacterias han desarrollado diversas estrategias para protegerse de los daños fotooxidativos. Así, *M. xanthus*, que en condiciones de oscuridad presenta una tonalidad amarilla (debido a la producción de un pigmento hidrosoluble no carotenoide) (Meiser *et al.*, 2006), en la luz adquiere una tonalidad rojiza debido a la síntesis de carotenoides (Martínez-Laborda *et al.*, 1990) (Figura 11). Los carotenoides protegen del daño ocasionado por las especies reactivas de oxígeno que se generan por la

exposición continuada a la luz.



Figura 11. Síntesis de carotenoides en *M. xanthus* en respuesta a la luz.

A lo largo de más de 30 años de estudio, el grupo en el que se ha realizado este trabajo, ha identificado un buen número de genes, tanto estructurales como reguladores, implicados en la respuesta a la luz. Estos genes se encuentran distribuidos en seis loci independientes (Elías-Arnanz *et al.*, 2008, 2011): los operones *carQRS* y *carDG*, los genes reguladores *ihfA* y *carF* (Fontes *et al.*, 2003; Galbis-Martínez *et al.*, 2008), el superoperón *carB-carA* (que agrupa genes estructurales y reguladores) y *crtlb*, el único gen estructural que no se encuentra agrupado en el superoperón *carB-carA* (Figura 12).



Figura 12. Modelo de organización y acción de los genes implicados en la respuesta a la luz de *M. xanthus.* Se muestran en rojo los genes reguladores y en negro los genes estructurales. Las flechas indican regulación positiva y las truncadas acciones negativas sobre promotores u otras proteínas. Modificada de Elías-Arnanz *et al.*, (2011).

En la compleja red reguladora de la respuesta a la luz destaca la interacción entre los productos de los genes *carQ*, un factor σ -ECF y *carR*, su factor anti- σ , que se expresan junto con carS (que determina un antirrepresor) en el operón carQRS. En la oscuridad, CarR secuestra a CarQ en la membrana, impidiendo así que ejerza su acción como factor σ . En presencia de luz azul, la fotoexcitación de la protoporfirina IX presente en la membrana genera oxígeno singlete (una especie muy reactiva del oxígeno) que, de algún modo, es detectado por proteína membranal CarF, provocando la inactivación de CarR, la liberación de CarQ y la activación de su propio promotor P_{QRS} y el de *crtlb*, P₁ (Elías-Arnanz *et al.*, 2011) (Figura 12). Curiosamente, la presencia de cobre en el medio también conduce a la inactivación de CarR, aunque de una manera independiente de CarF (Moraleda-Muñoz *et al.*, 2005). La acción de CarQ requiere del complejo CarD-CarG y, en PQRS, también de la proteína IHF (integration host factor) (Moreno et al., 2001), un factor arquitectónico que modula el empaquetamiento del cromosoma, la recombinación y la regulación de la transcripción (Browning et al., 2010) (Figura 12). Por otro lado, CarS, expresado en la luz a partir del operón carQRS contrarresta la acción ejercida por los represores parálogos CarA y CarH cifrados a partir del superoperón carB-carA. Ambas proteínas se unen a su operador en P_B en oscuridad bloqueando la transcripción. Pese a la similitud de ambas proteínas en su unión a la vitamina B₁₂ a través de su dominio C-terminal, solo CarH depende de ella (en su forma adenosilcobalamina o AdoB₁₂) para ejercer su actividad, como ya se ha comentado anteriormente (Figura 12) (Elías-Arnanz et al., 2008; Elías-Arnanz et al., 2011; Jost et al., 2015; Ortiz-Guerrero et al., 2011; Padmanabhan et al., 2017).

I.2.4 Factores σ en *M. xanthus*

M. xanthus presenta un número de factores σ inusualmente elevado (>50), en comparación con otros organismos procariotas, incluso teniendo en cuenta el gran tamaño de su genoma (Goldman *et al.*, 2006). Estos factores σ incluyen, en otros, el factor σ primario (σ^A), σ^{54} , σ^D y ~45 σ -ECF (Abellón-Ruiz *et al.*, 2014). Para una información más detallada sobre factores σ de *M. xanthus* puede consultarse Kroos & Inouye, (2008).

SigA (MXAN_5204) presenta una secuencia similar a *E. coli* σ^{70} salvo por la presencia de 100 aminoácidos extra en el extremo N-terminal (Inouye, 1990). Es el factor σ más abundante (Biran & Kroos, 1997) que se encarga del inicio de la transcripción durante el ciclo de vida vegetativo (Komano *et al.*, 1987) aunque también

se ha visto que activa un promotor que regula el desarrollo (Hao *et al.*, 2002). SigA es muy similar a *E. coli* σ^{70} y a *B. subtilis* σ^{43} en el reconocimiento de la región - 35, mientras que muestra diferencias en el reconocimiento de la región – 10, que presenta mayores diferencias respecto a la secuencia consenso y es más rica en pares GC (Biran & Kroos, 1997; Hao *et al.*, 2002).

SigD (MXAN_2957) muy parecido a SigA, compartiendo el 35% y 34% de identidad con el factor σ^A de *M. xanthus* y con el factor RpoS de *E. coli*, respectivamente (Ueki & Inouye, 1998). SigD carece de la región 1.1 que solo se encuentra en factores sigma primarios y homólogos a RpoS (Gruber & Bryant, 1997), y de gran parte de la región 3.1, que se conserva en todos los factores σ salvo en factores σ -ECF. SigD participa en la transcripción durante la fase estacionaria del crecimiento vegetativo, aunque también parece estar implicado en la transcripción de genes durante el desarrollo (Ueki & Inouye, 1998).

SigB (MXAN_3357), SigC (MXAN_6209) y SigE (MXAN_5870) son factores que presentan una alta similitud con los factores SigA y SigD, caracterizados por una secuencia consenso denominada caja RpoH, encontrada en factores σ específicos de choque térmico (Nakahigashi *et al.*, 1995), aunque se ha comprobado que estos factores no participan en dicha regulación (Apelian & Inouye, 1990; Ueki & Inouye, 2001), sino que participan en el desarrollo multicelular de *M. xanthus*. Así, SigB es esencial para la maduración de las esporas (Apelian & Inouye, 1990), SigC está implicada en la fase temprana de la formación de cuerpos fructíferos (Apelian & Inouye, 1993), y SigE regula tanto la fase vegetativa como el inicio del desarrollo multicelular (Ueki & Inouye, 1998).

SigF (MXAN_0785) dirige la transcripción de genes no identificados que afectan a la movilidad social y a la formación de cuerpos fructíferos (Ueki *et al.*, 2005).

SigG (MXAN_2473) es similar al factor σ FliA de *E. coli*, que está implicado en la biosíntesis flagelar (Liu & Matsumura, 1995). Forma parte de un operón que codifica los componentes de un sistema de secreción tipo III (Buttner & Bonas, 2002), por lo que se especula que quizás sea el encargado de regular la expresión de estos genes.

MXAN_1210, MXAN_2204 y MXAN_4535 presentan similitudes en sus secuencias con factores σ , aunque se desconoce la función que llevan a cabo.

Factores σ **tipo ECF:** el análisis del genoma predice hasta 45 factores σ -ECF, la mayoría de los cuales suelen estar acoplados traduccionalmente a genes que cifran proteínas transmembranales que pueden presentar un dominio ZAS típico de factores anti- σ . Asociados a algunos de estos factores σ -ECF suelen encontrarse una serie de proteínas quinasas de serina/treonina que podrían regular su función (Abellón-Ruiz *et al.*, 2014; Goldman *et al.*, 2006).

RpoN (MXAN_1061): como en muchas bacterias, *M. xanthus* presenta un solo factor σ^{54} (Goldman *et al.*, 2006) que, curiosamente, es esencial para la viabilidad celular (Garcia-Moreno *et al.*, 2009; Iniesta *et al.*, 2012; Keseler & Kaiser, 1997).

I.2.5 La familia de proteínas CarD-CdnL

I.2.5.1 El complejo regulador CarD-CarG

I.2.5.1.1 El operón carD-carG

Los genes *carD* y *carG* se cotranscriben a partir de un promotor dependiente del factor σ mayoritario, σ^A (Nicolás *et al.*, 1996; Peñalver-Mellado *et al.*, 2006). Se expresan de forma activa durante toda la fase vegetativa, presentando un incremento durante el inicio del desarrollo multicelular, seguido de un descenso pronunciado entre cinco y seis horas después (Nicolás *et al.*, 1994). La proteína CarD es necesaria para la regulación de casi el 1% de los genes en fase vegetativa (Galbis-Martínez *et al.*, 2004), de una manera que depende de CarG (Abellán-Martínez, 2006; Peñalver-Mellado *et al.*, 2006). No obstante, ni CarD ni CarG son esenciales para la viabilidad celular (Elías-Arnanz *et al.*, 2010).

El análisis de las secuencias de los genomas bacterianos depositados en las bases de datos reveló que *carD y carG* solo están presentes en el suborden *Cystobacterinea*e, que incluye a *M. xanthus*, *S. auriantiaca* y *A. dehalogenans* (Elías-Arnanz *et al.*, 2010; García-Heras *et al.*, 2009; Peñalver-Mellado *et al.*, 2006).

I.2.5.1.2 La proteína CarD

CarD es una proteína (316 aminoácidos; 34 kDa) con una arquitectura de dominios novedosa. Su dominio C-terminal, intrínsecamente desordenado, es similar a las proteínas eucarióticas de tipo HMGA. Dicho dominio está compuesto por una región acídica de unos cuarenta aminoácidos (rica en aspártico y glutámico) seguida de un dominio de unión al DNA formado por cuatro "ganchos AT" constituidos por la secuencia Arg-Gly-Arg-Pro embebida en un tramo, de unos 20 aminoácidos, rico en prolina y residuos básicos (Nicolas *et al.*, 1996; Padmanabhan *et al.*, 2001) (Figura 13). En comparación con las proteínas HMGA eucarióticas, CarD presenta un gancho AT adicional, y la región acídica se encuentra antes de los ganchos AT y no después como en las proteínas HMGA (Figura 13). Estos dominios tipo HMGA, tanto en CarD como

en las proteínas HMGA eucarióticas, presentan una alta afinidad y especificidad de unión por el surco menor del DNA de tramos ricos en pares AT y en fase (Padmanabhan *et al.*, 2001; Thanos & Maniatis, 1992). De hecho, CarD es capaz de reconocer *in vitro* el sitio de unión típico de la proteína HMGA1a humana en el promotor de IFN- β , y la proteína HMGA1a es capaz de unirse al promotor P_{QRS} de *M. xanthus*, y de reemplazar funcionalmente al dominio C-terminal de CarD *in vivo* (García-Heras *et al.*, 2009).





Además, CarD presenta un dominio N-terminal de unos 180 aminoácidos, con una estructura secundaria y terciaria bien definida (Nicolás *et al.*, 1996; Padmanabhan *et al.*, 2001). Este dominio, ausente en las proteínas HMGA eucarióticas, es esencial para su función *in vivo* (Cayuela *et al.*, 2003) y está constituido por dos subdominios: el subdominio N-terminal, que interacciona con la subunidad β de la RNAP y consigo mismo generando dímeros en solución (García-Heras *et al.*, 2009; Garcia-Moreno *et al.*, 2010; Peñalver-Mellado *et al.*, 2006), y el subdominio C-terminal, que interacciona con CarG (Figura 13) (Bernal-Bernal *et al.*, 2015). El dominio N-terminal de CarD define una familia de proteínas, denominada CarD_CdnL_TRCF (PF02559) por incluir, además de dicho dominio de CarD: (i) el dominio TRCF-RID de interacción con la subunidad β de la RNAP de los factores TRCF (<u>t</u>ranscription-<u>r</u>epair <u>c</u>oupling <u>f</u>actors), implicados en la reparación de errores durante la transcripción (Deaconescu *et al.*, 2006); (ii) las proteínas CdnL (<u>CarD N-t</u>erminus <u>l</u>ike), que se asemejan en tamaño (160 a 200 aminoácidos) y secuencia al dominio N-terminal de CarD y que carecen, por tanto, del dominio de unión al DNA (Bae *et al.*, 2015; Cayuela *et al.*, 2003; Elías-Arnanz *et al.*, 2010; Gallego-García *et al.*, 2014; Garcia-Moreno *et al.*, 2010; Stallings *et al.*, 2009). Mientras que las proteínas CdnL están ampliamente distribuidas en bacterias (una de las primeras en estudiarse fue la de *M. xanthus*), CarD solo se encuentra en mixobacterias y, exclusivamente, en el suborden Cystobacterineae, al que pertenece *M. xanthus* y especies relacionadas, como *Stigmatella auriantiaca* (con CarD_{Sa}) o *Anaeromyxobacter dehalogenans* (con CarD_{Ad}). Aunque los dominios N-terminal de CarD, CarD_{Sa} y CarD_{Ad} son muy similares, CarD_{Sa} presenta un gancho AT menos que CarD (Cayuela *et al.*, 2003) y CarD_{Ad} carece de los característicos ganchos AT y, en su lugar, presenta un alto contenido en K/A/P (lisina/alanina/prolina), típico de la región C-terminal (CTR) de la histona H1. Y, sin embargo, el dominio HMGA de CarD puede ser sustituido por el dominio tipo histona H1 de CarD_{Ad} sin que pierda funcionalidad en *M. xanthus*, quizás debido a la naturaleza intrínsecamente desestructurada de ambos tipos de dominios (García-Heras *et al.*, 2009).

Desde un punto de vista evolutivo, cabe contemplar al menos dos posibles hipótesis para la aparición de CarD. Dado que los dominios HMGA son abundantes en eucariotas y que en procariotas están confinados a un suborden específico de las mixobacterias (Cystobacterineae), CarD (y CarD_{sa}) podría haber surgido de la fusión de un módulo procariótico (similar a CdnL) con un módulo HMGA eucariótico adquirido mediante transferencia génica horizontal (Cayuela *et al.*, 2003; Elías-Arnanz *et al.*, 2010). Por otro lado, la existencia de un homólogo (CarD_{Ad}) con un dominio H1-CTR, en lugar del dominio HMGA, plantea una hipótesis alternativa. Varias especies bacterianas presentan dominios tipo H1-CTR, lo que ha llevado a postular un origen bacteriano para dicho dominio (Kasinsky *et al.*, 2001). De ser así, CarD_{Ad} podría haber surgido por la fusión de dos módulos procarióticos, y CarD por la adquisición de los "ganchos AT" por convergencia evolutiva a partir de un dominio ancestral de tipo H1 (Elías-Arnanz *et al.*, 2010; García-Heras *et al.*, 2009). En ese caso, la aparición de los "ganchos AT" debería haber ocurrido antes de que *M. xanthus* y *S. aurantiaca* se separaran, ya que están presentes en ambas especies.

I.2.5.1.3 La proteína CarG

CarG (170 aminoácidos, 19 kDa) no presenta similitud significativa con otras proteínas de las bases de datos, excepto por la presencia de un motivo (HEx₂Hx₂Gx₂HCx₄CxMx_{16/17}Cx₂C, siendo x cualquier aminoácido) típico de ciertas metaloproteasas dependientes de zinc (metzincinas) presentes en arqueas (y denominadas, por ello, arqueametzincinas), en la bacteria hipertermófila *Aquifex aeolicus*, y ampliamente distribuidas en eucariotas (Díaz-Perales *et al.*, 2005;

Peñalver-Mellado et al., 2006). En las metzincinas, la metionina localizada 20-40 aminoácidos después de la secuencia consenso HEx₂Hx₂Gx₂H/D participa en la formación de un giro tipo β1-4 denominado Met-turn (Gomis-Rüth, 2003), las tres histidinas unen un átomo de zinc, y el glutámico es esencial para la actividad proteolítica (Steele et al., 2000). La unión a zinc de las arqueametzincinas a través del motivo conservado ha sido confirmada a nivel estructural (Graef et al., 2012; Waltersperger et al., 2010). CarG muestra varios cambios en dicho motivo: la glicina conservada aparece reemplazada por un glutámico y, más importante, el glutámico clave para la actividad catalítica por una glutamina, por lo que CarG carece de actividad proteasa, aunque sí une zinc (Peñalver-Mellado et al., 2006). Estos cambios se observan también en las proteínas homólogas a CarG que, como CarD, están presentes solo en Cystobacterineae, aunque en la proteína CarG de A. dehalogenans (CarG_{Ad}), el glutámico implicado en la catálisis aparece sustituido por alanina, en vez de glutamina. La distribución filogenética de las argueametzincinas sugiere que este motivo surgió a partir de un ancestro común de arqueas y eucariotas y que su presencia en bacterias es probable que se deba a eventos de transferencia horizontal 2005). CarG habría divergido rápidamente de las (Díaz-Perales et al., arqueametzincinas tanto en secuencia (la similitud de secuencia es baja) como en función (pérdida de actividad proteasa) (Elías-Arnanz et al., 2010; Peñalver-Mellado et al., 2006).

Los ensayos espectrométricos y de ICP-A-ES (*Inductively Coupled Plasma-Atomic Emission Spectroscopy*) indican que CarG se une a dos átomos de zinc, uno de ellos a través de las tres histidinas de la secuencia consenso y el otro a través de sus cuatro cisteínas (Peñalver-Mellado, 2005; Peñalver-Mellado *et al.*, 2006). El análisis mutacional de las cisteínas ha determinado que estos residuos son esenciales para la estructura y la estabilidad de CarG (Elías-Arnanz *et al.*, 2010; García-Heras *et al.*, 2009; Peñalver-Mellado *et al.*, 2006), mientras que las mutaciones dobles en los residuos de histidina provocan un efecto mucho menos drástico (García-Heras, 2009; Peñalver-Mellado, 2005).

I.2.5.1.4 El complejo CarD-CarG

La interacción entre CarG y el dominio N-terminal de CarD genera un complejo estable (estequiometría CarD:CarG de 2:1) con una acción reguladora global en *M. xanthus* (Peñalver-Mellado *et al.*, 2006). Así, el complejo CarD-CarG está implicado en el control de la respuesta a la luz, el desarrollo multicelular y otros procesos que serán explicados detalladamente a lo largo de este trabajo. Muchos de estos procesos

celulares regulados por CarD-CarG se activan por factores σ -ECF que responden a determinados estímulos extracitoplásmicos (Abellón-Ruiz, 2012). A diferencia de CarD (y de la mayoría de los factores de transcripción), CarG carece de un dominio de unión al DNA. Consistente con esta observación, CarG solo se une al DNA *in vivo* e *in vitro* en presencia de CarD, por lo que debe ejercer su acción reguladora uniéndose indirectamente al DNA a través de su interacción con CarD (García-Heras *et al.*, 2009; Peñalver-Mellado *et al.*, 2006). Además de con el DNA, el complejo CarD-CarG interacciona, a través del dominio N-terminal de CarD, con la subunidad β de la RNAP (García-Heras, 2009; García-Heras *et al.*, 2013). Es posible que la acción reguladora del complejo ocurra de manera análoga a los "enhanceosomes" eucarióticos, en cuyo ensamblaje participan las proteínas HMGA, ya sea reclutando otros factores requeridos para la regulación de determinados promotores y/o estableciendo contactos con la maquinaria basal de transcripción (Elías-Arnanz *et al.*, 2010).

I.2.5.1.5 Funciones de carD-carG

I.2.5.1.5.1 Acción de carD y carG en la respuesta a la luz

Como se ha mencionado previamente, el complejo regulador CarD-CarG está implicado en la respuesta carotenogénica mediada por la luz azul. Los datos genéticos indican que el complejo se requiere directamente para la activación de los promotores P_{QRS} , P_{crtlb} e, indirectamente, para la activación de P_B (Figura 12). En P_{QRS} , un promotor que depende del factor σ -ECF CarQ, se ha demostrado que CarD se une a una secuencia constituida por dos tramos ricos en AT. En P_B , su efecto se ejerce a través de su acción positiva sobre la expresión de P_{QRS} , que conduce a la expresión de CarS (Elías-Arnanz *et al.*, 2010; Garcia-Heras *et al.*, 2013; Nicolás *et al.*, 1994; Padmanabhan *et al.*, 2001; Peñalver-Mellado *et al.*, 2006).

I.2.5.1.5.2 Acción de carD y carG en el desarrollo multicelular

El complejo CarD-CarG regula el desarrollo multicelular mediante la activación transcripcional de un grupo de genes dependientes parcial o totalmente de diferentes señales intercelulares que participan en la morfogénesis de los cuerpos fructíferos. O bien CarD (y CarG) es necesaria para que se produzcan estas señales, o participa en los mecanismos de activación génica que dependen de ellas (Nicolás *et al.*, 1994).

I.2.5.1.5.3 Acción de carD y carG en la fase vegetativa

Un estudio global mediante el uso del transposón Tn5 lac como sonda de promotores vegetativos en *M. xanthus* permitió estimar que en torno a un 1% de los genes está sujeto a regulación por CarD. Dicho estudio permitió identificar, entre otros, un locus cuya expresión se mostraba fuertemente regulada por CarD (posteriormente, se demostró que también lo estaba por CarG), al que se denominó ddvA (de CarDdependent very strong) (Galbis-Martínez, 2001; Galbis-Martínez et al., 2004; Peñalver-Mellado, 2005; Peñalver-Mellado et al., 2006). El análisis del sitio de inserción de Tn5 lac en ddvA reveló que el transposon se habría integrado interrumpiendo un gen (ddvA), acoplado transcripcional y traduccionalmente con el gen inmediatamente aguas arriba (ddvS), cuyo producto (DdvS) mostraba homología con factores σ-ECF. El estudio comparativo de la secuencia de DdvA puso de manifiesto la presencia de un dominio N-terminal de unos 70 aminoácidos similar a los dominios anti- σ de unión a zinc conocidos como ZAS (Abellón-Ruiz, 2012; Abellón-Ruiz et al., 2014), apuntando a que DdvS y DdvA constituyen una pareja σ -ECF/anti- σ , que ha sido objeto de estudio en esta tesis. Que la expresión de CarQ y CarR, otra pareja σ-ECF/anti-σ, esté sujeta también al control por CarD-CarG, llevó a estudiar si la expresión de otras parejas σ-ECF/anti- σ en *M. xanthus* (con ~45 factores σ -ECF) también dependen de CarD-CarG. Trece de las diecisiete parejas σ -ECF/anti- σ estudiadas, sorprendentemente, mostraron una dependencia total o parcial de CarD-CarG (Abellón-Ruiz, 2012; Abellón-Ruiz et al., 2014).

I.2.5.2 CdnL

CdnL (164 aminoácidos; 18,7 kDa), cuya secuencia se asemeja al dominio Nterminal de CarD (CarDNt), está constituida por dos dominios diferenciados: un dominio C-terminal, compacto y estructurado, y un dominio N-terminal, sensible a proteólisis (Figura 14A). A diferencia de CarD, CdnL es una proteína esencial para la viabilidad celular en *M. xanthus* (García-Moreno *et al.*, 2010) y, pese a la similitud de secuencia (Figura 14B), CdnL y CarDNt no son funcionalmente equivalentes. Así, la sustitución de CarDNt por CdnL genera una versión inactiva de CarD, probablemente porque CdnL es incapaz de interaccionar con CarG; y CarDNt no es capaz de suplir la falta de CdnL en su función esencial para la viabilidad celular a pesar de que, al igual que CdnL, CarDNt interacciona con la subunidad β de la RNAP. Por otro lado, mientras que CarD se ha implicado en la regulación de promotores dependientes de factores σ-ECF, CdnL parece participar como un activador de los promotores dependientes del factor σ mayoritario, tales como los promotores del RNA ribosómico (Gallego-García *et al.*, 2014). Aunque CdnL no se une directamente al DNA, pues carece de un dominio de unión al DNA, es capaz de estabilizar la formación de complejos abiertos de transcripción, al igual que su homólogo en *Mycobacterium tuberculosis* (Bae *et al.*, 2015; Srivastava *et al.*, 2013).



Figura 14. Comparación de las proteínas CarD y CdnL de *M. xanthus*. A) Comparación de dominios de las proteínas CarD y CdnL. B) Alineamiento de secuencia de CarDNt y CdnL.

I.3 Sistemas CRISPR-Cas

Las bacterias están expuestas a la entrada de material genético exógeno, debido a la infección por virus o a la incorporación de elementos genéticos móviles como plásmidos o transposones. Como consecuencia, han desarrollado un amplio rango de estrategias de defensa para favorecer la supervivencia celular, entre las que cabe mencionar (Koonin *et al.*, 2017; Plagens *et al.*, 2015):

- Modificaciones en la superficie de la bacteria debido a mutación, disminución de la expresión o bloqueo de receptores para fagos.
- Mecanismos de exclusión del fago a través de una infección previa que impide una infección por virus relacionados.

- iii. Sistemas de modificación-restricción que provocan el corte en secuencias específicas de DNA no metilado, como el DNA viral.
- iv. Defensa a través de los sistemas heredables CRISPR-Cas.
- v. Sistemas en los que participa la proteína pAgo, muy estudiada en eucariotas, que parece participar en la rotura de DNA mediada por DNA o RNA (Swarts *et al.*, 2015).
- vi. Sistemas abortivos, que anulan la infección de un virus induciendo la muerte de las células infectadas.

A continuación, y dado que parte de este trabajo doctoral versa sobre la regulación de un sistema CRISPR-Cas, se explican brevemente los aspectos más relevantes de estos sistemas de defensa.

I.3.1 Función de los sistemas CRISPR-Cas

Los sistemas CRISPR-Cas, muy en boga en la actualidad, fueron descubiertos hace unos 30 años, pero su función biológica fue una incógnita hasta mediados de la primera década del siglo XXI cuando tres grupos de investigación independientes relacionaron los CRISPR (*clustered regularly interspaced short palindromic repeats*) con secuencias presentes en bacteriófagos, sugiriendo una posible relación con la resistencia a virus (Bolotin *et al.*, 2005; Mojica *et al.*, 2005; Pourcel *et al.*, 2005). Poco después se demostró que, en efecto, los sistemas CRISPR-Cas constituyen un sistema de defensa heredable para proteger a las células procarióticas del ataque viral (Barrangou *et al.*, 2007; Barrangou & Horvath, 2017; Ishino *et al.*, 2018; Mojica & Montoliu, 2016; Mojica & Rodriguez-Valera, 2016).

Debido a la diversidad de estos sistemas CRISPR-Cas, su distribución en bacterias y su ubicuidad en arqueas, se ha propuesto que estos sistemas surgieron en un ancestro de arqueas y se han expandido a bacterias por transferencia horizontal (Makarova *et al.*, 2015; Makarova *et al.*, 2006).

I.3.2 Composición de los sistemas CRISPR-Cas

Pese a esta diversidad, los sistemas CRISPR-Cas presentan una estructura conservada compuesta por una región CRISPR y los genes *cas* (*CRISPR <u>as</u>sociated*).

I.3.2.1 La región CRISPR

El locus CRISPR está constituido por repeticiones palindrómicas cortas (20 a 50 pb) agrupadas y regularmente interespaciadas por secuencias cortas únicas de

material genético exógeno denominadas espaciadores (Figura 15) (Ishino *et al.*, 1987; Mojica *et al.*, 2000). En la mayoría de los casos, el locus CRISPR está situado en el cromosoma principal, pero también puede localizarse en plásmidos o megaplásmidos (Plagens *et al.*, 2015). El número de sistemas CRISPR-Cas, entre las especies bacterianas que los presentan, puede variar desde 1 a 18 en una misma célula (Ishino *et al.*, 2018). También es variable el número de repeticiones (normalmente <50, pero hay casos con hasta 587 repeticiones como en la myxobacteria *Haliangium ochraceum* (Marraffini & Sontheimer, 2010; Richter *et al.*, 2012), aunque para cada locus CRISPR la repetición suele tener la misma longitud y secuencia (Figura 15). Las repeticiones presentan secuencias cortas invertidas que pueden generar estructuras secundarias de tallo-bucle (*stem-loop*) (Richter *et al.*, 2012) o secuencias conservadas GAAA(C/G) en el extremo 3' (Sorek *et al.*, 2008), que juegan un papel importante en el procesamiento del transcrito de la región CRISPR (Richter *et al.*, 2012).





Por otro lado, las secuencias de los espaciadores son muy variables y de origen génico extracromosomal (fagos, plásmidos y otros elementos génicos móviles), aunque se han detetectado algunos espaciadores cuyas secuencias corresponden al propio genoma de la bacteria (Stern *et al.*, 2010). Los espaciadores son secuencias cortas (en torno a 40 pb, aunque pueden alcanzar 72 pb), cuyo tamaño para un mismo locus CRISPR varía mínimamente (1 o 2 pb) (Grissa *et al.*, 2007). La secuencia del genoma foráneo del que deriva el espaciador se denomina protoespaciador; en ocasiones, puede presentar una secuencia adyacente, conservada y corta, denominada PAM (*Protospacer-<u>a</u>ssociated <u>m</u>otif*) que es crucial para que las proteínas Cas diferencien el material genético propio del exógeno (Bhaya *et al.*, 2011; Van Orden *et al.*, 2017).

I.3.2.2 Los genes cas

Asociados al locus CRISPR se encuentran los genes *cas*, que cifran las proteínas Cas (*CRISPR <u>as</u>sociated*). Estas proteínas proporcionan la maquinaria

proteica requerida para que el sistema de inmunidad adquirida funcione. El número total de genes *cas*, así como su orientación y agrupamiento difiere entre los distintos sistemas CRISPR-Cas (Bhaya *et al.*, 2011), observándose grandes diferencias evolutivas entre las proteínas Cas de arqueas y bacterias (Makarova *et al.*, 2015).

Además de estos dos componentes principales, los sistemas CRISPR-Cas se caracterizan por presentar una región "líder" de 100-500 pb (generalmente rica en pares AT) localizada aguas arriba de la primera repetición. Dicha región puede incluir un promotor encargado de la transcripción de la región CRISPR y secuencias necesarias para la adquisición de nuevos espaciadores justo aguas abajo de la región líder, por lo que la posición de cada espaciador indica el orden temporal de adquisición (Bhaya *et al.*, 2011; Van Orden *et al.*, 2017).

I.3.3 Mecanismo de acción de los sistemas CRISPR-Cas

Los sistemas CRISPR-Cas constituyen un sistema inmune de defensa adquirida específico de secuencia, cuya acción se divide en tres fases características: adaptación, expresión e interferencia.

- La fase de adaptación consiste en el reconocimiento e incorporación del DNA invasor gracias a la actividad nucleasa de las proteínas Cas1 y Cas2, presentes en la mayoría de los sistemas CRISPR-Cas (Figura 16). El nuevo espaciador se integra aguas abajo de la secuencia líder, entre dos repeticiones, confiriendo memoria frente a nuevos ataques por el mismo material genético. Este mecanismo de integración es similar al de transposasas e integrasas virales (Arslan *et al.*, 2014; Nuñez *et al.*, 2015; Rollie *et al.*, 2015). En algunos sistemas CRISPR-Cas, la fase de adaptación requiere además de otras proteínas Cas (Bhaya *et al.*, 2011; Hille & Charpentier, 2016). Algunas proteínas Cas1 están fusionadas a una transcriptasa reversa (RT), cuya acción concertada permite adquirir material genético procedente de RNA (Silas *et al.*, 2016).
- 2. La fase de **expresión** consiste en la transcripción del locus CRISPR a partir de la región líder generando un precursor largo denominado pre-crRNA (Figura 16). Este precursor sufre un primer procesamiento en un sitio específico dentro de las repeticiones para generar los crRNAs, constituidos por una secuencia espaciadora flanqueada por secuencias parciales de las repeticiones. En algunos sistemas CRISPR-Cas, como los de tipo I o tipo III, se requiere de un procesamiento adicional para generar los crRNAs maduros que actuarán de guía frente a las secuencias exógenas invasoras (Bhaya *et al.*, 2011; Hille & Charpentier, 2016).



Figura 16. Mecanismo de acción de los sistemas CRISPR-Cas. Se representan de manera secuencial las fases de adaptación, biogénesis e interferencia.

3. Durante la interferencia, los crRNAs maduros bloquean una nueva invasión de ácidos nucleicos, asociándose a diferentes complejos efectores constituidos por varias proteínas Cas, o por proteínas multidominios, y guiándolos al material genético diana a degradar (Figura 16). En determinados sistemas CRISPR-Cas, el proceso de interferencia se basa en el reconocimiento de secuencias PAM asociadas a los protoespaciadores mientras que en otros la especificidad se la proporciona el crRNA maduro (Bhaya *et al.*, 2011; Hille & Charpentier, 2016).

I.3.4 Tipos de sistemas CRISPR-Cas

La clasificación de los sistemas CRISPR-Cas debería estar basada en sus relaciones evolutivas, pero debido a la divergencia en los genes *cas*, se ha generado una clasificación que incluye dos clases (I y II) en función de los genes que codifican los módulos que intervienen en la fase de interferencia. En los sistemas CRISPR-Cas de la clase I, que representan el 90% de los conocidos en bacterias y arqueas, participan complejos efectores formados por 4-7 proteínas Cas unidas con una

estequiometría desigual. El 10% restante pertenecen a la clase II, que requiere una sola proteína efectora multidominio exclusiva de bacterias (Ishino *et al.*, 2018; Makarova *et al.*, 2015; Shmakov *et al.*, 2017). A su vez, en cada clase se distinguen varios tipos. Así, la clase I incluye los tipos I, III y IV, mientras que la clase II incluye los II, V y VI. Además, cada tipo presenta múltiples subtipos (I: A-F y U; II: A-C; III: A-D; V: A-E y U; VI: A-C) en base a la organización del locus y la presencia de genes *cas* adicionales (Koonin *et al.*, 2017).

Clase I:

En los sistemas CRISPR-Cas de clase I, el complejo efector multiproteico que media el procesamiento y la interferencia puede ser de tipo Cascade (<u>C</u>RISPR-<u>associated complex for antiviral defence</u>) para los sistemas de tipo I, y Csm (<u>Mtube subtype cas</u>) o Cmr (<u>RAMP module cas</u>) para los sistemas de tipo IIIA y IIIB, respectivamente. Estos complejos efectores se basan en un esqueleto constituido por Cas5 y Cas7 (varios monómeros), ambas con un motivo RRM (<u>RNA recognition motif</u>) de unión a ácidos nucleicos perteneciente a la superfamilia RAMP (<u>Repeat-associated mysterious protein</u>) (Figura 17). Además de estas proteínas Cas, participan dos proteínas adicionales que se designan subunidad grande y subunidad pequeña, atendiendo a su diferente tamaño.



Figura 17. Composición y estructura tridimensional de los complejos efectores de los sistemas CRISPR-Cas de tipo I y III. a) y c) Composición y estructura de los módulos efectores de tipo I. b) y d) Composición y estructura de los módulos efectores de tipo III. Tomado de Makarova *et al.*, (2015).

En los sistemas de tipo III, la subunidad grande se corresponde con Cas10 (nucleasa de DNA de cadena doble para la fase de interferencia) mientras que en los de tipo I se corresponde con Cas8 (posible actividad nucleasa de DNA de cadena sencilla para la fase de interferencia) (Figura 17). La endonucleasa Cas6 parece ser la responsable del procesamiento del pre-crRNA a crRNAs en ambos casos, aunque solo

en algunos subitpos de tipo I forma parte del complejo efector; en los de tipo III, no suele encontrarse asociada al complejo y la actividad puede provenir en *trans* de otros sistemas dentro de la misma célula. A pesar de las diferencias en detalles estructurales entre ambos complejos efectores, su forma y arquitectura general son muy similares, lo que sugiere una evolución divergente a partir de un complejo efector ancestral (Koonin *et al.*, 2017; Makarova *et al.*, 2015).

- a) Tipo I: la evolución de estos sistemas CRISPR-Cas ha implicado el reclutamiento de proteínas adicionales a estos complejos efectores como es el caso de la proteína Cas3, una helicasa que tiene la capacidad de desenrollar DNA de doble cadena y dúplex RNA-DNA. A menudo, el dominio helicasa está fusionado a un dominio endonucleasa, que suele estar localizado en el extremo amino terminal de Cas3. Los diferentes subtipos que forman el tipo I difieren en la compleja estructura de estos módulos efectores.
- b) Tipo III: en estos sistemas el complejo efector se denomina Csm o Cmr y difieren en el gen que cifra la subunidad pequeña (*csm2* en los de tipo III-A o *cmr5* en los de tipo III-B). Estos sistemas presentan una composición y organización muy diversa debido a duplicaciones, deleciones o inserciones de dominios que pueden estar implicados en la fase de interferencia. Algunos subtipos como los III-A presentan las proteínas Cas1 y Cas2, necesarias para la integración de nuevos espaciadores, mientras que otros sistemas como los de tipo III-B aprovechan otros sistemas CRISPR-Cas presentes en el mismo genoma. Además se han identificado otros subtipos, los sistemas III-C y III-D, que también requieren la actuación de los genes *cas1* y *cas2* en *trans* (Makarova *et al.,* 2015). Recientemente se ha observado que, durante la fase de interferencia, el módulo efector reconoce el RNA diana provocando su rotura, y permite la escisión del DNA a través de la actividad nucleasa de Cas10 (Pyenson & Marraffini, 2017).
- c) Tipo IV: estos sistemas consisten en complejos formados por los productos de cas5, cas7 y csf1, que suelen encontrarse en plásmidos. Carecen de los genes cas1 y cas2, y no están en las proximidades de un locus CRISPR, por lo que se especula que podrían ser módulos móviles que usan crRNAs de otros loci CRISPR (Makarova et al., 2015).

Clase 2:

Estos sistemas CRISPR-Cas son menos abundantes y se basan en un módulo efector constituido por una sola proteína que presenta varios dominios catalíticos.

- a) Tipo II: es el más simple en cuanto al número de genes se refiere. Su seña de identidad es *cas9*, que determina una proteína multifunción con dominios nucleasa conservados (HNH y RuvC), que interviene en las fases de adaptación, procesamiento del pre-crRNA e interferencia. Además de *cas9*, estos sistemas presentan *cas1* y *cas2*, y la mayoría codifica un pequeño RNA trans-activador de crRNA (tracrRNA) que es complementario a las repeticiones del mismo locus. Este tracrRNA se une a las repeticiones de los pre-crRNAs, junto a la proteína Cas9, y es reconocido por la endorribonucleasa III celular que corta el pre-crRNA y genera una forma intermedia, que tras sufrir una nueva modificación genera un crRNA maduro (con un espaciador en el extremo 5' y una repetición en el extremo 3'). El híbrido tracrRNA:crRNA guía a la proteína efectora Cas9 para que introduzca un corte en el DNA exógeno (Hille & Charpentier, 2016; Ishino *et al.*, 2018; Makarova *et al.*, 2015; Plagens *et al.*, 2015).
- b) Tipo V: se caracterizan por la presencia de un gen *cas12* adyacente, dependiendo del subtipo, a *cas1*, *cas2* y al locus CRISPR. Como Cas9, Cas12 presenta un dominio nucleasa RuvC y, en el subtipo V-B, Cas12b actúa usando tracrRNA de manera análoga a Cas9. Sin embargo, en el subtipo V-A, basta con Cas12a (antes denominada Cpf1) para generar el pre-crRNA y cortar el DNA (Koonin *et al.*, 2017; Makarova *et al.*, 2015).
- c) Tipo VI: en estos sistemas destaca la presencia de la proteína Cas13 que, aunque carece del característico motivo nucleasa, presenta un par de dominios HEPN (<u>high eukaryotes and prokaryotes nucleotide-binding</u>) asociados a ribonucleasas (RNasas). El descubrimiento reciente de este tipo, que actúa frente a RNA, está proporcionando nuevas oportunidades para la explotación con fines aplicados de los sistemas CRISPR-Cas (Abudayyeh *et al.*, 2016; Ishino *et al.*, 2018; Koonin *et al.*, 2017).

I.3.5 Regulación de sistemas CRISPR-Cas

A pesar de los numerosos estudios acerca de los mecanismos y aplicaciones biotecnológicas de los sistemas CRISPR-Cas, apenas se conoce su regulación, que puede ocurrir a nivel de la región CRISPR o de los genes *cas* (Patterson *et al.*, 2017). En algunos casos, que se detallan a continuación, se ha podido observar regulación de la expresión de un sistema CRISPR-Cas:

- a) Se ha observado que Serratia y Pseudomonas, dos bacterias Gram negativas, activan la expresión de sus sistemas CRISPR-Cas en función de la concentración de señales extracelulares ("quorum sensing") secretadas al medio. Así, cuando la densidad celular es elevada, se promueve tanto la interferencia como la adaptación, a modo de prevención frente a un ataque por material genético exógeno (Høyland-Kroghsbo *et al.*, 2016; Patterson *et al.*, 2016).
- b) Sistemas de dos componentes como VicRK en *Streptococcus mutans* o BaeSR en *E. coli* activan la expresión de los genes *cas* al detectar cambios en la integridad de la membrana (Patterson *et al.*, 2017).
- c) Represión de la expresión de un sistema CRISPR-Cas a través de la unión de H-NS (<u>h</u>istone-like <u>n</u>ucleoid-<u>s</u>tructuring) (Westra et al., 2010) y/o de LRP (<u>l</u>eucine-responsive <u>r</u>egulatory <u>p</u>rotein) a regiones situadas aguas arriba y abajo del inicio de la transcripción (Medina-Aparicio et al., 2011).
- d) Regulación a partir de variaciones en las rutas metabólicas del hospedador, como un descenso de la glucosa intracelular que promueve la formación de AMPc, cofactor de la proteína CRP que activa la expresión de un sistema CRISPR-Cas ante la entrada de material genético exógeno. Además, la proteína CRP está modulada por el gen del metabolismo de la galactosa *galM* (Patterson *et al.*, 2015). Otro ejemplo es la activación de un sistema CRISPR-Cas por el regulador transcripcional LeuO, ante el descenso de los niveles de aminoácidos provocados por la escasez de nitrógeno debido a la infección por un fago (Medina-Aparicio *et al.*, 2011).
- e) Otro grupo de reguladores implicados en la expresión de estos sistemas se corresponde con los productos de genes adyacentes a la región CRISPR, tales como las proteínas Dev, que regulan negativamente la expresión de un sistema CRISPR-Cas en *M. xanthus* (Rajagopalan & Kroos, 2017; Rajagopalan *et al.*, 2015; Viswanathan *et al.*, 2007) o la proteína Csa3a de *Sulfolobus islandicus*, que activa la expresión de los genes *cas* (Liu *et al.*, 2015).

I.3.6 Sistemas CRISPR-Cas en M. xanthus

La secuenciación y anotación del genoma de *M. xanthus* permitió identificar tres loci CRISPR-*cas* concentrados en una misma región del genoma: CRISPR1, CRISPR2-3 (con dos regiones CRISPR contiguas asociadas a los mismos genes *cas*), y CRISPR4) (Figura 18) (Goldman *et al.*, 2006).



Figura 18. Sistemas CRISPR-Cas de *M. xanthus.* Entre paréntesis, el número de espaciadores de cada región CRISPR. Modificada de Bernal-Bernal *et al.*, (2018).

El locus CRISPR1, con 57 espaciadores, correspondería a un sistema CRISPR-Cas de tipo I-C, en base a las características de los siete genes *cas* adyacentes (Figuras 18 y 19). El segundo sistema CRISPR-Cas, que incluye dos loci CRISPR (CRISPR2 con 22 espaciadores y CRISPR3 con 8), correspondería a un sistema CRISPR-Cas de tipo I-A (Figuras 18 y 19) (Rajagopalan & Kroos, 2017). La expresión de los genes *cas* asociados, incluidos los que determinan las proteínas DevT, DevR, y DevS, parece estar sujeta a activación por el proceso de desarrollo multicelular (Viswanathan *et al.*, 2007). Por último, el locus CRISPR4 (con 52 espaciadores), objeto de estudio en este trabajo doctoral, por sus genes *cas* adyacentes constituiría un sistema CRISPR-Cas de tipo III-B (Figuras 18 y 19). Como se explicará en profundidad a lo largo de esta tesis, este sistema CRISPR-Cas está controlado por la pareja σ -ECF/anti- σ DdvS/DdvA, cuyos genes forman parte del mismo locus (Abellón-Ruiz, 2012). Además, según estudios de otro grupo de investigación, este sistema CRISPR-Cas ha sido relacionado con el desarrollo multicelular (Wallace *et al.*, 2014).



Figura 19. Esquema de los cuatro loci CRISPR-cas de *M. xanthus*. Modificado de Bernal-Bernal *et al.*, (2018)

I.4 Objetivos

Trabajos realizados previamente en nuestro laboratorio pusieron de manifiesto que CarD, a pesar de su similitud con CdnL, es funcionalmente diferente, y que su acción se centra en la regulación de determinados factores σ -ECF de *M. xanthus.* Para tratar de entender la base molecular de estas diferencias, así como profundizar en la comprensión sobre el modo de acción de este complejo en los factores σ -ECF se propusieron una serie de objetivos.

- (a) Identificar y analizar la relación estructura-función del subdominio N- terminal de CarDNt y los aminoácidos implicados en la interacción con la subunidad β de la RNAP, así como su análisis comparativo con CdnL.
- (b) Estudiar el subdominio C-terminal de CarDNt que participa en la interacción con CarG, a diferencia de CdnL.
- (c) Examinar seis posibles parejas σ-ECF/NZASD-eSTK anti-σ y estudiar la dependencia de CarD-CarG de una de ellas.
- (d) Determinar la localización *in vivo* del complejo CarD-CarG, así como examinar si la asociación de CarD con el DNA requiere de CarG.
- (e) Profundizar en la comprensión del sistema CRISPR4-Cas de *M. xanthus* y en su regulación por DdvS/DdvA y CarD-CarG.

II. MATERIALES Y MÉTODOS

II.1 Estirpes y plásmidos

II.1.1 Estirpes de *Myxococcus xanthus*

Las estirpes de *M. xanthus* utilizadas en este trabajo derivan directa o indirectamente de DK101 (cepa silvestre que presenta variación de fase en la pigmentación de las colonias y puede ser amarilla o morena en oscuridad) y son capaces de sintetizar carotenoides previa activación por la luz. En nuestro laboratorio se usan habitualmente dos estirpes silvestres, DK1050 y DK1622. La estirpe DK1050 es una variante espontánea de DK101 que presenta una coloración amarilla estable en la oscuridad (Ruiz-Vázquez & Murillo, 1984) y contiene la mutación *pilQ1* que afecta a la movilidad social. La estirpe DK1622 se obtuvo a partir de DK101 por transducción de un fragmento que contiene la versión silvestre de *pilQ*, de modo que esta estirpe es capaz de generar cuerpos fructíferos simétricos y es la estirpe habitual para experimentos de fructificación (Kaiser, 1979; Wall *et al.*, 1998).

En la **Tabla 1** se detallan las estirpes de *M. xanthus* utilizadas en este trabajo, indicando su descripción y origen. Aquellas que han sido generadas en nuestro laboratorio se nombran con las letras MR seguidas de un número de aislamiento. Aclaraciones sobre el genotipo de algunas estirpes se incluyen en la sección de Resultados).

Estirpe	Descripción	Origen o fuente
Relacionadas con el dominio N-terminal de CarD: estructura y función		
MR1316	∆ddvA	Abellón-Ruiz <i>et al</i> ., 2014
MR1317	∆ddvA∆carD	Abellón-Ruiz <i>et al</i> ., 2014
MR1389	P _{ddvSA} ∷ <i>lacZ</i> ; ∆ <i>ddvA</i> ; Km ^R	pMR3890 x MR1316
MR1390	P _{ddvSA} ∷ <i>lacZ</i> ; ∆ <i>ddvA</i> ∆ <i>carD</i> ; Km ^R	pMR3890 x MR1317
MR1900	∆carD	Cayuela <i>et al</i> ., 2003
MR1901	carD reinsertardo en su sitio endógeno	pMR2696 x MR1900
MR1904	P _{QRS} ∷lacZ; ∆carD	pDAH217 x MR1900
MR1905	P _{QRS} ::/acZ;	pDAH217 x MR1901
MR2222	carDNt	pMR3389 x MR1900

Tabla 1. Estirpes de M. xanthus

Estirpe	Descripción	Origen o fuente	
MR2223	P _{QRS} :: lacZ ; carDNt	pDAH217 x MR2222	
MR2240	∆carDNt	pMR2768 x MR1900	
MR2241	P _{QRS} ∷ lacZ ; ∆carDNt	pDAH217 x MR2240	
MR2257	carD(F41A)	pMR3687 x MR1900	
MR2258	carD(M54A)	pMR3688 x MR1900	
MR2259	carD(P56A)	pMR3689 x MR1900	
MR2280	carD(W92A)	pMR4114 x MR1900	
MR2281	carD(L100A)	pMR4115 x MR1900	
MR2282	carD(T129A)	pMR4116 x MR1900	
MR2283	carD(K130A/R132A)	pMR4157 x MR1900	
MR2284	carD(K93A/R95A/R97A)	pMR4158 x MR1900	
MR2292	P _{QRS} :: <i>lacZ; carD(F41A)</i> ; Km ^R	pDAH217 x MR2257	
MR2293	P _{QRS} ∷ <i>lacZ</i> ; <i>carD(M54A</i>); Km ^R	pDAH217 x MR2258	
MR2294	P _{QRS} ∷ <i>lacZ</i> ; <i>carD(P56A</i>); Km ^R	pDAH217 x MR2259	
MR2295	P _{QRS} ∷ <i>lacZ</i> ; <i>carD(W</i> 92A); Km ^R	pDAH217 x MR2280	
MR2296	P _{QRS} ∷ <i>lacZ</i> ; <i>carD(L100A)</i> ; Km ^R	pDAH217 x MR2281	
MR2297	P _{QRS} :: <i>lacZ</i> ; <i>carD(T129A</i>); Km ^R	pDAH217 x MR2282	
MR2298	P _{QRS} :: <i>lacZ</i> ; <i>carD(K130A/R132A)</i> ; Km ^R	pDAH217 x MR2283	
MR2299	P _{QRS} :: <i>lacZ</i> ; <i>carD(K93A/R95A/R97A)</i> ; Km ^R	pDAH217 x MR2284	
MR2755	carD(R95A)	pMR4253 x MR1900	
MR2756	∆ddvA carDNt	pMR3389 x MR1317	
MR2757	P _{QRS} ∷ <i>lacZ; carD(R95A)</i> ; Km ^R	pDAH217 x MR2755	
MR2758	P _{ddvSA} ∷ <i>lacZ</i> ; ∆ <i>ddvA carDNt</i> ; Km ^R	pMR3890 x MR2756	
MR2759	∆ddvA ∆carDNt	pMR2768 x MR1316	
MR2785	P _{ddvSA} ∷ <i>lacZ</i> ; ∆ <i>ddvA</i> ∆ <i>carDNt</i> ; Km ^R	pMR3890 x MR2759	
Relacionadas con los factores σ -ECF y su dependencia del complejo CarD-CarG			
DK1050	Estirpe silvestre	Ruiz-Vázquez & Murillo, 1984	
DK1622	Estirpe silvestre	Kaiser, 1979	
Estirpe	Descripción	Origen o fuente	
---------	---	--	--
MR1517	∆ <i>carG;</i> DK1622	Peñalver-Mellado <i>et al</i> ., 2006	
MR1910	∆ <i>carD</i> ; DK1050	Peñalver-Mellado <i>et al</i> ., 2006	
MR1911	∆ <i>carG</i> ; DK1050	Peñalver-Mellado <i>et al</i> ., 2006	
MR2650	P _{ddvSA} ::/acZ 2PR::ddvSddvA ₍₁₋₉₀₎ ; Km ^R	pMR3797 x DK1050	
MR2651	P _{ddvSA} :: <i>lacZ</i> ; 2PR:: <i>ddvS</i> ; Km ^R	pMR3796 x DK1050	
MR2652	P _{ddvSA} ::/acZ; 2PR::ddvSddvA ₍₁₋₇₀₎ ; Km ^R	pMR3798 x DK1050	
MR2681	P ₁₇₀₉₋₁₀ ::/ <i>acZ</i> ; Km ^R	pMR4021 x DK1622	
MR2682	P ₁₇₀₉₋₁₀ :: <i>lacZ</i> ; 2PR:: <i>MXAN_1709</i> ; Km ^R	pMR4042 x DK1622	
MR2683	P ₁₇₀₉₋₁₀ :: <i>lacZ</i> ; 2PR:: <i>MXAN_1709</i> ; Δ <i>carD</i> ; Km ^R	pMR4042 x MR1900	
MR2684	P ₁₇₀₉₋₁₀ :: <i>lacZ</i> ; 2PR:: <i>MXAN_1709</i> ; Δ <i>carG</i> ; Km ^R	pMR4042 x MR1517	
MR3225	P ₅₅₀₆₋₀₇ :: <i>lacZ</i> ; Km ^R	pMR4378 x DK1050	
MR3226	P ₅₅₀₆₋₀₇ :: <i>lacZ</i> ; 2PR:: <i>MXAN</i> _5507; Km ^R	pMR4379 x DK1050	
MR3227	P ₅₅₀₆₋₀₇ :: <i>lacZ</i> ; 2PR:: <i>MXAN_</i> 5507; Δ <i>carD</i> ; Km ^R	pMR4379 x MR1910	
MR3228	P ₅₅₀₆₋₀₇ :: <i>lacZ</i> ; 2PR:: <i>MXAN</i> _5507; Δ <i>carG</i> ; Km ^R	pMR4379 x MR1911	
	Relacionadas con el modo de acción del complejo CarD-CarG		
MR2751	CarD-Flag (sitio endógeno)	pMR4196 x MR1900	
MR2954	CarG-Flag (sitio endógeno)	pMR4481 x MR1517	
MR2992	P _{van} :: <i>flag-car</i> Q; Km ^R	pMR4517 x DK1622	
MR2993	P _{van} ::flag-carQ; ΔcarD; Km ^R	pMR4517 x MR1900	
MR2994	P _{van} ∷flag-carQ; ΔcarG; Km ^R	pMR4517 x MR1517	
Relacio	nadas con el control multifactorial de un sistema CRI DdvS/DdvA y el complejo regulador global (SPR-Cas por la pareja CarD-CarG	
MR1388	P _{ddvSA} :: /acZ ; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)	
MR1389	P _{ddvSA} :: /acZ ; ΔddvA; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)	
MR1390	P _{ddvSA} :: lacZ ; ΔddvA ΔcarD; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)	
MR1391	P _{ddvSA} :: /acZ; ΔddvA ΔcarG; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)	
MR1392	Р _{<i>мхал_7291</i>} :: /acZ ; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)	
MR1393	P _{MXAN_7291} :: /acZ ; ΔddvA; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)	

II. Materiales y métodos

Estirpe	Descripción	Origen o fuente
MR1394	P _{MXAN_7291} ::lacZ; ΔddvA ΔcarD; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR1395	P _{MXAN_7291} ::lacZ; ΔddvA ΔcarG; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2378	P _{MXAN_7283} ∷lacZ; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2379	P _{MXAN_7283} ::lacZ; ΔddvA; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2380	Р _{<i>мхал_7286</i>} :: /acZ ; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2381	P _{MXAN_7286} ::lacZ; ΔddvA; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2387	P _{MXAN_7286} ::lacZ; ΔddvA ΔcarD; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2388	P _{MXAN_7283} ::lacZ; ΔddvA ΔcarD; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2395	P _{MXAN_7286} ::lacZ; ΔddvA ΔcarG; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2396	P _{MXAN_7283} ::lacZ; ΔddvA ΔcarG; Km ^R	(Bernal-Bernal <i>et al</i> ., 2018)
MR2565	Δ <i>ddvA</i> Δ <i>cas6;</i> DK1050	pMR4179 x MR1316
MR2590	ΔddvA	pMR3037 x DK1622
MR2598	ΔddvA Δcas6; DK1622	pMR4179 x MR2590
MR2766	∆ddvA ∆carD	pMR2603 x MR2590
MR2767	∆ddvA ∆carG	pMR2874 x MR2590
MR2777	Líder:: <i>lacZ</i> ; WT (sitio heterólogo); Km ^R	pMR4334 x DK1622
MR2778	Líder:: <i>lacΖ</i> ; Δ <i>ddvA</i> (sitio heterólogo); Km ^R	pMR4334 x MR2590
MR2779	Líder:: <i>lacΖ</i> ; Δ <i>ddvA</i> Δ <i>carD</i> (sitio heterólogo); Km ^R	pMR4334 x MR2766
MR2780	Líder:: <i>lacZ</i> ; Δ <i>ddvA</i> Δ <i>carG</i> (sitio heterólogo); Km ^R	pMR4334 x MR2767
MR2781	Líder:: <i>lacZ</i> ; WT (sitio endógeno); Km ^R	pMR4351 x DK1622
MR2782	Líder:: <i>lacZ</i> ; Δ <i>ddvA</i> (sitio endógeno); Km ^R	pMR4351 x MR2590
MR2783	Líder:: <i>lacΖ</i> ; Δ <i>ddvA</i> Δ <i>carD</i> (sitio endógeno); Km ^R	pMR4351 x MR2766
MR2784	Líder:: <i>lacΖ</i> ; Δ <i>ddvA</i> Δ <i>carG</i> (sitio endógeno); Km ^R	pMR4351 x MR2767
MR2793	<i>cas6::lacZ</i> ; WT; Km ^R	pMR4384 x DK1622
MR2794	<i>cas6::lacZ</i> ; Δ <i>ddvA</i> ; Km ^R	pMR4384 x MR2590
MR2795	cas6:: <i>lacZ</i> ; ΔddvA ΔcarD; Km ^R	pMR4384 x MR2766
MR2796	cas6:: <i>lacZ</i> ; ΔddvA ΔcarG; Km ^R	pMR4384 x MR2767
MR2797	Líder(Mut):: <i>lacZ</i> ; WT (sitio heterólogo); Km ^R	pMR4385 x DK1622

Estirpe	Descripción	Origen o fuente
MR2798	Líder(Mut):: <i>lacZ</i> ; Δ <i>ddvA</i> (sitio heterólogo); Km ^R	pMR4385 x MR2590
MR2799	Líder(Mut):: <i>lacZ</i> ; Δ <i>ddvA</i> Δ <i>carD</i> (sitio heterólogo); Km ^R	pMR4385 x MR2766
MR2851	Líder(Mut):: <i>lacZ</i> ; ∆ <i>ddvA</i> ∆ <i>carG</i> (sitio heterólogo); Km ^R	pMR4385 x MR2767
MR2852	Líder(Mut):: <i>lacZ</i> ; WT (sitio endógeno); Km ^R	pMR4402 x DK1622
MR2853	Líder(Mut):: <i>lacZ</i> ; Δ <i>ddvA</i> (sitio endógeno); Km ^R	pMR4402 x MR2590
MR2854	Líder(Mut):: <i>lacZ</i> ; ∆ <i>ddvA</i> ∆ <i>carD</i> (sitio endógeno); Km ^R	pMR4402 x MR2766
MR2855	Líder(Mut):: <i>lacZ</i> ; ∆ <i>ddvA</i> ∆ <i>carG</i> (sitio endógeno); Km ^R	pMR4402 x MR2767
MR2862	ΔddvA Δcas6 P _{van} ::cas6; Km ^R	pMR4350 x MR2565
MR2863	ΔCRISPR4	pMR4386 x DK1622
MR2864	ΔddvA ΔCRISPR4	pMR4386 x MR2590
MR2959	∆ddvA P _{IPTG} ∷ddvA; Tc ^R	pMR4315 x MR2586
MR2962	∆ddvA ∆cas6 PIPTG∷ddvA; Tc ^R	pMR4409 x MR2959
MR2967	∆ddvA ∆cas6 P _{IPTG} ∷ddvA P _{van} ∷cas6; Km ^R Tc ^R	pMR4493 x MR2962
MR2971	P _{van} ::flag-ddvS; WT; Km ^R	pMR4513 x DK1622
MR2972	P _{van} ∷flag-ddvS; ΔcarD; Km ^R	pMR4513 x MR1900
MR2973	P _{van} ∷flag-ddvS; ΔcarG; Km ^R	pMR4513 x MR1517
MR3062	∆ddvA ∆cas6 P _{IPTG} ::ddvA P _{van} ::cas6a; Km ^R Tc ^R	pMR4552 x MR2962
MR3063	∆ <i>ddvA</i> ∆ <i>cas6</i> P _{van} ∷ <i>cas6a</i> ; Km ^R	pMR4552 x MR2565
MR3078	∆devT	pMR4567 x DK1622

El sitio heterólogo hace referencia al segmento de 1.38 kb, entre la posición 8498672 y 8500048 del genoma o a la región de un 1.31 kb ubicada entre los *MXAN18-19*. P_{IPTG}: promotor inducible por IPTG (isopropil- β -tiogalactopiranósido). P_{van}, promotor inducible por vanilato. 2PR: promotor fuerte dependiente del factor σ mayoritario de *M. xanthus*. Km^R: resistencia a kanamicina. Tc^R: resistencia a oxitetraciclina.

II.1.2 Estirpes de E. coli

La **Tabla 2** recoge las estirpes de *E. coli* utilizadas en este trabajo. Las estirpes DH5 α y JM109 son deficientes en recombinación (*recA*⁻), propiedad que favorece la estabilidad de los plásmidos y las convierte en hospedadores adecuados en los experimentos de clonación. La estirpe BL21(DE3) posee en su cromosoma una copia del gen de la polimerasa de RNA del fago T7 bajo el control del promotor *lacUV5*, activable por IPTG. Es además defectiva para la proteasa *lon* y para la proteasa *ompT* de la membrana externa, características que favorecen la expresión y posterior

purificación de proteínas. La estirpe BTH101 posee el genotipo adecuado para el ensayo del doble híbrido bacteriano (DHB).

Tabla	2.	Estirp	es de	• E.	coli	usadas	en	este	trabaio
I UNIU		Louip	00 40		0011	acaaac	U	0010	il asajo

Estirpes de <i>E. coli</i> usadas en este trabajo						
Estirpes	Origen o fuente					
DH5α Hanahan, 1983						
BL21(DE3) Novagen						
JM109 Yanisch-Perron <i>et al.</i> , 1985						
BTH101 Karimova <i>et al.</i> , 2000						

La **Tabla 3** recoge los plásmidos utilizados en este trabajo, indicando sus marcadores de selección, utilidad y origen.

Plásmido	Descripción / utilidad	Origen o fuente				
Plásm	Plásmidos empleados para estudiar el dominio N-terminal de CarD: función y estructura					
pDAH217	Vector que introduce P _{QRS} :: <i>lacZ</i> en <i>M. xanthus</i> ; Km ^R	Hodgson DA, 1993				
pKT25	Vector que permite generar fusiones al extremo C-terminal del fragmento T25 de CyaA para su uso en ensayos de doble híbrido bacteriano; Km ^R	Karimova <i>et al.</i> , 2000				
pUT18C	Vector que permite generar fusiones al extremo C-terminal del fragmento T18 de CyaA para su uso en ensayos de doble híbrido bacteriano; Amp ^R	Karimova <i>et al.</i> , 2000				
pTYB12	Vector que permite la sobreexpresión de proteínas con una cola N-terminal de inteína; Amp ^R	NE Biolabs				
pMR2603	Vector que presenta los segmentos aguas arriba y abajo del gen <i>carD</i> para su deleción. Útil para estudios de complementación; Km ^R	Cayuela <i>et al</i> ., 2003				
pMR2696	Derivado de pMR2603 con <i>carD</i> ; Km ^R	Cayuela <i>et al</i> ., 2003				
pMR2768	Vector que introduce <i>carD</i> (180-316) (sitio endógeno); Km ^R	Cayuela <i>et al</i> ., 2003				
pMR2880	pUT18c con la fusión <i>carG-T18</i> ; Km ^R	Peñalver- Mellado <i>et al</i> ., 2006				
pMR2881	pKT25 con la fusion <i>T25-carD</i> ; Amp ^R	Peñalver- Mellado <i>et al</i> ., 2006				
pMR3389	Derivado de pMR2603 con <i>carD</i> ₍₁₋₁₇₉₎ ; Km ^R	Este trabajo				
pMR3397	Derivado de pTYB12 que sobreexpresa CarD ₍₁₋₁₇₉₎ ; Amp ^R	Este trabajo				

Tabla 3. Plásmidos utilizados en este trabajo

Plásmido	Descripción / utilidad	Origen o fuente	
pMR3412	pUT18c con la fusión <i>T18-Mxβ₍₁₉₋₁₄₈₎</i> ; Amp ^R	Este trabajo	
pMR3521	Derivado de pTYB12 con CarD ₍₆₁₋₁₇₉₎ ; Amp ^R	Este trabajo	
pMR3585	pUT18c con la fusión <i>T18-Mxβ₍₁₉₋₁₄₈₎ (D122A)</i> ; Amp ^R	Este trabajo	
pMR3586	pUT18c con la fusión <i>T18-Mxβ₍₁₉₋₁₄₈₎ (V123A)</i> ; Amp ^R	Este trabajo	
pMR3587	pUT18c con la fusión <i>T18-Mxβ₍₁₉₋₁₄₈₎ (K124A)</i> ; Amp ^R	Este trabajo	
pMR3588	pUT18c con la fusión <i>T18-Mxβ₍₁₉₋₁₄₈₎ (E125A)</i> ; Amp ^R	Este trabajo	
pMR3633	Derivado de pTYB12 con <i>carD₍₁₋₇₂₎;</i> Amp ^R	Este trabajo	
pMR3687	Derivado de pMR2603 con <i>carD(F41A)</i> ; Km ^R	Este trabajo	
pMR3688	Derivado de pMR2603 con <i>carD(M54A</i>); Km ^R	Este trabajo	
pMR3689	Derivado de pMR2603 con <i>carD(P56A)</i> ; Km ^R	Este trabajo	
pMR3890	Vector que introduce P _{ddvSA} :: <i>lacZ</i> en <i>M. xanthus;</i> Km ^R	Abellón-Ruiz <i>et al</i> ., 2014	
pMR4114	IR4114 Derivado de pMR2603 con <i>carD(W92A)</i> ; Km ^R		
pMR4115	IR4115 Derivado de pMR2603 con <i>carD(L100A)</i> ; Km ^R		
pMR4116	R4116 Derivado de pMR2603 con <i>carD(T129A)</i> ; Km ^R		
pMR4157	Derivado de pMR2603 con <i>carD(K130A/R132A)</i> ; Km ^R	Este trabajo	
pMR4158	Derivado de pMR2603 con <i>carD(K93A/R95A/R97A)</i> ; Km ^R	Este trabajo	
pMR4173	pKT25 con la fusión <i>T25-carDNt(K130A/R132A)</i> ; Km ^R	Este trabajo	
pMR4174	pKT25 con la fusión <i>T25-carDNt(K93A/R95A/R97A)</i> ; Km ^R	Este trabajo	
pMR4253	Derivado de pMR2603 con <i>carDNt(R95A)</i> ; Km ^R	Este trabajo	
Plásmido	os empleados para el estudio de las parejas σ-ECF/anti-σ o quinasa de serina/treonina en el factor anti-σ	con un dominio	
pMR2792	pKT25 con la fusión <i>T25-carR</i> ; Km ^R	Galbis-Martínez <i>et al.</i> , 2008	
pMR2858	pUT18C con la fusión <i>T18-carQ</i> ; Amp ^R	Galbis-Martínez <i>et al.</i> , 2008	
pMR2931	Vector usado para estudios de topología; Km ^R	Galbis-Martínez <i>et al</i> ., 2008	
pMR3190	 Vector sonda de promotores para la clonación de fragmentos aguas arriba del gen <i>lacZ</i> sin promotor, y la clonación del factor sigma determinado aguas abajo de dos promotores ribosómicos de <i>rrnD</i>. Permite la integración heteróloga en el cromosoma de <i>M. xanthus</i>; Km^R 		
pMR3217	verivado de pMR3190 con P _{ddvSA} Abellón-R et al., 201		

II. Materiales y métodos

Plásmido	Descripción / utilidad	Origen o fuente
pMR3277	pUT18C con la fusión <i>T18-ddvS</i> ; Amp ^R	Abellón-Ruiz <i>et al</i> ., 2014
pMR3280	pKT25 con la fusión <i>T25-ddvA;</i> Km ^R	Abellón-Ruiz et al., 2014
pMR3740	Derivado de pMR2931 con <i>ddvA₍₁₋₉₅₎;</i> Km ^R	Este trabajo
pMR3749	Derivado de pMR2931 con <i>ddvA</i> (1-865); Km ^R	Este trabajo
pMR3751	Derivado de pMR2931 con <i>carR₍₁₋₇₁₎;</i> Km ^R	Este trabajo
pMR3752	Derivado de pMR2931 con <i>carR₍₁₋₂₂₁₎;</i> Km ^R	Este trabajo
pMR3753	Derivado de pMR2931 con <i>ddvA₍₁₋₉₄₆₎;</i> Km ^R	Este trabajo
pMR3754	Derivado de pMR2931 con <i>ddvA₍₁₋₉₉₁₎</i> ; Km ^R	Este trabajo
pMR3756	pKT25 con la fusión <i>T25-ddvA₍₁₋₇₀₎;</i> Km ^R	Este trabajo
pMR3757	pKT25 con la fusión <i>T25-ddvA₍₁₋₉₅₎</i> ; Km ^R	Este trabajo
pMR3758	Derivado de pMR2931 con <i>carR₍₁₋₁₀₀₎;</i> Km ^R	Este trabajo
pMR3759	Derivado de pMR2931 con <i>carR</i> (1-131); Km ^R	Este trabajo
pMR3760	Derivado de pMR2931 con <i>carR₍₁₋₁₆₂₎;</i> Km ^R	Este trabajo
pMR3768	pKT25 con la fusión <i>T25-ddvA_(96–991);</i> Km ^R	Este trabajo
pMR3783	Derivado de pMR2931 con <i>carR₍₁₋₁₈₈₎;</i> Km ^R	Este trabajo
pMR3784	Derivado de pMR2931 con <i>carR₍₁₋₂₁₈₎;</i> Km ^R	Este trabajo
pMR3788	pKT25 con la fusión <i>T25-carR₍₁₋₄₄₎;</i> Km ^R	Este trabajo
pMR3789	pKT25 con la fusión <i>T25-carR</i> (1-71); Km ^R	Este trabajo
pMR3790	pKT25 con la fusión <i>T25-carR_(190–221);</i> Km ^R	Este trabajo
pMR3796	Derivado de pMR3217 con <i>ddvS</i> ; Km ^R	Este trabajo
pMR3797	Derivado de pMR3217 con <i>ddvSddvA₍₁₋₉₅₎</i> ; Km ^R	Este trabajo
pMR3798	Derivado de pMR3217 con <i>ddvSddvA</i> (1-70); Km ^R	Este trabajo
pMR3854	pUT18c con la fusión <i>T18-ddvA_(96–991);</i> Amp ^R	Este trabajo
pMR3977	pKT25 con la fusión <i>T25-pkn8₍₁₋₇₀₎;</i> Km ^R	Este trabajo
pMR3994	pKT25 con la fusión <i>T25-MXAN_</i> 2738 ₍₁₋₇₀₎ ; Km ^R	Este trabajo
pMR3995	pUT18c con la fusión <i>T18-MXAN_</i> 2737; Amp ^R	Este trabajo
pMR3996	pKT25 con la fusión <i>T25-MXAN_6009₍₁₋₇₀₎;</i> Km ^R	Este trabajo
pMR3997	pUT18c con la fusión <i>T18-MXAN_6010</i> ; Amp ^R	Este trabajo

Plásmido	Descripción / utilidad	Origen o fuente		
pMR3998	pKT25 con la fusión <i>T25-MXAN_7082₍₁₋₇₀₎;</i> Km ^R	Este trabajo		
pMR3999	pUT18c con la fusión <i>T18-MXAN_7083</i> ; Amp ^R	Este trabajo		
pMR4000	pKT25 con la fusión <i>T25-MXAN_7162₍₁₋₇₀₎;</i> Km ^R	Este trabajo		
pMR4001	pUT18c con la fusión <i>T18-MXAN_7161</i> ; Amp ^R	Este trabajo		
pMR4021	Derivado de pMR3190 llevando promotor P ₁₇₀₉₋₁₀ ; Km ^R	Este trabajo		
pMR4042	Derivado de pMR4021 llevando la sobreexpresión de MXAN_1709; Km ^R	Este trabajo		
pMR4658	pKT25 con la fusión <i>T25-MXAN_5507</i> ; Km ^R	Este trabajo		
pMR4736	pUT18c con la fusión <i>T18-MXAN_5506</i> ; Amp ^R	Este trabajo		
pMR4737	pKT25 con la fusión <i>T25-MXAN_5507₍₁₋₇₀₎;</i> Km ^R	Este trabajo		
pMR4738	Derivado de pMR3190 con la región promotora de <i>MXAN_5506</i> P ₅₅₀₆₋₀₇ ; Km ^R	Este trabajo		
pMR4739	Derivado de pMR4738 con la sobreexpresión de <i>MXAN_5506</i> ; Km ^R	Este trabajo		
	Plásmidos empleados para estudiar el modo de acción complejo regulador CarD-CarG	del		
pMR2875	Portador del locus <i>carG</i> , con la deleción de <i>carG</i> . Permite generar versiones de CarG e introducirlas en <i>M. xanthus</i> por recombinación homóloga; Km ^R	Peñalver- Mellado, 2005		
pMR4196	Derivado de pMR2603 con <i>carD-flag</i> ; Km ^R	Este trabajo		
pMR4481	Derivado de pMR2875 con <i>carG-flag</i> ; Km ^R	Este trabajo		
pMR4517	['] Derivado de pMR3679 con fl <i>ag-carQ</i> ; Km ^R Es			
Plási CRISPR	nidos empleados para estudiar el control multifactorial de - Cas por la pareja DdvS/DdvA y el complejo regulador glo	un sistema bal CarD-CarG		
pET28b	Expresión de proteínas bajo el control de Plac; Km ^R	Novagen		
pMR3130	Derivado de pET28b con CarG; Km ^R	Peñalver- Mellado <i>et al.,</i> 2006		
pMR3183	Vector usado para generar fusiones transcripcionales al gen <i>lacZ</i> . Contiene un fragmento de 1.38 kb para su integración en el genoma de <i>M. xanthus</i> ; Km ^R	Iniesta <i>et al.,</i> 2012		
pMR3229	Derivado de pTYB12 con CarD; Amp ^R	García-Heras <i>et al.</i> , 2009		
pMR3232	 Vector usado para generar fusiones transcripcionales al gen <i>lacZ</i> que se integra en el genoma de <i>M. xanthus</i> en su sitio Abellón-F 			
pMR3487	Vector que permite la expresión de una proteína de interés bajo un promotor inducible por IPTG que se integra en el genoma de <i>M. xanthus</i> a través de un fragmento de 1.38 kb; Km ^R	Iniesta <i>et al.,</i> 2012		

Plásmido	Descripción / utilidad	Origen o fuente
pMR3675	Derivado de pMR3183 con 530 pb que incluye $P_{\mbox{\scriptsize MXAN}_{-7283}};$ \mbox{Km}^{R}	Abellón-Ruiz J
pMR3676	Derivado de pMR3183 con 530 pb que incluye $P_{\textit{MXAN}_7286};$ \textit{Km}^{R}	Abellón-Ruiz J
pMR3679	Vector que permite la expresión de una proteína de interés bajo un promotor inducible por vanilato. Se integra en el genoma de <i>M. xanthus</i> a través de un fragmento de 1.38 kb; Km ^R	Iniesta <i>et al.,</i> 2012
pMR3690	Vector que permite la expresión de una proteína de interés bajo un promotor inducible por vanilato. Se integra en el genoma de <i>M. xanthus</i> a través de un fragmento de 1.31 kb (<i>MXAN_0018-MXAN-0019</i>); Km ^R	Iniesta <i>et al.,</i> 2012
pMR3891	Derivado de pMR3183 con 489 pb que incluye $P_{\textit{MXAN}_7291};$ \textit{Km}^{R}	Abellón-Ruiz J
pMR4179	Plásmido usado para generar la deleción de cas6; Km ^R	Este trabajo
pMR4315	Derivado de pMR3487 con P _{IPTG} ∷ <i>ddvA</i> ; Tc ^R	Este trabajo
pMR4334	Derivado de pMR3183 con Líder; Km ^R	Este trabajo
pMR4350	Derivado de pMR3679 con P <i>van</i> ::cas6; Km ^R	Este trabajo
pMR4351	Derivado de pMR3232 con un fragmento de ~1kb que incluye a Líder; ${\rm Km}^{\rm R}$	Este trabajo
pMR4384	Derivado de pMR3232 con un fragmento de ~1kb situado aguas arriba de Líder; $\rm Km^R$	Este trabajo
pMR4385	Derivado de pMR3183 con Líder mutado en una región -35 de un posible promotor σ^{A} ; Km^{\text{R}}	Este trabajo
pMR4386	Plásmido usado para generar la deleción de la región CRISPR4; Km ^R	Este trabajo
pMR4402	Derivado de pMR3232 con un fragmento de ~1kb que incluye a Líder con una triple mutación en una región -35 de un posible promotor σ^A ; Km ^R	Este trabajo
pMR4493	Derivado de pMR3690 con P _{van} ∷ <i>cas6</i> ; Km ^R	Este trabajo
pMR4513	Derivado de pMR3679 con fl <i>ag-ddvS</i> ; Km ^R	Este trabajo
pMR4552	Derivado de pMR3690 con <i>cas6a</i> del sistema CRISPR2/3- Cas; Km ^R	Este trabajo
pMR4567	Pásmido usado para generar la deleción completa del gen <i>devT</i> ; Km ^R	Rajagopalan <i>et al.,</i> 2017

II.1.3 Oligonucleótidos

La **Tabla 4** recoge los oligonucleotidos utilizados en este trabajo, indicando su nombre, secuencia, temperatura de hibridación y su aplicación.

Tabla 4. Oligonucleotidos usados en este trabajo.

Nombre	Secuencia (5' - 3')	Tm	Aplicación
33EMSA5	GACAGCGTCCCTTGCTGCTC	68	EMSA: promotor <i>ddvSA</i>
33EMSA3	CAAGCTCTTCAATTTTCGCAAG	66	EMSA: promotor <i>ddvSA</i>
7291EMSA5	TTGCTCGGGGTGGCCTTGCT	75	EMSA: promotor MXAN_7291
7291EMSA3	TATGGTCTCCTTCGCCCTAA	63	EMSA: promotor MXAN_7291
7286EMSA5	AGAGCCTTGCTGTGGCTGGA	69	EMSA: promotor MXAN_7286
7286EMSA3	GGTGGGAAGTGGAAGTGCCA	69	EMSA: promotor MXAN_7286
7283EMSA5	AATACCGCTGTCCGGAGCCC	71	EMSA: promotor MXAN_7283
7283EMSA3	CTCCTCGTGGTCTCATCTCT	61	EMSA: promotor MXAN_7283
rnk fow ChIP	GGGTCGCCTTTTGGGG	51	ChIP: promotor <i>rnk</i>
rnk rev ChIP	TCTAACTGCCTCGCCTGCT	53	ChIP: promotor <i>rnk</i>
fow 6483 ChIP	CCCATCCGTTCCATTTGACT	52	ChIP: promotor MXAN_6483
rev 6483 ChIP	TGCGGGCCCATGTCTTAT	50	ChIP: promotor MXAN_6483
rev rpoE1 ChIP	CCGGCCAGGGAAACATATT	51	ChIP: promotor rpoE1
fow rpoE1 ChIP	GGTTCGGATGGCGACATC	53	ChIP: promotor <i>rpoE1</i>
qPCR DdvSF	GCTAAAGGCCCCGGAAATT	58	ChIP: promotor <i>ddvS</i>
qPCR DdvSR	AATAGGAAAACGCAGGTCATCTG	55	ChIP: promotor <i>ddvS</i>
P7291F	CCCTCGAAGCCGCTGTAG	55	ChIP: promotor MXAN_7291
P7291R	TCGCCCTAAAGAACGATGGA	52	ChIP: promotor MXAN_7291
P7286F	ACCGAGTCCTTGCTGAATTCA	52	ChIP: promotor MXAN_7286
P7286R	AATCACGCAAGACATGGCAAT	50	ChIP: promotor MXAN_7286
P7283F	CGGACTGCCCATCCAACT	53	ChIP: promotor MXAN_7283
P7283R	ACGGGAATGGTTCTGGAAATT	50	ChIP: promotor MXAN_7283
PDqPCRfw	TCGCAACCCCGTGACTTT	59	ChIP: promotor PQRS
PDqPCRrv	CGCTCGCGCTTCTCAAG	56	ChIP: promotor P _{QRS}
qPCR intra 1R	GCCTTGCCAGCGACGTT	52	Región intragénica ChIP
qPCR intra 1F	GCCGTCACGTCGGTCTTG	55	Región intragénica ChIP
rev rpoE1 RTPCR	CGTGCGCTCCTTCACCTT	53	RTPCR rpoE1
fow rpoE1 RTPCR	TCTTCAAGATTCTCACCAACACCTT	54	RTPCR rpoE1

rev 6483 RTPCR	CATCTTCCCCAGGTCATTGAAG	55	RTPCR MXAN_6483
fow 6483 RTPCR	CGCCGTTCATCGGTCTGT	53	RTPCR MXAN_6483
rnk fow RTPCR	CGTCGTCGCTTCGGAAGA	53	RTPCR rnk
rnk rev RTPCR	CGATTGTTCGTCCTCGAAGAAC	55	RTPCR rnk
1-83	GGCCTCGGGGCGCGCACACG	66	RT-PCR MXAN_7283
2-82	ACAGCGCGCGCCACCAGAAG	60	RT-PCR MXAN_7282
3-82	CAAGGCGCTGCCGCTGGCGG	64	RT-PCR MXAN_7282
4-81	CCGTGCAGCCTGCTCGGGCG	64	RT-PCR MXAN_7281
5-81	GCCCGTGAAGCTCACCCCGG	62	RT-PCR MXAN_7281
6-80	ACCGCTGACCGCCATGGCGT	60	RT-PCR MXAN_7280
7-80	TACTTCTTCGAGAAACACAG	48	RT-PCR MXAN_7280
8-79	GCAGCACGCCCTTGAGCGAC	60	RT-PCR MXAN_7279
9-79	GAGGAGAGCCTGCCCGCGGA	62	RT-PCR MXAN_7279
10-78	GTCAGGTCTTCCAACAGGTG	54	RT-PCR MXAN_7278
11-78	CCGGGGTTGTCGGAGCCGTC	62	RT-PCR MXAN_7278
12-77	AATCGTCTGGCTCGGGGTGC	58	RT-PCR MXAN_7277
13-77	TGGCGCCGGAGGCGCGTGAG	64	RT-PCR MXAN_7277
14-76	GAATACGCCCAGGTCCAAGC	56	RT-PCR MXAN_7276
15-76	CGCGTCCCGGGTGCGGGGC	66	RT-PCR MXAN_7276
16-1esp	TGCATGAGTCGAAAGAGATT	48	RT-PCR 1er espaciador CRISPR4
17-2esp	TCGCGCGCCCGTGGACGTGG	64	RT-PCR 2do espaciador CRISPR4
85 fow	ATGGGAGACCGCTTTCCGTCG	58	RT-PCR MXAN_7285
84 rev	ACCACGTAGGCCCCCGCTTC	60	RT-PCR MXAN_7284
84 fow	TGCTGCTCATCGTGGGCGGTG	60	RT-PCR MXAN_7284
83 rev	CGTGCGTATTCAAGGCACGC	56	RT-PCR MXAN_7283
CRISPR3 Esp End	TGACAGCGCCCCATCCGCTTGCCGTGGTGCGTCAGCA	74	Sonda Northern blot Espaciador 22 CRISPR3
131	TGCATGAGTCGAAAGAGATTTTGAAGCGCCGGAC	70	Sonda Northern blot Espaciador 52 CRISPR4
132	GTCTGGCAACACCAGAATGGCGGAACCGACTA	68	Sonda Northern blot Espaciador 1 CRISPR4
133	CTTAACTTCCGTGTTCGGGATGGGAACGGGTGGGAC	69	Sonda Northern blot 5S

II.2 Medios de cultivo

II.2.1 Medios ricos para *M. xanthus*

CTT (Bretscher & Kaiser, 1978)
 Bacto-casitona (Difco) 10 mg/ml
 Tampón Tris-HCl (pH 7,6) 10 mM
 MgSO₄ 8 mM
 Tampón fosfato (pH 7,6) 1 mM
 Antes de la esterilización en autoclave, se ajustó el pH a 7,6.
 Para la preparación de medio sólido se utilizó Bacto-agar Difco (1,5%).

Para la preparación de agar de cobertera (CTTSA) se utilizó el mismo agar al 0,75%.

Las soluciones de antibióticos, indicadas a continuación en la **Tabla 5**, fueron esterilizadas por filtración a través de un filtro Millipore de 0,22 µm de diámetro de poro y añadidas al medio CTT una vez esterilizado y enfriado hasta 50 °C.

Antibiótico	Stock (concentración y solvente)	Concentración final en medio líquido/sólido
Kanamicina	10 mg/ml en agua	40 µg/ml
Oxitetraciclina	10 mg/ml en 50% etanol	10 µg/ml

* CTT + Gal

CTT más galactosa (Panreac) al 1%. Una vez esterilizado el medio CTT y enfriado hasta 50 °C, se añadió la disolución de galactosa, preparada en agua y esterilizada por separado en autoclave.

* CTT + X-Gal

CTT más 40 μg/ml de 5-Br-4-Cl-3-indolil-β-D-galactósido (X-Gal). La disolución de X-Gal se preparó a una concentración de 25 mg/ml en dimetilformamida o en dimetilsufóxido (DMSO) justo antes de su adición al medio.

II.2.2 Medios de fructificación para *M. xanthus*

*	★ TPM (Bretscher & Kaiser, 1978)		
	Tampón Tris-HCI (pH 7,6)	10 mM	
	MgSO ₄	8 mM	
	Tampón fosfato (pH 7,6)	1 mM	

Antes de la esterilización en autoclave, se ajustó el pH a 7,6. Para la preparación de medio sólido se utilizó Bacto-agar Difco (1,5%).

* CF

Tampón Tris-HCI (pH 7,6)	10 mM	
MgSO ₄	8 mM	
Tampón fosfato (pH 7,6)	1 mM	
(NH ₄) ₂ SO ₄	0,2 mg / ml	
Casitona	150 µg / ml	
Piruvato sódico	1 mg / ml	
Sodio citrato tri-básico	2 mg / ml	
Bacto-agar Difco	1,5%	

Se ajustó el pH a 7,6 antes de su esterilización en autoclave.

II.2.3 Medios ricos para E. coli

*	Luria	(LB)	preparado	según (Miller.	1972))
•		_ _/	proparado	eegan,	(1

Extracto de levadura	0,5%
Bacto-triptona	1%
NaCl	1%

Antes de esterilizar en autoclave, se ajustó el pH a 7,2-7,4.

Las soluciones de antibióticos, indicadas a continuación en la **Tabla 6**, fueron esterilizadas por filtración a través de un filtro Millipore de 0,22 µm de diámetro de poro y añadidas al medio una vez esterilizado y enfriado hasta 50 °C.

Tabla 6. Antibióticos	empleados en E	. coli.
-----------------------	----------------	---------

Antibiótico	Stock (concentración y solvente)	Concentración final en medio líquido/sólido
Kanamicina	10 mg/ml en agua	40 µg/ml
Oxitetraciclina	10 mg/ml en 50% etanol	10-12 µg/ml
Ampicilina	100 mg/ml en agua	100 µg/ml

* LB + X-Gal

LB más 40 μg/ml de 5-Br-4-Cl-3-indolil-β-D-galactósido (X-Gal) preparado en DMSO.

II.3 Condiciones de cultivo

En general, los cultivos en medio líquido se realizaron mediante inoculación de una colonia en un matraz de 50 ml con 10 ml de medio nutritivo. Para los cultivos a mayor escala, se utilizaron matraces de 250 ml, 500 ml o 5 l con 50 ml, 80 ml o 1 l de medio nutritivo, respectivamente. En el caso de volúmenes pequeños, la incubación tuvo lugar en un agitador orbital, a 300 rpm y 33 °C para *M. xanthus* y 250 rpm y 37 °C para *E. coli*. En el caso de los cultivos de 1 l, la inoculación se realizó en un agitador orbital de la marca New Brunswick Scientific, a 150 rpm. La inoculación en medio nutritivo sólido tras la transformación de *M. xanthus* se efectuó con ayuda de 2,5-4 ml de agar de cobertera. Las cajas Petri se incubaron a 33 °C para *M. xanthus* y 37 °C para *E. coli*.

El crecimiento, salvo que se indique lo contrario, tuvo lugar en condiciones de oscuridad. Cuando fue necesario, los cultivos en medio líquido o en cajas Petri se iluminaron colocándolos a unos 50 cm de una batería de 6 lámparas fluorescentes de 20 W. El flujo total de fotones incidentes, medido con un detector LI-189 y un sensor LI-100 (ambos LI-COR Inc. Lincoln, NE, USA) fue de aproximadamente 41 µmol s⁻¹m⁻². El flujo correspondiente a la banda de azul, medido con un filtro de máxima transmitancia entre 400 y 460 nm, fue de 12 ± 1 µmol s⁻¹m⁻².

II.4 Inducción del desarrollo multicelular de M. xanthus

Típicamente, se centrifugó 1 ml de cultivo de *M. xanthus* en fase exponencial de crecimiento durante 2 minutos, a 13000 rpm y temperatura ambiente. Tras realizar un lavado con TPM para eliminar los restos del medio CTT, las células se resuspendieron en un volumen de TPM necesario para dejar el cultivo a una OD₅₅₀~6. Del cultivo concentrado se depositaron gotas de 15 µl sobre cajas de TPM o CF. Una vez que las gotas fueron absorbidas por el agar, las cajas se incubaron a 33 °C. El tiempo de incubación varió según el experimento, desde unas pocas horas hasta varios días. Para la visualización y toma de fotografías de los cuerpos fructíferos se usó una lupa Zeiss y una cámara ProgRes C3.

II.5 Tampones y soluciones

II.5.1 Electroforesis de DNA

TAE 1X (Sambrook & Russell, 2001)

40 mM Tris-acetato, 1 mM EDTA (pH 8,0)

- **TBE 1X** (Sambrook & Russell, 2001)
 89 mM Tris-borato, 2 mM EDTA (pH 8,0)
- * Tampón de carga 10X

Suministrado por Fermentas

* Marcador de tamaño 100 pb y 1 kb

Suministrado por ThermoFisher Scientific

II.5.2 Electroforesis de proteínas

* Tampón de carga/ruptura para geles de poliacrilamida-SDS 1X

50 mM Tris-HCI (pH 6,8), 10% glicerol, 3% SDS, 3% β -mercaptoetanol, 0,3% azul de bromofenol

* Tampón de electroforesis para geles de poliacrilamida-SDS

50 mM Tris, 192 mM glicina, 0,1% SDS

***** Solución de tinción Coomassie

PageBlue Protein Staining Solution de Thermo Scientific

* Geles de poliacrilamida-SDS

Preparados según (Sambrook et al., 1989)

II.5.3 Electroforesis de RNA

TAE 1X (Sambrook & Russell, 2001)

40 mM Tris-acetato, 1 mM EDTA (pH 8,0), H₂O tratada con DEPC

II.5.4 Purificación de proteínas

* Tampón A

150 mM NaCl, 50 mM tampón fosfato pH 7,5, 5 mM β -mercaptoetanol

* Tampón B

500 mM NaCl, 20 mM Tris-HCl pH 8,5, 1 mM EDTA, 1 mM $\beta\text{-}$ mercaptoetanol

* Tampón C

150 mM NaCl, 20 mM HEPES pH 7,5, 5 mM β -mercaptoetanol, 0,1 mM ZnCl_2

* Tampón D

100 mM NaCl, 20 mM HEPES pH 7,5, 5 mM β -mercaptoetanol, 0,1 mM ZnCl_2

II.5.5 Detección de proteínas mediante anticuerpos específicos

* Tampón fosfato salino (PBS)

10 mM Na₂HPO₄, 1,8 mM KH₂PO4, 140 mM NaCl, 3,3 mM KCl, pH 7,5

* Tampón fosfato salino con Tween (PBS-T)

PBS, 0,1% Tween 20

* Tampón de bloqueo (PBS-TB)

PBS-T, 5% leche

* Tampón de equilibrado

30 mM Glicina, 48 mM Tris, 0,037% SDS, 20% metanol

* Tampón de transferencia I para el ánodo

0,3 M Tris, 10% metanol, pH 10,4

* Tampón de transferencia II para el ánodo

25 mM Tris, 10% metanol, pH 10,4

* Tampón de transferencia para el cátodo

25 mM Tris, 40 mM ácido aminocapróico, 20% metanol, pH 9,4

II.5.6 Determinación de la actividad β-galactosidasa

Tampón Z (Kroos & Kaiser, 1987)

40 mM NaH₂PO₄, 60 mM Na₂HPO₄, 10 mM KCI,1 mM MgSO₄

★ Solución Na₂CO₃

 $250 \text{ mM} \text{ Na}_2\text{CO}_3$

* Solución B1

 $40 \text{ mM} \text{ CuSO}_4 5 \text{H}_2 0$

* Solución B2

70 mM tartrato sódico

* Solución agarosa-X-Gal

0,5 M Tampón fosfato potásico pH 7,0, 6% dimetilformamida, 0,1% SDS, 40 μ g/ml X-Gal, 7 mM β -mercaptoetanol, 5 mg/ml agarosa "Low Melting"

II.5.7 Inmunoprecipitación de cromatina

* Tampón fosfato sódico

845 mM Na₂HPO₄, 155 mM NaH₂PO₄, pH 7,6

* Solución de lisis (TES buffer)

100 mM NaCl, 1 mM EDTA pH 8,0, 10 mM Tris pH 7,5, 2,2 mg/ml lisozima, 25X cOmplete[™] Protease Inhibitor Cocktail

* ChIP buffer

167 mM NaCl, 1,2 mM EDTA pH 8,0, 16,7 mM Tris-HCl pH 8,1, 1,1% Triton X-100, 1X cOmplete[™] Protease Inhibitor Cocktail

* ChIP buffer + SDS

ChIP buffer, 0,01% SDS

* Low Salt Immune Complex Wash Buffer

150 mM NaCl, 20 mM Tris-HCl pH 8,1, 2 mM EDTA pH 8,0, 1% Tritón X-100, 0,1% SDS

* High Salt Immune Complex Wash Buffer

500 mM NaCl, 20 mM Tris-HCl pH 8,1, 2 mM EDTA pH 8,0, 1% Tritón X-100, 0,1% SDS

* LiCI Immune Complex Wash Buffer

0,25 M LiCl, 1% NP-40, 1% deoxicolato, 1 mM EDTA pH 8,0, 10 mM Tris-HCl pH 8,1

* TE

10 mM Tris-HCl pH 8,0, 1 mM EDTA pH 8,0

* Elution Buffer

1% SDS, 0,1 M NaHCO₃. Recién preparado en cada uso.

II.5.8 Northern blot

II.5.8.1 Para análisis de RNAs pequeños

* Tampón de carga 2X

95% formamida, 18 mM EDTA pH 8,0, 0,025% SDS, 0,25 mg/ml xylene cyanol, 0,25 mg/ml bromophenol blue

* Tampón de electroforesis (TBE 0,5X)

40 mM Tris-Cl, pH 8,3, 45 mM ácido bórico, 1 mM EDTA pH 8,0

* Marcador de tamaño RNAs

Decade[™] Markers Systems de Invitrogen

* Tampón de hibridación

7% SDS, 40% formamida, 0,3 M NaCl, 0,05 M fosfato sódico pH 7,0, 1X Denhardt's solution

* Denhardt's solution

0,05 g/ml Ficoll 400, 0,05 g/ml polivinilpirrolidina (PVP), 0,05 g/ml BSA. Tras su preparación requiere de filtración.

* 20X SSC

3 M NaCl, 300 mM citrato trisódico, pH 7,0

* Tampón de lavado

2X SSC, 0,2% SDS

II.5.8.2 Para análisis de RNAs totales

* Mezcla de Northern blot

1,3X MOPS, 64,5% formamida, 8,4% formaldehido

* Tampón de electroforesis RNAs totales

1X MOPS, 6,6% formaldehido

* Tampón de prehibridación

0,9 M NaCl, 1% SDS, 100 µg/ ml DNA salmón

* Tampón de hibridación

0,9 M NaCl, 0,01% dextran sulfato, 1% SDS, 50 µg/ ml DNA salmón

* Tampón de lavado (2X SSC \rightarrow 0.1X SSC)

2X SSC, 0,1% SDS

II.6 Extracción de DNA

Para el aislamiento del DNA cromosómico de *M. xanthus* se utilizó el kit de purificación de DNA genómico de *Wizard*TM (Promega). En algunos casos este aislamiento se llevó a cabo usando la matriz *InstaGene* de Bio-Rad, siguiendo el protocolo facilitado por esta casa. El aislamiento de DNA plasmídico de *E. coli* se realizó mediante el uso del kit de purificación de DNA plasmídico *Illustra* de GE Healthcare o *QIAprep*TM *Spin Miniprep* de QIAGEN.

II.7 Transformación

II.7.1 Transformación de E. coli

La obtención de células competentes de *E. coli* y su transformación con 1-10 ng de DNA por choque térmico se efectuaron según Sambrook & Russell, (2001). Cuando fue preciso obtener un alto número de transformantes, se utilizó la técnica de

electroporación (Sambrook & Russell, 2001), empleándose cubetas de electroporación de 0,1 cm y los siguientes parámetros: 1,5 kV, 25 μ F y 200 Ω .

II.7.2 Transformación de M. xanthus

Se realizó mediante electroporación siguiendo el procedimiento descrito en Kashefi & Hartzell, (1995). Para preparar células "electrocompetentes" de M. xanthus se centrifugaron (2 min, 13000 rpm) 3 ml de un cultivo en fase exponencial de crecimiento (DO_{550~}0,7–1,1). El precipitado celular se lavó tres veces con 1,5 ml de agua bidestilada estéril y se resuspendió finalmente en 40 µl de agua bidestilada. Las células electrocompetentes se mezclaron con el DNA correspondiente (0,5–1 µg de DNA) y la mezcla, una vez transferida a una cubeta de electroporación de 0,1 cm, se sometió a un pulso eléctrico con el electroporador Bio-Rad Gene Pulser II (0,65 kV, 25 μ F y 400 Ω). Inmediatamente después del pulso eléctrico se añadió 1 ml de CTT a la cubeta, se mezcló bien y se transfirió a un matraz de 50 ml que contenía 5 ml de CTT. Las células se cultivaron 5 horas y se sembraron a distintas diluciones en cajas de CTT con el antibiótico al que confiere resistencia el plásmido introducido. Dichas cajas se incubaron a 33 °C hasta que aparecieron los transformantes, normalmente a los 5 días de incubación. Las colonias obtenidas se purificaron en cajas de CTT con el antibiótico correspondiente. En todos los casos, una vez obtenidos los transformantes, se comprobó la correcta integración del plásmido mediante PCR o secuencación.

II.8 Tratamiento enzimático del DNA

Los tratamientos enzimáticos más usuales realizados al DNA fueron:

- a) Digestión con enzimas de restricción (GE Healthcare), que se realizó siguiendo las indicaciones del suministrador.
- b) Desfosforilación de extremos 5´ con la enzima rAPid Alkaline Phosphatase (Roche) para evitar la "autoligación" de los vectores. Se realizó de acuerdo con las especificaciones del suministrador.
- c) Ligación de fragmentos de DNA, que se realizó en tampón de ligación con ATP (1 mM), y una unidad de ligasa T4 de New England Biolabs. La cantidad de DNA en la mezcla varió según el experimento. Las mezclas de ligación se incubaron a 25 °C durante 2 horas o a 16 °C durante toda la noche.

II.9 Amplificación de DNA mediante PCR

Se llevó a cabo en un termociclador Eppendorf a partir de 15–30 ng de DNA molde, con 20 pmoles de oligonucleótidos, los cuatro dNTPs (200 µM cada uno), DMSO (5%, dada la riqueza en pares G+C del DNA de *M. xanthus*), tampón de reacción y 1,5 U de *Pfu* (Promega). En las reacciones con fines diagnósticos se utilizaron bolitas para PCR *PuReTaqTM Ready-To-GoTM* (GE Healthcare), que contienen los dNTPs, la polimerasa de DNA Taq y su tampón de reacción. Los productos de PCR se purificaron con el kit *High PureTM PCR Product* (Roche). En algunos casos, el producto de PCR se sometió a electroforesis en gel de agarosa de bajo punto de fusión (Pronadisa), recortándose la banda de interés y purificándola con el mismo kit *High PureTM PCR Product Purification* (Roche) según las especificaciones del fabricante.

II.10 Secuenciación de DNA

Los fragmentos de DNA a secuenciar se amplificaron por PCR y se clonaron en un plásmido. Los productos de PCR se purificaron con el kit *High Pure[™] PCR Product Purification* (Roche) y los plásmidos con el kit *QIAprep[™] Spin Miniprep* de QIAGEN, o el kit *illustra[™] plasmid Prep Mini Spin* de GE Healthcare. La secuenciación se realizó en el Servicio de Secuenciación de la Universidad de Murcia.

II.11 Extracción, cuantificación y análisis de la calidad de RNA

II.11.1 Para ensayos de Northern blot

Para el aislamiento de RNA de *M. xanthus* en fase vegetativa se incubó la estirpe de interés hasta una OD_{550} ~0,5-0,7 en 10 ml de CTT y se recogieron 5 ml de cultivo.

Para la extracción de RNA, se resuspendió el precipitado celular en 1 ml de TRIzol[®] Reagent y se incubó durante 10 minutos a 65 °C. Se le añadieron 0,2 ml de cloroformo y se mezcló durante 15 segundos mediante inversión. Tras incubar a temperatura ambiente durante 5 minutos, las muestras se centrifugaron a máxima velocidad durante 15 minutos a 4 °C. La fase superior se precipitó con 0,5 ml de isopropanol y se congeló a -80 °C durante toda la noche. A continuación, la muestra descongelada se centrifugó a máxima velocidad durante 30 minutos a 4 °C. El pellet se lavó con 1 ml de etanol 75%, se mezcló mediante vórtex y se centrifugó a máxima velocidad durante 5 minutos a 4 °C. El pellet se secó durante 10-15 minutos al aire y

se resuspendió en 50 μ l de H₂O tratada con DEPC (H₂O DEPC) para la obtención de RNAs totales o en 300 μ l para el análisis de los RNAs pequeños.

Para la obtención de RNAs pequeños, la muestra se incubó durante 10 minutos a 65 °C y se le añadieron 40 μ l de 50% polietilenglicol 8000 y 50 μ l de NaCl 4 M. Tras mezclar bien se incubó durante 30 minutos a 4 °C y se centrifugó a 10000 rpm durante 10 minutos a 4 °C. Así, se obtuvieron aproximadamente 300 μ l de sobrenadante que contiene tRNAs, RNAs pequeños y 5S, que se precipitaron con 0,1 volumen de acetato sódico 3 M pH 5,0 y tres volúmenes de etanol 100% y se incubaron a -20 °C durante toda la noche. Tras centrifugar a 10000 rpm durante 30 minutos a 4 °C se lavaron con 600 μ l de etanol 80% y se volvieron a centrifugar a 10000 rpm durante 10 minutos a 4 °C. Se realizó un lavado adicional con etanol 80% bajo las mismas condiciones y se dejó secar el pellet durante 10-15 minutos al aire. Finalmente, el pellet se resuspendió con 23 μ l H₂O DEPC.

Para la extracción de RNA a partir de cuerpos fructíferos, las células crecidas en cultivos de 10 de CTT OD₅₅₀~0,6 fueron lavadas y resuspendidas en 2 ml de TPM y alicuoteadas en gotas de 20 µl en 4 placas de CF. Tras incubar a 33 °C el tiempo deseado, los cuerpos fructíferos fueron recogidos con un palillo en un tubo que contenía 1 ml de TPM y el equivalente a 0,5 ml de bolitas de cristal de 0,1 mm de diámetro para su lisis con un Mini-beadbeater (BioSpec) 3X (30 segundos ON/30 segundos OFF). Tras centrifugar se recuperó el sobrenadante y se procedió a la extracción de RNAs pequeños tal y como se ha descrito anteriormente.

La cuantificación del RNA se realizó mediante un espectrofotómetro *BioPhotometer* de Eppendorf a partir de una dilución 1:50 de la muestra en cubetas con un paso de luz de 8,5 mm. La integridad del RNA se determinó sometiendo una alícuota de 1 µl de RNA a electroforesis en gel de agarosa al 1%. La presencia de las bandas de los distintos tipos de RNA ribosómicos es un buen indicador de la integridad del RNA.

II.11.2 Para ensayos de RT-PCR o qPCR

La extracción de RNA para estudios de RT-PCR o qPCR se realizó siguiendo el protocolo descrito anteriormente hasta la obtención de RNAs totales. Para eliminar completamente el DNA genómico, el RNA extraído se trató con 4 U *Turbo DNasa* de Ambion durante 3 horas a 37 °C en presencia de 1X *Buffer Turbo DNasa* y 20 U *Protector RNase Inhibitor* (Roche) Una vez terminado el tratamiento, se realizó la purificación con el kit *PureLink*® *RNA Mini* de Ambion para eliminar tanto los restos de DNasa como del tampón, eluyendo en un volumen final de 40 µl.

El RNA se cuantificó mediante la medida de la absorbancia a 260 nm en un espectrofotómetro *NanoDrop* de Thermo Scientific. La calidad del RNA se comprobó mediante la electroforesis de 1 µl de RNA purificado en gel de agarosa al 1%. Además, las medidas en el *NanoDrop* ofrecen dos parámetros de calidad. Por un lado, el cociente de absorbancia 260/280 indica posibles contaminaciones por proteínas o fenol. Valores en torno a 2 de este cociente indican una alta pureza de la muestra de RNA (valores inferiores indican la presencia de contaminantes). Por otro lado, el cociente de absorbancia 260/230 permite detectar la presencia de otro tipo de contaminantes que absorben a 230, y valores entre 2-2,2 son signo de una alta pureza de la muestra.

II.12 Procedimiento de mutagénesis

II.12.1 Mutagénesis dirigida

Las mutaciones puntuales en el DNA se obtuvieron mediante síntesis química de los genes (GenScript) o mediante la técnica de extensión de productos de PCR solapantes (Ho *et al.*, 1989), y posterior clonación en el vector adecuado. Para la PCR solapante, primero se obtienen dos productos de PCR a partir de dos parejas de oligonucleotidos, que presentan la mutación de interés en uno de sus extremos y que solapan en la región mutada. En un segundo paso, los fragmentos obtenidos previa purificación mediante electroforesis en gel de agarosa, se mezclan en cantidades equimoleculares y se someten a una reacción de extensión y amplificación en presencia de los oligonucleótidos de los extremos. Mediante secuenciación del DNA se comprueba la presencia de la mutación de interés en el producto resultante y la ausencia de otras mutaciones adicionales que pudieran haberse generado durante el proceso de amplificación.

II.12.2 Generación de estirpes de *M. xanthus* portadoras de una deleción

Para la obtención de mutaciones por deleción completa de un gen, en primer lugar se amplificaron por PCR las regiones (~1 kb) inmediatamente aguas arriba y aguas abajo del gen correspondiente, y se clonaron en los plásmidos pBJ113 o pBJ114 (diferenciados en la orientación de la región de multiclonación). La utilidad de este tipo de vectores reside en que, además de contener un marcador que confiere resistencia a kanamicina, incluyen un gen que provoca sensibilidad a galactosa (gen *galK*). La estirpe silvestre de *M. xanthus* se transformó mediante electroporación con el plásmido portador de la deleción y varios de los merodiploides resistentes a kanamicina resultantes se incubaron durante varias generaciones en ausencia del antibiótico con el objetivo de permitir la pérdida del plásmido por recombinación homóloga. Los segregantes haploides resultantes se seleccionaron en cajas de CTT suplementadas con galactosa al 1%. Varias de las colonias crecidas en este medio se sometieron a análisis mediante PCR para diagnosticar la presencia o ausencia de la deleción (Figura 20).



Figura 20. Esquema del proceso de obtención de una estirpe portadora de una deleción en *M. xanthus*.

II.13 Técnicas electroforéticas

II.13.1 Electroforesis de DNA en gel de agarosa

La separación de fragmentos de DNA se realizó mediante electroforesis en geles horizontales de agarosa (Pronadisa o Sigma-Aldrich) de distintas concentraciones, preparados en tampón TAE 1X. El DNA se visualizó mediante tinción con bromuro de etidio o *RealSafe Nucleic Staining Solution* de Promega e iluminación con luz ultravioleta. La purificación de fragmentos de restricción se realizó tras su separación en un gel de agarosa de bajo punto de fusión (Pronadisa). Las porciones de agarosa con el DNA de interés se recortaron del gel con un bisturí y el DNA se purificó con el kit *High PureTM PCR Product Purification* (Roche).

II.13.2 Electroforesis de proteínas en geles de poliacrilamida-SDS

La separación en función del tamaño de las proteínas presentes en diferentes muestras se realizó mediante electroforesis en geles verticales de poliacrilamida-SDS, utilizando equipos de Bio-Rad, y siguiendo el método de (Laemmli, 1970). Los marcadores de tamaño molecular utilizados (*Spectra Multicolor Broad Range Protein Ladder*) fueron suministrados por Thermo Scientific. Antes de ser cargadas, las muestras se incubaron durante 5-10 minutos a 100 °C. La electroforesis se realizó durante una hora a 200 V.

II.13.3 Electroforesis de DNA en gel de acrilamida no desnaturalizante

Este tipo de electroforesis permite analizar cambios en la movilidad electroforética que ocurren cuando una proteína se une a un fragmento de DNA. Para la elaboración de un gel (1,5 mm de grosor) de 50 ml al 4% de poliacrilamida los componentes utilizados fueron los siguientes: 6,6 ml de acrilamida:bisacrilamida 30% (BIO-RAD) y 43,4 ml TBE 0,5X. Tras mezclar estos componentes, se filtró la solución y se añadió persulfato amónico 25% (133 μ l) y TEMED (25 μ l). Una vez polimerizado, el gel se transfirió a un tanque de electroforesis vertical que contenía tampón TBE 0,5X, se precorrió (30 minutos, 200 V, 30 mA) y, tras cargar las muestras, se realizó la electroforesis (1–2 horas, 200V, 30 mA).

II.13.4 Electroforesis de RNA en gel de acrilamida desnaturalizante

Este tipo de electroforesis se utilizó para el análisis de RNAs pequeños. El gel para la electroforesis consta de: 14,53 ml de poliacrilamida (stock 40%, 19:1

acrilamida:bisacrilamida BIO-RAD), 16,8 g de urea y 2 ml TBE 10X en un volumen final de 40 ml. Se filtra la solución y se añade persulfato amónico 10% (320 µl) y TEMED (42,6 µl). Una vez polimerizado, el gel se transfiere a un tanque de electroforesis vertical que contiene tampón TBE 0,5X, se precorre durante 45 minutos (300 V, 30 mA) y, tras cargar las muestras, se realiza la electroforesis durante 30 minutos (80V) seguido de 2 horas 30 minutos (300V, 30 mA).

II.13.5 Electroforesis de RNA en gel de agarosa

Para el análisis de RNAs totales mediante Northern blot se requiere la electroforesis de dicho RNA en un gel de agarosa al 1,2%. La preparación de este gel requiere de 1,5 g de agarosa, 12,5 ml de MOPS 10X y 22,5 ml de formaldehido 37% en un volumen final de 125 ml. El gel se corre en tampón de electroforesis que contiene 1X MOPS, 6,6% formaldehido a 80 V durante 15 minutos para favorecer la ordenación de las moléculas, seguido de 120 V durante 1 hora y 45 minutos.

II.14 Expresión y purificación de proteínas

II.14.1 Expresión y purificación de CarD

La expresión en *E. coli* y la posterior purificación de la proteína CarD de *M. xanthus* se llevó a cabo mediante el sistema IMPACT-CN de New England Biolabs. Para la expresión de la proteína CarD en *E. coli* se utilizó el vector pTYB12 que presenta una etiqueta de Inteína aguas arriba del sitio de clonación de *carD*. El plásmido resultante se transformó en la estirpe BL21(λ DE3). Para purificar la proteína, se inocularon células de la estirpe transformada en 10 ml de LB suplementado con ampicilina y el cultivo se incubó a 37 °C hasta que el cultivo alcanzó una OD₆₀₀~0,6. En ese momento, se transfieron los 10 ml de cultivo a 1 litro LB fresco con ampicilina y se incubó a 37 °C hasta alcanzar una OD₆₀₀~0,6. A continuación, las células se transfirieron durante 20 horas a 18 °C en presencia de 1 mM IPTG. Las células se recogieron por centrifugación y el precipitado celular se almacenó a -70 °C hasta su utilización.

El proceso de purificación comienzó con la resuspensión del precipitado en 30 ml de tampón B (500 mM NaCl, 20 mM Tris-HCl pH 8,5, 1 mM EDTA, 1 mM βmercaptoetanol) con 1 mM benzamidina, 1 mM de fluoruro de fenilmetilsufonilo (PMSF) y 1 pastilla de 50X *cOmplete*[™] *EDTA free Protease Inhibitor Cocktail* (Roche). Las células se lisaron en una prensa French y, mediante centrifugación (40 min, 11000 rpm, 4 °C) se eliminaron los restos celulares y proteínas insolubles. A continuación, la fracción soluble se sometió a una cromatografía de afinidad usando una resina de quitina-agarosa, previamente equilibrada con 30 ml de tampón B. Tras un lavado con 10 volumenes de tampón B con 1 M NaCl para eliminar uniones inespecíficas a la resina, se añadió a la columna un volumen de tampón B con 50 mM DTT que cubría la totalidad de la resina y se incubó durante 20 horas a 16 °C. Pasado este tiempo se eluyeron diversas fracciones para obtener la proteína libre de inteína. Las fracciones se cargaron en un gel SDS-PAGE y aquéllas que presentaban la proteína CarD se dializaron en tampón A en tres pases de 8 horas. Finalmente, se realizó un último paso de purificación mediante una columna de intercambio iónico MonoS y se concentró con Amicon CO 10000 MW antes de su congelación a -70 °C en 50% glicerol.

II.14.2 Expresión y purificación de CarD₁₋₇₂ y CarDNt₆₁₋₁₇₉

El proceso de expresión en *E. coli* y la purificación de la proteína CarD₁₋₇₂, así como de la versión CarD₆₁₋₁₇₉ es similar al descrito anteriormente para la proteína CarD. Para la expresión y purificación de la proteína se siguieron las mismas condiciones referidas para CarD salvo que para éstas no fue necesario el uso de la columna de intercambio iónico, sino que se hizo un paso de purificación mediante cromatografía de exclusión en una columna HiLoad 16/60 Superdex200. Las proteínas purificadas fueron concentradas mediante Amicon CO 10000 MW y congeladas a - 70 °C. Las identidades de ambas proteínas fueron confirmadas mediante espectrometría de masas en el Servicio de Apoyo a la Investigación (SAI) de la Universidad de Murcia.

II.14.3 Expresión y purificación de CarG

El proceso de purificación de CarG de *M. xanthus* se llevó a cabo mediante el sistema pET de Novagen. El plásmido que presenta CarG en el vector de expresión pET28b se introdujo por transformación en la estirpe de *E. coli* BL21 (λDE3). Para la purificación de CarG se partió de 1 litro de cultivo crecido hasta una OD₆₀₀~0,6 inducido con 1 mM de IPTG durante toda la noche, a 25 °C, en presencia de 1 mM de ZnCl₂. Las células se recogieron por centrifugación y se guardaron a -70 °C hasta su uso. Para la purificación de la proteína, se descongeló a 4 °C y se resuspendió en 30 ml de tampón C con 1 mM de benzamidina, 1 pastilla de *cOmplete*™ *EDTA-free Protease Inhibitor Cocktail* (Roche) y 1 mM de fluoruro de fenilmetilsulfonilo (PMSF). Las células se lisaron mediante dos pases en una prensa French. Los restos celulares y proteínas insolubles se eliminaron por centrifugación (40 min, 11000 rpm, 4 °C) y la fracción

eliminar el DNA por centrifugación (20 minutos, 13000 rpm, 4 °C) la proteína se purificó parcialmente mediante la precipitación con sulfato de amonio al 55% y se dializó (tres pases de horas) frente a tampón D. Posteriormente se pasó por dos columnas de intercambio iónico a las que CarG no se une (fosfocelulosa y MonoQ) y por último por una última fase de purificación mediante cromatografía de exclusión en una columna HiLoad 16/60 Superdex200. La proteína pura se concentró con Amicon CO 10000 MW y se congeló a -70 °C en 50% glicerol.

II.14.4 Purificación de la polimerasa de RNA de M. xanthus

Se usó un stock de proteína de polimerasa de RNA de *M. xanthus* disponible en el laboratorio (López-Rubio *et al.*, 2004).

II.15 Detección de proteínas mediante anticuerpos específicos (*Western-blot*)

La obtención de extractos de células vegetativas de *M. xanthus* para la detección de proteínas con anticuerpos específicos se realizó incubando el cultivo hasta una OD_{550} ~0,7. Las células presentes en 1 ml del cultivo se recogieron por centrifugación y se conservaron a -80 °C. El precipitado celular se descongeló a 4 °C y se resuspendió en 300 µl de una solución que contiene tampón Z, mezcla de inhibidores de proteasas 1X (*cOmplete*TM *EDTA-free Protease Inhibitor Cocktail* (Roche), 1 mM de PMSF y 1 mM de benzamidina.

Las proteínas separadas por electroforesis en geles de poliacrilamida-SDS se detectaron con anticuerpos específicos mediante la técnica de *Western* (Towbin *et al.*, 1979; Burnette, 1981). Para ello, se transfirieron a una membrana de PVDF (*Inmobilon-P* de MilliporeTM) mediante un sistema semiseco, utilizando el equipo *Trans-Blot*® *Turbo*TM *Blotting System* de Bio-Rad. Para seguir la transferencia visualmente, se utilizaron los marcadores de tamaño molecular coloreados *Recombinant Proteins Molecular Weight Markers Full-range Rainbow* (Amersham Biosciences). La transferencia se realizó durante 30 minutos a 25V y 1,0 A. Una vez realizada la transferencia, la membrana se sumergió durante 5 horas en tampón PBS-T con 5% de leche desnatada como agente bloqueante (PBS-TB). A continuación, se realizaron 3 lavados de 5-15 minutos en tampón PBS-T. Después, la membrana se incubó durante ~16 horas en presencia de 5-10 µl del anticuerpo primario diluido en 20 ml de PBS-TB. Tras eliminar el exceso de anticuerpo, aplicando tres nuevos lavados con PBS-T, la membrana se incubó durante una hora en una dilución 1:1,000 del anticuerpo secundario (conjugado a peroxidasa de rábano) en 20 ml de PBS-TB. El exceso de

anticuerpo secundario se eliminó mediante tres lavados con PBS-T. Los lavados y las incubaciones con los anticuerpos se llevaron a cabo con agitación suave y constante, y a temperatura ambiente para periodos cortos (1-2 horas) o 4 °C para periodos largos (~16 horas). Para el revelado se utilizó el sistema *ECL Prime* (GE Healthcare). Tras añadir el reactivo que contiene el sustrato quimioluminiscente para la peroxidasa conjugada al anticuerpo, la membrana se expuso a una película *Amersham Hyperfilm*TM *ECL* (GE Healthcare) durante un período de tiempo variable entre 1 y 5 minutos.

II.16 Ensayo de cuantificación relativa mediante RT-PCR en tiempo real (qRT-PCR)

El RNA de la estirpe objeto de estudio se obtuvo según el protocolo descrito en el apartado II.11.2. Para obtener el cDNA se incubaron 2 µg de RNA, 2,5 µg de hexanucleótidos aleatorios y 2,5 mM dNTPs mezclados durante 5 minutos a 65 °C en un volumen de hasta 13 µl con agua MiliQ tratada con DPEC al 0,2%. Transcurrida la incubación se pasaron los eppendorfs a hielo y se añadió tampón de la transcriptasa inversa (1X), 5 mM DTT, 40 U de Protector RNase Inhibitor (Roche) y 10 U de SuperScript[®] IV Reverse Transcriptase de Invitrogen. La mezcla se mantuvo 10 minutos a 23 °C, 10 minutos a 50 °C y finalmente 10 minutos a 80 °C. Puesto que la reacción de retrotranscripción es lineal, el rendimiento fue de 2 µg de cDNA total. Los oligonucleótidos para la detección de RNAs mensajeros específicos se diseñaron usando Primer3Plus (programa con licencia GNU) (Untergasser et al., 2007), siguiendo las opciones estándar que ofrece el programa. Las reacciones de amplificación y detección se llevaron a cabo en el termociclador StepOne™ de Applied Biosystems, con capacidad para placas de 48 pocillos (MicroAmp© Optical 48-Well Reaction Plate). La detección del "amplicón" se realizó con el fluoróforo SYBR[©] Green I, una molécula que emite fluorescencia al unirse a DNA de doble cadena. La reacción de amplificación contenía: *iTag™ Universal SYBR[®] Green Supermix* de BIO-RAD, 2 µI de cDNA (concentración variable de 100 ng a 0,01 ng), oligonucleótidos a una concentración final entre 50 y 400 nM y agua hasta un volumen final de 20 µl. Las condiciones del termociclador fueron las siguientes: 10 minutos a 95 °C para activar la *iTaq*™ DNA polymerase (contenida en la Master Mix) y a continuación 40 ciclos de 15 segundos a 95 °C y 1 minuto a 60 °C.

Para elegir qué método de cuantificación relativa debía usarse, se comprobó la eficiencia de amplificación de cada uno de los amplicones generados con cada pareja de oligonucleótidos. Para ello se construyó una recta patrón de amplificación

realizando diluciones seriadas del cDNA y depositando las siguientes cantidades (ng) en los pocillos de las placas: 100, 10, 1, 0,1, 0,01 (cada reacción se llevó a cabo por triplicado). Tras la reacción de amplificación, se representó en el eje de abscisas la cantidad de cDNA y en el eje de ordenadas los valores de Ct (ciclo de amplificación en el que la fluorescencia captada por el detector supera la línea basal durante la fase exponencial de amplificación). Los valores de Ct que se consideran adecuados son los comprendidos entre 10 y 30, ya que fuera de este intervalo la variabilidad de los resultados es muy alta. Una vez representados los datos, se calcula el factor de ajuste de los mismos a la recta obtenida, R², que debe estar entre 0,978 y 0,99 para que el ajuste sea bueno. Finalmente, se calcula la eficiencia de la amplificación (E), usando para ello la fórmula siguiente:

 $E = (10^{(-1/pendiente)}-1) \times 100$

En las muestras con una concentración de mensajero especialmente baja, el factor aplicado en las diluciones seriadas fue 1:2, si bien se mantuvo la cantidad de 100 ng de cDNA como primer punto de la recta patrón.

Independientemente del método de cuantificación relativa utilizado, se requiere un gen de referencia (control endógeno) cuya expresión sea constante a lo largo del ciclo celular, para normalizar el nivel de expresión de las muestras de cada ensayo. En este trabajo se utilizó el gen *sigA*, que mostró una eficiencia superior al 98%.

Una vez realizados todos los análisis previos, se dispusieron en los pocillos las -Normalización de la muestra problema: Ct gen diana - Ct gen endógeno = ∆Ct -Normalización de la muestra control: Ct gen diana - Ct gen endógeno = ∆Ct

-Diferencia entre problema y control: Δ Ct muestra problema - Δ Ct muestra control = $\Delta\Delta$ Ct distintas muestras. Para cada una de ellas se midió tanto el gen problema como el gen endógeno, ambos por triplicado. Seguidamente se procedió al análisis de los datos. En todos los casos analizados, el valor de eficiencia estuvo entre el 95% y 100%, lo cual permitió realizar la cuantificación relativa con el método $\Delta\Delta$ Ct (Pfaffl, 2001). Para ello se usaron las siguientes fórmulas:

Al usar un fluoróforo de unión inespecífica al DNA se realiza una curva de disociación del producto de PCR para comprobar si aparecía más de un pico, que indicaría la amplificación de más de un fragmento o la hibridación entre los cebadores. Esta curva se realizó en el mismo termociclador que la reacción de amplificación utilizando un programa estándar.

II.17 Ensayo de RT-PCR

Para la realización de los experimentos de RT-PCR se partió de RNA extraído tal y como se describe en el apartado II.11.2, que se procesó para obtener cDNA como se ha detallado en el apartado anterior II.16. Una vez obtenido el cDNA se realizó una reacción de amplificación usando bolitas para PCR *PuReTaqTM Ready-To-GoTM* (GE Healthcare), a las que se les añadió 5% DMSO, 200 nM de oligonucleótidos, 25 ng/µl de cDNA o su equivalente en RNA para obtener un control negativo de contaminación por DNA. Se ajustó con H₂O hasta un volumen final de 25 µl y se sometió a un programa de 30 a 35 ciclos de amplificación según las parejas de oligonucleótidos. Los productos de PCR obtenidos se cargaron en un gel de agarosa al 1,5%. La distribución de las bandas obtenidas fue tomada en un transiluminador *Gel Logic 200 Imaging System* de Kodak.

II.18 Ensayo de doble híbrido bacteriano (DHB)

El sistema de doble híbrido bacteriano en *E. coli* (*BACTH, Bacterial Adenylate Cyclase Two-Hybrid System,* de Euromedex) permite estudiar la interacción entre dos proteínas. Este sistema se basa en que el dominio catalítico de la ciclasa de adenilato de *Bordetella pertussis* consta de dos fragmentos complementarios, T25 y T18, que son inactivos cuando están físicamente separados pero activos cuando se fusionan a polipéptidos que interaccionan físicamente. La interacción puede seguirse mediante la activación del gen chivato *lacZ* y la consiguiente producción de la enzima β -galactosidasa (Figura 21). Dicha actividad enzimática puede ser analizada tanto cualitativamente (en placas con X-Gal, un sustrato cromogénico de la β -galactosidasa), como cuantitativamente siguiendo el protocolo descrito en el siguiente apartado. Este sistema permite la fusión de las proteínas de interés tanto al extremo C-terminal (plásmidos pKT25 y pUT18C) como al extremo N-terminal (plásmidos pKNT25 y pUT18) de estos dos fragmentos.



Figura 21. Esquema del sistema del doble híbrido bacteriano. Complejo activo debido a la interacción física de las proteínas A y B, unidas traduccionalmente a los fragmentos T25 y T18 (arriba), y los fragmentos T25 y T18 inactivos por separado (abajo).

II.19 Ensayo de β-galactosidasa

La actividad β -galactosidasa de una estirpe (de *M. xanthus* o *E. coli*) se estimó cualitativamente mediante crecimiento en cajas del medio adecuado suplementado con X-Gal (40 µg/ml), o en placas sin X-Gal a las que posteriormente se le añadió una solución de agarosa-X-Gal. Ambos protocolos permiten observar una coloración azul cuya intensidad depende de la actividad β -galactosidasa de la estirpe.

II.19.1 Ensayo de β-galactosidasa en *M. xanthus*

Para cuantificar la actividad β -galactosidasa específica en una estirpe de *M. xanthus*, ésta se inoculó en medio líquido CTT más el antibiótico correspondiente. Cuando el cultivo alcanzó la densidad celular deseada, se tomaron muestras de 1 ml que se centrifugaron, se resuspendieron en tampón Z y se almacenaron a -70 °C hasta su análisis. Cuando la expresión de β -galactosidasa era dependiente de la iluminación con luz azul, el cultivo inicial se incubó en oscuridad y, una vez alcanzada una OD₅₅₀~0.4, se dividió en dos cultivos, uno de los cuales se mantuvo en condiciones de oscuridad y el otro se incubó con luz azul. Tras 8 o 14 horas de incubación en luz u oscuridad, se tomaron muestras de 1 ml de cada cultivo, se centrifugaron y se resuspendieron en tampón Z, almacenándose a -20 °C.

Las muestras se descongelaron a temperatura ambiente justo antes de ser analizadas y se colocaron en hielo. La rotura de las células se realizó mediante tratamiento con SDS/cloroformo (Griffith & Wolf, 2002) para las muestras crecidas en fase vegetativa, mientras que para las células crecidas en fase de desarrollo la rotura se llevó a cabo añadiendo a cada muestra unos 100 µg de bolas de silicato de 0,1 mm de tamaño, como abrasivo para la rotura de los cuerpos fructíferos. El ensayo de actividad β-galactosidasa se realizó a 33 °C según Miller (1972), empleando 50-200 µl del extracto celular. La concentración de proteínas se midió según Lowry, *et al.*, (1951). Salvo que se indique lo contrario, la actividad β-galactosidasa específica se expresa en nanomoles de o-nitrofenol (generado por rotura enzimática del sustrato ONPG) producidos por minuto y miligramo de proteína, de acuerdo con la ecuación:

Actividad específica = (213 x A₄₂₀) / (t x ml x (mg/ml))

donde 213 es el coeficiente de extinción del σ -nitrofenol y convierte la densidad óptica medida a 420 nm en los ensayos de actividad en nanomoles de σ -nitrofenol producidos; *ml* es el volumen del extracto utilizado en la determinación de la actividad; *mg/ml* es la concentración de proteínas de la muestra; y *t* representa el tiempo de incubación en minutos a 33 °C antes de detener la reacción.

Para medir actividad β -galactosidasa de células en desarrollo, se inoculó la estirpe deseada en medio líquido CTT más el antibiótico correspondiente y cuando el cultivo alcanzó una OD₅₅₀~0,6-0,8 se centrifugó durante 3 minutos a 12500 rpm a temperatura ambiente. Las células precipitadas se resuspendieron en el volumen de tampón TPM necesario para dejar la muestra a una OD₅₅₀~6,0 y se depositaron gotas de 10 µl en placas de CF. Una vez que las gotas se absorbieron en el agar, las cajas se incubaron a 33 °C durante 4, 8, 12 y 24 horas, y las muestras se recogieron con un palillo, se mezclaron con 1 ml de buffer Z y se almacenaron a -20 °C. Para su análisis, las muestras fueron descongeladas, colocadas en un eppendorf con bolitas de cristal de 0,1 mm de diámetro en un volumen equivalente a 0,5 ml y lisadas con un Minibeadbeater (BioSpec) 3X (60 segundos ON/60 segundos OFF). El sobrenandante fue analizado siguiendo el protocolo descrito previamente.

II.19.2 Ensayo de β-galactosidasa en *E. coli*

Para medir la activación del gen chivato *lacZ* en el sistema del doble híbrido bacteriano, la estirpe de interés se inoculó en medio rico LB más los antibióticos necesarios y se incubó a 30 °C durante toda la noche. El cultivo se diluyó 20 veces en el mismo medio y se añadió IPTG hasta 0,5 mM para inducir la expresión de las

proteínas híbridas. Este cultivo se incubó a 30 °C hasta que alcanzó la fase estacionaria (aproximadamente 16 horas). La toma de muestras y el análisis cuantitativo de la actividad β -galactosidasa se realizaron siguiendo el mismo procedimiento descrito para *M. xanthus*.

II.20 Ensayo de retraso en la movilidad electroforética (EMSA)

Los ensayos de cambio en la movilidad electroforética se realizaron usando como sondas fragmentos de DNA de entre 170 y 304 pb con las regiones de interés, indicadas más adelante. Los distintos fragmentos se obtuvieron por PCR empleando un oligonucleótido marcado previamente en su extremo 5´ con [γ^{32} P]ATP y quinasa de polinucleótido de T4 de Amersham Biosciences. Tras la amplificación, la radiactividad no incorporada se eliminó mediante centrifugación en columna de Sephadex G-25 (*DNA grade*, Amersham Biosciences) y el fragmento se purificó con el kit *High Pure*TM *PCR Product Amplification* (Roche).

La reacción de unión se realizó en un volumen de 20 µl que contenía 1 nM del fragmento de DNA marcado (~10000 cpm) y las cantidades necesarias de las proteínas, en un tampón de reacción (70 mM KCl, 25 mM Tris pH 8,0, 1 mM DTT, 10 mM MgCl₂, 50 ng/µl de BSA, 10% glicerol), con 1 µg de poli dG-dC, como competidor no específico. Las muestras fueron incubadas (30 minutos, 37 °C). Tras dicha incubación, cuando fue necesario, se añadió 1 µg de heparina y se incubó durante 5 minutos a 37 °C. Las muestras se sometieron entonces a electroforesis en un gel de acrilamida no desnaturalizante al 4% (1-2 horas, 200 V, 30 mA). Las posiciones de las bandas se detectaron, tras secar el gel, mediante autorradiografía en película Biomax MS de Kodak (1 o 2 horas exposición a -70 °C).

II.21 Northern blot

II.21.1 Northern blot de RNAs pequeños

Los experimentos de *Northern blot* usando RNAs pequeños se realizaron a partir de RNA extraído tal y como se detalla en el apartado II.11.1. Cada muestra a analizar contenía 10 μ g de RNA en H₂O tratada con DEPC y 10 μ l de tampón de carga 2X en volumen final de 20 μ l. La mezcla se calentó (95 °C, 3 minutos) para eliminar estructuras secundarias del RNA y a continuación se dejó 5 minutos en hielo. Las muestras se cargaron en un gel desnaturalizante de poliacrilamida al 15% (19:1 acrilamida:bisacrilamida) preparado como se indica en el apartado II.13.4 utilizando como tampón de electroforesis TBE 0,5X. Una vez finalizada la electroforesis, se

transfirió el RNA a una membrana de nitrocelulosa (Amershan Hybond N+, de GE Healthcare) mediante electrotransferencia a 250 mA constantes durante 45 minutos. (Semi-Dry Electroblotting Unit, Sigma), usando como tampón de transferencia TBE 0,5X. El RNA se fijó a la membrana mediante radiación ultravioleta (0,120 J/cm²). La prehibridación de la membrana se realizó con 20 ml de tampón de prehibridación con 0,1 mg/ml de DNA de salmón durante 2 horas a 30 °C. Posteriormente se le añadió la sonda radiactiva previamente desnaturalizada (10 minutos, 95 °C). Se incubó durante toda la noche a 30 °C y a la mañana siguiente se procedió a realizar el lavado de la membrana a 50 °C con una solución 2X SSC y 0,2% SDS. Una vez lavada la membrana se puso en contacto con películas Biomax MS de Kodak[®] a -70 °C durante 1-5 horas, según el experimento.

II.21.2 Northern blot de RNAs total

Para el análisis de RNAs totales mediante Northern blot se procedió con la extracción de RNA total como se indica en el apartado II.11.1. Cada muestra contenía 10 µg de RNA en un máximo de 10 µl al que se le añadieron 30 µl de la mezcla de Northern blot (1,3X MOPS, 64,5% formamida, 8,4% formaldehido). La mezcla se calentó (65 °C, 10 minutos) para eliminar estructuras secundarias del RNA y a continuación se le añadieron 4,5 µl de tampón de carga. Las muestras se cargaron en un gel de agarosa 1,2% preparado como se refiere en el apartado II.13.5. Una vez acabada la electroforesis, el gel se lavó durante 1,5 horas con agua, cambiándosela cada 15 minutos. El RNA se transfirió por capilaridad durante toda la noche a una membrana de Nylon cargada positivamente (Amershan Hybond N+, de GE Healthcare) usando como tampón 20X SSC y se fijó a la membrana mediante radiación ultravioleta (0,120 J/cm²). La prehibridación/hibridación de la sonda marcada se hizo de manera similar a la descrita anteriormente en el apartado II.21.1 excepto que la temperatura de prehibridación/hibridación/lavados fue de 65 °C. El lavado de la membrana fue realizado usando tampones 2X SSC con 0,1% SDS en los dos primeros lavados, 1X SSC con 0,1% SDS y 0,5X SSC con 0,1% SDS en los dos siguientes lavados. A continuación, la membrana se puso en contacto con películas Biomax MS de Kodak $^{\scriptscriptstyle (\! 8\!)}$ a -70 °C durante 24 horas.

II.22 Inmunoprecipitación de cromatina-qPCR (ChIP-qPCR)

A 50 ml de cultivo de *M. xanthus* crecido hasta la mitad de la fase logarítmica $(OD_{550}\sim0,7)$ en CTT, se añadió rifampicina (25 µg/ml) y se mantuvo a 33 °C, 300 rpm durante 30 minutos para bloquear la unión de las proteínas a su sitio de unión en el

DNA. Para realizar el entrecruzamiento se añadió 1% formaldehido (Sigma) en presencia de 10 mM tampón fosfato sódico (pH 7,6) y se incubó durante 10 minutos a temperatura ambiente con agitación suave, para fijar las interacciones proteína-DNA y proteína-proteína. A continuación, el cultivo se pasó a hielo durante 30 minutos y se centrifugó (7800 rpm, 10 minutos, 4 °C). Se lavó el precipitado celular con 25 ml de PBS (Sigma) en frío y se volvió a centrifugar bajo las mismas condiciones. Finalmente, el precipitado celular se resuspendió en 1 ml de PBS. Tras centrifugar (13000 rpm, 5 minutos a 4 °C) y eliminar completamente el sobrenadante, se congeló a -80 °C hasta su uso. Para ChIP, el precipitado fue descongelado en hielo, resuspendido en 470 µl de solución de lisis (10 mM Tris pH 7,5, 1 mM EDTA, 100 mM NaCl, 2,2 mg/ml lisozima, 25X cocktail inhibidor de proteasas (Roche)) e incubado a temperatura ambiente durante 30 minutos con agitación suave (100 rpm). Posteriormente se añadió 550 µl de ChIP buffer (16,7 mM Tris pH 8,1, 1,2 mM EDTA, 167 mM NaCl, 1,1% Triton X-100, 1X cocktail inhibidor de proteasas (Roche)) y tras invertir varias veces se incubó a 37 °C durante 10 minutos. 300 µl de la muestra lisada se sonicaron durante 60 ciclos (30 segundos de sonicación/30 segundos de descanso) en un Bioruptor®Plus UCD-300 de Diagenode para obtener fragmentos de DNA de un tamaño de ~0,5 Kb. La muestra se centrifugó (13000 rpm, 5 minutos, 4 °C) y el sobrenadante fue transferido a un tubo limpio que se centrifugó nuevamente (13000 rpm, 5 minutos, 4 °C). Del sobrenadante obtenido de esta última centrifugación, se tomaron 20 µl como muestra sin inmunoprecipitar o "input" y al restante (~280 µl) se le añadieron 325 µl de ChIP buffer con 0.01% SDS v 5 ul de anticuerpo monoclonal ANTI-FLAG[®] M2 (F3165) de Sigma-Aldrich para proteínas etiquetadas con el epítopo FLAG, o de anticuerpo policional anti-CarD (Padmanabhan et al., 2001). Tras incubar durante toda la noche a 4 °C en rotación se le añadieron 30 µl de bolitas magnéticas Dynabeads® Protein A (Life Technologies) previamente lavadas con PBS y tratadas con 1 mg/ml de BSA. La muestra junto al anticuerpo γ las bolitas fueron incubadas durante 4-6 horas a 4 °C en rotación. Las bolitas fueron entonces lavadas durante 15 minutos a 4 ºC con rotación en el siguiente orden: (a) 1 ml de tampón de baja sal (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8,1, 150 mM NaCl); (b) 1 ml del mismo tampón a alta sal (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8,1, 500 mM NaCl); (c) 1 ml de tampón de lavado con LiCl (0,25 M LiCl, 1% NP-40, 1% SDS, 1 mM EDTA, 10 mM Tris pH 8,1); (d) dos lavados con 1 ml de TE buffer (10 mM Tris pH 8,0, 1 mM EDTA). La muestra fue eluída en dos tandas de 100 µl de buffer de elución (1% SDS, 0,1 M NaHCO₃) durante 10 minutos a 65 °C. A las muestras eluídas y las de "input" se le añadió 5 µl proteinasa K (20 µg/µl) y se incubó primero a 42 °C durante 2 horas para degradar las proteínas unidas al DNA, seguido de 65 °C durante 6 horas para eliminar

el entrecruzamiento del DNA (consigo mismo y con las proteínas que no hayan sido degradadas por el tratamiento anterior) (Figura 22). El DNA fue aislado con el kit *High Pure*TM *PCR Product Amplification* (Roche) y cuantificado mediante PCR cuantitativa (qPCR) a partir de curvas patrón realizadas con diluciones seriadas del "input" usando los oligonucleótidos que amplifican las regiones de interés. El enriquecimiento para cada promotor fue expresado bien como el porcentaje del input o como el ratio de DNA para un promotor examinado respecto al observado en una región intragénica (control negativo).



Figura 22. Representación esquemática de las principales etapas de la inmunoprecipitación de cromatina.

II.23 Chromatin Immunoprecipitation Sequencing (ChIP-Seq)

Para la identificación de los sitios de unión globales de una determinada proteína se realizaron estudios de ChIP-Seq. La toma de muestra y la inmunoprecipitación de cromatina se realizaron de forma similar a la descrita en el apartado II.22 salvo que para estos estudios globales se parte de 80 ml de cultivo crecido hasta una OD₅₅₀~0,7 y tratados con rifampicina. El entrecruzamiento y los lavados fueron realizados tal y como se detalla anteriormente. Así, a 450 µl de muestras procesadas y lisadas se le añadieron 550 µl de ChIP buffer con SDS y se incubaron con 90 µl de bolitas magnéticas que habían sido lavadas con PBS y tratadas con 1 mg/ml BSA para eliminar las uniones inespecíficas de los complejos DNA-
proteína a las bolitas. Posteriormente, a la mezcla sin bolitas se le añadieron 5 µl del anticuerpo correspondiente y se incubó en rotación a 4 °C durante toda la noche. Los complejos DNA-proteínas-anticuerpos fueron capturados al adicionar de nuevo 90 µl de bolitas magnéticas, tras incubar durante 5 horas a 4 °C en rotación. Las muestras fueron lavadas como se detalla anteriormente y el eluído fue tratado con 5 µl de proteinasa K (20 mg/ml) durante 2 horas a 45 °C y 6 horas a 65 °C (Figura 22).

El DNA se purificó siguiendo las instrucciones del kit *High Pure*TM *PCR Product Amplification* (Roche) y posteriormente se cuantificó mediante *Qubit 3.0 Fluorometer* de ThermoFisher Scientific usando *Qubit dsDNA HS Assay Kit* de ThermoFisher Scientific que permite la detección de muy bajas cantidades de DNA. A partir de 10 ng de DNA se procedió a realizar la librería usando *TruSeq ChIP Library Preparation Kit* de Illumina según las indicaciones recomendadas por el fabricante. Las librerías fueron enviadas para su secuenciación al Laboratoire de Quimie Bactérienne CNRS, Marsella (Francia).

II.24 Análisis bioinformático

La búsqueda regular de proteínas similares a las diferentes parejas σ -ECF/ anti-o se llevó a cabo mediante el programa BLASTP (Altschul et al., 1997) del NCBI (https://www.ncbi.nlm.nih.gov/). Los análisis de comparación de secuencias se realizaron mediante el programa ClustalW2 (Larkin *et al.*, 2007) del EMBL-EBI (https://www.ebi.ac.uk/). Se ha utilizado el programa BoxShade (https://embnet.vitalit.ch/sofware/Box form.html) para analizar el grado de conservación o similitud de las proteínas de interés. Las predicciones de la topología en la membrana para los factores anti-o CarR y DdvA fue realizada mediante los programas informáticos Tmpred (https://embnet.vital-it.ch/software/TMPRED form.html), SOSU (http://harrier.nagahama-i-bio.ac.jp/sosui/) y TMHMM (http://www.cbs.dtu.dk/services/ TMHMM/). El contexto genómico y los operones objeto de estudio en esta tesis fueron analizados software Artemis (http://www.sanger.ac.uk/ usando el resources/software/Artemis).

III. RESULTADOS

III.1 El dominio N-terminal de CarD: función y estructura

El contenido incluido en este apartado deriva del artículo publicado por Bernal-Bernal *et al.*, en PLoS One (2015), 10(3): e0121322). Los estudios estructurales mediante RMN se realizaron en colaboración con la Dra. María Ángeles Jiménez López perteneciente al Departamento de Química Física Biológica del Instituto de Química-Física "Rocasolano", CSIC, Madrid.

Tal y como se comentó en la Introducción, el Grupo de Investigación en el que se ha desarrollado esta tesis ha estudiado en profundidad la proteína CarD, un factor transcripcional con una novedosa arquitectura de dominios (Figuras 13 y 23) que actúa como un regulador global en *M. xanthus*. Dichos estudios indican que CarD presenta un dominio C-terminal (CarDCt), desestructurado y de unión al DNA, similar a las proteínas eucarióticas de tipo HMGA. CarDCt está constituido por una región acídica rica en aspártico y glutámico, seguida de cuatro repeticiones ricas en aminoácidos básicos (Arg-Gly-Arg-Pro) denominadas ganchos AT (Figura 23). En cambio, el dominio N-terminal de CarD (CarDNt) presenta una estructura terciaria definida y es el dominio involucrado en la interacción con la subunidad β de la RNAP (β RNAP), la dimerización y la interacción con CarG, una proteína de unión a Zn incapaz de unirse al DNA, con la que forma un complejo transcripcional que regula la acción de determinados factores σ-ECF en *M. xanthus* (Figura 23) (Abellón-Ruiz *et al.*, 2014; Elías-Arnanz *et al.*, 2001; Peñalver-Mellado *et al.*, 2006).



Figura 23. Alineamiento de las proteínas CarD y CdnL. La línea discontinua verde indica el dominio N-terminal de CarD. La flecha azul acota la región implicada en la interacción con la β RNAP. Los triángulos invertidos con relleno rojo representan los aminoácidos mutados para el estudio de la interacción con la β RNAP o con CarG. Los rectángulos rojos indican las cuatro repeticiones ricas en aminoácidos básicos presentes en CarD. Debajo de la secuencia se indica para CdnL, mediante flechas negras las láminas β del subdominio de interacción con la β RNAP, con un rectángulo negro la hélice 3₁₀. y mediante rectángulos grises las hélices α .

III.1.1 El dominio N-terminal de CarD complementa parcialmente la falta de CarD

Estudios previos mostraron que la ausencia del dominio CarDNt provoca una pérdida total de la función de CarD in vivo, indicando que el dominio de unión al DNA de CarD, CarDCt, por sí solo es incapaz de realizar la función de CarD in vivo (Cayuela et al., 2003), a pesar de que in vitro se une al DNA con la misma afinidad y especificidad que la proteína completa (Padmanabhan et al., 2001). Quedaba pendiente estudiar el comportamiento del dominio CarDNt solo, para poder determinar así la contribución del dominio C-terminal tipo HMGA a la función de CarD. Para ello, se estudió la actividad in vivo de CarDNt en dos promotores dependientes de factores σ -ECF regulados por CarD: el promotor P_{QRS} (Elías-Arnanz et al., 2010; Elías-Arnanz et al., 2011; García-Heras et al., 2013) y el promotor P_{ddvSA} (Abellón-Ruiz et al., 2014). El primero, P_{QRS}, participa en la respuesta carotenogénica a la luz azul, y está regulado por el factor σ -ECF CarQ y su factor anti-σ CarR, y por el complejo CarD-CarG. Se generó una estirpe con CarDNt en el sitio endógeno, en sustitución de CarD, y se comparó su fenotipo con el de la estirpe silvestre y el de las estirpes mutantes $\Delta carD$ (con la deleción completa de *carD*) y $\Delta carDNt$ (con la deleción solo de *carDNt*). A nivel cualitativo, es posible analizar el fenotipo mediante el color de los cultivos (a partir de gotas depositadas en medio sólido). Así, mientras que la estirpe silvestre presenta color amarillo cuando se incuba en la oscuridad y rojizo cuando crece en la luz (al activarse la producción de carotenos), las estirpes $\Delta carDNt$ y $\Delta carD$ permanecen amarillas tanto en oscuridad como en luz (el color amarillo se debe a un pigmento no carotenoide). La estirpe carDNt, portadora de la deleción del dominio HMGA, presentó una coloración anaranjada en la luz, indicando una activación de la carotenogénesis, aunque a un nivel menor que la observada en la estirpe silvestre (Figura 24A). Para cuantificar el defecto de dicha estirpe, se analizó la expresión del promotor PQRS fusionado al gen chivato lacZ (PQRS::lacZ), tras 14 horas de crecimiento en la oscuridad o en la luz. En comparación con las estirpes *\(\Delta carD\)* o $\Delta carDNt$, que mostraron niveles muy bajos de actividad β -galactosidasa tanto en oscuridad como en la luz, la estirpe carDNt mostró cierta fotoinducción del promotor P_{QRS}, aunque en torno a unas cuatro veces menor que la observada con la estirpe silvestre (Figura 24B).

El segundo promotor estudiado para examinar la actividad de CarDNt *in vivo*, P_{ddvSA} , depende de CarD-CarG, y de la pareja formada por el factor σ -ECF DdvS y su factor anti- σ DdvA. Como se desconoce el estímulo desencadenante de la inactivación de DdvA y que conduciría a la liberación de DdvS, se generaron las estirpes en un fondo genético $\Delta ddvA$, que carece del factor anti- σ DdvA. El análisis de la expresión del

promotor P_{ddvSA} fusionado a *lacZ* (P_{ddvSA} ::*lacZ*) mostró resultados similares a los observados en P_{QRS} . Así, en la estirpe $\Delta ddvA$ carDNt se observó un descenso de la actividad β -galactosidasa entre tres y cuatro veces respecto a la estirpe $\Delta ddvA$, que expresa la proteína CarD completa. Nuevamente, las estirpes con la deleción completa de *carD* o con la deleción de *carDNt* mostraron niveles despreciables de actividad β -galactosidasa (Figura 24C).

En su conjunto, los resultados indican que mientras CarDNt es esencial para la función de CarD, CarDCt es parcialmente dispensable. Su presencia, no obstante, potencia de manera significativa la actividad de CarD, permitiendo que las células de *M. xanthus* respondan de manera óptima al estrés luminoso y al estrés, desconocido, que provoca la inactivación de DdvA.



Figura 24. Comparación del efecto de la eliminación de los dominios de CarD versus la eliminación de la proteína completa en *M. xanthus*. A) Fenotipo de las estirpes de *M. xanthus* indicadas crecidas tanto en oscuridad como en luz durante distintos tiempos (horas). B) Expresión de P_{QRS}::/*acZ* en las estirpes indicadas crecidas en oscuridad y en luz durante 14 horas. C) Expresión de P_{ddvSA}::/*acZ* en las estirpes indicadas crecidas en oscuridad hasta una OD₅₅₀: 0.8. La obtención de la estirpe CarDNt y su caracterización preliminar forman parte del Trabajo de Fin de Máster realizado por Gema García Martínez (Bernal-Bernal *et al.*, 2015).

III.1.2 Disección molecular de las interacciones del dominio N-terminal de CarD con la RNAP y con CarG

El auge de la secuenciación de genomas bacterianos (y su análisis) reveló que CarDNt forma parte de una gran familia denominada CarD_CdnL_TRCF (PF02559), a la que también pertenece CdnL, y que incluye un grupo de proteínas que interaccionan con la RNAP (Cayuela *et al.*, 2003; García-Moreno *et al.*, 2010; Stallings *et al.*, 2009). Tanto CdnL como CarD coexisten en *M. xanthus* y se ha demostrado que, pese a la similitud de su secuencia, son proteínas funcionalmente diferentes: ambas presentan un subdominio N-terminal conservado de interacción con β RNAP (Figura 23) (García-Moreno *et al.*, 2010) pero solo CarD es capaz de interaccionar con CarG a través de su dominio N-terminal (García-Heras *et al.*, 2009; Peñalver-Mellado *et al.*, 2006), formando un complejo que regula la acción de determinados factores σ-ECF (Abellón-Ruiz *et al.*, 2014).

Uno de los objetivos de esta tesis ha sido el estudio de la base molecular de la interacción de CarD con la RNAP y su análisis comparativo con CdnL, así como el análisis de la interacción de CarD con CarG, para comprender mejor la acción reguladora sobre la transcripción realizada por CarD, y la base de sus diferencias con CdnL.

III.1.2.1 Análisis estructural de CarDNter y su comparación con CdnL

El programa PSIPRED predice una estructura secundaria similar para CarDNt y CdnL, que muestran un grado de identidad y similitud del 35% y 58%, respectivamente (ver Figura 14, apartado 1.2.5.2). Sin embargo, la comparación de sus estructuras secundarias usando dicroísmo circular en el UV lejano indicaron un contenido en hélices α del 28% para CarDNt, inferior al observado para CdnL (~50%) (Figura 25A). Como se comentó en la Introducción, CarDNt está constituido por dos subdominios funcionalmente diferentes (Figura 23): un subdominio N-terminal que abarca hasta el aminoácido 72, CarD₁₋₇₂, que interacciona con un subdominio de β RNAP (Mx β_{19-148} ; aminoácidos 19 a 148); y un subdominio C-terminal entre los aminoácidos 61 a 179, CarD₆₁₋₁₇₉, que interacciona con CarG (Bernal-Bernal *et al.*, 2015). Debido a las dificultades para la obtención de la estructura de CarDNt a alta resolución usando resonancia magnética nuclear (RMN) (por problemas de señal/ruido debido a la presencia de monómeros y dímeros, y su intercambio en solución), se llevaron a cabo nuevos intentos usando los dos subdominios de CarDNt por separado. Para ello, se purificaron tanto CarD₁₋₇₂ como CarD₆₁₋₁₇₉, pero este último no se mostró muy estable, en consonancia con experimentos previos de proteólisis limitada (Padmanabhan *et al.*, 2001) y a diferencia de lo observado con el dominio equivalente de CdnL, que es un dominio estable y compacto (Gallego-García *et al.*, 2014). Así pues, a pesar de las similitudes en sus secuencias, CarD y CdnL presentan diferencias estructurales y funcionales significativas, especialmente entre el subdominio C-terminal de CarDNt y el dominio C-terminal de CdnL.



Figura 25. Análisis estructural de CarDNt. A) Comparación de las estructuras secundarias de CarDNt y CdnL mediante dicroísmo circular en el UV-lejano. B) Estructura terciaria de CarD₁₋₇₂ determinada por RMN. C) Superposición de las estructuras de CarD₁₋₇₂ y de CdnLNt. En rosa se muestra CarD₁₋₇₂ y en verde CdnLNt (Bernal-Bernal *et al.*, 2015).

Dado el comportamiento indeseado de CarD₆₁₋₁₇₉, solo ha sido posible determinar la estructura del subdominio CarD₁₋₇₂ por RMN. Los resultados obtenidos han permitido concluir que dicho subdominio está constituido por cinco láminas β antiparalelas que comprenden los residuos 15-17, 23-34, 37-46, 51-56 y 65-67 y que adoptan una estructura en "sándwich retorcido" parecida a los dominios tipo Tudor; las láminas β se disponen con una topología β 5- β 1- β 2- β 3- β 4 y, entre β 4- β 5, se observa una hélice 3₁₀ (Figura 25B). La superposición de la estructura determinada para CarD₁₋₇₂ con la del domino N-terminal de CdnL indica un buen solapamiento de ambas estructuras (Figura 25C), lo que apoya la idea de que la diferencia principal entre CarDNt y CdnL radica en la región C-terminal de ambas proteínas.

III.1.2.2 Análisis de las mutaciones en el subdominio de interacción con la RNAP

Como se ha comentado, CarDNt interacciona con el subdominio Mx β_{19-148} de β RNAP (Figura 26E), como lo hace también la proteína CdnL. Además, varios residuos en CdnL (F36, M49 o P51; Figura 23) implicados en su interacción con Mx β_{19-148} están conservados en CarDNt (Figura 23) y ocupan posiciones similares en la estructuras determinadas por RMN (Figura 26A) (Gallego-García *et al.*, 2014). Los ensayos de doble híbrido bacteriano (DHB) indicaron que la mutación a alanina de estos residuos en CarDNt (F41, M54 o P56; Figura 26A) no afecta a la interacción con CarG (Figura 26B), pero sí a la interacción con Mx β_{19-148} (Figura 26C). Estos datos sugieren que residuos equivalentes en CarD y en CdnL median la interacción con la subunidad β de la RNAP.



Figura 26. Análisis mutacional de la interacción entre CarDNt y β RNAP. A) En verde, los aminoácidos mutados en CarD₁₋₇₂; en violeta, los aminoácidos equivalentes en CdnL. B) Análisis cualitativo mediante DHB de la interacción entre CarG y los mutantes de CarDNt. C) Análisis cuantitativo mediante DHB de la interacción entre los mutantes de CarDNt y la β RNAP. D) Análisis cuantitativo mediante DHB de la interacción entre los mutantes de β RNAP. D) Análisis cuantitativo mediante DHB de la interacción entre los mutantes de β RNAP. D) Análisis cuantitativo mediante DHB de la interacción entre los mutantes de β RNAP. MAP. D) Análisis cuantitativo mediante DHB de la interacción entre los mutantes de β RNAP. MAP. D) Análisis cuantitativo mediante DHB de la interacción de β RNAP de diferentes bacterias. Para *M. xanthus:* Mx β_{19-148} , *T. thermophilus:*Tt β_{10-133} , *E. coli:* Ec β_{19-142} y *M. tuberculosis:* Mt β_{39-166} (Bernal-Bernal *et al.*, 2015).

Para determinar si β RNAP utiliza los mismos residuos para interaccionar con CarDNt y CdnL, se realizaron mutaciones en un tramo conservado en las subunidades β de distintas bacterias (Figura 26E) que ha sido implicado en la interacción de β RNAP con CdnL y algún otro miembro de la familia (Gallego-García *et al.*, 2014; Gallego-García *et al.*, 2012; Weiss *et al.*, 2012). Al igual que en la interacción de Mx β_{19-148} con CdnL (Gallego-García *et al.*, 2014), fueron las mutaciones a alanina de los aminoácidos V123 y E125, pero no las de D122 y K124, las que afectaron a la interacción con CarDNt (Figura 26D). En su conjunto, los datos indican que la interacción de CarDNt y de CdnL con la β RNAP se establece a través de residuos conservados en superficies equivalentes.

A continuación, se examinó el efecto de las mutaciones F41A, M54A o P56A en la función de CarD en *M. xanthus*. Tras comprobar mediante *Western blot* que estas mutaciones no afectaban a la estabilidad de las proteínas en *M. xanthus* (Figura 27A), se examinó el fenotipo de color de estos mutantes, así como la actividad de la fusión P_{QRS} ::*lacZ*, en la oscuridad y en la luz. En los ensayos cualitativos se observó que los mutantes adquirieron una coloración anaranjada en la luz, similar a la de la estirpe silvestre (Figura 27B); sin embargo, los ensayos cuantitativos revelaron niveles de expresión de P_{QRS} ::*lacZ* entre un 60%-70% de los observados en la estirpe silvestre (Figura 27C). Estos datos sugieren que las mutaciones que eliminan la interacción con la RNAP disminuyen pero no eliminan completamente la actividad de CarD en *M. xanthus*, a diferencia de lo observado para las mutaciones equivalentes con CdnL, que provocaron un defecto mucho más acusado (Gallego-García *et al.*, 2014).



Figura 27. Efecto *in vivo* de las mutaciones en residuos de CarD implicados en la interacción con la RNAP. A) Análisis mediante *Western blot* de las estirpes mutantes de CarD (F41A, M54A o P56A). B) Fenotipo de las estirpes de *M. xanthus* crecidas tanto en oscuridad como en luz durante las horas indicadas. C) Expresión de P_{QRS} ::*lacZ* de estirpes crecidas en oscuridad y en luz durante 14 horas (Bernal-Bernal *et al.*, 2015).

III.1.2.3 Análisis de las mutaciones en el subdominio de interacción con CarG

La ausencia del dominio CarDNt provoca la falta de función de CarD, pero la eliminación de la interacción con la RNAP solo afecta parcialmente a la función, por lo que cabe esperar un papel importante para la región $CarD_{61-179}$ de CarDNt, que interacciona con CarG, pero no con la RNAP. De hecho, el dominio equivalente en CdnL (que no interacciona con CarG) es crucial para su función, y en él se identificaron varios residuos hidrofóbicos expuestos al solvente (W88, M96, F125), así como residuos básicos ubicados en la misma región (R128/K129 y R90/R91/R93), que afectaban a su función en *M. xanthus* (Gallego-García *et al.*, 2014). Dado que la mayoría de estos residuos se conservan en el subdominio CarD₆₁₋₁₇₉ (los aminoácidos equivalentes son W92, L100, T129, K130/R132 y K93/R95/R97; Figura 23), para profundizar en el papel de este subdominio en la función de CarD se realizaron las mutaciones simples W92A, L100A, T129A, la mutación doble K130A/R132A, y la mutación triple K93A/R95A/R97A,

y se analizó su efecto en la interacción con βRNAP y con CarG, así como su efecto *in vivo* en *M. xanthus*.

Los resultados obtenidos con las versiones mutadas en los experimentos de DHB indicaron que los residuos correspondientes no están implicados en la interacción con la RNAP, tal y como cabía esperar (Figura 28A). Para estudiar el efecto de las mutaciones en *M. xanthus*, se sustituyó la versión silvestre de CarD por las distintas versiones mutadas. Una vez obtenidas, las estirpes mutantes se transformaron con el plásmido portador de la fusión Pors::lacZ. El análisis mediante Western blot de los extractos celulares de *M. xanthus* indicó que las mutaciones en cuestión no afectan a la estabilidad de la proteína (Figura 28B). En el análisis cualitativo de la respuesta a la luz no se observó ningún efecto claro de las mutaciones, excepto para el triple mutante (K93A/R95A/R97A), que se mostró tan afectado como el mutante con la deleción de carD (Figura 28C). En las medidas de expresión de PQRS:: lacZ, las estirpes mutantes presentaron niveles de actividad β -galactosidasa en la luz entre el 60-90% de los observados para la estirpe silvestre, con la excepción nuevamente de la estirpe portadora de la triple mutación (K93A/R95A/R97A), que mostró niveles de fotoinducción tan bajos como la estirpe $\Delta carD$ (Figura 28D). Estos resultados, que se asemejan a los observados para las mutaciones equivalentes en CdnL (Gallego-García et al., 2014), ponen de manifiesto un papel clave para los residuos básicos K93/R95/R97 (R90/R91/R93 en CdnL). Acorde con un papel común de estos aminoácidos en ambas proteínas, el mutante K93A/R95A/R97A no se vio afectado en la interacción con CarG (Figura 28E), una interacción que es exclusiva de CarD.



Figura 28. Análisis mutacional de residuos conservados en el subdominio C-terminal (CarD₆₁₋₁₇₉) de CarDNt. A) Análisis mediante DHB de la interacción entre los mutantes de CarD y β RNAP. B) *Western blot* a partir de los extractos celulares de *M. xanthus* que expresan las versiones mutantes de CarD indicadas. C) Estudio del fenotipo de estirpes mutantes en placas incubadas en oscuridad o expuestas a la luz (3 o 6 horas). D) Medida de la actividad β -gal de P_{QRS}::*lacZ* en estirpes crecidas en oscuridad y en luz. E) Ensayo de interacción entre CarG y la versión K93A/R95A/R97A de CarDNt mediante DHB (Bernal-Bernal *et al.*, 2015).

La comparación de CarD con otras proteínas de la familia permitió identificar que, de los tres residuos básicos K93, R95 y R97, la arginina de la posición 95 es la más conservada. Por ello, nos planteamos estudiar si la mutación a alanina solo de este aminoácido (R95A) sería suficiente para anular la función de CarD en *M. xanthus*. Siguiendo la estrategia descrita anteriormente, se generó una estirpe de *M. xanthus* portadora de dicha versión mutada y se analizó su fenotipo cualitativa y cuantitativamente. El análisis del fenotipo de color puso de manifiesto un defecto severo en la fotoinducción de la carotenogénesis en el mutante R95A, aunque menos acusado que el mostrado por el triple mutante K93A/R95A/R97A (Figura 29A). Las medidas de la expresión del promotor P_{QRS} ::/*acZ* confirmaron los resultados anteriores, mostrando un claro descenso de la expresión en la estirpe que presenta la mutación R95A en comparación con la estirpe silvestre, aunque inferior al observado para el triple mutante (K93A/R95A/R95A/R97A) (Figura 29B).



Figura 29. Efecto de la mutación R95A en la proteína CarD en *M. xanthus.* A) Fenotipo de color de gotas de cultivo crecidas en medio sólido en oscuridad y luz. B) Medida de la expresión del promotor P_{QRS} ::*lacZ* (actividad β -galactosidasa) a partir de estirpes crecidas en oscuridad o expuestas a la luz.

III.2 Los factores σ-ECF y su dependencia del complejo CarD-CarG

Parte del contenido de este apartado deriva del artículo publicado por Abellón-Ruiz, Bernal-Bernal *et al.*, en Environmental Microbiology (2014) 16(8), doi:10.1111/1462-2920.12386.

El análisis del genoma de *M. xanthus* puso de manifiesto la presencia de hasta 45 posibles factores σ -ECF. De 19 de estos factores σ -ECF estudiados, 13 de ellos se mostraron parcial o completamente dependientes del complejo CarD-CarG (Abellón-Ruiz *et al.*, 2014). Para la gran mayoría de los factores σ -ECF de *M. xanthus* estudiados se predice una regulación negativa por la acción de un factor anti- σ (producto de un gen adyacente al del factor σ -ECF), con un dominio de unión a zinc característico (ZAS, de *zinc anti-\sigma domain*) en su extremo N-terminal (NZASD), y una hélice transmembranal.

Una excepción notable a esta arquitectura, como se describe en el siguiente apartado, es CarR (Abellón-Ruiz *et al.*, 2014). Con el objetivo de continuar la caracterización de las parejas σ -ECF/anti- σ y su dependencia de CarD-CarG, en esta tesis se prosiguió el estudio de las parejas CarQ/CarR y DdvS/DdvA, y se abordó el análisis de otras seis parejas que resultan novedosas porque el posible factor anti- σ presenta un dominio NZASD seguido de un dominio quinasa de serina/treonina (eSTK).

III.2.1 Los factores anti- σ CarR y DdvA: topología en la membrana e interacción con sus respectivos factores σ -ECF

El grupo de investigación en el que se ha desarrollado este trabajo ha estudiado en profundidad las parejas σ-ECF/anti-σ CarQ/CarR y DdvS/DdvA, cuya acción está regulada por el complejo CarD/CarG (Abellón-Ruiz *et al.*, 2014; Elías-Arnanz *et al.*, 2010, 2011; García-Heras *et al.*, 2013). El análisis mediante programas bioinfomáticos como *SOSUI* o *TMpred* (disponibles en el servidor "*BCM Search Launcher*") predice hasta seis hélices transmembranales para CarR y una o dos (una de ellas con baja probabilidad) para DdvA (Figura 30).



Figura 30. Predicción de la topología en la membrana de los factores anti- σ CarR y DdvA.

Para determinar el número real de hélices transmembranales y la topología en la membrana se utilizó una aproximación experimental que ha funcionado con éxito para otras proteínas de membrana de *M. xanthus* (Galbis-Martínez *et al.*, 2008; Kyungyun & Zusman, 1999). La estrategia consiste en realizar fusiones traduccionales de distintos fragmentos del gen de interés al gen *lacZ* (en el plásmido pKY481), cuya expresión está bajo el control de un promotor que se expresa bien en *E. coli*. En las proteínas de fusión generadas, la enzima β -galactosidasa quedaría ligada al extremo C-terminal en cada una de las versiones. La base del ensayo es que la β -galactosidasa solo es activa cuando se encuentra localizada en el interior celular, ya que el elevado pH del periplasma inhibe su actividad. Así pues, la presencia o ausencia de actividad de las distintas construcciones en *E. coli* sirve para determinar la orientación de cada dominio en la membrana.

En el caso de CarR, se generaron siete versiones: seis que comprendían desde el extremo N-terminal hasta cada una de las seis posibles hélices transmembranales predichas por los programas bioinformáticos, y una última que comprendía el gen *carR* completo (salvo el codón STOP) (Figura 31A). En placas con X-Gal se observó coloración azul solo para las construcciones con los fragmentos hasta el segundo, cuarto y sexto dominio transmembranal, y con la proteína CarR completa (Figura 31B).



Figura 31. Topología en la membrana del factor anti-*σ* **CarR.** A) Plásmidos generados para estudiar la topología. B) Fenotipo en medio sólido con X-Gal. C) Topología de CarR en la membrana confirmada experimentalmente.

Estos resultados validan las predicciones de seis hélices transmembranales realizadas por los programas informáticos, de tal modo que tanto el extremo amino como el carboxilo de CarR quedarían en el citoplasma (Figura 31C).

En los típicos factores anti- σ con una sola hélice transmembranal y un dominio NZASD, dicho dominio se sitúa en el citoplasma y es el responsable de secuestrar en la membrana al factor σ-ECF. La falta de un dominio ZAS en CarR y su topología en la membrana, muy inusual para un factor anti- σ , sugieren un modo diferente de interacción del anti- σ con su factor σ -ECF. No obstante, y dado que tanto el dominio N- como Cterminal de CarR se encuentran en el citoplasma, cualquiera de los dos (o ambos) podrían mediar la interacción CarQ-CarR demostrada previamente (Browning et al., 2003; Elías-Arnanz et al., 2011). Por ello, se analizó la posible interacción mediante DHB entre CarQ y diferentes regiones de CarR (Figura 32A): (i) la región N-terminal hasta justo antes o después del primer dominio transmembranal; (ii) la región que incluye la última hélice transmembranal seguida de la región C-terminal. Como se observa en la Figura 32B-C, no se detectó interacción entre CarQ y cualquiera de los fragmentos de CarR ensayados. Estos resultados indican que, por si solas, ni la región N-terminal ni la C-terminal de CarR interaccionan con CarQ. De participar una u otra región de CarR, o ambas, en la interacción con CarQ debe haber, no obstante, otras regiones necesarias para que ello ocurra.



Figura 32. Ensayo de interacción entre los dominios del factor anti- σ CarR y el factor σ -ECF CarQ. A) Parejas analizadas en los ensayos de DHB. B) Fenotipo en medio sólido con X-Gal. C) Medida de la expresión del gen *lacZ* (actividad β -galactosidasa) en los ensayos de interacción.

Como se ha comentado anteriormente, los programas predicen la existencia de una o dos hélices transmembranales en el caso de DdvA. Por ello, para estudiar su topología en la membrana, se generaron versiones de *ddvA* que permitiesen expresar desde el extremo N-terminal de la proteína hasta: (i) cada una de las dos hélices transmembranales predichas; (ii) una posición (coincidiendo con un punto de corte para *Bam*HI) entre ambos segmentos transmembranales; (iii) el final de la proteína (sin incluir el codón STOP) (Figura 33A). En medio con X-Gal, ninguna de las construcciones generadas produjo una coloración azul al ser introducida en *E. coli*, lo que indica que en los cuatro casos el gen *lacZ* debió quedar orientado hacia el espacio periplásmico (Figura 33B). Estos resultados son consistentes con la presencia de un único segmento transmembranal en DdvA, que colocaría el dominio NZASD en el citoplasma y el dominio C-terminal, de casi 900 aminoácidos, en el espacio extracitoplásmico (Figura 33C).



Figura 33. Topología del factor anti-o DdvA. A) Plásmidos generados para estudiar la topología. B) Fenotipo en medio sólido con X-Gal. C) Topología de DdvA en la membrana confirmada experimentalmente.

Cabría esperar que el dominio NZASD de DdvA, de unos 70 aminoácidos, fuese el responsable de interaccionar con DdvS para mantenerla secuestrada en la membrana en ausencia de la señal (de naturaleza desconocida), y que el dominio C-terminal fuese el que media la percepción de dicha señal. Para determinar experimentalmente, mediante el sistema del DHB, qué parte de DdvA está implicada en la interacción con DdvS, se generaron tres versiones truncadas de DdvA (Figura 34A) constituidas por: (i) el dominio NZASD (DdvA₁₋₇₀); (ii) el dominio NZASD seguido de la región transmembranal (DdvA₁₋₉₅); (iii) el dominio C-terminal (DdvA₉₆₋₉₉₁). Tanto los análisis cualitativos en medio con X-Gal (Figura 34B) como las medidas de actividad β-

galactosidasa específica (Figura 34C) pusieron de manifiesto que el dominio NZASD de DdvA, ya sea con o sin la región transmembranal, es el responsable de la interacción con DdvS. Así, las estirpes que expresaban dicho dominio y DdvS mostraron niveles de activación del gen chivato *lacZ* (intensidad del color azul y actividad β -galactosidasa específica) similares a los de la estirpe que expresaba la proteína DdvA completa y DdvS (usada como control).



Figura 34. Ensayo de interacción entre los dominios del factor anti- σ DdvA y el factor σ -ECF DdvS. A) Parejas analizadas en los ensayos de DHB. B) Fenotipo en medio sólido con X-Gal. C) Medida de la expresión del gen *lacZ* (actividad β -galactosidasa) en los ensayos de interacción.

Como se comentó en la Introducción, la acción negativa de un factor anti- σ sobre su factor σ-ECF es estequiométrica. Por ello, si se provoca la sobreexpresión del factor σ -ECF, sin la sobreexpresión concomitante de su factor anti- σ , las moléculas de factor σ-ECF resultantes provocarían una activación de sus genes diana de una manera independiente de la presencia del estímulo correspondiente. En experimentos previos del grupo se había demostrado que, en efecto, la sobreexpresión de DdvS bajo el control de dos promotores fuertes en tándem provocaba una activación de su propio promotor (P_{ddvSA}), medida por la expresión de un gen chivato lacZ situado aguas abajo de un fragmento conteniendo dicho promotor (Abellón-Ruiz et al., 2014). Si, como indican los experimentos de DHB, el dominio NZASD de DdvA es el responsable de la interacción con DdvS, cabría esperar que el efecto activador causado por la sobreexpresión de DdvS se viese eliminado por la sobreexpresión simultánea del dominio NZASD de DdvA. Para ello, se clonaron fragmentos que permiten sobreexpresar ddvS seguido de la región N-terminal de ddvA, sola o con la hélice transmembranal, bajo el control de dos promotores fuertes en tándem, en un plásmido portador de la fusión transcripcional P_{ddvSA}:: *lacZ* (Figura 35A). Tras la integración de dichos plásmidos por recombinación homóloga en la estirpe silvestre de *M. xanthus*, se analizó cualitativa (Figura 35B) y cuantitativamente (Figura 35C) la expresión del gen chivato *lacZ*. La detección de actividad β -galactosidasa solo para la estirpe control, que sobreexpresa únicamente *ddvS*, indica que para anular los efectos de la sobreexpresión de DdvS en *M. xanthus* basta con la coexpresión del dominio NZASD de DdvA.



Figura 35. La sobreexpresión del dominio NZASD de DdvA anula los efectos de la sobreexpresión de DdvS en *M. xanthus*. A) Representación esquemática de las construcciones generadas. Los plásmidos resultantes se incorporan en el genoma de *M. xanthus* por recombinación homóloga a través de un fragmento de 1.38 kb. B) Fenotipo en medio sólido con X-Gal. C) Medida de la actividad transcripcional de la fusión P_{ddvSA} ::*lacZ* (actividad β -galactosidasa específica).

Algunos factores anti-σ, como AsiA (Gilmore *et al.*, 2010) o el propio CarR (Galbis-Martínez *et al.*, 2008), son capaces de interaccionar consigo mismos. Para analizar si ocurre lo mismo con DdvA, se llevaron a cabo estudios de interacción mediante el sistema del DHB (Figura 36A). Tanto los datos cualitativos como cuantitativos revelaron que DdvA es capaz de interaccionar consigo mismo (DdvS, por el contrario, dio resultados negativos) (Figura 36B y C). El análisis con versiones truncadas de DdvA (Figura 36A) permitió asignar la capacidad de oligomerización al segmento transmembranal.



Figura 36. Análisis de la posible oligomerización de DdvA y de DdvS. A) Parejas analizadas en los ensayos de DHB. B) Fenotipo en medio sólido con X-Gal. C) Medida de la expresión del gen *lacZ* (actividad β -galactosidasa) en los ensayos de interacción.

III.2.2 Dependencia de CarD-CarG de otras parejas σ-ECF/anti-σ

Además de CarQ y DdvS, al menos otros once factores σ -ECF de *M. xanthus* dependen parcial o totalmente de CarD y CarG para su actividad (Abellón-Ruiz *et al.*, 2014). Estos resultados revelan un vínculo sin precedentes entre el singular complejo regulador formado por CarD-CarG y un buen número de factores σ -ECF de *M. xanthus*. En este trabajo se ha estudiado la posible dependencia de CarD-CarG de algunos factores σ -ECF que no se habían analizado previamente.

III.2.2.1 Una pareja σ-ECF/anti-σ similar a DdvS/DdvA

La búsqueda de homólogos de *ddvA* en el genoma de *M. xanthus* permitió identificar un gen (*MXAN_5507*) cuyo producto (975 aminoácidos; de tamaño similar a DdvA) presenta una identidad del 36% y una similitud del 49% con DdvA. Dado que el gen inmediatamente aguas arriba (*MXAN_5506*) determina uno de los posibles factores σ -ECF de *M. xanthus*, cabe esperar que se trate de una pareja σ -ECF/anti- σ . El análisis de la secuencia de *MXAN_5507* predice un dominio ZAS en su extremo N-terminal, similar al observado en DdvA. Para probar si se trata de una pareja σ -ECF/anti- σ , se analizó la interacción de MXAN_5506 con la proteína completa MXAN_5507 o con solo su NZASD. Los datos mostraron que, en efecto, MXAN_5506 y MXAN_5507 interaccionan físicamente en el sistema del DHB, y que dicha interacción está mediada por el dominio NZASD, como sucede con la pareja DdvS/DdvA (Figura 37A).

La similitud entre DdvA y MXAN_5507, y entre DdvS y MXAN_5506, sugiere un origen evolutivo común de las dos parejas como consecuencia, probablemente, de una duplicación génica. Por ello, se examinó la posibilidad de que exista una interacción cruzada entre las dos parejas mediante el sistema del DHB. Sin embargo, no se observó interacción ni del factor anti- σ completo ni de su dominio NZASD con el factor σ -ECF de la otra pareja (Figura 37B). Así pues, pese al parecido, cada anti- σ reconoce específicamente a su factor σ -ECF.



Figura 37. Estudio de la pareja σ -ECF/anti- σ MXAN_5506/MXAN_5507. A) Ensayo de interacción entre MXAN_5507 o su NZASD y su posible pareja MXAN_5506. B) Análisis de la posible interacción cruzada entre las parejas DdvS/DdvA y MXAN_5506/MXAN_5507. C) Efecto de la sobreexpresión de MXAN_5506 sobre su propio promotor en diferentes fondos genéticos.

Como ocurre con varios factores σ -ECF y sus anti- σ en *M. xanthus*, incluidos DdvS-DdvA, la pareja MXAN_5506/MXAN_5507 podría depender del complejo CarD-CarG. Para estudiarlo se utilizó la misma aproximación descrita en la Figura 35. Se generaron cuatro estirpes de *M. xanthus* con la región promotora de *MXAN_5506/MXAN_5507*, P₅₅₀₆₋₀₇, fusionada al gen *lacZ*. A una de ellas, la estirpe control (fondo genético silvestre), no se le introdujo la construcción para sobreexpresar MXAN_5506 mientras que a las otras tres sí (fondos genéticos silvestre, $\Delta carD$, $\Delta carG$). Curiosamente, la fusión P₅₅₀₆₋₀₇::*lacZ* mostró niveles significativos de expresión en la estirpe control, que fueron similares a los observados en la estirpe que sobreexpresa MXAN_5506 (Figura 37C). Esto puede ser debido a que, en las condiciones de cultivo utilizadas, el factor σ -ECF MXAN_5506, o una parte de este, se encuentre de forma libre

activando la transcripción de sus promotores diana (entre los que se incluiría su propio promotor). En cualquier caso, la disminución de unas 20 veces en la expresión de P₅₅₀₆₋₀₇::*lacZ* observada en un fondo genético sin CarD o CarG apunta a que se trata de otro factor σ -ECF cuya actividad depende del complejo CarD-CarG (Figura 37C).

III.2.2.2 Parejas σ -ECF/anti- σ con un dominio quinasa de serina/treonina en el factor anti- σ

Entre los 45 posibles σ -ECF predichos en *M. xanthus*, seis de ellos destacan por su asociación con posibles factores anti- σ para los que los programas bioinformáticos predicen una arquitectura inusual: un dominio NZASD seguido de un dominio quinasa de serina/treonina (eSTK) y al menos una hélice transmembranal (Figura 38). Exceptuando la pareja de genes MXAN 0068/MXAN 0070, que se encuentra separada por un gen, en los otros cinco casos el gen para el factor σ -ECF y el gen para el factor anti-o se encuentran adyacentes, estando situado aguas arriba el gen para el factor o-Curiosamente, ECF en tres de los casos (Figura 38). las parejas MXAN 1709/MXAN 1710 ٧ MXAN 2737/MXAN 2738, por un lado, V MXAN_6009/MXAN_6010 y MXAN_7082/MXAN_7083, por otro, se asemejan entre sí tanto en lo que se refiere a la organización génica como en el tamaño y secuencia de los correspondientes productos génicos. Así, la pareja MXAN 1709/MXAN 1710 presenta hasta un 50% de identidad y un 66% de similitud con la pareja MXAN 2737/MXAN 2738; y la pareja MXAN 6009/MXAN 6010 un 67% de identidad y un 77% de similitud con la pareja MXAN 7082/MXAN 7083. Por otro lado, es posible que MXAN 7161, más pequeño que los otros σ-ECF predichos y con el dominio σ2 truncado, no sea capaz de funcionar como factor σ .



Figura 38. Posibles parejas σ -ECF/anti- σ con un factor anti- σ atípico con un dominio ZAS y quinasa de serina/treonina.

En este trabajo se ha demostrado que el dominio NZASD de DdvA y de MXAN_5507 es el que media la interacción entre dichos factores anti- σ y sus respectivos factores σ -ECF. Por ello, nos planteamos estudiar si en el caso de los factores anti- σ con un dominio eSTK, su dominio NZASD sería también el responsable de interaccionar con el correspondiente σ -ECF. Tras realizar las construcciones pertinentes, se ensayó la posible interacción mediante el sistema del DHB (Figura 39A). Cinco de las seis parejas analizadas mostraron niveles muy altos de actividad β -galactosidasa, lo que apoya la hipótesis de que se trate de parejas σ -ECF/anti- σ . La excepción fue la pareja MXAN_0068/MXAN_0070, que produjo niveles de actividad β -galactosidasa tan bajos como los controles negativos (estirpes que expresan solo uno de los dos miembros de la pareja) (Figura 39B).



Figura 39. Estudio de la interacción de las posibles parejas σ -ECF/anti- σ atípicas. En rosa se representa los factores σ -ECF mientras que en verde se muestra el motivo NZASD de los factores anti- σ .

El alto grado de similitud entre las parejas MXAN_1709/MXAN_1710 y MXAN_2737/MXAN_2738, y MXAN_6009/MXAN_6010 y MXAN_7082/MXAN_7083 podría facilitar las interacciones cruzadas entre las diferentes parejas. Para comprobarlo, se realizaron ensayos de DHB partiendo de las construcciones generadas previamente. Sin embargo, no se observó interacción en ninguna de las combinaciones ensayadas entre miembros de parejas distintas (Figura 40). Estos resultados nuevamente ilustran que la interacción entre un factor anti- σ y su factor σ -ECF es muy específica.



Figura 40. Ensayo de posibles interacciones cruzadas entre parejas σ -ECF/anti- σ con una alta homología.

Como paradigma de este tipo de parejas σ -ECF/anti- σ se eligió para su análisis más detallado la pareja MXAN_1709/MXAN_1710, ya que el posible factor anti- σ se corresponde con la proteína Pkn8, que ha sido estudiada previamente en *M. xanthus*. Se ha descrito que Pkn8 es una eSTK de membrana que participa en una cascada de regulación fosforilando a la proteína citoplásmica Pkn14, que a su vez regula negativamente a MrpC, un activador de la expresión de *fruA* (implicado en el desarrollo multicelular) (Nariya & Inouye, 2005, 2006). Para determinar si la pareja MXAN_1709/*pkn8* era dependiente del complejo CarD-CarG, al igual que varias otras parejas σ - ECF/anti- σ analizadas, se utilizó la estrategia ilustrada en la Figura 41. Básicamente, se generó un plásmido portador de la región promotora aguas arriba de *MXAN_1710* (*pkn8*), P₁₇₁₀₋₀₉, fusionada al gen *lacZ*, que se introdujo en la estirpe silvestre (control de no sobreexpresión). Por otro lado, al plásmido con la fusión P₁₇₁₀₋₀₉::*lacZ* se le insertó el gen para el factor σ -ECF MXAN_1709 aguas abajo de dos promotores fuertes en tándem, y se introdujo en la estirpe silvestre de *M. xanthus*, y en los fondos genéticos con la deleción de *carD* o de *carG* (Figura 41).



Figura 41. Esquema de la estrategia utilizada para analizar la actividad del promotor P₁₇₁₀₋₀₉ y su dependencia de CarD y CarG.

Como se aprecia en la Figura 42A, el análisis del fenotipo de los transformantes tanto en fase vegetativa como durante el desarrollo multicelular, así como la medida de actividad β -galactosidasa en ambas fases del ciclo de vida de *M. xanthus* (Figura 42B), indicaron que el factor σ -ECF no depende del complejo CarD-CarG.



Figura 42. Estudio de la dependencia de CarD-CarG en el factor σ -ECF MXAN_1709. A) Esquema de las estirpes generadas y su fenotipo en placas de CTT y TPM. B) Medida de actividad β -galactosidasa durante la fase vegetativa y durante el desarrollo.

III.3 Modo de acción del complejo regulador CarD-CarG

La similitud de CarD con las proteínas eucarióticas HMGA, su forma de actuar en complejo con una proteína que no se une directamente al DNA (CarG) y su capacidad de interaccionar con la subunidad β de la RNAP han llevado a proponer que el complejo CarD-CarG podría funcionar a modo de un "enhanceosome" bacteriano, reclutando otros factores requeridos para la regulación y/o la maquinaria basal de transcripción (Elías-Arnanz *et al.*, 2010). Para profundizar en el modo de acción del complejo CarD - CarG, determinar cómo regula la acción de los factores σ -ECF, e identificar nuevos promotores diana, se llevaron a cabo estudios *in vivo* de localización de dichas proteínas en el genoma mediante la técnica de inmunoprecipitación de cromatina.

III.3.1 Análisis y validación de los sitios de unión del complejo CarD-CarG

Hasta la realización de este trabajo, los únicos promotores que se conocían cuya expresión depende del complejo CarD-CarG dependen también de un factor σ -ECF (Abellón-Ruiz *et al.*, 2014). No obstante, la acción reguladora global del complejo CarD - CarG podría estar mediada por su demostrada acción positiva sobre un buen número de factores σ -ECF, a la vez que por su acción sobre otros promotores no dependientes de un factor σ -ECF. Dado que la técnica de inmunoprecipitación de cromatina (ChIP), asociada a secuenciación masiva (ChIP-Seq) o PCR cuantitativa

(ChIP-qPCR), permite identificar los sitios en el DNA a los que se une (directa o indirectamente) una proteína *in vivo*, se llevaron a cabo este tipo de experimentos con CarD y CarG (durante una estancia corta en Marsella en el laboratorio del Dr. Emanuele Biondi). Aunque disponemos de anticuerpos monoclonales anti-CarD se generó, no obstante, una estirpe de *M. xanthus* con *carD* sustituido en el sitio nativo por una versión que permite la expresión de CarD con la etiqueta Flag en su extremo carboxilo, para su inmunoprecipitación con anticuerpos comerciales. Del mismo modo, para permitir la inmunoprecipitación de CarG (para la cual carecemos de anticuerpos), se generó una estirpe de *M. xanthus* que expresa CarG con la etiqueta Flag también en su extremo carboxilo. Antes de realizar los experimentos de ChIP, se comprobó que las proteínas de fusión se expresan de manera estable en *M. xanthus* mediante *Western blot* (Figura 43A) y que son funcionalmente activas, mediante el análisis de la respuesta carotenogénica a la luz (Figura 43B).



Figura 43. Comprobación de la expresión y funcionalidad CarD-Flag y CarG-Flag. A) Análisis de la expresión de CarD-Flag y CarG-Flag mediante *Western blot* con anticuerpos anti-Flag a partir de extractos celulares de las estirpes correspondientes de *M. xanthus*. A la izquierda, se muestran los tamaños de las proteínas usadas como marcadores de tamaño. B) Estudio de la funcionalidad de CarD-Flag y CarG-Flag mediante el análisis del fenotipo de color en respuesta a la luz de las estirpes correspondientes de *M. xanthus*.

Para la búsqueda global de los sitios de unión de CarD se realizó la inmunoprecipitación de cromatina tanto con el anticuerpo anti-Flag como con el anticuerpo monoclonal anti-CarD 138/6 disponible en el laboratorio. El análisis de los datos de ChIP-Seq reveló perfiles de unión prácticamente idénticos tras la inmunoprecipitación de CarD con cualquiera de los dos anticuerpos, y permitió la identificación de más de 100 posibles sitios distribuidos a lo largo de todo el genoma (Figura 44), consistente con su acción reguladora global. Los resultados del análisis ChIP-Seq para CarG, que no se une directamente al DNA, fueron muy similares a los obtenidos con CarD (Figura 44), lo que refuerza la noción de que ambas proteínas funcionan siempre juntas y de que CarG se localiza en el DNA a través de su interacción con CarD.



Figura 44. Perfiles de unión de CarD y CarG en el genoma de *M. xanthus*.

El estudio más detallado de los datos de ChIP-Seq permitió determinar que la gran mayoría de los sitios de unión de CarD-CarG se encuentran en regiones intergénicas, que deben incluir las regiones promotoras de los genes correspondientes. Curiosamente, bajo las condiciones de oscuridad y fondo silvestre en las que se realizó el experimento, no se observó la presencia de CarD-CarG en las secuencias promotoras dependientes del complejo mejor estudiadas, las que dependen de los factores σ -ECF CarQ y DdvS, y a las que CarD sí se une *in vitro;* tampoco se observó enriquecimiento del complejo en los otros promotores dependientes de factores σ -ECF cuya actividad depende también de CarD-CarG. (Abellón-Ruiz *et al.*, 2014; García-Heras *et al.*, 2013). Estos resultados sugieren que, en lo que se refiere a los promotores regulados por factores σ -ECF y por CarD-CarG, es necesario que el promotor se encuentre activo para detectar el enriquecimiento del complejo en dicho promotor (véase el siguiente apartado).

Para validar mediante ChIP-qPCR los resultados obtenidos, se eligieron tres posibles sitios de unión de CarD-CarG de entre los identificados mediante ChIP-Seq: (i) la región promotora de *MXAN_3919* o *rnk* (regulator of nucleoside diphosphate kinase), un gen que normalmente cifra una proteína de interacción con la RNAP muy similar a los factores Gre; (ii) la región promotora de *MXAN_4147* (*rpoE1*), que cifra un factor σ -ECF parcialmente dependiente del complejo CarD-CarG y cuya expresión está mediada por dos promotores, uno reconocido por el factor mayoritario σ^A y otro por el propio RpoE1 (Abellón-Ruiz *et al.*, 2014; Ward *et al.*, 1998); (iii) la región promotora de

MXAN_6483, que cifra una proteína perteneciente a la familia de canales de protones MotA/TolQ/ExbB. Para la inmunoprecipitación de cromatina con los anticuerpos anti- Flag se utilizó la estirpe que expresa CarD-Flag y, como controles, la estirpe silvestre (que expresa CarD sin la etiqueta) y la estirpe que carece de CarD ($\Delta carD$), en relación con una región intragénica a la que no se espera que se una CarD (localizada en el primer *orf* del operón *carB*). Solo se observó un enriquecimiento (de unas 3-4 veces) en las tres regiones analizadas para la estirpe que expresa CarD-Flag (Figura 45), lo que confirma los resultados obtenidos en los experimentos de ChIP-Seq.



Figura 45. Validación mediante ChIP-qPCR de la unión de CarD a las tres regiones promotoras indicadas.

Una vez demostrado que CarD se asocia *in vivo* a estos promotores, se estudió mediante qRT-PCR su posible regulación por CarD y CarG, comparando su expresión en la estirpe silvestre versus las estirpes que carecen de CarD o CarG. Como se puede observar en la Figura 46, la expresión de *rnk* o *rpoE1* disminuyó aproximadamente a la mitad en ausencia de *carD* o *carG*, lo que sugiere que el complejo CarD-CarG regula positivamente la expresión de estos dos promotores. En cambio, la expresión de *MXAN_6483* no se vio disminuida, sino más bien algo aumentada, por la ausencia de CarD o CarG. Es posible que en esta región el complejo CarD-CarG ejerza una ligera acción negativa o que, más que un papel regulador, juegue un papel arquitectónico, modulando la estructura del DNA, de manera análoga a sus homólogos eucarióticos (García-Heras *et al.*, 2009; Kasinsky *et al.*, 2001; Thanos & Maniatis, 1992).



Figura 46. Análisis de la dependencia de CarD y CarG mediante qRT-PCR de los genes indicados.

III.3.2 Estudio de la asociación in vivo de CarQ y CarD al promotor PQRS

La activación del promotor P_{QRS} por el factor σ -ECF CarQ y por CarD-CarG es un paso clave en la respuesta a la luz de M. xanthus. Para confirmar que la acción positiva de CarQ sobre PQRS es una acción directa, se realizaron ensayos de ChIP-qPCR. Como no disponemos de anticuerpos anti-CarQ lo suficientemente específicos, se generaron estirpes de M. xanthus que expresan CarQ con la etiqueta Flag en su extremo N-terminal, bajo un promotor inducible por vanilato, y en un fondo silvestre, $\Delta carD$ o $\Delta carG$ (Figura 47A). Para comprobar que la proteína de fusión se comporta normalmente, se estudió la capacidad de las estirpes correspondientes de responder o no a la luz mediante análisis del fenotipo de color en medio con vanilato (para inducir la expresión de Flag-CarQ). Dado que solo se observó una activación de la respuesta carotenogénica en el fondo silvestre y en presencia de vanilato (Figura 47B), cabe concluir que la proteína de fusión es funcional y mantiene su dependencia del complejo CarD-CarG. Para analizar la asociación de CarQ al promotor P_{QRS} in vivo, se comparó el enriquecimiento en dicho promotor para la estirpe con la fusión Flag- CarQ, en fondo genético silvestre, crecida en presencia o ausencia de vanilato. Los resultados obtenidos mediante ChIP-qPCR mostraron un enriquecimiento de CarQ en P_{QRS} solo en el cultivo crecido en presencia del inductor (Figura 47C), lo que apoya que CarQ se une a dicho promotor y ejerce, por tanto, un papel directo en la activación de su expresión.



Figura 47. Generación de la estirpe que expresa *Flag-CarQ* y análisis de su comportamiento. A) Esquema del plásmido que expresa *flag-carQ* bajo el control del promotor inducibe por vanilato (P_{van}) y de su integración en *M. xanthus* por recombinación homóloga. B) Estudio de la activación de la respuesta carotenogénica por Flag-CarQ en diferentes fondos genéticos. C) Ensayo de ChIP-qPCR para estudiar la localización de Flag-CarQ en P_{QRS} .

In vitro, la proteína CarD se une a dos pequeños tramos ricos en pares AT situados aguas arriba del promotor P_{QRS} (García-Heras *et al.*, 2013). Sin embargo, en los estudios de ChIP-Seq realizados a partir de cultivos crecidos en la oscuridad (Figura 44), condiciones en las que el promotor P_{QRS} se encuentra inactivo, no se observó enriquecimiento ni de CarD ni de CarG en dicho promotor. Dado que la expresión en la luz de P_{QRS} es absolutamente dependiente de CarD-CarG, podría ocurrir que, *in vivo*, el complejo solo se localize en el promotor cuando este se encuentra activo. Por ello, utilizando la estirpe que expresa CarD-Flag descrita anteriormente, se realizaron estudios de ChIP-Seq a partir de células crecidas en la luz. En estas condiciones sí se observó un enriquecimiento de CarD en la región promotora de P_{QRS} (Figura 48A). También se observó la localización de CarD en P_{QRS} en células iluminadas cuando el análisis se llevó a cabo mediante ChIP-qPCR. Así, en comparación con una región intragénica a la que no se une CarD, se observó un enriquecimiento en P_{QRS} en las muestras de la estirpe que expresa CarD-Flag crecida en luz respecto de la misma estirpe crecida en oscuridad, que mostró valores de enriquecimiento similares a los

observados para la estirpe que expresa CarD sin la etiqueta (Figura 48B). Estos resultados indican que CarD solo se une a P_{QRS} cuando el promotor se encuentra activado.



Figura 48. Análisis de la asociación *in vivo* de CarD a la región promotora P_{QRS} . A) Estudio mediante ChIPseq de la asociación de CarD-Flag a P_{QRS} a partir de células crecidas en condiciones de iluminación. Solo se muestra la región del genoma que incluye a P_{QRS} B) Comparación de la unión *in vivo* de CarD-Flag a P_{QRS} en la luz versus la oscuridad; WT: estirpe silvestre que expresa CarD sin la etiqueta Flag.

III.3.3 Análisis del efecto de CarG sobre la unión de CarD al DNA in vivo

A pesar de que CarD por si sola es capaz de unirse al DNA in vitro a través de su dominio C-terminal de tipo HMGA, y a la β RNAP a través de su dominio N-terminal, in vivo es incapaz de realizar ninguna de sus acciones reguladoras conocidas si no está et al., 2009; Peñalver-Mellado presente CarG (García-Heras et al., 2006). Desconocemos qué papel cumple CarG, que lo hace imprescindible en el complejo. Por analogía con los "enhanceosomes" eucarióticos, CarG podría funcionar como una proteína adaptadora, a modo de "puente" entre CarD y otros factores/otras subunidades de la RNAP (Elías-Arnanz et al., 2010), pues aunque CarD interacciona directamente con βRNAP, dicha interacción no es estrictamente necesaria para su función (véase apartado III.1.2.2). De ejercer un papel de "puente" entre CarD y la maquinaria basal de transcripción, cabría esperar que CarG estabilizase la unión de CarD a sus promotores diana in vivo. Para estudiarlo, se analizó primero el efecto de CarG sobre la unión de CarD a los promotores Prnk, PrpoE1 y PMXAN_6483 mediante ChIP-qPCR. Usando el anticuerpo monoclonal anti-CarD 138/6 para la inmunoprecipitación de la cromatina, se analizó el enriquecimiento de CarD en dichos promotores en la estirpe silvestre en comparación con las estirpes mutantes $\Delta carG$ y $\Delta carD$, esta última usada como control negativo de unión de CarD a las regiones promotoras. Los resultados obtenidos mostraron enriquecimiento de CarD en los promotores examinados en la estirpe silvestre, pero no en la estirpe $\Delta carG$, que produjo valores similares a los del control negativo $\Delta carD$ (Figura 49A). Un efecto parecido asociado a la falta de CarG se observó también en el promotor P_{QRS}: la unión de CarD a P_{QRS} observada en la estirpe silvestre solo en la luz desapareció en la estirpe que carece de CarG, al igual que en la estirpe control portadora de la deleción de *carD* (Figura 49B). Así pues, aunque *in vitro* CarD es capaz de unirse al DNA en ausencia de CarG, *in vivo* CarD no se une eficazmente al DNA a menos que esté presente CarG, lo que explica que CarG juegue un papel esencial en el complejo. Quizás este papel consista en servir de intermediario de la interacción entre CarD y la maquinaria basal de transcripción, y estabilizar de ese modo la unión de CarD al DNA.



Figura 49. Efecto de CarG sobre la unión de CarD al DNA *in vivo*. A) Análisis mediante ChIP-qPCR de la unión de CarD a los promotores indicados y en los fondos genéticos indicados. B) Análisis mediante ChIP-qPCR de la unión de CarD al promotor P_{QRS} en los fondos genéticos indicados y en la luz versus la oscuridad.

III.4 Control multifactorial de un sistema CRISPR-Cas por la pareja DdvS/DdvA y el complejo regulador global CarD-CarG

El contenido de este apartado deriva del artículo publicado por Bernal-Bernal *et al.*, en Nucleic Acids Research (2018), 46(13) doi: 10.1093/nar/gky475.

Como se ha comentado anteriormente, CarQ/CarR y DdvS/DdvA, cuya acción está regulada por el complejo CarD/CarG, han sido las parejas σ -ECF/anti- σ mejor estudiadas por el grupo de investigación donde se ha desarrollado esta tesis. Para llevar a cabo su acción activadora de la transcripción, DdvS requiere la inactivación previa de DdvA por una señal extracitoplásmica cuya naturaleza ignoramos. Por ello, con el objetivo de conocer el regulón activado por DdvS, y dependiente de CarD y CarG, se realizaron análisis transcriptómicos comparando la estirpe silvestre con las estirpes mutantes $\Delta ddvA$, $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$. Así, se identificaron 20 genes activados directa o indirectamente por DdvS y a su vez dependientes de CarD-CarG (Tabla 7).

MXAN locus	Gen	ΔddvA vs. WT	∆ddvA ∆carD vs. ∆ddvA	ΔddvA ΔcarG vs. ΔddvA
181 <mark>0</mark>	RDD family protein	3,5	-10,54	-3,61
1891	Hypothetical protein	2,69	-5,43	-2,96
1892	Putative ser/thr protein kinase	3,16	-9,49	-3,33
1893	Hypothetical protein	3,25	-7,52	-3,34
1894	DNA-binding protein	2,21	-3,05	-2,22
7276	CRISPR associated Cas6 family protein	17,53	-3,25	-2,82
7277	CRISPR associated RAMP Cmr6 family protein	25,10	-3,43	-3,09
7278	CRISPR associated RAMP Cmr5 family protein	6,66	-2,88	-2,64
7280	CRISPR associated RAMP Cmr3 family protein	33,94	-3,40	-3,51
7281	CRISPR associated RAMP Cmr2 family protein	18,32	-3,42	-3,62
7282	CRISPR associated RAMP Cmr1 family protein	65,45	-4,38	-4,09
7283	CRISPR associated protein	66,61	-4,46	-4,25
7284	Hypothetical protein	21,04	-5,27	-5,36
7285	ThiF family protein	58,51	-7,15	-6,56
7286	Hypothetical protein	122,96	-7,31	-8,73
7287	Hypothetical protein	8,51	-2,44	-2,32
7288	DdvA	8,35	-2,92	-2,06
7289	DdvS	204,52	-5,75	-4,67
7290	Hypothetical protein	285,32	-5,39	-4,15
7291	Hypothetical protein	572,45	-4,66	-7,32

Tabla 7. Lista de genes activados en la estirpe $\Delta ddvA$ y reprimidos en $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$. Enmarcados en azul se muestran los genes que pertenecen al locus relacionado con un sistema CRISPR-Cas.
Curiosamente, la mayoría de estos genes forman parte de un mismo locus que cifra un sistema CRISPR-Cas de tipo III-B situado en las inmediaciones de *ddvS/ddvA* (Tabla 7 y Figura 50); los genes restantes pertenecen a dos loci diferentes ubicados en otra posición del genoma y cifran proteínas hipotéticas (MXAN_1891 y MXAN_1893), una posible quinasa de Ser/Thr (MXAN_1892), una posible proteína de unión al DNA (MXAN_1894), y una proteína de la familia RDD (MXAN_1810; denominada así por presentar una Arg y dos Asp conservados formando parte de una hélice transmembranal) (Tabla 7 y Figura 50) (Abellón-Ruiz, 2012).



Figura 50. Distribución en el genoma de los loci activados por DdvS y dependientes de CarD y CarG. En rojo se representa el factor σ -ECF DdvS, en verde se representa el factor anti- σ DdvA y en azul los genes *cas* asociados al sistema CRISPR4-Cas.

III.4.1 Promotores dependientes de DdvS y CarD-CarG en el locus CRISPR-Cas

El sistema CRISPR-Cas identificado como dependiente de DdvS y de CarD- CarG es uno de los tres sistemas CRISPR-Cas presentes en el genoma de *M. xanthus* (Figura 18 y 19). Dicho sistema está constituido por ocho genes *cas/cmr* seguidos de una región líder y de la región CRISPR (CRISPR4), que consta de 52 espaciadores y 53 repeticiones. Aguas arriba de *ddvS-ddvA*, y entre *ddvS-ddvA* y los genes *cas*, se encuentran dos (*MXAN_7291-MXAN_7290*) y cuatro genes (*MXAN_7287-MXAN_7284*), respectivamente, que cifran proteínas hipotéticas y cuya expresión también depende de DdvS y CarD-CarG (Tabla 7 y Figura 51).



Figura 51. Esquema del sistema CRISPR4-Cas. Con flechas amarillas se indican los cuatro promotores dependientes de DdvS y de CarD-CarG identificados. La flecha negra indica un posible promotor dependiente del factor σ mayoritario en la región líder.

Para elucidar cómo DdvS regula la expresión de estos genes, se realizó una búsqueda *in silico* de promotores dependientes de DdvS en esta región del genoma usando como criterio de búsqueda la secuencia $GTAAn_{16}CGT$ (donde "n" indica cualquier nucleótido). Así, se identificaron tres nuevos promotores, que presentan regiones -35 y -10 similares a P_{ddvSA} , localizados aguas arriba de *MXAN_7291*, *MXAN_7286* y *MXAN_7283* (Figuras 51 y 52).

La identificación mediante 5' RACE del sitio de inicio de la transcripción permitió confirmar que las regiones identificadas *in silico* realmente se corresponden con las regiones -35 y -10 de los promotores. Además, mediante qRT-PCR se comprobó que la expresión de los genes *MXAN_7291*, *MXAN_7286* y *MXAN_7283*, al igual que la de *ddvS*, depende de DdvS y de CarD-CarG (Figura 52; Abellón-Ruiz, 2012).



Figura 52. Alineamiento de las regiones -35 y -10 de los promotores dependientes de DdvS e identificación del sitio de inicio de la transcripción mediante 5' RACE. En la parte inferior se muestra una representación Weblogo de los promotores descritos.

Como prueba adicional de que las regiones identificadas se corresponden con promotores cuya actividad depende de DdvS y de CarD-CarG, cada una de las regiones se clonó aguas arriba del gen chivato *lacZ* y la fusión correspondiente (P::*lacZ*) se introdujo en la estirpe silvestre y en los fondos genéticos mutantes $\Delta ddvA$, $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$. El análisis de la actividad β -galactosidasa, usando la fusión P_{ddvSA}::*lacZ* como control, sirvió para confirmar que los tres nuevos promotores son dependientes de DdvS y de CarD-CarG, pues solo se expresan en la estirpe $\Delta ddvA$ (Figura 53).



Figura 53. Medida de la actividad β -galactosidasa de los promotores P_{MXAN_7291}, P_{ddvSA}, P_{MXAN_7286} y P_{MXAN_7283} en los fondos genéticos indicados.

Para comprobar si la regulación por DdvS y CarD-CarG de los cuatro promotores es una acción directa se realizaron estudios de ChIP-qPCR. Para ello hubo que generar un plásmido que permitiera expresar, bajo el control de un promotor inducible por vanilato, DdvS con la etiqueta Flag en su extremo N-terminal. Ello permite expresar Flag-DdvS de manera condicional e independiente del estímulo extracitoplásmico. Dicho plásmido se introdujo en la estirpe silvestre y en las estirpes mutantes $\Delta carD$ y $\Delta carG$, en las que se integra por recombinación homóloga (Figura 54A). La correcta expresión de la proteína Flag-DdvS en las estirpes resultantes se analizó mediante *Western blot* con anticuerpos comerciales anti-Flag (Figura 54B), y su funcionalidad y dependencia del complejo CarD-CarG mediante análisis por *Northern blot* de los crRNAs correspondientes a la región CRISPR4 (Figura 54C; véase apartado III.4.4).



Figura 54. Estudio de la estirpe que expresa condicionalmente *flag-ddvS.* A) Esquema del plásmido generado y de su integración por recombinación homóloga en *M. xanthus.* B) Estudio de la expresión de Flag-DdvS mediante *Western blot.* C) Análisis de la funcionalidad de Flag-DdvS y de su dependencia de CarD-CarG mediante *Northern blot* de los crRNAs producidos por la región CRISPR4, a partir de muestras de las estirpes indicadas crecidas en presencia del inductor vanilato a 0,5 mM durante una hora; control de carga: 5S rRNA. D) Análisis mediante ChIP-qPCR de la unión de Flag-DdvS a los cuatro promotores indicados a partir de muestras crecidas en presencia de vanilato.

El análisis mediante ChIP-qPCR (inmunoprecipitando Flag-DdvS con los anticuerpos anti-Flag) a partir de muestras crecidas en presencia de vanilato y tratadas con rifampicina para atrapar los complejos de la RNAP en dichos promotores, mostró un enriquecimiento de DdvS en los cuatro promotores, de unas 5-7 veces, en relación a muestras crecidas en ausencia de vanilato y sometidas al mismo tratamiento (Figura 54D).

Para estudiar si la regulación de los cuatro promotores por CarD-CarG responde a una acción directa del complejo, se realizaron estudios mediante ChIP-qPCR a partir de la estirpe silvestre y de las estirpes mutantes $\Delta ddvA$, $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$, inmunoprecipitando solo CarD (con los anticuerpos monoclonales anti-CarD). Los resultados obtenidos mostraron un enriquecimiento de CarD en los cuatro promotores en la estirpe $\Delta ddvA$ respecto a la estirpe $\Delta ddvA \Delta carD$, usada como control negativo. Sin embargo, no se observó unión de CarD a dichos promotores ni en la estirpe silvestre, que no expresa los promotores por no disponer de DdvS activo, ni en la estirpe $\Delta ddvA$ $\Delta carG$ que tampoco los expresa por falta de CarG (Figura 55). Estos resultados concuerdan con los descritos en los apartados III.3.2 y III.3.3 para P_{QRS}, en el que tampoco se observó unión de CarD cuando el promotor no se expresa, bien porque CarQ se encuentra inactivado en la membrana por CarR (en la oscuridad), o bien porque CarG no está presente. En cualquier caso, los resultados permiten concluir que la acción reguladora positiva de CarD-CarG sobre los promotores dependientes de DdvS está mediada por una acción directa.



Figura 55. Análisis de la asociación *in vivo* de CarD a los promotores dependientes de DdvS.

A través de su dominio C-terminal, CarD se une a regiones ricas en pares AT (Abellón-Ruiz *et al.*, 2014; García-Heras *et al.*, 2013). Dado que el análisis de las regiones promotoras dependientes de DdvS permite identificar en ellas tramos ricos en pares AT (Figura 56A), que podrían mediar la unión del complejo CarD-CarG, se estudió mediante EMSA la unión *in vitro* de CarD, o de CarD-CarG, a fragmentos de DNA de 170 pb que incluyen los elementos -35 y -10 de los promotores y las secuencias ricas en pares AT. Como se observa en la Figura 56B, CarD o el complejo CarD-CarG fueron capaces de unirse a los cuatro promotores ensayados, apareciendo en el caso de P_{ddvSA} y P_{MXAN_7286} dos bandas retrasadas que apuntan a dos posibles sitios de unión de CarD en estas regiones.



Figura 56. Secuencia de las regiones promotoras dependientes de DdvS y estudio de la unión de CarD-CarG a dichas regiones. A) Secuencia de cada uno de los promotores activados por DdvS y dependientes de CarD-CarG. En negrita se indican las regiones -35 y - 10 de los diferentes promotores y subrayados los tramos con ≥ 3 AT. B) EMSA mostrando la unión de CarD o CarD-CarG a los cuatro promotores.

En conjunto, estos resultados confirman la existencia de cuatro promotores dependientes de DdvS, CarD y CarG en la proximidad del locus CRISPR4-*cas*. Uno de dichos promotores, $P_{MXAN_{7283}}$, se localiza inmediatamente aguas arriba del grupo de genes *cas/cmr*.

III.4.2 La región líder de CRISPR4 tiene una actividad promotora despreciable

Los resultados mostrados hasta el momento indican que la expresión de los genes *cas* del sistema CRISPR4-Cas depende de DdvS y del complejo CarD-CarG, pero no informan sobre cómo se lleva a cabo la transcripción de la región CRISPR4. En otros sistemas CRISPR-Cas se ha descrito que la región CRISPR se expresa a partir de un promotor dependiente del factor σ mayoritario situado en la región líder. El análisis del segmento líder de CRISPR4, que abarca 304 pb, permitió identificar una secuencia TTGACA separada por 17 pb de una secuencia TTTATG que podrían corresponder, respectivamente, al elemento - 35 y -10 de un promotor dependiente del factor mayoritario σ^A de *M. xanthus* (Figura 57A). Para comprobarlo, en primer lugar, se estudió *in vitro*, mediante EMSA, si la holoenzima RNAP- σ^A de *M. xanthus* es capaz de

reconocer la región líder silvestre o dicha región con el TTG del posible elemento -35 mutado a CCA. Para la realización del ensayo, el DNA marcado se incubó con la holoenzima y, a continuación, se añadió heparina para eliminar los complejos cerrados e inespecíficos, y dejar solo los complejos abiertos (RP_o). Como se observa en la Figura 57B, pudo detectarse una banda retrasada con el fragmento líder silvestre, que desapareció con el fragmento mutado, lo que sugiere que el líder contiene un promotor dependiente de la RNAP- σ^A .



Figura 57. Análisis de la región líder. A) Secuencia de la región líder. En negrita se indican las posibles regiones -35 y -10. B) Ensayo de retraso en gel con la RNAP- σ^A y la región líder silvestre (WT) o mutada (Mut) en la posible región -35. RP₀: complejo abierto.

Como complemento a los experimentos *in vitro*, se examinó *in vivo* la expresión del posible promotor en la región líder. Para ello, se generaron plásmidos con la región líder silvestre o la versión mutada en la región -35 fusionadas al gen *lacZ* (Líder::*lacZ* o Líder(Mut)::*lacZ*, respectivamente) (Figura 58A). Dichos plásmidos se introdujeron en una estirpe silvestre y en las estirpes mutantes $\Delta ddvA$, $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$, en las que se integran por recombinación homóloga en un sitio heterólogo. Sorprendentemente, teniendo en cuenta que la RNAP- σ^A sí fue capaz de reconocer el posible promotor de la región líder *in vitro*, las cuatro estirpes mostraron niveles muy bajos de actividad β-galactosidasa (Figura 58B). Puesto que la ausencia de actividad promotora podría ser debida a la falta de algún elemento crítico en el sitio heterólogo, se generaron plásmidos que permiten integrar las fusiones Líder::*lacZ* o Líder(Mut)::*lacZ* (Figura 58A) en el sitio endógeno mediante recombinación homóloga en los fondos genéticos silvestre, $\Delta ddvA$, $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$. Tal y como se observa en la figura 58B, solo se detectó actividad promotora significativa para la fusión Líder::*lacZ* en la estirpe $\Delta ddvA$; además, dicha actividad se mantuvo a niveles similares para la versión

Líder(Mut)::*lacZ*. Así pues, el posible promotor dependiente de σ^A identificado en la región líder no parece contribuir de manera significativa a la expresión de CRISPR4.



Figura 58. Expresión de la región líder silvestre o mutada en un sitio heterólogo versus en el sitio endógeno. A) Estrategia seguida para la generación de las estirpes. B) Medida de la actividad β -galactosidasa de las estirpes indicadas.

III.4.3 La región CRISPR4 se expresa a partir del operón cas

Puesto que la región líder carece de actividad promotora, la expresión de CRISPR4 podría tener su origen en el promotor situado aguas arriba de los genes *cas*, P_{MXAN_-7283} . De ser así, DdvS, CarD y CarG coordinarían la expresión de los genes *cas* con la transcripción de la región CRISPR4. Para examinarlo, se generó un plásmido que presenta el gen *lacZ* inmediatamente aguas abajo de *cas*6 (el gen *cas* que precede a la región líder) y se integró en el sitio endógeno mediante recombinación homóloga en los diferentes fondos génicos (silvestre, $\Delta ddvA$, $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$) (Figura 59A). Como se muestra en la Figura 59B, solo se observó expresión del gen *lacZ* en la estirpe $\Delta ddvA$, y no en el fondo silvestre, el $\Delta ddvA \Delta carD$ o el $\Delta ddvA \Delta carG$, lo que sugiere que la transcripción dependiente de DdvS, CarD y CarG va más allá del propio *cas*6.



Figura 59. Expresión de la fusión *cas6::lacZ***.** A) Estrategia seguida para generar la estirpe de *M. xanthus* con la fusión *cas6::lacZ* en el sitio endógeno. B) Medida de la actividad β -galactosidasa de *cas6::lacZ* en los diferentes fondos genéticos indicados.

La mayoría de los genes cas del sistema CRISPR4-Cas muestran solapamiento de los codones de inicio y fin (cmr1-cmr2, cmr3-cmr4, cmr4-cmr5, cmr5-cmr6 y cmr6cas6) o una separación de solo 2 (cmr2-cmr3) o 4 pb (MXAN 7283-cmr1). Ello sugiere que constituyen un operón, que cabría esperar que se transcribiera a partir de P_{MXAN 7283} como un mRNA policistrónico (Figura 60A). Para confirmarlo, se analizó RNA total extraído de la estirpe silvestre y de la estirpe $\Delta ddvA$ mediante RT-PCR con parejas de oligonucleótidos que amplifican las uniones entre los diferentes genes cas. Además, para comprobar si el transcrito incluye la región CRISPR4, se usaron parejas de cebadores que permitirían amplificar un fragmento entre cas6 y el primer o segundo espaciador de la región CRISPR4 (Figura 60A). Como cabía esperar, no se observó amplificación en las reacciones control, en las que se utilizó el RNA sin retrotranscribir como molde (Figura 60B). Por otro lado, el análisis de los productos amplificados permitió corroborar que los genes cas no se expresan en la estirpe silvestre en las condiciones estándar de cultivo utilizadas en el laboratorio, pero sí lo hacen en la estirpe ΔddvA (Figura 60B). La aparición a partir de dicha estirpe de productos amplificados con todas las parejas de oligonucleótidos utilizadas apoya la hipótesis de que los genes cas se transcriben, de manera dependiente de DdvS, como un mRNA policistrónico que, además, incluye la región CRISPR4.



Figura 60. Análisis de la expresión de un mRNA policistrónico a partir de P_{MXAN_7283} y su extensión hasta la región CRISPR4. A) Esquema de los genes *cas* y de la región CRISPR4. Entre paréntesis se indica el tamaño en pb de los fragmentos (1 al 9) amplificados por PCR. B) Ensayo de RT-PCR de las uniones de los diferentes genes *cas*, y de *cas6* y el primer o segundo espaciador de CRISPR4, en un fondo silvestre o $\Delta ddvA$. El signo "+" indica que la PCR se realizó con cDNA como molde, mientras que el signo "–" indica que la PCR se realizó com RNA no retrotranscrito como molde.

La transcripción dependiente de DdvS de la región CRISPR4 podría provenir no solo de P_{MXAN_7283} sino también de los otros promotores dependientes de DdvS ubicados aguas arriba. Ello implicaría la formación de un transcrito que englobaría, además de la información de los genes *cas*, la de los genes situados aguas arriba. Para comprobarlo, se analizaron mediante RT-PCR las uniones de los genes *MXAN_7285-MXAN_7284*, y *MXAN_7284-MXAN_7283*, usando como control positivo el fragmento de 400 pb que se obtiene al amplificar la unión entre *MXAN_7283* y *MXAN_7282* (fragmento 1 del experimento anterior). La aparición de productos amplificados en las tres reacciones, nuevamente solo en la estirpe $\Delta ddvA$ (Figura 61), apunta a que la transcripción de los genes *cas*, y probablemente la del locus CRISPR4, podría también proceder de los promotores dependientes de DdvS situados aguas arriba de P_{MXAN_7283}.



Figura 61. Análisis mediante RT-PCR de la expresión de un mRNA policistrónico entre los genes cas y los genes situados aguas arriba. El signo "+" indica que la PCR se realizó con cDNA como molde, mientras que el signo "-" indica que la PCR se realizó con RNA no retrotranscrito como molde.

III.4.4 DdvS, CarD y CarG controlan la producción de los crRNAs

Como se ha comentado en la Introducción, la transcripción de la región CRISPR. generalmente a partir de la región líder, genera un pre-crRNA que es procesado por las proteínas Cas a pequeñas moléculas o crRNA maduros (cada uno de las cuales incluye solo uno de los espaciadores). Para examinar la formación del transcrito de gran tamaño (~12 kb) que tendría su origen en el promotor $P_{MXAN 7283}$ e incluiría el operón cas y la región CRISPR4, o transcritos aún de mayor tamaño procedentes de los promotores situados agua arriba, se analizó el RNA total mediante Northern blot a partir de: la estirpe silvestre, las estirpes mutantes $\Delta ddvA$, $\Delta ddvA$ $\Delta carD$, $\Delta ddvA$ $\Delta carG$, y una estirpe que expresa *ddvA* bajo un promotor inducible por IPTG (P_{IPTG}::*ddvA*) en un fondo genético ∆ddvA. El análisis del RNA usando una sonda (~1 kb) que abarca desde el espaciador 52 al 37 reveló una señal intensa en los carriles correspondientes a la estirpe $\Delta ddvA$ y a la estirpe $\Delta ddvA$ P_{IPTG}:: ddvA crecida en ausencia de IPTG (Figura 62). Más que una banda definida, se observó una señal distribuida a lo largo del carril (>4 a <0,2 kb), que mostró su máxima intensidad en la zona correspondiente a los fragmentos de menor tamaño. Estos resultados sugieren que el transcrito inicial, que se produce de una manera dependiente de DdvS y CarD-CarG, se procesa rápidamente para producir los crRNAs.

La señal intensa no se observó en el resto de las estirpes examinadas (silvestre, $\Delta ddvA \Delta carD$, $\Delta ddvA \Delta carG$, y $\Delta ddvA P_{IPTG}$::ddvA crecida en presencia de IPTG), en las que sí se pudo detectar una banda tenue de ~3-4 kb (Figura 62), que podría corresponder al pre-crRNA, expresado a partir del promotor débil dependiente de σ^A situado en la región líder. Para comprobarlo, se generaron estirpes portadoras de la deleción de CRISPR4 (Δ CRISPR4) en un fondo genético silvestre o Δ *ddv*A. El análisis del RNA total extraído a partir de dichas estirpes también permitió detectar la banda de ~3-4 kb (Figura 62), lo que indica que se trata de una señal inespecífica, que no se corresponde con el pre-crRNA. Por otro lado, que la señal intensa a lo largo del carril tampoco se detectase en la estirpe Δ *ddv*A P_{IPTG}::*ddv*A crecida en presencia de IPTG indica que DdvA puede actuar en *trans*.



Figura 62. Análisis mediante *Northern blot* del RNA total obtenido a partir de las estirpes indicadas usando una sonda que hibrida desde el espaciador 52 al 37 de CRISPR4. 23S rRNA: control de carga obtenido por hidridación de la misma membrana con una sonda específica del rRNA 23S. A la izquierda se muestran los tamaños de los fragmentos de RNA usados como marcadores de tamaño El signo "+" indica presencia de IPTG y el signo "-" indica ausencia de IPTG.

Para analizar específicamente la producción de los crRNAs, se aislaron los RNAs pequeños a partir de las estirpes anteriores y se llevó a cabo un ensayo de *Northern blot* usando sondas que hibridan en el espaciador 52 (el más cercano al líder y el último que debió haberse incorporado en la región CRISPR4) o en el espaciador 1 (el más alejado del líder y el primero que debió haberse incorporado en la región CRISPR4). El análisis con ambas sondas permitió identificar en la estirpe $\Delta ddvA$, y en la estirpe $\Delta ddvA$ P_{IPTG}::ddvA crecida en ausencia de IPTG, una banda intensa de ~40 nt y otras bandas más tenues de ~35 y ~45 nt. La ausencia de dichas bandas en la estirpe $\Delta ddvA$ Δ CRISPR4 demuestra que se trata de crRNAs derivados específicamente de la región CRISPR4 (Figura 63). Tampoco se observaron los crRNAs en el resto de estirpes

examinadas (silvestre, $\Delta ddvA \Delta carD$, $\Delta ddvA \Delta carG$, y $\Delta ddvA P_{IPTG}$::ddvA crecida en presencia de IPTG), acorde con la falta de transcripción del locus *cas* y de la región CRISPR4 en dichas estirpes. Así pues, la formación de los crRNAs a partir de la región CRISPR4 es absolutamente dependiente de DdvS y del complejo CarD-CarG.



Figura 63. Análisis de los crRNAs producidos por las estirpes indicadas usando una sonda que hibrida en el espaciador 52 y otra en el espaciador 1 de la región CRISPR4. 5S rRNA: control de carga obtenido por hidridación de la misma membrana con una sonda específica del rRNA 5S. A la izquierda se muestran los tamaños de los fragmentos de RNA usados como marcadores de tamaño.

III.4.5 La producción de los crRNAs derivados de la región CRISPR4 depende de la proteína Cas6

Uno de los genes *cas* del sistema CRISPR4-Cas determina una proteína homóloga a Cas6 que, como se comentó en la Introducción, es una endonucleasa necesaria para la formación de los crRNAs, pues se encarga de cortar el pre-crRNA en cada una de las repeticiones que separan los espaciadores (Carte *et al.*, 2008). En los sistemas CRISPR de tipo III-B, como es el caso del sistema CRISPR4-Cas objeto de estudio, se ha descrito que Cas6 funciona como una proteína independiente, sin necesidad de formar un complejo con otras proteínas Cas (Pyenson & Marraffini, 2017). Para estudiar el papel de la proteína homóloga a Cas6 en la producción de los crRNAs y determinar si es capaz de actuar en *trans*, se generó una estirpe que expresa *cas6* bajo un promotor inducible por vanilato en un fondo genético $\Delta ddvA \Delta cas6$ (Figura 64A). En relación a la estirpe $\Delta ddvA$, el análisis mediante *Northern blot* puso de manifiesto la desaparición de los crRNAs en la estirpe $\Delta ddvA \Delta cas6$ y su reaparición en la estirpe $\Delta ddvA \Delta cas6$ P_{van}:: *cas6* crecida en presencia de vanilato (Figura 64B). Así pues, los resultados indican que Cas6 es indispensable para la formación de los crRNAs y que es posible suministrar su actividad en *trans*, lo que apoya su capacidad de actuar como una proteína independiente.



Figura 64. Papel de Cas6 en la formación de los crRNAs. A) Representación esquemática de la construcción de la estirpe que expresa *cas6* bajo el control del promotor inducible por vanilato (P_{van}) en un fondo genético $\Delta ddvA \ \Delta cas6$; el plásmido se integra por recombinación homóloga a través del fragmento *MXAN_18-19* (Iniesta *et al.,* 2012). B) Análisis mediante *Northern blot* de la formación de los crRNAs en las estirpes indicadas. 5S rRNA: control de carga obtenido por hidridación con una sonda específica del rRNA 5S. A la izquierda se muestran los tamaños de los fragmentos de RNA usados como marcadores de tamaño.

Cuando se estudió la producción del RNA largo, a partir de RNA total y usando la sonda que hibrida desde el espaciador 52 al 37, se observó una señal distribuida a lo largo del carril para la estirpe $\Delta ddvA \Delta cas6 P_{van}$::*cas6* crecida con vanilato, similar a la observada con la estirpe control $\Delta ddvA$ (Figura 65). Acorde con el papel de Cas6 en la formación de los crRNAs, con la estirpe $\Delta ddvA \Delta cas6$ y la $\Delta ddvA \Delta cas6 P_{van}$::*cas6* crecida en ausencia de vanilato no se observó el incremento de intensidad correspondiente a los fragmentos de menor tamaño (Figura 65). Curiosamente, esta disminución en los RNAs pequeños no se vio acompañada de un incremento de señal en los fragmentos de mayor tamaño, quizás porque la falta de Cas6 provoca la inestabilidad del pre-crRNA o el descenso en su expresión.



Figura 65. Análisis del papel de Cas6 mediante *Northern blot* de RNA total obtenido a partir de las estirpes indicadas usando una sonda que hibrida desde el espaciador 52 al 37 de CRISPR4. 23S rRNA: control de carga obtenido por hidridación de la misma membrana con una sonda específica del rRNA 23S. A la izquierda se indican los tamaños de los fragmentos de RNA usados como marcadores de tamaño.

La búsqueda de proteínas similares a Cas6 en el genoma de *M. xanthus* permitió identificar una proteína homóloga que forma parte del sistema CRISPR2/3-Cas de tipo I-C. Este sistema incluye los genes *devTRS*, cuya expresión se activa durante el desarrollo multicelular (Rajagopalan & Kroos, 2017; Rajagopalan *et al.*, 2015; Viswanathan *et al.*, 2007). Dicha proteína Cas6, a la que denominamos Cas6a, presenta una cola adicional de 27 aminoácidos en su extremo N-terminal y comparte un 66% de identidad con la proteína Cas6 del sistema CRISPR4-Cas (Figura 66).

Cas6(CRISPR4)	1VTSLDLLYPVRGGPVPLD GYAL SALCTRVPA
Cas6a(CRISPR2/3)	1 LARSVGWRRSALITVSGCVILNWWGNLMVFVDLLPPVCGGPVPLD AYLL SALSRHLPA
consensus	1***.** ***************************
Cas6(CRISPR4)	34 LERIDI GVETINGERAVGDIN HIGRGTIRINGSPEVVPLITSTPLTPPNVAGREVVIGA
Cas6a(CRISPR2/3)	61 LERSDMGVESINGVSNTREIN YLGRGTMRINGPIBAVATITPTVSAPTETAGR
consensus	61 **** * *** *** *** *** *** *** * * *
Cas6(CRISPR4)	94 EVILALISEASSIJEAR IVTEKHALDAASELREVSRSIJEALACDARPVI GRRRIVRIAGKRV
Cas6a(CRISPR2/3)	121 ESILIALEEVPSIJEAR IVTEKHAMDEAAEVAASRAJJEALEVOATLKVGRRRIVRITGKKV
consensus	121 * *.**** ******.********* * * ** ****. * .******
	GhGxxxxxGhG
Cas6 (CRISPR4)	154 VGYAMEINIGVSGEDSIRVQAQEIGGRRHMCCEMEINPERPVVRGVVPQIGEIARAA
Cas6a (CRISPR2/3)	181 VGFAMEINIGTSAEHSIRVQEQCMCGRRHMCCEMEINPEGRAARVQSHCKAA
consensus	181 **.******.*.*

Figura 66. Alineamiento de las dos proteínas Cas6 de *M. xanthus*. En rojo se indica el posible sitio activo y en violeta se muestra la región de reconocimiento de RNA.

Además, el análisis de ambos sistemas CRISPR-Cas reveló que sus repeticiones solo difieren en dos posiciones y que sus regiones líderes presentan hasta un 74% de identidad (Figura 67).

CRISPR	Secuencia de las repeticiones (tamaño en pb)	Lider		
			Tamaño	% GC
2	GTGCTCAACGCCTTTCGGCATCACGGCGAGCGGGACC (3	37)	80	63.75
3	GTGCTCAACGCCTTTCGGCATCACGGCGAGCGGGAC (3	36)	248	59.27
4	GTGCTCAACGCCTTCCGGCATCACGGCGAGCGGCAC (3	36)	304	58.55

Figura 67. Alineamiento de las repeticiones de las regiones CRISPR2 y CRISPR3 del sistema CRISPR2/3-Cas y de las repeticiones de la región CRISPR4 del sistema CRISPR4-Cas. En rojo se indican las dos posiciones en las que difieren las secuencias.

En base a estas similitudes se estudió si la proteína Cas6a, actuando en *trans*, es capaz de procesar el pre-crRNA generado por el sistema CRISPR4-Cas. Como puede observarse en la Figura 68, la ausencia de crRNAs propia de la estirpe $\Delta ddvA$ $\Delta cas6$ se vio rescatada por la expresión de Cas6a bajo el promotor P_{van}. Estos resultados indican que Cas6a es, en efecto, capaz de procesar el pre-crRNA del sistema CRISPR4-Cas, consistente con el gran parecido entre las dos proteínas Cas6 y entre las repeticiones de ambos sistemas CRISPR-Cas. Además, indican que Cas6a, como Cas6, puede actuar en *trans*.



Figura 68. Formación de los crRNAs de la región CRISPR4 mediada por Cas6a. A) Representación esquemática de la construcción de la estirpe que expresa *cas6a* bajo el control del promotor inducible por vanilato (P_{van}) en un fondo genético $\Delta ddvA \Delta cas6$; el plásmido se integra por recombinación homóloga a través del fragmento *MXAN_18-19* (Iniesta *et al.*, 2012). B) Análisis mediante *Northern blot* de la formación de los crRNAs en las estirpes indicadas. 5S rRNA: control de carga obtenido por hidridación con una sonda específica del rRNA 5S. A la izquierda se muestran los tamaños de los fragmentos de RNA usados como marcadores de tamaño.

III.4.6 La expresión del sistema CRISPR4 no está activada por el desarrollo

Los resultados obtenidos con Cas6a y las similitudes indicadas entre el sistema CRISPR4-Cas (de tipo III-B) y el CRISPR2/3-Cas (de tipo I-A) abren la posibilidad de que exista una comunicación cruzada entre los dos sistemas CRISPR-Cas, un escenario que requeriría que ambos se encontraran activos simultáneamente. Dado que el sistema CRISPR2/3-Cas se activa durante el desarrollo multicelular que conduce a la formación de los cuerpos fructíferos (Rajagopalan & Kroos, 2017; Rajagopalan *et al.*, 2015; Viswanathan *et al.*, 2007), se analizó si el sistema CRISPR4-Cas también se activa bajo esas condiciones. Para ello, en primer lugar, se comparó la formación de cuerpos fructíferos a las 24, 48 y 72 horas entre la estirpe silvestre y la estirpe mutante $\Delta ddvA$, usando como control negativo de fructificación la estirpe $\Delta carD$, ya que la falta de CarD

afecta al desarrollo multicelular (Nicolás *et al.*, 1994). No se observó diferencia entre la estirpe silvestre y el mutante $\Delta ddvA$ en cuanto a la velocidad de fructificación o el aspecto de los cuerpos fructíferos, si bien la estirpe $\Delta ddvA$ produjo un número de cuerpos fructíferos inferior, quizás debido al exceso de DdvS como consecuencia de su expresión constitutiva en la estirpe $\Delta ddvA$ y la posible competición por la RNAP con otros factores σ que actúan durante el desarrollo multicelular (Figura 69). Así pues, la expresión constitutiva de DdvS no interfiere de manera notable con el proceso de fructificación.



Figura 69. Producción de cuerpos fructíferos en la estirpe silvestre y las estirpes mutantes $\Delta ddvA$ y $\Delta carD$ a los tiempos indicados durante el desarrollo multicelular.

Para determinar si el sistema CRISPR4-Cas se expresa durante el desarrollo, se emplearon las estirpes con la fusión Líder::*lacZ* introducida en el sitio endógeno en los fondos genéticos silvestre, $\Delta ddvA$, $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$. El análisis cualitativo en presencia de X-gal mostró expresión del gen *lacZ* solo en la estirpe $\Delta ddvA$, que produjo coloración azul tanto en fase vegetativa (0 h) como en diferentes momentos del desarrollo multicelular (12, 24, 48 y 72 horas) (Figura 70A). Para analizar de manera cuantitativa la expresión del gen *lacZ*, se tomaron muestras de las mismas estirpes durante las primeras 24 horas de desarrollo (0, 4, 8, 12 y 24 horas), utilizando la estirpe $\Omega 4435$::Tn5-*lac* ($\Omega 4435$) como control positivo, pues es sabido que la actividad β -galactosidasa derivada del transposón Tn5-*lac* asociado a dicha inserción se incrementa a partir de las 12 horas de desarrollo (Kroos *et al.*, 1986). La estirpe $\Delta ddvA$ mostró niveles elevados de actividad β -galactosidasa durante el proceso de desarrollo, similares a los observados durante la fase vegetativa que, como cabía esperar, disminuyeron notablemente en las estirpes $\Delta ddvA \Delta carD$ y $\Delta ddvA \Delta carG$ (Figura 70B). Por otro lado, no se observó expresión en la estirpe silvestre, lo que apunta a que la



falta de nutrientes (la señal que induce el desarrollo multicelular) no es la señal que provoca la inactivación de DdvA e induce la expresión del sistema CRISPR4-Cas.

Figura 70. Análisis cualitativo y cuantitativo de la expresión de la fusión Líder::*lacZ* durante el desarrollo en los fondos genéticos indicados. A) Ensayo cualitativo de la expresión de Líder::*lacZ* durante el desarrollo en medio de fructificación TPM. B) Medida de la actividad β -galactosidasa de Líder::*lacZ* durante el desarrollo.

Por último, como ensayo adicional, se analizó la producción de los crRNAs durante el desarrollo de la estirpe silvestre y la estirpe $\Delta ddvA$, usando la sonda que hibrida en el espaciador 52 del sistema CRISPR4-Cas. Mientras que con la estirpe $\Delta ddvA$ se detectaron los crRNAs desde la fase vegetativa hasta las 48 h de desarrollo, con la estirpe silvestre no se observaron a ninguno de los tiempos examinados (Figura 71, panel izquierdo).

Como se ha comentado previamente, varios estudios indican que la expresión del sistema CRISPR2/3-Cas de *M. xanthus* se induce durante el desarrollo a niveles bajos, que aumentan significativamente cuando se elimina la autoregulación negativa ejercida por los genes *devTRS* (Rajagopalan & Kroos, 2017; Rajagopalan *et al.*, 2015; Viswanathan *et al.*, 2007). Sin embargo, hasta la fecha no se había examinado la producción de los crRNAs correspondientes. Cuando se analizó, usando una sonda que hibrida con el espaciador 22 del sistema CRISPR2/3-Cas (el más cercano a la región líder), no llegó a detectarse ninguna señal a los tiempos examinados durante el desarrollo (tampoco, como cabía esperar, durante la fase vegetativa) (Figura 71, panel derecho). Puesto que los resultados negativos podrían deberse a los niveles bajos de inducción durante el desarrollo en el fondo genético silvestre, se analizaron también los crRNAs tras eliminar la autoregulación negativa mediante deleción de *devT*.



Figura 71. Análisis de los crRNAs producidos durante el desarrollo por los sistemas CRISPR4-Cas y CRISPR2/3-Cas. Panel izquierdo: análisis mediante *Northern blot* de la formación de los crRNAs del sistema CRISPR4-Cas en las estirpes indicadas. Panel derecho: análisis mediante *Northern blot* de la formación de los crRNAs del sistema CRIPR2/3-Cas en las estirpes indicadas. 5S rRNA: control de carga obtenido por hidridación con una sonda específica del rRNA 5S. A la izquierda se muestran los tamaños de los fragmentos de RNA usados como marcadores de tamaño.

En las muestras del mutante $\Delta devT$ sí se observó hibridación de la sonda, aunque las señales fueron débiles y el tamaño de los crRNAs detectados (~90 nt, ~150 nt y superiores) fue mayor que el observado para los crRNAs del sistema CRISPR4-Cas (Figura 71). Este mayor tamaño podría deberse a diferencias en el procesamiento final de los crRNAs del sistema CRISPR2/3-Cas en relación al sistema CRISPR4-Cas, a que el procesamiento es incompleto en el fondo mutante $\Delta devT$ o a una desestabilización de los crRNAs maduros provocada por la falta de DevT (un homólogo de la proteína Cas8a1, que probablemente forma parte del complejo de interferencia). Ni la deleción de devT provocó la aparición de los crRNAs del sistema CRISPR4-Cas ni la de ddvA la aparición de los crRNAs del sistema CRISPR2/3-Cas (Figura 71), lo que indica que cada una de ellas regula específicamente el sistema al que se encuentra asociado el gen correspondiente. En conjunto, los resultados permiten concluir que el desarrollo multicelular no activa la expresión del sistema CRISPR4-Cas ni promueve la producción de sus crRNAs. En consecuencia, parece improbable que exista comunicación cruzada entre el sistema CRISPR4-Cas y el CRISPR2/3-Cas, al menos, en las condiciones ensayadas.

IV. DISCUSIÓN

IV. 1 El dominio N-terminal de CarD: subdominios e interacciones

Las proteínas reguladoras CarD y CdnL, ambas pertenecientes a la familia CarD_CdnL_TRCF de proteínas de unión a la RNAP, son funcionalmente diferentes en *M. xanthus* a pesar del parecido entre CarDNt y CdnL. CarD está presente solo en las mixobacterias, no es esencial (Elías-Arnanz *et al.*, 2010) y regula junto a CarG la acción de determinados factores σ -ECF (Abellón-Ruiz *et al.*, 2014). CdnL, en cambio, está ampliamente extendida en bacterias (Bae *et al.*, 2015; Gallego-García *et al.*, 2017; Gallego-García *et al.*, 2014; Garcia-Moreno *et al.*, 2010), es esencial para la viabilidad celular en la mayoría de las especies analizadas y actúa estabilizando los complejos de transcripción abiertos en promotores dependientes del factor σ mayoritario (Bae *et al.*, 2015; Gallego-García *et al.*, 2014). Los estudios con CdnL y con CarD llevados a cabo por el grupo de investigación en el que se ha realizado este trabajo están permitiendo obtener una mejor comprensión del modo de acción molecular de CdnL y CarD, de la función de sus dominios proteicos, y del conjunto de genes sujetos a su acción reguladora.

La arquitectura de dominios de CarD, única en el mundo procariótico, la hace especialmente interesante. Su dominio CarDCt, ausente en CdnL, muestra un parecido sorprendente con las proteínas eucarióticas de tipo HMGA, que funcionan como factores arquitectónicos capaces de modular la topología del DNA. Acorde con este parecido, la sustitución de CarDCt por la proteína HMGA humana genera una proteína tan funcional in vivo como la propia CarD (García-Heras et al., 2009). Además, in vitro, CarD muestra la misma especificidad de unión, al surco menor de secuencias ricas en pares AT, que sus homólogos eucarióticos (García-Heras *et al.*, 2009). Acorde con ello, la mutación de los tramos ricos en pares AT en el promotor fotoinducible Pors (García-Heras et al., 2013) o en el promotor P_{ddvSA} anulan la unión *in vitro* de CarD a las correspondientes regiones promotoras (Abellón-Ruiz et al., 2014). Que la falta de CarDCt provoque solo una pérdida parcial de actividad en los promotores analizados hasta la fecha (P_{ORS} y P_{ddvSA}) sugiere que la función *in vivo* de dicho dominio es la de maximizar la actividad de CarD. En el caso de las proteínas HMGA eucarióticas, sus interacciones con el DNA, con varios factores proteicos y con la maquinaria basal de transcripción favorecen de manera sinérgica el ensamblaje cooperativo de los "enhanceosomes" (Merika & Thanos, 2001; Thanos & Maniatis, 1992; Yie et al., 1997). De manera análoga, es posible que la unión de CarD al DNA a través de su dominio tipo HMGA tenga un efecto sinérgico sobre la actividad parcial de CarDNt, quizás favoreciendo una conformación particular del DNA en el promotor diana o estabilizando la unión del complejo CarD-CarG a sus promotores diana a través de su interacción con regiones ricas en pares de bases AT.

Al menos en P_{QRS} , el cambio de posición del sitio de unión de CarD produce un efecto negativo sobre la activación del promotor (García-Heras *et al.*, 2013), lo que sugiere que su acción reguladora está mediada por la interacción con la maquinaria basal de transcripción, ya sea directamente o a través de otros factores. Es probable que este papel resida en el dominio N-terminal de CarD, CarDNt, por el cual CarD interacciona tanto con CarG como con la subunidad β de la RNAP. Así, si la falta de CarDCt reduce la actividad de CarD, la falta de CarDNt genera una proteína completamente inactiva (Cayuela *et al.*, 2003), lo que refuerza la importancia de las interacciones proteína-proteína en la función de CarD.

Las interacciones proteína-proteína conocidas en las que participa CarD recaen en dos subdominios distintos de CarDNt: la interacción con β RNAP en el subdominio Nterminal (CarD₁₋₇₂), y la interacción con CarG en el subdominio C-terminal (CarD₆₁₋₁₇₉). La estructura tipo Tudor del subdominio CarD₁₋₇₂, con cinco láminas β antiparalelas, determinada en este trabajo en colaboración con la Dra. María Ángeles Jiménez (IQFR-CSIC), es muy similar a la observada para el dominio equivalente de CdnL de *M. xanthus* (Gallego-García *et al.*, 2014), de *T. thermophilus* (Gallego-García *et al.*, 2012), o de *M. tuberculosis* (Bae *et al.*, 2015; Gulten & Sacchettini, 2013; Kaur *et al.*, 2014; Srivastava *et al.*, 2013), y para el dominio de interacción con la RNAP (RID) de los factores TRCF (Deaconescu *et al.*, 2006; Westblade *et al.*, 2010). Dicha similitud estructural y el conocimiento de la estructura del complejo formado por el dominio RID del factor TRCF de *T. thermophilus* y β RNAP nos han llevado a identificar los aminoácidos candidatos a estar implicados en la interacción CarDNt- β RNAP.

Los resultados obtenidos en este trabajo han permitido confirmar la implicación de los aminoácidos F41, M54 y P56 en la interacción de CarD con β RNAP y a la vez desvelar que, sorprendentemente, la falta de dicha interacción, debida a la mutación de cualquiera de dichos residuos, solo afecta ligeramente a la función reguladora de la proteína *in vivo*, a diferencia de lo que ocurre con CdnL. La diferencia entre CarD y CdnL en el efecto de las mutaciones podría deberse a que CarD participa en al menos dos interacciones adicionales, con CarG y con el DNA, que se mantienen intactas en los mutantes y que podrían suplir en cierto modo la falta de interacción con β RNAP. CdnL, sin embargo, al ser incapaz de interaccionar con CarG y carecer de un dominio de unión al DNA, es esperable que dependa de manera más estricta de su interacción con la RNAP para llevar a cabo su acción reguladora (García-Moreno *et al.*, 2010). Por otra parte, las mutaciones en los aminoácidos V123 o E125, localizados en un tramo conservado de β RNAP que ha sido implicado en la interacción con CdnL o TRCF-RID, también eliminan la interacción con CarDNt. En conjunto, los resultados del análisis

mutacional indican que, más allá de pequeñas diferencias, las superficies de contacto están muy conservadas.

La coexistencia en *M. xanthus* de tres proteínas (TRCF, CdnL y CarD) que interaccionan con la misma región de β RNAP plantea la posibilidad de que se produzca una competición entre ellas por la unión a la RNAP. Cabe esperar que dicha competición esté determinada por las diferentes afinidades de cada una de ellas por β RNAP, por las especificidades por sus promotores diana (dependientes de factores σ -ECF versus dependientes del factor σ^A , por ejemplo), por sus interacciones con otros factores, y por las condiciones fisiológicas concretas en las que se encuentren creciendo las células. Así, a pesar del parecido estructural entre CarD₁₋₇₂ y el subdominio equivalente de CdnL (CdnLNt), y de que ambas interaccionan de manera similar con β RNAP, solo CdnL es capaz de estabilizar la formación de los complejos abiertos formados por la RNAP- σ^A (Gallego-García *et al.*, 2014).

A diferencia de CarD₁₋₇₂, no ha sido posible determinar la estructura a alta resolución del módulo CarD₆₁₋₁₇₉, que participa en la interacción con CarG. Se trata de un módulo muy susceptible a la degradación por proteasas, lo que sugiere que por sí mismo no adopta un plegamiento definido, a diferencia de su equivalente en CdnL (CdnLCt) que es compacto, resistente a proteasas y con una estructura terciaria bien definida (Gallego-García et al., 2014). Y, sin embargo, las mutaciones en una serie de residuos básicos conservados entre CarD₆₁₋₁₇₉ y CdnLCt que, en CarD₆₁₋₁₇₉ no están implicados en la interacción con CarG, afectan tan drásticamente a la función de CarD como a la de CdnL. Así, las mutaciones de los aminoácidos K93, R95, y R97 de CarD provocaron un fenotipo de falta de función similar al de la deleción completa del gen, al igual que las mutaciones de los residuos equivalentes (R90, R91, R93) de CdnL que, además, eliminaron su actividad estabilizadora de los complejos abiertos de transcripción (Gallego-García et al., 2014). En estudios con la proteína CdnL de M. tuberculosis se ha sugerido que dicha región básica media la unión débil e inespecífica al DNA, ya que su eliminación por mutagénesis anula la unión al DNA *in vitro* (Gulten & Sacchettini, 2013) y provoca una reducción en el crecimiento celular in vivo (Garner et al., 2014). Sin embargo, ni CarDNt ni CdnL de M. xanthus son capaces de unirse al DNA in vitro (Gallego-García et al., 2014). Por lo tanto, cualquiera que sea la función que desempeña esta región básica, se trata de una función clave para la actividad de todas las proteínas de la familia en las que se ha analizado. Que en CdnL se requiera la región básica para su actividad estabilizadora de los complejos abiertos de transcripción apunta a una posible interacción de los aminoácidos básicos, más que con el DNA libre, con el DNA de la burbuja de transcripción en los complejos con RNAP-o^A o, en el caso de

CarD, en los complejos formados en los promotores dependientes de factores σ -ECF. La determinación de la estructura tridimensional de los correspondientes complejos de transcripción permitiría dar respuesta a estas y otras cuestiones aún por resolver.

Desde un punto de vista evolutivo, que CarDNt y CdnL presenten una organización de dominios similar, compartan una estructura semejante en el subdominio N-terminal de interacción con la RNAP y unas regiones básicas en el subdominio C-terminal que resultan cruciales para su funcionalidad, sugiere que ambas derivan de un ancestro común. La adquisición en el caso de CarD del dominio C-terminal de tipo HMGA (ya sea por transferencia horizontal o por convergencia evolutiva) y de la capacidad de interaccionar con CarG debe haber permitido la diversificación funcional que caracteriza a CarD.

IV. 2 Acción reguladora global del complejo CarD-CarG

Aunque CarD (y posteriormente, CarG) se identificó por su papel en la respuesta a la luz y el proceso de desarrollo multicelular (Nicolás *et al.*, 1994), el análisis genético posterior sugirió una acción reguladora global (Galbis-Martínez *et al.*, 2004). Dicha acción global podría estar mediada, al menos en parte, por el hecho de que un buen número de factores σ -ECF de *M. xanthus* dependen total o parcialmente del complejo CarD-CarG para su actividad (Abellón-Ruiz *et al.*, 2014). No obstante, los datos de los análisis de ChIP-Seq realizados en este trabajo indican la unión de CarD a más de 100 sitios distribuidos a lo largo de todo el genoma, situados en regiones intergénicas *a priori* no relacionadas con promotores dependientes de factores σ -ECF. Por otro lado, la coincidencia entre los perfiles de unión obtenidos para CarG, que no se une directamente al DNA, y los obtenidos para CarD, indican una acción conjunta en todos y cada uno de los sitios identificados en el genoma de *M. xanthus*, consistente con su funcionamiento en forma de complejo.

Entre los nuevos genes identificados que parecen estar sujetos a regulación por CarD-CarG se encuentra el gen *rnk*, que cifra una proteína similar al dominio C-terminal de los factores Gre, factores que se unen a la RNAP y promueven la elongación. En *E. coli*, se ha visto que la proteína homóloga se une a la RNAP *in vitro* y compite con los factores Gre, lo que sugiere una posible actividad anti-Gre para Rnk (Lamour *et al.*, 2008). Por otro lado, aunque se observó localización de CarD-CarG en la región promotora de *MXAN_6483*, dicha unión no parece ir asociada a cambios en los niveles de expresión del gen. Cabe la posibilidad de que la unión del complejo en este y algunos otros sitios identificados en este trabajo esté relacionada con un papel estructural del genoma, más que regulador, similar al que también ejercen las proteínas eucarióticas

HMGA en algunos de sus sitios diana. El análisis bioinformático más detallado y los experimentos *in vivo* e *in vitro* con una muestra más amplia de las regiones identificadas permitirá una mejor comprensión de la acción global del complejo CarD-CarG.

Un aspecto inesperado de los estudios de ChIP-Seq y ChIP-gPCR ha sido la observación de que en los promotores P_{ORS} y P_{ddvSA} , ambos dependientes de factores σ-ECF, solo se observa enriquecimiento de CarD-CarG cuando dichos promotores se encuentran activos (en la luz y en ausencia de DdvA, respectivamente). Estos resultados sugieren que, pese a la presencia de un dominio de unión al DNA en CarD y de regiones ricas en pares AT en ambos promotores, a las que CarD se une in vitro, la asociación estable del complejo a dichos promotores in vivo solo se produce cuando están activos. Además, aunque CarG no se une directamente al DNA, su presencia es indispensable para que CarD se localice en los promotores P_{QRS} y P_{ddvSA} in vivo, lo que explicaría que CarD, por sí sola, carezca de actividad reguladora. CarG podría participar estabilizando el complejo entre CarD-CarG y la RNAP o, a través de la interacción con otras proteínas reguladoras que permitan el reconocimiento de dichas regiones promotoras. Aunque mediante el sistema del doble-híbrido no se ha observado interacción de CarG con ninguna de las subunidades de la RNAP, ni con el factor σ -ECF CarQ (García-Heras et al., 2013), no se puede descartar que dicha interacción sí ocurra una vez formado el complejo CarD-CarG, constituyendo un complejo análogo a los complejos multiproteicos o "enhanceosomes" que dirigen la transcripción en células eucarióticas. En estos complejos, CarD-CarG podrían actuar no solo como factores transcripcionales sino también como elementos auxiliares que favorecen la estabilización y el ensamblaje de reguladores específicos, como sucede con las proteínas HMGA (Thanos & Maniatis, 1992; Yie et al., 1997). Cabe esperar que la actividad del complejo y su localización in vivo en el DNA esté determinada por la contribución sinérgica de las diversas interacciones, proteína-DNA y proteína-proteína, en las que participan sus dos componentes.

IV.3. Factores anti- σ en *M. xanthus*: topología en la membrana y asociación de dominios

Una de las estrategias utilizadas por las bacterias para adaptarse a los cambios en las condiciones medioambientales es el uso de factores σ -ECF para la activación de la expresión de grupos específicos de genes en función del estímulo concreto al que se enfrentan (Gaballa *et al.*, 2018). En general, los factores σ -ECF están regulados negativamente por contacto directo con un factor anti- σ , que suele ser una proteína transmembranal cifrada por un gen contiguo que a menudo se coexpresa con el factor

MXAN 7455

consensus

71 .

σ-ECF. Hasta un 40% de estos factores anti-σ presentan un dominio NZASD, que ha sido estudiado en profundidad en el factor anti-σ RsrA de *Streptomyces coelicolor* (Rajasekar *et al.*, 2016) o ChrR de *Rhodobacter sphaeroides* (Greenwell *et al.* 2011), entre otros. Los dominios NZASD tienen en común el motivo HisX₃CysX₂Cys (HCC) como núcleo principal de coordinación de Zn, al que se suma una histidina o cisteína a 23-26 aminoácidos del núcleo de coordinación (motivos HHCC o CHCC).

		Hxxx <mark>exx</mark> e
MXAN_0070	1	MNSLHRPPM-S-ACPDENLAAVSCSAPVGKVAVVEA
MXAN_0948	1	VICVICVICVICVIC-VIC-VIC
MXAN_1513	1	AEQ YLCALAPEKAAS EAHTLE CEP ARL QEEAL
MXAN_1660	1	VPAVKGFLPRASELAEGGGMTSKSC-DGYQLATEMSRHCALDVR_KEEAEQHAATOPEOQRAHALSAE
MXAN_1710	1	MFGKPERVAESSPM-Q-SCPQETTITDL ACLISDAHRAQVLAHVETCADORWAIGAGVG
MXAN_2029	1	MRGWSSTKPCNDCSAPTGPEV-SMARHEHDSL ALAACODDADASGR EAHVATCSECAQADEQVRQ
MXAN_2394	1	ACALEDDALG IQ DEGOEGORQAHAASKE
MXAN_2738	1	M-GCPSETTISD ECAIPDAHRTRULAUEGER QEHIALGAS
MXAN_3960	1	
MXAN_4148	1	DE FOPE RVD ET USGAD RRRAEEERO
MAAN_4315	1	WEALSPYWERISA
MXAN_4001	1	
MXAN 4986	1	
MXAN 5507	1	
MXAN 6009	1	DPENALSGANCALDPEGMTT RO
MXAN 6460	1	DE SPE TOH REHIRGCSPOVDFIRTYRA
MXAN 6682	1	VRNREHALMRWYERHIDASTTRSLC SSHOSAR LR AACR GTRYERWAQ
MXAN 6758	1	VEPMAGNPAC-ERFVPMUSPYYDCOUSSGORVNYER <mark>HUAAC</mark> RD <mark>O</mark> TGRAADLRA
MXAN_7082	1	ECQUSSERAQQVRAHVAGCAECRSVIAALSS
MXAN_7162	1	M_AC-LDENTFVALMCDPPARATEVDALTODA
MXAN_7215	1	GAQPP <mark>HITHC</mark> TTCRARLEQLRT
MXAN_7288*	1	ADCOMPPMOAQAFGQ
MXAN_7327	1	AD USPE ATR RA AQA PA QHAUSURA
MXAN_7455	1	FSCR APP ERR RE PD AR RERYERQLL
	_	
MXAN_0070	54	SSSLPGAVVVEREAEAP
MXAN_0948	44	ELLASRALGGSASVAEPPPAKVTPLR
MXAN_1513	47	SEQLHEVANTYAPAVHVLRPARW
MXAN_1660	68	
MXAN_1710	59	TQAPS-SAPTAPG
MXAN_2029	67	SRAV
MXAN_2394	47	LSLVALP-ALSPQEKTELQELPRR
MXAN_2738	45	AAEAV-PPPVEGPLVTTGTQLSRYVVQ
MXAN_3960	43	MEQDLYRLADPLPPPTLVASVMARVSAE
MXAN_4148	46	QQALRRAARHSVSGMRAPASLRAG
MXAN_4315	48	RQALRAEGSAEQQVPPELEERV
MXAN_4661	49	VRATMSQL-SVEPAPDAGLESLM
MXAN_4732	60	TVQRL-QRVERE
MXAN_4986	45	ELLGFHVMGDPRVLAHEQEDAPETW
MXAN_5507	44	ELLADDALNSPGTERLGATSPWRNKS
MXAN_6009	45	ESP-P-PLVDTGLLVSGARVGRIVVLG-
MXAN_6460	4/	TPGLCKKALVAKMPKEVSEKLTEF
MYAN 6759	20	
MYAN 7092	23	
MXAN_7162	40	
MYAN 721F	40	
MXAN 7288*	29	TOTICRCYTERHCDVDTOWHALDOND
MXAN 7327	45	ERCWMACRARRM-DARDALDEALESR

Figura 72. Alineamiento de los 70 primeros aminoácidos de los factores anti-σ con motivo NZASD de *M. xanthus*. En rojo y azul se indican las posibles histidinas y cisteínas, respectivamente, que podrían participar en la coordinación de Zn. En violeta se muestra el núcleo de coordinación de Zn. El asterisco en MXAN_7288 indica que éste se corresponde con DdvA.

47 ARLDPSAPDARTR----LA----RGLGLETR-----

De los ~45 posibles factores σ -ECF que presenta *M. xanthus*, la mayoría sin caracterizar, muchos están asociados a un posible factor anti- σ con un dominio NZASD. En estos factores anti- σ , el que podría ser el cuarto aminoácido del núcleo de coordinación (His o Cys) se sitúa a 19-26 aminoácidos del motivo HisX₃CysX₂Cys (Figura 72).

En este trabajo se han estudiado nueve parejas σ -ECF/anti- σ de *M. xanthus* (Figura 73), ocho de las cuales presentan un dominio NZASD en el factor anti- σ . De los dos factores anti-σ estudiados con mayor profundidad, CarR y DdvA, solo DdvA (MXAN 7288) contiene el dominio NZASD (con un motivo CHCC; Figura 72). Los resultados de este trabajo han permitido determinar que DdvA adopta la topología clásica de los factores anti-σ, con una única hélice transmembranal que sitúa el dominio NZASD en el citoplasma y el dominio C-terminal en la parte periplásmica. Además, acorde con el papel establecido para los dominios NZASD en los estudios con los factores anti-o RsrA o ChrR, se ha demostrado que el dominio NZASD de DdvA es necesario y suficiente para la interacción con el factor σ-ECF DdvS y para su inactivación. CarR, sin embargo, carece del dominio NZASD y presenta una topología en la membrana con seis hélices transmembranales que sitúan tanto su extremo amino como carboxilo en el citoplasma. La ausencia de un dominio NZASD, la topología membranal atípica de CarR, y los resultados obtenidos en este trabajo indican un modo diferente de interacción con su factor σ-ECF, en el que ni la región N-terminal ni la Cterminal son suficientes por sí solas para que se produzca la interacción. Cabe esperar, por tanto, que los estudios futuros sobre la pareja CarQ-CarR lleven al descubrimiento de nuevos aspectos sobre los mecanismos que median la interacción específica entre un factor σ -ECF y su correspondiente anti- σ , y los que conducen a la inactivación del anti- σ y la liberación del σ -ECF.



Figura 73. Distribución en el genoma de *M. xanthus* de las parejas σ -ECF/anti- σ estudiadas en este trabajo. En rojo se muestra la pareja *ddvS/ddvA*, en azul *carQ/carR*, en negro las parejas σ -ECF/anti- σ con un dominio eSTK en el factor anti- σ y en verde la pareja de secuencia similar a *ddvS/ddvA*.

Por otro lado, seis posibles factores anti- σ de *M. xanthus* estudiados en este trabajo presentan una asociación novedosa de un dominio NZASD con un dominio Ser/Thr quinasa de tipo eucariótico (eSTK), seguido de una hélice transmembranal predicha y de una extensa región C-terminal que podría actuar como dominio sensor (Figura 38). Dichos dominios eSTKs conservan todos los residuos que se han establecido como críticos en otras STKs, con la excepción del que aparece en MXAN_7162, que carece de algunos de los residuos importantes (Figura 74, página siguiente). Es posible que la pareja constituida por este último y el posible σ -ECF defectuoso (MXAN_7161) haya degenerado hasta convertirse en una pareja no funcional (que, sin embargo, aún mantiene la capacidad de interaccionar a través del dominio NZASD).

Aunque en los últimos años se han identificado eSTKs en muchas especies bacterianas, la primera eSTK procariótica (Pkn1) se descubrió precisamente en *M. xanthus* (Muñoz-Dorado *et al.*, 1991). Posteriormente, el análisis de su genoma permitió la identificación de hasta 99 eSTKs, una expansión que parece haber ocurrido en el linaje de las mixobacterias y que se ha especulado que estaría relacionada con su complejo estilo de vida (Goldman *et al.*, 2006; Perez *et al.*, 2008). Al igual que las STKs eucarióticas, las eSTKs procarióticas participan en cascadas de regulación en las que a menudo interaccionan con otras eSTKs, que a su vez interaccionan con otras proteínas cuya actividad regulan. Para ello, debe producirse la activación inicial por autofosforilación de la eSTK en respuesta a la unión a efectores alostéricos, cambios en la localización subcelular, u otros factores (Dworkin, 2015; Pereira *et al.*, 2011).

PkA	1	GTGSFGRVMLVKHKESG
PknB	1	GFGGMSEVHLARDLRLH
MXAN 6009	1	PPLVDTGLLVSGARVGRYVVLGLLGEGGMGRVHAAYDPELD
MXAN 7082	1	SFAGVGPLAPGTRVGRYVVLGLIGEGGMGRVHVAYDPELD
MXAN 7162	1	DDDPPGATTPLTNSRPPSGKGPSLERGTAVGRYLVLERLASGGMGVVYSAYDPELD
MXAN 0068	1	LLOGVIGMGCMGVVYAADDLRLG
MXAN 1710	1	WALGAGVGTQAPSSAPTAPGEEANSQPLPLGARLSRYVVRECIGAGAMGIVYAADDPELG
MXAN 2738	1	IGRGAMGEVYAAHDPELD
consensus	1	
PkA	38	NHYAMKILDKOKVVKLKOIEHTLN <mark>E</mark> KRILOAVNEPELVKLEESEKDNSNLYMV
PknB	35	RDVAVKVLRADLARDPSFYLRFRFAONAAALNHPATVAVYDTGEAETPAGPLPYTV
MXAN 6009	42	RKVALKI, NI ARI GEDSLAGARORI EREARTMARI, SHPHVAQI HDVGEFOGOLFI.V
MXAN 7082	41	RKVALKLEVERFHEDSLALARORLERFARTMAKLSHPHTASLHDVGEYOGOLFLV
MXAN 7162	57	REVALKILE VAALGI – EAEOGRAHILHEAOAMARVSHPHVVPVVDVGTEGPOVEIT
MXAN 0068	24	REVALKILRPLIAGAODOGREELLREAOAMARISHPNVLPLELGTAEGHDELA
MXAN 1710	61	REVALEVING THE SOURCE AND THE REPART OF THE REPART. OF THE REPART OF THE REPART. OF THE REPART OF THE REPART. OF THE REPART OF THE REPART OF THE REPART. OF THE REPART OF THE REPART OF THE REPART OF
MXAN 2738	43	REVALET LEEDER REPAIRS AND A CONTRACT AND A CONTRAC
consensus	61	* * * * * *
consensus	01	···· · · · · · · · · · · · · · · · · ·
		Sitio catalítico
		*
PkA	91	MEYVAGGEMESHLERIGEESEPHAREYAAOTVLTEEYLHSLDLIYEDLKPENLLIDOO
PknB	92	MEYVDGVTLRDTVHTEGPMTPKRATEVTADACOALNESHONGTIHRDVKPANIMISAT
MXAN 6009	98	MELVEGGTLERBUA-EKPETERETLARETOA-ADGLAATHALGTIHEDEKPONULLTED
MXAN 7082	97	MEFLEGGTLRRWLT-EOPRPLREVLERFROAADGLAASHALGTVHRDFKPDNULLTKG
MXAN 7162	112	MELVDARTLRPWLK-DAPRTWROVLALFLDAGRGLAAAHAAGVVHGDFKPENLLVGOD
MXAN 0068	78	MEWVDGTTLADWLR-ERERPWRDVLDVFLAAGAGLAAAHRAGVVHRDFKPSNVLVGRD
MXAN 1710	114	MELVECTTLAEWMK-E-BRPWKEVLBVFLEAGKGLAAAHAAGLVHRDFKPANALIGKD
MXAN 2738	96	LELVEGASLAEWLE-A-PRPWODVVRVFVDAGRGLAAAHAAGLVHRDFKPGNVLVGKD
consensus	121	* * * * * * * * * *
00110011040		· · · · · · · · · · · · · · · · · · ·
		Their a manager a
		Union a magnesio
PKA	149	GYIQVIDEGFARRVKGRTWTLCGTPEYHAPEIIL
PknB	150	NAVKVMDFGFARAIADSGNSV <mark>T</mark> QTAAVIGTAQYISPEQAR
MXAN_6009	155	GQVRIT <mark>DFGL</mark> ANAAVGMGPPREDAAIDSCEAPL <mark>T</mark> VTGAFLGTPAYGAPEQLR
MXAN_7082	154	GLVRIT <mark>DFGL</mark> ANATLVPGTAAPGTVSPASL <mark>T</mark> VTGTLLGTLAYCAPEDLR
MXAN_7162	169	GRVRVT <mark>D</mark> F <mark>G</mark> LARNANPLDDVSPLAGGTPAYMAP <mark>E</mark> QME
MXAN 0068	135	GRVRVT <mark>DFG</mark> LARRGTGESPEAPAGVSEGASKASML <mark>T</mark> EWGQAAGTPAYMSP <mark>E</mark> QRA
MXAN_0068 MXAN 1710	135 170	GRVRVT <mark>DFGL</mark> ARRGTGESPEAPAGVSEGASKASML <mark>T</mark> EWGQAAGTPAYMSPEDRA GRVFVT DFG LARLLQQEEGASH-QEHIEAP-VTPTGRL <mark>T</mark> RTGQLLGTPAYIAPELVR
MXAN_0068 MXAN_1710 MXAN_2738	135 170 152	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPPOLO
MXAN_0068 MXAN_1710 MXAN_2738	135 170 152 181	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYMSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEOLQ
MXAN_0068 MXAN_1710 MXAN_2738 consensus	135 170 152 181	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYMSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEQLQ ***.*.
MXAN_0068 MXAN_1710 MXAN_2738 consensus	135 170 152 181	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYMSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEOLQ ***.*
MXAN_0068 MXAN_1710 MXAN_2738 consensus	135 170 152 181	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEDLQ ****.*
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB	135 170 152 181 183	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEOLQ ***.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLCULYEVLTGEPPFTGDSPVSVAYOHVEDPIPPSARHEGISADLD
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009	135 170 152 181 183 190	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEDLQ ***.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFFGDSPVSVAYQHVREDPIPPSARHEGLSADLD
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7022	135 170 152 181 183 190 207	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEOLQ ***.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK CE-HCDARGDOFFCVALYEALNGQRPFFCFTERELL-SOMDRAWDEEBACWDAWDE
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082	135 170 152 181 183 190 207 203	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEDLQ ***.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7162	135 170 152 181 183 190 207 203 206	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEDLQ ****** SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7162 MXAN_7168	135 170 152 181 183 190 207 203 206 189	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ ****** SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710	135 170 152 181 183 190 207 203 206 189 225	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEORA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ ****.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPFFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDOFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDOFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDOFSFCVALYEALNGGRPFAGATREELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDFFSCVALHEALHGARPFQGETLQEVV-LAAQQGRMSPPKREVKVPTRVR
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738	135 170 152 181 183 190 207 203 206 189 225 212	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYNAPEDLQ ****.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDEFSFCVALHEALHGRPFQGETLQEVV-LAAQQGRMSPPKREVKVPTRVR GH-GVDARSDQFSFCVALYEALHGVRPFEGRTLEALG-QAAREGRIRAPERESKIPARVR
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738 consensus	135 170 152 181 183 190 207 203 206 189 225 212 241	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPEDLQ ***.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTIFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDEFSFCVALHEALHGRPFQGETLQEVV-LAAQQGRMSPPKREVKVPTRVR GH-GVDARSDQFSFCVALYEALNGVRPFEGRTLEALG-QAAREGRIRAPERESKIPARVR ****.
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738 consensus	135 170 152 181 183 190 207 203 206 189 225 212 241	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ ***.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFFGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTIFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDEFSFCVALHEALHGERPFGLATTPHELL-TEVRAGKVRPAPRGTHVPFWLR GR-AVDARDDYSFCVALHEALHGERPFGLA
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_7162 MXAN_1710 MXAN_2738 consensus	135 170 152 181 183 190 207 203 206 189 225 212 241	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARRGTGESPEAPAGVEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEDLQ ****.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLA
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA	135 170 152 181 183 190 207 203 206 189 225 212 241 238	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARRGTGESPEAPAGVEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPEDLQ ****** SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTIFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDQFSFCVALYEALNGVPFEGRTLEALG-QAAREGRIRAPERESKIPARVR *.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.*.
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_7162 MXAN_2738 consensus PkA PknB	135 170 152 181 183 190 207 203 206 189 225 212 241 238 249	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ *****. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPERAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDQFSFCVALYEALNGVFFEGRTLEALG-QAAREGRIRAPERESKIPARVR *.*.*.*.*.*.*.*.*
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_7162 MXAN_6009	135 170 152 181 183 190 207 203 206 189 225 212 241 238 249 263	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082	135 170 152 181 183 1900 207 203 206 189 225 212 241 238 249 263 259	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ ****.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDQFSFCVALYEALNGVFFEGRTLEALG-QAAREGRIRAPERESKIPARVR ***.*.*.*.* DLLRNLLQVDLTKRFGNLKNGVN
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162	135 170 152 181 183 190 207 203 206 203 206 203 206 225 212 241 238 249 265	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYNAPEDLQ ****.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVALYEALNGQRPFEGKTREALL-EQVARKAVREPRAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDEFSFCVALHEALHGRPFQGETLQEVV-LAAQQGRMSPPKREVKVPTRVR GH-GVDARSDQFSFCVALYEALNGVRPFEGRTLEALG-QAAREGRIRAPERESKIPARVR **.*.*.*.*.*. DLLRNLLQVDLTKRFGNLKNGVN
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7082 MXAN_7082 MXAN_0068 PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7062 MXAN_7162 MXAN_0668	135 170 152 181 183 190 207 203 206 212 241 238 249 263 259 265 234	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARRGTGESPEAPAGVEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYMAPEDLQ ****.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR GR-AVDARGDQYSFCVALHEALHGERPFHATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GQ-RADARSDEFSFCVALHEALHGERPFHAA
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710	135 170 152 181 183 190 207 203 206 207 203 206 212 241 238 249 263 259 263 259 263 234 234	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ *****.*. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVYSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFSFCVALYEALNGQRPFEGKTREALL-EQVARKAVRPERAGVPAWLR AHAPADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPHLR GQ-RADARSDQFSFCVALYEALNGVFFEGRTLEALG-QAAREGRIRAPERESKIPARVR *.*.*.*.*.*.*.*.*.*. DLLRNLLQVDLTKRFGNLKNGVN
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7162 MXAN_7162 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_7162 MXAN_7170 MXAN_2738	1355 1700 1522 1811 1833 1900 207 2033 2066 2032 2041 2388 2499 2633 2599 2655 2344 2599 2655 2343 2599 2655 2343 2700	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ *****. SK-GYNKAVDWWALGVLIYEMAAGYPPFFADQPIQIY-EKIVSGKVRFPSHFSSDLK GD-SVDARSDVSLGCVLYEVLTGEPPFTGDSPVSVAYQHVREDPIPPSARHEGLSADLD GE-RGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVALYEALNGQRPFAGATREELL-ASMQRQEVRPEQRGVPSWLK GE-HGDARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGERPGHLAASAAPRTSSVPRHLR GQ-RADARSDQFAFCVTLFEALYGERPFAATTPHELL-TEVRAGKVRPAPRGTHVPPWLR GR-AVDARGDQYSFCVALHEALHGRAPFQGETLQEVV-LAAQQGRMSPPKREVKVPTRVR GH-GVDARSDQFSFCVALYEALNGVRPFEGRTLEALG-QAAREGRIRAPERESKIPARVR **.*****************************
MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738 consensus PkA PknB MXAN_6009 MXAN_7082 MXAN_7082 MXAN_7162 MXAN_7162 MXAN_0068 MXAN_1710 MXAN_2738	135 170 152 181 183 190 207 203 206 189 225 212 241 238 259 265 234 263 259 265 234 283 270	GRVRVTDFGLARRGTGESPEAPAGVSEGASKASMLTEWGQAAGTPAYNSPEDRA GRVFVTDFGLARLLQQEEGASH-QEHIEAP-VTPTGRLTRTGQLLGTPAYIAPELVR GRVRVTDFGLARPSHRGVSGHAAPASLDTVTAPASLVDSPLTHSGALLGTPAYIAPELQ

Figura 74. Alineamiento del dominio eSTK de PkA de *Mus musculus* **y PknB de** *M. tuberculosis* **con los dominios eSTK de los factores anti-σ estudiados de** *M. xanthus.* **Los motivos catalíticos son representados en cajas. Sombreados en azul se muestran los aminoácidos conservados. Los asteriscos rojo y verde indican el aspártico catalítico y la treonina fosforilada, respectivamente. En amarillo se muestra la arginina que precede al aspártico catalítico que se requiere para la fosforilación de la treonina del bucle de activación.**

Los posibles factores anti- σ con dominios eSTK de *M. xanthus* se encuentran acoplados a genes para posibles factores σ-ECF, con los que en este trabajo se ha observado que interaccionan físicamente a través de su dominio NZASD (a excepción de la pareja MXAN 0068/MXAN_0070, para la que no se ha podido observar interacción). Estos resultados apuntan, por tanto, a que se trata de verdaderas parejas σ-ECF/anti-σ (con la posible excepción de MXAN 7161/MXAN 7162). Por otro lado, el alto grado de identidad y similitud entre varias de estas parejas sugiere un origen evolutivo común por duplicaciones génicas, un fenómeno relativamente frecuente en M. xanthus. Que no se observe interacción cruzada entre las diferentes parejas analizadas indica que, a pesar del gran parecido entre ellas, han divergido lo suficiente para que las interacciones sean específicas de pareja. Es probable, por ello, que cada pareja se haya especializado en la detección de una determinada señal externa, para la activación de un subconjunto específico de genes de respuesta, y que no haya redundancia de función. De ser así, la generación de mutantes sencillos en los genes para los factores anti- σ o su sobre-expresión, o la sobre-expresión de los correspondientes factores σ-ECF, podría aportar información sobre la función que desempeña cada una de las parejas. Además, de demostrarse in vitro que los posibles factores anti-o presentan actividad eSTK, podría también analizarse el efecto, sobre la señalización y el fosfoproteoma, de eliminar exclusivamente el dominio eSTK.

Hasta la fecha, de los factores anti- σ con dominios eSTK, solo se ha estudiado con cierto grado de profundidad MXAN_1709 (Pkn8). Consistente con las predicciones bioinformáticas, Pkn8 es una proteína de membrana que muestra actividad eSTK y que, además, está implicada en el desarrollo multicelular y la esporulación a través de una compleja ruta en la que participa fosforilando a la eSTK citosólica Pkn14 (Figura 75) (Nariya & Inouye, 2005, 2006). Como dichos estudios se realizaron antes de nuestra observación de que podría funcionar como un posible factor anti- σ , no se ha contemplado el papel que podría tener en dicha cascada de señalización la interacción de parejas σ -ECF/anti- σ de *M. xanthus* analizadas, la pareja MXAN_1710/Pkn8 no depende del complejo CarD-CarG, por lo que es improbable que el papel de CarD-CarG sobre el desarrollo multicelular ocurra a través de una acción reguladora sobre Pkn8 (Nicolás *et al.*, 1994; Peñalver-Mellado *et al.*, 2006).



Figura 75. Regulación del desarrollo multicelular de *M. xanthus.* MrpC controla la expresión de *fruA* regulando la formación de cuerpos fructíferos y la esporulación. Este factor transcripcional, sufre una doble regulación. Por un lado, durante la fase vegetativa, la fosforilación por las eSTK Pkn14/8 disminuyen la afinidad de MrpC por el promotor de *fruA*, y por otro, ante señales de ausencia de nutrientes, MrpA/MrpB activan la transcripción de *mrpC*. Tomada de Pereira *et al.*, (2011).

Desentrañar la estrategia de regulación de estas novedosas parejas σ -ECF/antiσ presentes en *M. xanthus* podría ampliar el conocimiento sobre los mecanismos celulares de detección de señal y de respuesta para afrontar los cambios en el medio extracitoplásmico. En base al trabajo realizado con varias eSTKs membranales, fundamentalmente de M. tuberculosis, se ha propuesto un modelo de activación mediante dimerización, inducida por la unión del dominio extracitoplasmático a un ligando (Figura 76). Cabe especular un modo de activación similar para los factores NZASD/eSTK de M. xanthus, todos ellos con un extenso dominio C-terminal que debería estar situado en el espacio periplásmico, en disposición de percibir la presencia de determinadas señales, aún por determinar. Quizás, los cambios conformacionales que acompañan a la activación del eSTK por autofosforilación se trasmitan de algún modo al dominio NZASD, provocando la liberación del factor σ -ECF. Alternativamente, la autofosforilación podría provocar la interacción con otra proteína que, una vez fosforilada, sea la responsable de la liberación del σ-ECF; ello podría ocurrir, entre otros mecanismos posibles, por activación mediante fosforilación de una proteasa o de un factor anti-anti-σ. En este sentido, cabe señalar que PknD, una de las nueve eSTKs transmembranales codificadas en el genoma de M. tuberculosis, una vez activada en respuesta a cambios de osmolaridad fosforila al posible factor anti- σ Rv0516c/OprA, lo permite que el factor σ alternativo SigF se una a la RNAP y active la expresión del propio Rv0516c/OprA y de genes implicados en la estructura y metabolismo de la pared celular (Greenstein *et al.*, 2007; Hatzios *et al.*, 2013).



Figura 76. Modelos de activación para eSTKs. A) Dos o más monómeros pueden dimerizar para unir un solo ligando a través de su dominio extracelular. B) Una eSTK activada puede fosforilar una proteína diana o fosforilar a una quinasa soluble que participe en otra cascada. Tomada de Pereira *et al.*, (2011).

IV.4. Sistemas CRISPR-Cas en M. xanthus y su regulación

Los resultados de este trabajo y de otros previos del grupo conducen al siguiente modelo de regulación del sistema CRISPR4-Cas (Figura 77), uno de los tres sistemas CRISPR-Cas de M. xanthus. La percepción de una señal extracitoplásmica, de naturaleza desconocida, por el factor anti-o DdvA provoca su inactivación y la consiguiente liberación del factor σ-ECF DdvS. Una vez libre, DdvS se une al núcleo catalítico de la RNAP y junto con CarD-CarG activa mediante una acción directa la expresión de cuatro promotores situados en las inmediaciones de ddvS-ddvA (Figura 77). Más que a partir de un promotor en la región líder, la región CRISPR4 se expresa por la transcripción procedente de los genes cas situados aguas arriba, lo que genera un largo pre-crRNA que es procesado a través de Cas6 y el complejo Cmr a pequeños crRNAs maduros (Figura 77). Como la expresión de los genes cas está regulada por DdvS/DdvA y el complejo CarD-CarG, también lo está la expresión de la región CRISPR4 y, en consecuencia, la producción de los crRNAs. Ello proporciona un mecanismo muy eficaz para la expresión coordinada de los distintos componentes del sistema CRISPR4-Cas solo en presencia de la señal que conduce a su activación. Esta situación contrasta con la observada para varios otros sistemas CRISPR-Cas descritos, en los que los genes cas y la región CRISPR se expresan de forma constitutiva e independiente, a partir de promotores (aguas arriba de los genes cas y en la región líder) reconocidos por el factor σ mayoritario.



Figura 77. Modelo de regulación del sistema CRISPR4-Cas.

De qué manera DdvA detecta un cambio en las condiciones externas, así como la naturaleza de dicha señal, son algunas de las incógnitas sobre este complejo sistema de regulación que quedan por elucidar. Como se ha comentado con anterioridad, DdvA presenta una sola hélice transmembranal, a través de la cual podría interaccionar con otros monómeros de DdvA para formar un complejo multimérico. Consistente con el parecido entre la región citoplásmica de DdvA, que precede a esta hélice, y los dominios NZASD (Figura 78), en este trabajo se ha demostrado que dicho dominio es necesario y suficiente para interaccionar con DdvS e impedir su actividad. Orientado hacia el espacio periplásmico se dispone el extenso dominio C-terminal, de ~900 aminoácidos, que cabría esperar fuese el responsable de percibir la señal. Este dominio presenta tres módulos en tándem de repeticiones de tetratricopéptidos (TPR) (Figura 78), constituidas principalmente por aminoácidos hidrofóbicos, que suelen estar implicadas en las interacciones proteína-proteína (Blatch & Lässle, 1999) y en la formación de complejos multiproteicos con funciones celulares relevantes (Cerveny et al., 2013). En bacterias patógenas, por ejemplo, se ha visto que este tipo de repeticiones cumplen funciones asociadas a virulencia (Edqvist et al., 2006). El módulo TPR de DdvA, situado próximo a la hélice transmembranal, podría servir como plataforma de interacción con alguna otra proteína que participe en el proceso transducción de la señal que conduce a la liberación de DdvS.

MSSSPCDQLQSFADGDLPPMEAQAFGQHLADCEKCQVELTRLLQLDQLGRGYIERHGPVI <mark>IPWHALPRNR</mark>WMAGGFALASMVAVAFILLGTLRP<mark>Q</mark>APRALPELWAQPQRTLEARVTYLAA DRYRRPPQVMMGGAAARSGPNRSLALLSELEKRGDLHQLAVAYLATGVPEPSSAF MRSD<mark>LRWQ</mark>SADVLCDLGVAHYVASKPLDAARATEELREALRLFDTVLAM<mark>QPGH</mark>VQALWNF LVYRDLGLPLSAMKDLTEFEHRETDEGWRSEARDRRARLSSTLRRKERWLAADQTGADL INRGAQELARALTFVDVPLLRRDFYHAVRARTSSTDVLALLPLAERLDASVGSGTVLADY VHQVAARDFSRRAPLAEQYARLISGRIPESEQDALLQRFLTSDETDLALGALAHVMQRLP AYASELVRRTQHDEDPWFRVLGLQAQAMLERQQEHYKEALAPLEQALDICRRERLVYRCI FIENDLSHVKSWLFRVNAAAQHARDGLALARPNQWDLEGVMLQALGNVARQAADVTLGRA YYGEALLMAEGDKWSTRNIHQNLAHLAIWALELDEARASLDRAMDTGLPLTQHGVAALVD VARTRRSPRDALMVEQALAREPGNTPGQRAYAKFLHGRILVEVDPARGRMLLDEAIRQAE ALPLDDVSAAHARAYSYTSLIFADADTGDFIAALARFGAELGFETPARCVLGLTADTERS LLVARGAQGQLLSAYVPLRSSRFEAASMEGAVPPEMLAALQACTLVDVLARPPLQGRSGL LPPGIAWRYRTRAAAPPPPAGPGTHLVVNEVRYSEERNEVPLQWLPRTAPGAEARFLRDL AATPTQVLEAISTATEIDLATHGKVDPDSNFAYLLLAPGADGRDTLFEDSIRASQLTGAP ${\tt LVVL}$ ASLAVAVRDERLQWLSAGGDSE<mark>WVNAVLVFE</mark>

Figura 78. Secuencia de DdvA con sus dominios funcionales. En rojo se resalta el dominio NZASD. En marrón se muestra la región transmembranal. En magenta se indica el dominio TPR y en verde el dominio CHAT.

Además del módulo TPR, la región extracitoplasmática de DdvA presenta, en su extremo C-terminal (Figura 78), un dominio CHAT (Caspase HetF Associated with Tprs). Este tipo de dominios, muy extendidos en bacterias, están relacionados con unas proteasas eucarióticas denominadas separinas que participan en la segregación cromosómica, y que las células eucarióticas podrían haber adquirido durante las primeras fases de su evolución (Aravind & Koonin, 2002). En procariotas, los dominios CHAT están ejemplificados por la proteína HetF de cianobacterias, que muestra una actividad proteasa probablemente implicada en diversas rutas de señalización, incluida la formación de los heterocistos (Aravind & Koonin, 2002; Klemenčič *et al.*, 2015; Risser & Callahan, 2008). La presencia de un dominio CHAT en la extensa cola extracitoplásmica de DdvA podría estar relacionada con un mecanismo de inactivación de DdvA por la supuesta actividad proteasa de dicho dominio, que podría activarse por la presencia de alguna señal externa por determinar. Una posibilidad muy atractiva que está siendo explorada es que el dominio C-terminal de DdvA esté directamente implicado en el reconocimiento las partículas virales. También es posible que el sistema reconozca la perturbación en la membrana que podría acompañar a la entrada del material genético exógeno.

La búsqueda de parejas similares a DdvS/DdvA en el genoma de *M. xanthus* permitió identificar la pareja MXAN_5506/MXAN_5507 (Figura 73), cuyo factor anti- σ comparte con DdvA un alto grado de identidad y similitud, así como los dominios NZASD, TPR y CHAT descritos anteriormente. Todas estas similitudes sugieren, como en el caso de los factores anti- σ asociados a eSTKs, un origen evolutivo por duplicación
génica. Por otro lado, aunque ambas parejas muestran una dependencia del complejo CarD-CarG, no se observa interacción cruzada entre ellas. Así pues, los fenómenos de duplicación y posterior divergencia evolutiva parecen haber jugado un papel importante en el desarrollo y especialización de una compleja red de percepción de estímulos externos basados en el uso de parejas σ -ECF/anti- σ que, a buen seguro, desempeñan una función clave en la adaptación de las células de *M. xanthus* a los cambios del medio externo.

El sistema CRISPR4-Cas comparte con el sistema CRISPR2/3-Cas, otro de los sistemas CRISPR-Cas de *M. xanthus*, una similitud notable en las repeticiones que componen la región CRISPR y en la secuencia de la región líder (Figura 79).

a)										
CRISPR2/3-Cas		GT	GCTCAAC	CGCCTT	TCGGCA	TCACGG	CGAGCG	GGAC (36	βpb)	
CRISPR4-Cas		GT(**;	GCTCAA(* * * * * * *	CGCCTT * * * * * *	CCGGCA ****	TCACGG(* * * * * * * *	CGAGCG(* * * * * * *	G <mark>C</mark> AC (36	s pb)	
b)										
,	CRISPR3 Líder CRISPR4 Líder consenso	1 1 2 1	AGGCTTGC	CGGTGGA	GCAGGGG	CCTTGCAG	GAACGTA	GGACCGGG	GGCTGTG	GACGG GGTGC
	CRISPR3 Líder CRISPR4 Líder consenso	8 61 61	GACAATCA GACATTCA **** ***	AGCACCT T <mark>GCACCT</mark> *****	'CGCCGTG 'CGCCGTG ******	GCGGAGGC GTGGAGGC * ***** -35	TCGCGAT CTGTGCC * *	GG <mark>C</mark> GAAAG GGTGAAAT ** ****	T-GGCGGA CGTTGAAA 	ATTCCC ATCCTC ** * *
	CRISPR3 Líder CRISPR4 Líder consenso	67 5 121 0 121	ICGTTTCA CGAGTTCA . ****	TGGGGGT TGGGGTT *****	'GAGGCTC 'GGGGCTC *. ****	TTTGACAG TTTGACAG * <mark>*****</mark> *	GTGAATA GTGAATA ******	GGTGCATG GGTGCATG * * * * * * * *	ATCGATG ATTTATG ** ***	AAAAGC AAAAAC
	CRISPR3 Líder CRISPR4 Líder consenso	127 181 181	AGTGGTT AGTGATT ****.**	CATGGGG CATGGGG ******	TT <mark>G</mark> GCGA TTT <mark>TG</mark> GAA ** * *	TTCTGGGT TCTTGGGT * *****	AGGGCCT(AGGGCCT(* * * * * * * *	GGGATGGC GGGA <mark>GG</mark> CG * * * * *	CAACCTG(CAACCTG(******	GGGGAA GGGCCG ***
	CRISPR3 Líder CRISPR4 Líder consenso	187 241 241	AGGGGGGAG AGGGGGGGA	GTGCT GGTGCTT *	TGAAAGG GAAAAGA . ****.	GGTTTGTA CGGGA <mark>GTA</mark> * ***	AG <mark>G</mark> TGCC(AGTTGCC(** ****	GGAA <mark>T</mark> TGC GGAAC <mark>TGC</mark> **** ***	TGGGGGA TGGGAGA ****	FTCAGG FTCAGA
	CRISPR3 Líder CRISPR4 Líder consenso	245 301 301	GGCT GGCT * * * *							

Figura 79. Similitudes de los sistemas CRISPR4-Cas y CRISPR2/3-Cas. A) Alineamiento de las repeticiones de los dos sistemas. B) Alineamiento de las secuencias líder de las regiones CRISPR4 y CRISPR2/3.

Aunque ambos sistemas CRISPR-Cas presentan un posible promotor dependiente del factor σ mayoritario en la región líder, en este trabajo se ha demostrado que, al menos el del sistema CRISPR4-Cas presenta una actividad despreciable. Por otro lado, la actividad del promotor (P_{dev}) que dirige la expresión de los genes *cas* del sistema CRISPR2/3-Cas también está regulada, aunque por la entrada en la fase de

desarrollo multicelular (Viswanathan *et al.*, 2007). Además, datos recientes (no mostrados) derivados de nuestro trabajo han puesto de manifiesto que la expresión del sistema CRISPR1, cuya región líder no comparte similitud con la de los otros dos sistemas, tampoco es constitutiva. Así pues, la expresión regulada de los tres sistemas CRISPR-Cas de *M. xanthus* permite a las células evitar el gasto energético que supondría la continua transcripción de las largas regiones CRISPR y el peligro potencial para la integridad del propio genoma derivado de la expresión constitutiva de las nucleasas Cas (Patterson *et al.*, 2017).

Otra característica común de los sistemas CRISPR4-Cas y CRISPR2/3-Cas es la presencia de un gen *cas6* en ambos y la similitud de secuencia entre sus productos génicos. Acorde con esta observación y con el conocido papel de las proteínas Cas6 en la generación de los crRNAs maduros (Charpentier, *et al.*, 2015; Hochstrasser & Doudna, 2015; Sefcikova *et al.*, 2017), la eliminación del gen *cas6* del sistema CRISPR4-Cas provocó la desaparición de los correspondientes crRNAs; y su producción se vio rescatada por la expresión de cualquiera de los dos genes *cas6* en un sitio heterólogo. Estos resultados sugieren que, en efecto, las dos proteínas funcionan como endonucleasas Cas6, y que son funcionalmente intercambiables y capaces de actuar *en trans*, como módulos independientes.

Otro aspecto a destacar es la presencia en el primer gen cas "cmr1-like" del sistema CRISPR4-Cas de un fragmento de 79 nucleótidos que coincide (salvo en 14 posiciones) con una secuencia situada entre las regiones CRISPR3 y CRISPR2. Todas estas similitudes sugieren que ambos sistemas comparten un origen evolutivo y que podría existir una intercomunicación entre ellos, siempre que existan condiciones en las que los dos sistemas estén activos a la vez. Sin embargo, mientras que el sistema CRISPR4-Cas está regulado por DdvS/DdvA y CarD-CarG, el sistema CRISPR2/3-Cas parece estar regulado por el desarrollo y por los genes dev (devT, devR y devS), que controlan negativamente su propio promotor P_{dev} (Viswanathan *et al.*, 2007). En este trabajo se ha analizado, por primera vez, la producción de crRNAs maduros a partir del sistema CRISPR2/3-Cas y hubo que eliminar el regulador negativo DevT para observar una señal correspondiente a los crRNAs. Estos resultados indican que, per se, el proceso de desarrollo multicelular no provoca una activación significativa del sistema, o que lo hace de manera muy transitoria, en cualquier caso, insuficiente para la detección de los crRNAs. Además, el tamaño de los crRNAs observados en el fondo genético $\Delta devT$, superior al habitual de ~40 nucleótidos, sugiere un procesamiento incompleto, quizás debido a que la falta de DevT genera un complejo efector defectuoso. DevT presenta similitud con Cas8a, y DevR y DevS con Cas7a y Cas5a, respectivamente, por

lo que se ha sugerido que podrían funcionar juntos como parte de un complejo Cascade (Rajagopalan *et al.*, 2015).

Por otro lado, durante el desarrollo multicelular no se observó la producción de los crRNAs del sistema CRISPR4-Cas ni en el fondo silvestre ni en el *\devT*, lo que indica que la falta de nutrientes no es la señal que dispara la inactivación de DdvA y la expresión de dicho sistema. Tampoco se observaron los crRNAs del sistema CRISPR2/3-Cas en la estirpe $\Delta ddvA$. A la vista de estos resultados y, pese a las similitudes subrayadas, no cabe argumentar a favor de que exista una comunicación cruzada entre ambos sistemas. No obstante, no se puede descartar que ambos sistemas estén interconectados a través de un posible papel común en la defensa contra material genético exógeno. Así pues, los sistemas CRISPR-Cas de tipo III-B coexisten con los sistemas de tipo I en muchas bacterias, incluida M. xanthus, y estudios recientes con Marinomonas mediterranea han puesto de manifiesto que un sistema de los de tipo III-B, generalmente más promiscuos y tolerantes a mutaciones en el protoespaciador, puede usar no solo sus crRNAs sino también los de un sistema de tipo I (I-F). De este modo, los sistemas de tipo III-B pueden servir de refuerzo, como una línea de defensa adicional, para contrarrestar el posible escape al ataque por los sistemas de tipo I (Silas et al., 2017). Aunque, de momento, no se hayan encontrado condiciones en las que se expresen tanto el sistema CRISPR2/3-Cas (de tipo I-A) como el CRISPR4-Cas (de tipo III-B) también podría existir dicha intercomunicación entre sistemas en M. xanthus. A favor de la posible cooperación entre ambos sistemas está la observación de que la mayoría de aislados naturales de *M. xanthus* presentan tanto el sistema de tipo III-B como el de tipo I-A (Wielgoss et al., 2016).

Espaciador CRISPR2/3	5'-TGACAGCGCCCCATCCGCTTGCCGTGGTGCGTCAGCA-3'
Fago Mx8 <i>int</i>	5'-TGACAGCGCCCCATCCGCTTGCCGTGGTGCGTCAGCA-3'

Figura 80. Homología de un espaciador de la región CRISPR2/3 con un fragmento del gen de la integrasa del fago Mx8.

La búsqueda de secuencias similares a cada uno de los espaciadores de los sistemas CRISPR-Cas de *M. xanthus* permitió observar que un espaciador del sistema CRISPR2/3-Cas coincide con la secuencia del gen de la integrasa (*intP*) del bacteriófago Mx8 (Figura 80), por lo que se ha especulado sobre su posible papel en la protección de las células frente a la lisogenia por Mx8 durante la fase de desarrollo (Viswanathan *et al.*, 2007). Por lo demás, no se ha observado homología de ninguno de los espaciadores con ninguna secuencia depositada en las bases de datos, si bien hasta la fecha solo está disponible la secuencia del genoma del fago Mx8. No obstante, se han

aislado otros bacteriófagos de *M. xanthus*, como Mx1, Mx4 o Mx9, cuyas secuencias (en proceso de obtención) podrían desvelar homología con alguno de los espaciadores de cualquiera de los tres sistemas CRISPR-Cas. Mientras tanto, cabe esperar que los estudios en progreso sobre la capacidad infectiva de dichos mixofagos en distintos fondos genéticos (silvestre, con deleción de *ddvA* y/o la región CRISPR4, entre otros) aporten información sobre el papel de dichos sistemas y la posible relación entre ellos en la defensa frente a los mixofagos conocidos.

IV.5 ¿Otros genes de defensa en el locus CRISPR4-Cas?

Las bacterias mantienen una lucha constante contra los bacteriófagos que las infectan. Ambos organismos están en un continuo estado de coevolución desarrollando numerosas estrategias de supervivencia y estableciendo un equilibrio dinámico que permite su coexistencia. Entre estas estrategias se encuentran las modificaciones en el sistema de acceso al hospedador, los sistema de restricción-modificación (R-M), las modificaciones anti-CRISPR-Cas, o los mecanismos de infección abortiva (Samson *et al.*, 2013). Es común encontrar los genes encargados de llevar a cabo dicha defensa agrupados en determinadas regiones del genoma, formando las llamadas islas de defensa (Koonin *et al.*, 2017; Makarova *et al.*, 2011). Por ello, genes de función desconocida que se encuentren próximos a un sistema CRISPR-Cas son serios candidatos a estar también implicados en la inmunidad bacteriana (Koonin, 2018). Así, se han identificado hasta 10 nuevos sistemas de defensa, con dominios nucleasas y helicasas. Nueve de estos sistemas están implicados en la defensa frente a fagos mientras que uno parece participar en la defensa frente a plásmidos (Doron *et al.*, 2018).

El regulón activado por DdvS y dependiente de CarD-CarG incluye dos grupos de genes de función desconocida situados en la vecindad del sistema CRISPR4-Cas. Por un lado, se encuentran los genes *MXAN_7291* y *MXAN_7290*, localizados inmediatamente aguas arriba de *ddvS/ddvA* (Figura 81) y, por otro, los genes *MXAN_7287-MXAN_7284*, localizados aguas abajo (Figura 81). Al menos en las condiciones estándar de cultivo de laboratorio, la deleción conjunta de los genes *MXAN_7287* a *MXAN_7284* no tuvo ningún efecto sobre la producción de los crRNAs de la región CRISPR4 (Bernal-Bernal *et al.*, 2018), lo que sugiere que cualquiera que sea su función no parece estar directamente ligada a dicho proceso.



Figura 81. Genes de función hipotética asociados al sistema CRISPR4-Cas. Los posibles operones formados por los genes de función hipotética (flechas gruesas blancas) aparecen recuadrados con línea discontinua.

El análisis bioinformático indica la presencia de dominios conservados en la mayoría de los productos de los dos grupos de genes. Así, las proteínas MXAN 7291 y MXAN 7286 presentan cierta similitud con los dominios nucleotidiltransferasa del tipo encontrado en las sintasas de 2,5 oligoadenilato (OAS) eucarióticas, que producen oligonucleótidos lineales de adenina con enlaces 2', 5' (en vez del 3', 5' habitual) entre al menos tres nucleótidos (en MXAN 7286 algunos de los aminoácidos implicados en dicha función no están conservados). Curiosamente, las OAs están relacionadas con la inmunidad innata en vertebrados frente a virus de RNA, al participar en la fase inicial de la infección activando a la RNAsa L, que degrada el RNA viral y celular (Choi et al., 2015). En este sentido, cabe resaltar el hallazgo muy reciente de que, en sistemas CRISPR-Cas de tipo III-A, se utiliza una forma circular de oligonucleótidos de adenina (cOA) como un segundo mensajero para la activación de un sistema de degradación no específica de RNA (Kazlauskiene et al., 2017; Niewoehner et al., 2017; Rouillon et al., 2018). En este caso, el reconocimiento del mRNA viral a través del complejo efector estimula la actividad sintasa de cOA de Cas10 y, una vez producido, el cOA se asocia con la proteína Cas Csm6 y activa su capacidad de degradar RNA de manera inespecífica (Figura 82). De este modo, si falla la última línea de defensa, y se produce RNA viral en abundancia, la célula puede cometer "suicidio" (por degradación de su propio RNA) e impedir la producción de nuevas partículas virales (Amitai & Sorek, 2017). El sistema CRISPR4-Cas no solo presenta Cas10 sino también las posibles nucleotidiltransferasas, lo que abre la interesante posibilidad de que, además de la señalización por oligonucleótidos de adenina cíclicos, en M. xanthus se produzca también una señalización por su forma lineal.

Por otro lado, la proteína MXAN_7290 presenta un dominio CHAT, similar al observado en el extremo C-terminal de DdvA, que como se ha comentado anteriormente está relacionado con una actividad proteasa. MXAN_7285 presenta un dominio típico de la familia de las enzimas activadoras de proteínas similares a ubiquitina y miembros de la familia bacteriana ThiF/MoeB/HesA, que catalizan la activación de pequeñas proteínas a través de la formación de un enlace tioéster de alta energía dependiente de ATP. En MXAN_7284 se predice un dominio MPN (Mpr1/Pad1 N-terminal), también

llamado JAB, que contiene un motivo propio de una metaloenzima con actividad isopeptidasa; su función es liberar la ubiquitina de proteínas ubiquitinadas preparadas para la degradación, participando en las rutas de recambio proteico en eucariotas. Estos dominios se han encontrado en arqueas, bacterias y fagos, donde se cree que presentan una función similar (Melissa & Echalier, 2014). Por otro lado, otra familia de proteínas bacterianas presenta los dominios JAB similares a RadC, que se cree que funcionan como nucleasas. En arqueas halofílicas, estos dominios JAB suelen estar asociados a nucleotidiltransferasas. Por último, la proteína MXAN_7287 no presenta motivos o dominios funcionales conocidos.



Figura 82. Producción de cOA en los sistemas CRISPR-Cas de tipo III. 1) El reconocimiento del mRNA del fago por el complejo efector-crRNA activa la degradación de RNA. 2) La proteína Cas10 promueve el corte de DNA. 3) Cas10 a través de sus dominios Palm sintetiza cOA. 4) La difusión de cOA activa a Csm6, que promueve la rotura inespecífica de RNA. Tomado de Amitai & Sorek, (2017).

Así, la presencia de dominios implicados en la defensa frente a RNA viral (MXAN_7291 o MXAN_7286), en metabolismo de proteínas o de ácidos nucleicos (MXAN_7284) o en la activación de pequeñas proteínas (MXAN_7285), en los productos de los genes en la vecindad del locus CRISPR4-Cas, apunta a que podrían ejercer una acción defensiva complementaria a la proporcionada por dicho sistema. Cabe esperar que los estudios en marcha con distintos mixofagos y estirpes de *M. xanthus* portadoras de deleciones en los genes en cuestión permitan confirmar o descartar su posible papel en la defensa frente al ataque viral.

V. CONCLUSIONES

Los resultados derivados de este trabajo permiten establecer las siguientes conclusiones:

- Mientras que CarDNt es esencial para la función de CarD, CarDCt es parcialmente dispensable. Su presencia, no obstante, potencia la actividad de CarD permitiendo que las células de *M. xanthus* respondan de manera óptima al estrés luminoso y al estrés, desconocido, que provoca la inactivación del factor anti-σ DdvA.
- La estructura determinada por RMN para el subdominio N-terminal de CarDNt, CarD₁₋₇₂, que interacciona con el dominio β1 de βRNAP, es muy similar a la observada para CdnL. Dicha estructura consiste en un sándwich β retorcido con cinco láminas β antiparalelas, β5-β1-β2-β3-β4, y una hélice 3₁₀ entre β4-β5, similar a una estructura de tipo Tudor.
- La superficie de interacción con βRNAP está conservada en CarD y CdnL. Sin embargo, las mutaciones que eliminan la interacción con la RNAP disminuyen pero no eliminan completamente la actividad de CarD en *M. xanthus*, a diferencia de lo observado para las mutaciones equivalentes en CdnL.
- 4. El dominio C-terminal de CarDNt presenta una región expuesta al solvente de aminoácidos no polares y básicos que no están implicados en la interacción con CarG. La mutación triple de los residuos básicos K93, R95, y R97 anula la función de CarD, al igual que sucede con la proteína CdnL.
- 5. El factor anti-σ DdvA presenta una sola hélice transmembranal precedida de un dominio NZASD que es suficiente y necesario para secuestrar en la membrana al factor σ-ECF DdvS e inactivarlo. En cambio, el factor anti-σ CarR presenta una topología en la membrana con seis hélices transmembranales que orientan las regiones N- y C-terminal hacia el citoplasma pero, por sí solas, ninguna de las dos regiones interaccionan con CarQ.
- Los genes MXAN_5506/MXAN_5507 constituyen una pareja σ-ECF/anti-σ que presenta homología con la pareja DdvS/DdvA y cuya actividad depende del complejo CarD-CarG.
- Los miembros de cinco de las seis posibles parejas σ-ECF/anti-σ cuyo factor anti-σ presenta un dominio NZASD asociado a un dominio eSTK interaccionan entre sí, lo que apoya la hipótesis de que se trata de verdaderas parejas σ-ECF/anti-σ. Pese al parecido, cada anti-σ reconoce específicamente a su factor σ-ECF.

- 8. La colocalización de CarD y CarG *in vivo* refuerza la idea de que ambas proteínas funcionan siempre juntas y de que CarG se localiza en el DNA a través de su interacción con CarD.
- CarD se une a las regiones promotoras P_{QRS} o P_{ddvSA} solo cuando están activas y en presencia de CarG, que podría actuar como un puente entre CarD y la maquinaria basal de transcripción, y estabilizar de ese modo la unión de CarD al DNA.
- 10. La expresión del sistema CRISPR4-Cas de *M. xanthus* está regulada por la pareja σ-ECF/anti-σ DdvS/DdvA y por el complejo regulador global CarD-CarG.
- 11. La acción de DdvS y CarD-CarG sobre los cuatro promotores dependientes de DdvS situados en el locus CRISPR4-Cas es una acción directa.
- 12. La región CRISPR4 no se expresa a partir del promotor dependiente de σ^A identificado en la región líder sino a partir de la transcripción procedente de los genes *cas* situados aguas arriba, lo que genera un largo pre-crRNA. Como la expresión de los genes *cas* está regulada por DdvS/DdvA y el complejo CarD-CarG, también lo está la expresión de la región CRISPR4 y, en consecuencia, la producción de los crRNAs.
- 13. La proteína homóloga a Cas6 del sistema CRISPR4 es necesaria para la producción de los crRNAs maduros y puede ser suplida en *trans* por la proteína Cas6 del sistema CRISPR2-3-Cas de *M. xanthus*.
- 14. El desarrollo multicelular no es la señal que activa la expresión del sistema CRISPR4-Cas ni promueve la producción de los correspondientes crRNAs.

VI. BIBLIOGRAFÍA

Abellán-Martínez, M. (2006). Expresión del locus ddvA de Myxococcus xanthus y su

dependencia de los genes carD, sud-2 y sud-5. Tesis Doctoral, Universidad de Murcia.

Abellón-Ruiz, J. (2012). Identificación de factores sigma-ECF en *Myxococcus xanthus* y estudio de su dependencia del complejo CarD-CarG.

Abellón-Ruiz, J., Bernal-Bernal, D., Abellán, M., Fontes, M., Padmanabhan, S., Murillo, F. J., & Elías-Arnanz, M. (2014). The CarD/CarG regulatory complex is required for the action of several members of the large set of *Myxococcus xanthus* extracytoplasmic function σ factors. *Environmental Microbiology*, *16*, 2475-2490.

Åberg, A., Fernández-Vázquez, J., Cabrer-Panes, J. D., Sánchez, A., & Balsalobre, C. (2009). Similar and divergent effects of ppGpp and DksA deficiencies on transcription in *Escherichia coli*. *Journal of Bacteriology*, *191*, 3226-3236.

Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., ... Zhang, F. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*, *353*.

Alhadid, Y., Chung, S. Y., Lerner, E., Taatjes, D. J., Borukhov, S., & Weiss, S. (2017). Studying transcription initiation by RNA polymerase with diffusion-based single-molecule fluorescence. *Protein Science*, *26*, 1278-1290.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, *25*, 3389-3402.

Amitai, G. & Sorek, R. (2017). Intracelullar signaling in CRISPR-Cas defense. *Science*, 357, 550-551.

Ansari, A. Z., Bradner, J. E., & O'halloran, T. V. (1995). DNA-bend modulation in a repressor-to-activator switching mechanism. *Nature*, *374*, 370-375.

Anthony, J. R., Newman, J. D., & Donohue, T. J. (2004). Interactions between the *Rhodobacter sphaeroides* ECF sigma factor, σE, and its anti-sigma factor, ChrR. *Journal of Molecular Biology*, *341*, 345-360.

Apelian, D., & Inouye, S. (1990). Development-specific *σ*-factor essential for late-stage differentiation of *Myxococcus xanthus*. *Genes and Development*, *4*, 1396-1403.

Apelian, D., & Inouye, S. (1993). A new putative sigma factor of *Myxococcus xanthus*. *Journal of Bacteriology*, *175*, 3335-3342.

Aravind, L., & Koonin, E. V. (2002). Classification of the caspase-hemoglobinase fold:

Detection of new families and implications for the origin of the eukaryotic separins. *Proteins: Structure, Function and Genetics*, *46*, 355-367

Arnold, J. W., & Shimkets, L. J. (1988). Cell surface properties correlated with cohesion in *Myxococcus xanthus*. *Journal of Bacteriology*, *170*, 5771-5777.

Arslan, Z., Hermanns, V., Wurm, R., Wagner, R., & Pul, Ü. (2014). Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system. *Nucleic Acids Research*, *42*, 7884-7893.

Bae, **B.**, **Chen**, **J.**, **Davis**, **E.**, **Leon**, **K.**, **Darst**, **S. A.**, **& Campbell**, **E. A.** (2015). CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *eLife*, *4*, 08505.

Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., & Darst, S. A. (2015). Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *eLife*, *4*, 08504.

Barne, K. A., Bown, J. A., Busby, S. J. W., & Minchin, S. D. (1997). Region 2.5 of the *Escherichia coli* RNA polymerase σ70 subunit is responsible for the recognition of the «extended -10» motif at promoters. *EMBO Journal*, *16*, 4034-4040.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., ... Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*,1709-1712.

Barrangou, R., & Horvath, P. (2017). A decade of discovery: CRISPR functions and applications. *Nature Microbiology*, *2*, 1-9.

Basu, R. S., Warner, B. A., Molodtsov, V., Pupov, D., Esyunina, D., Fernández-Tornero, C., ... Murakami, K. S. (2014). Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. *The Journal of Biological Chemistry*, *289*, 24549-24559.

Berleman, J. E., & Kirby, J. R. (2009). Deciphering the hunting strategy of a bacterial wolfpack. *FEMS Microbiology Reviews*, *33*, 942-957.

Bernal-Bernal, D., Abellón-Ruiz, J., Iniesta, A. A., Pajares-Martínez, E., Bastida-Martínez, E., Fontes, M., ... Elías-Arnanz, M. (2018). Multifactorial control of the expression of a CRISPR-Cas system by an extracytoplasmic function σ /anti- σ pair and a global regulatory complex. *Nucleic Acids Research*, *46*, 6726-6745. Bernal-Bernal, D., Gallego-García, A., García-Martínez, G., García-Heras, F., Jiménez, M. A., Padmanabhan, S., & Elías-Arnanz, M. (2015). Structure-function dissection of *Myxococcus xanthus* CarD N-terminal domain, a defining member of the CarD_CdnL_TRCF family of RNA polymerase interacting proteins. *PloS One*, *10*, e0121322.

Bhaya, D., Davison, M., & Barrangou, R. (2011). CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. *Annual Review of Genetics*, *45*, 273-297.

Biran, D., & Kroos, L. (1997). In vitro transcription of *Myxococcus xanthus* genes with RNA polymerase containing σ A, the major sigma factor in growing cells. *Mol Microbiol*, *25*, 463-472.

Blatch, G. L., & Lässle, M. (1999). The tetratricopeptide repeat: A structural motif mediating protein-protein interactions. *BioEssays*, *21*, 932-939.

Bolotin, A., Quinquis, B., Sorokin, A., & Dusko Ehrlich, S. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*, 2551-2561.

Bortoluzzi, A., Muskett, F. W., Waters, L. C., Addis, P. W., Rieck, B., Munder, T., ... O'Hare, H. M. (2013). *Mycobacterium tuberculosis* RNA polymerase-binding protein A (RbpA) and its interactions with sigma factors. *Journal of Biological Chemistry*, *288*, 14438-14450.

Borukhov, S., & Nudler, E. (2003). RNA polymerase holoenzyme: Structure, function and biological implications. *Current Opinion in Microbiology*, *6*, 93-100.

Bowden, M. G., & Kaplan, H. B. (1998). The *Myxococcus xanthus* lipopolysaccharide O-antigen is required for social motility and multicellular development. *Molecular Microbiology*, *30*, 275-284.

Bown, J. A., Owens, J. T., Meares, C. F., Fujita, N., Ishihama, A., Busby, S. J. W., & Minchin, S. D. (1999). Organization of open complexes at *Escherichia coli* promoters: Location of promoter DNA sites close to region 2.5 of the ??70 subunit of RNA polymerase. *Journal of Biological Chemistry*, 274, 2263-2270.

Braun, V., & Mahren, S. (2005). Transmembrane transcriptional control (surface signalling) of the *Escherichia coli* Fec type. *FEMS Microbiology Reviews*, *29*, 673-684.

Bretscher, A. P., & Kaiser, D. (1978). Nutrition of *Myxococcus xanthus*, a fruiting myxobacterium. *Journal of Bacteriology*, *133*, 763-768.

Brooks, B. E., & Buchanan, S. K. (2008). Signaling mechanisms for activation of extracytoplasmic function (ECF) σ factors. *Biochimica et Biophysica Acta - Biomembranes*, *1778*, 1930-1945.

Brown, N. L., Stoyanov, J. V, Kidd, S. P., & Hobman, J. L. (2003). The MerR family of transcriptional regulators. *FEMS Microbiology Reviews*, 27, 145-163.

Browning, D. F., & Busby, S. J. W. (2004). The regulation of bacterial transcription initiation. *Nature Reviews Microbiology*, *2*, 57-65.

Browning, D. F., & Busby, S. J. W. (2016). Local and global regulation of transcription initiation in bacteria. *Nature Reviews Microbiology*, *14*, 638-650.

Browning, D. F., Grainger, D. C., & Busby, S. J. W. (2010). Effects of nucleoidassociated proteins on bacterial chromosome structure and gene expression. *Current Opinion in Microbiology, 13*, 773-780.

Browning, D. F., Whitworth, D. E., & Hodgson, D. A. (2003). Light-induced carotenogenesis in *Myxococcus xanthus:* Functional characterization of the ECF σ factor CarQ and anti- σ factor CarR. *Molecular Microbiology*, *48*, 237-251.

Bulyha, I., Hot, E., Huntley, S., & Søgaard-Andersen, L. (2011). GTPases in bacterial cell polarity and signalling. *Current Opinion in Microbiology*, *14*, 726-733.

Burnete, W. N. (1981)."Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate--polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal Biochem, 112,* 195-203.

Bustamante, V. H., Santana, F. J., Calva, E., & Puente, J. L. (2001). Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression. *Molecular Microbiology*, *39*, 664-678.

Buttner, D., & Bonas, U. (2002). Getting across--bacterial type III effector proteins on their way to the plant cell. *EMBO J*, *21*, 5313-5322.

Callaci, S., & Heyduk, T. (1998). Conformation and DNA binding properties of a single-

stranded DNA binding region of σ 70 subunit from *Escherichia coli* RNA polymerase are modulated by an interaction with the core enzyme. *Biochemistry*, *37*, 3312-3320.

Campbell, E. A., Greenwell, R., Anthony, J. R., Wang, S., Lim, L., Das, K., ... Darst, S. A. (2007). A Conserved Structural Module Regulates Transcriptional Responses to

Diverse Stress Signals in Bacteria. Molecular Cell, 27, 793-805.

Campbell, E. A., Westblade, L. F., & Darst, S. A. (2008). Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. *Current Opinion in Microbiology*, *11*, 121-7.

Campbell, E. A, Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., ... Darst, S. A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Molecular Cell*, 9, 527-539.

Carte, J., Wang, R., Li, H., Terns, R. M., & Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes and Development*, *22*, 3489-3496.

Cayuela, M. L., Elías-Arnanz, M., Peñalver-Mellado, M., Padmanabhan, S., & Murillo, F. J. (2003). The Stigmatella aurantiaca homolog of *Myxococcus xanthus* high-mobility-group A-type transcription factor CarD: Insights into the functional modules of CarD and their distribution in bacteria. *Journal of Bacteriology*, *185*, 3527-3537.

Cerveny, L., Straskova, A., Dankova, V., Hartlova, A., Ceckova, M., Staud, F., & Stulik, J. (2013). Tetratricopeptide repeat motifs in the world of bacterial pathogens: Role in virulence mechanisms. *Infection and Immunity*, *81*, 629-635.

Chaba, R., Grigorova, I. L., Flynn, J. M., Baker, T. A., & Gross, C. A. (2007). Design principles of the proteolytic cascade governing the σ^{E} -mediated envelope stress response in *Escherichia coli*: Keys to graded, buffered, and rapid signal transduction. *Genes and Development*, *21*, 124-136.

Charpentier, E., Richter, H., van der Oost, J., & White, M. F. (2015). Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiology Reviews*, *39*, 428-41

Choi, U. Y., Kang, J. S., Hwang, Y. S. ahng, & Kim, Y. J. (2015). Oligoadenylate synthase-like (OASL) proteins: dual functions and associations with diseases. *Experimental & Molecular Medicine*, 47:e144.

Cox, M. M., O'Donnell, M., & Doudna, J. (2012). *Molecular Biology. Principles and Practice.*

Daniels, D., Zuber, P., & Losick, R. (1990). Two amino acids in an RNA polymerase σ factor involved in the recognition of adjacent base pairs in the -10 region of a cognate promoter. *Proceedings of the Nationall Academy of Sciences USA*, *87*, 8075-8079.

Davis, M. C., Kesthely, C. A., Franklin, E. A., & MacLellan, S. R. (2016). The essential activities of the bacterial σ factor. *Canadian Journal of Microbiology*, 99, 1-36.

Dawid, W. (2000). Biology and global distribution of myxobacteria in soils. *FEMS Microbiology Reviews*, *24*, 403-427.

Deaconescu, A. M., Chambers, A. L., Smith, A. J., Nickels, B. E., Hochschild, A., Savery, N. J., & Darst, S. A. (2006). Structural basis for bacterial transcription-coupled DNA repair. *Cell*, *124*, 507-520.

Díaz-Perales, A., Quesada, V., Peinado, J. R., Ugalde, A. P., Álvarez, J., Suárez, M. F., ... López-Otín, C. (2005). Identification and characterization of human archaemetzincin-1 and -2, two novel members of a family of metalloproteases widely distributed in archaea. *Journal of Biological Chemistry*, *280*, 30367-30375.

Dombroski, A. J., Walter, W. A., & Gross, C. A. (1993). Amino-terminal amino acids modulate sigma-factor DNA-binding activity. *Genes & Development*, *7*, 2446-2455.

Dombroski, A. J., Walter, W. A., Record, M. T., Slegele, D. A., & Gross, C. A. (1992). Polypeptides containing highly conserved regions of transcription initiation factor σ 70 exhibit specificity of binding to promoter DNA. *Cell*, *70*, 501-512.

Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., ... Sorek, R. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, pii: eaar4120

Dove, S. L., Darst, S. a, & Hochschild, A. (2003). Region 4 of sigma as a target for transcription regulation. *Molecular Microbiology*, *48*, 863-874.

Dworkin, J. (2015). Ser/Thr phosphorylation as a regulatory mechanism in bacteria. *Current Opinion in Microbiology, 24*, 47-52.

Dworkin, M. (1962). Nutritional requirements for vegetative growth of *Myxococcus xanthus. Journal of Bacteriology*, *84*, 250-7.

Ebright, R. H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *Journal of Molecular Biology*, *304*, 687-698.

Edqvist, P. J., Bröms, J. E., Betts, H. J., Forsberg, Å., Pallen, M. J., & Francis, M. S. (2006). Tetratricopeptide repeats in the type III secretion chaperone, LcrH: Their role in substrate binding and secretion. *Molecular Microbiology*, *59*, 31-44.

Eiamphungporn, W., & Helmann, J. D. (2008). The *Bacillus subtilis* σ^{M} regulon and its contribution to cell envelope stress responses. *Molecular Microbiology*, 67, 830-48.

Elías-Arnanz, M., & Murillo, F. J. (1991a). Induction of germination in *Myxococcus xanthus* fruiting body spores. *Journal of General Microbiology*, *137*, 381-388.

Elías-Arnanz, M., & Murillo, F. J. (1991b). Mutations affecting germination in *Myxococcus xanthus. Journal of General Microbiology*, *137*, 389-397.

Elías-Arnanz, M., Padmanabhan, S., & Murillo, F. J. (2010). The regulatory action of the myxobacterial CarD/CarG complex: a bacterial enhanceosome? *FEMS Microbiology Reviews*, *34*, 764-778.

Elías-Arnanz, M., Padmanabhan, S., & Murillo, F. J. (2011). Light-dependent gene regulation in nonphototrophic bacteria. *Current Opinion in Microbiology*, *14*, 128-135.

Faure, L. M., Fiche, J. B., Espinosa, L., Ducret, A., Anantharaman, V., Luciano, J.,

... **Mignot, T. (2016).** The mechanism of force transmission at bacterial focal adhesion complexes. *Nature*, *539*, 530-535.

Feklistov, A. (2013). RNA polymerase: In search of promoters. *Annals of the New York Academy of Sciences*, *1293*, 25-32.

Feklístov, A., Sharon, B. D., Darst, S. A., & Gross, C. A. (2014). Bacterial σ factors: A historical, structural, and genomic perspective. *Annual Review of Microbiology*, *68*, 357-376.

Fioravanti, A., Fumeaux, C., Mohapatra, S. S., Bompard, C., Brilli, M., Frandi, A., ... Biondi, E. G. (2013). DNA binding of the cell cycle transcriptional regulator GcrA depends on N6-Adenosine methylation in *Caulobacter crescentus* and other Alphaproteobacteria. *PLoS Genetics*, *9*, e1003541.

Fontes, M., Galbis-Martínez, L., & Murillo, F. J. (2003). A novel regulatory gene for light-induced carotenoid synthesis in the bacterium *Myxococcus xanthus*. *Molecular Microbiology*, *47*, 561-571.

Francez-Charlot, A., Kaczmarczyk, A., Fischer, H.-M., & Vorholt, J. A. (2015). The general stress response in Alphaproteobacteria. *Trends in Microbiology*, 23, 164-171.

Fu, J., Gnatt, A. L., Bushnell, D. A., Jensen, G. J., Thompson, N. E., Burgess, R. R., ... Kornberg, R. D. (1999). Yeast RNA polymerase II at 5 Å resolution. *Cell*, *98*, 799-810.

Gaballa, A., Guariglia-Oropeza, V., Dürr, F., Butcher, B. G., Chen, A. Y., **Chandrangsu, P., & Helmann, J. D. (2018).** Modulation of extracytoplasmic function (ECF) σ factor promoter selectivity by spacer region sequence. *Nucleic Acids Research*, *46*, 134-145.

Galbis-Martínez, L., Galbis-Martínez, M., Murillo, F. J., & Fontes, M. (2008). An antiantisigma factor in the response of the bacterium *Myxococcus xanthus* to blue light. *Microbiology*, *154*, 895-904.

Galbis-Martínez, M., Fontes, M., & Murillo, F. J. (2004). The high-mobility group A-type protein CarD of the bacterium *Myxococcus xanthus* as a transcription factor for several distinct vegetative genes. *Genetics*, *167*, 1585-1595.

Gallego-García, A., Iniesta, A. A., González, D., Collier, J., Padmanabhan, S., & Elías-Arnanz, M. (2017). *Caulobacter crescentus* CdnL is a non-essential RNA polymerase-binding protein whose depletion impairs normal growth and rRNA transcription. *Scientific Reports*, 7:43240.

Gallego-García, A., Mirassou, Y., Elías-Arnanz, M., Padmanabhan, S., & Jiménez, M. A. (2012). NMR structure note: N-terminal domain of *Thermus thermophilus* CdnL. *Journal of Biomolecular NMR*, *53*, 355-363.

Gallego-García, A., Mirassou, Y., Garcia-Moreno, D., Elías-Arnanz, M., Angeles Jimenez, M., & Padmanabhan, S. (2014). Structural Insights into RNA Polymerase Recognition and Essential Function of *Myxococcus xanthus* CdnL. *PLoS One*, *9*, e108946.

García-Heras, F. (2009). Un singular complejo transcripcional bacteriano que presenta similitudes con los complejos de transcripción eucarióticos. *Tesis Doctoral, Universidad de Murcia.*

García-Heras, F., Abellón-Ruiz, J., Murillo, F. J., Padmanabhan, S., & Elías-Arnanz, M. (2013). High-mobility-group a-like CarD binds to a DNA site optimized for affinity and position and to RNA polymerase to regulate a light-inducible promoter in *Myxococcus xanthus*. *Journal of Bacteriology*, *195*, 378-88.

García-Heras, F., Padmanabhan, S., Murillo, F. J., & Elías-Arnanz, M. (2009). Functional equivalence of HMGA- and histone H1-like domains in a bacterial transcriptional factor. *Proceedings of the National Academy of Sciences USA*, *106*, 13546-51. García-Moreno, D., Abellón-Ruiz, J., Garcia-Heras, F., Murillo, F. J., Padmanabhan, S., & Elías-Arnanz, M. (2010). CdnL, a member of the large CarD-like family of bacterial proteins, is vital for *Myxococcus xanthus* and differs functionally from the global transcriptional regulator CarD. *Nucleic Acids Research*, *38*, 4586-4598.

García-Moreno, D., Polanco, M. C., Navarro-Avilés, G., Murillo, F. J., Padmanabhan, S., & Elías-Arnanz, M. (2009). A vitamin B₁₂-based system for conditional expression reveals *dksA* to be an essential gene in *Myxococcus xanthus*. *Journal of Bacteriology*, *191*, 3108-3119.

Garner, A. L., Weiss, L. A., Manzano, A. R., Galburt, E. A., & Stallings, C. L. (2014). CarD integrates three functional modules to promote efficient transcription, antibiotic tolerance, and pathogenesis in mycobacteria. *Molecular Microbiology*, *93*, 682-697.

Gilmore, J. M., Bieber Urbauer, R. J., Minakhin, L., Akoyev, V., Zolkiewski, M., Severinov, K., & Urbauer, J. L. (2010). Determinants of affinity and activity of the antiσ factor AsiA. *Biochemistry*, *4*9, 6143-6154.

Goldman, B. S., Nierman, W. C., Kaiser, D., Slater, S. C., Durkin, A. S., Eisen, J. A., ... Kaplan, H. B. (2006). Evolution of sensory complexity recorded in a myxobacterial genome. *Proceedings of the National Academy of Sciences USA*, *103*, 15200-15205.

Gómez-Santos, N., Pérez, J., Sánchez-Sutil, M. C., Moraleda-Muñoz, A., & Muñoz-Dorado, J. (2011). Core from *Myxococcus xanthus* is a Copper-Dependent RNA polymerase σ factor. *PLoS Genetics*, 7:e1002106

Gomis-Rüth, F. X. (2003). Structural aspects of the metzincin clan of metalloendopeptidases. *Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology*, *24*, 157-202

Goutam, K., Gupta, A. K., & Gopal, B. (2017). The fused SnoaL_2 domain in the *Mycobacterium tuberculosis* σ factor σ^{J} modulates promoter recognition. *Nucleic Acids Research*, *45*, 9760-9772.

Graef, C., Schacherl, M., Waltersperger, S., & Baumann, U. (2012). Crystal structures of archaemetzincin reveal a moldable substrate-binding site. *PLoS One*, *7*, e43863.

Greenstein, A. E., MacGurn, J. A., Baer, C. E., Falick, A. M., Cox, J. S., & Alber, T. (2007). *M. tuberculosis* Ser/Thr protein kinase D phosphorylates an anti-anti- σ factor homolog. *PLoS Pathogens*, *3*, e49.

Greenwell, R., Nam, T. W., & Donohue, T. J. (2011). Features of *Rhodobacter sphaeroides* ChrR required for stimuli to promote the dissociation of σ^{E} /ChrR complexes. *Journal of Molecular Biology*, *407*, 477-491.

Gribskov, M., & Burgess, R. R. (1986). σ factors from *E. coli*, *B. subtilis*, phage SP01, and phage T4 are homologous proteins. *Nucleic Acids Research*, *14*, 6745-6763.

Gries, T. J., Kontur, W. S., Capp, M. W., Saecker, R. M., & Record, M. T. (2010). Onestep DNA melting in the RNA polymerase cleft opens the initiation bubble to form an unstable open complex. *Proceedings of the National Academy of Sciences USA*, *107*, 10418-10423.

Griffith, K. L., & Wolf, R. E. (2002). Measuring β-galactosidase activity in bacteria: Cell growth, permeabilization, and enzyme assays in 96-well arrays. *Biochemical and Biophysical Research Communications*, *290*, 397-402.

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, *8*, 172.

Gruber, T. M., & Bryant, D. A. (1997). Molecular systematic studies of eubacteria, using σ 70-type σ factors of group 1 and group 2. *Journal of Bacteriology*, *179*, 1734-1747.

Gulten, G., & Sacchettini, J. C. (2013). Structure of the *Mtb* CarD/RNAP beta-Lobes Complex Reveals the Molecular Basis of Interaction and Presents a Distinct DNA-Binding Domain for *Mtb* CarD. *Structure*, *21*, 1859-1869.

Haakonsen, D. L., Yuan, A. H., & Laub, M. T. (2015). The bacterial cell cycle regulator

gcrA is a σ^{70} cofactor that drives gene expression from a subset of methylated promoters. *Genes and Development*, 29, 2272-2286.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *Journal of Molecular Biology*, *166*, 557-580.

Hao, T., Biran, D., Velicer, G. J., & Kroos, L. (2002). Identification of the Ω4514 regulatory region, a developmental promoter of *Myxococcus xanthus* that is transcribed *in vitro* by the major vegetative RNA polymerase. *Journal of Bacteriology, 184*, 3348-3359.

Hastie, J. L., Williams, K. B., Bohr, L. L., Houtman, J. C., Gakhar, L., & Ellermeier, C. D. (2016). The anti- σ factor RsiV is a bacterial receptor for lysozyme: co-crystal structure determination and demonstration that binding of lysozyme to RsiV is required for σ V activation. *PLoS Genetics*, *12*, e1006287.

Hatzios, S. K., Baer, C. E., Rustad, T. R., Siegrist, M. S., Pang, J. M., Ortega, C., ... Bertozzi, C. R. (2013). Osmosensory signaling in *Mycobacterium tuberculosis* mediated by a eukaryotic-like Ser/Thr protein kinase. *Proceedings of the National Academy of Sciences USA*, *110*, e5069-77.

Haugen, S. P., Berkmen, M. B., Ross, W., Gaal, T., Ward, C., & Gourse, R. L. (2006). rRNA Promoter Regulation by Nonoptimal Binding of σ region 1.2: An Additional Recognition Element for RNA Polymerase. *Cell*, *125*, 1069-1082.

Haugen, S. P., Ross, W., & Gourse, R. L. (2008). Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nature Reviews Microbiology*, *6*, 507-19.

Haugen, S. P., Ross, W., Manrique, M., & Gourse, R. L. (2008). Fine structure of the promoter-sigma region 1.2 interaction. *Proceedings of the National Academy of Sciences USA*, *105*, 3292-3297.

Heinrich, J., Hein, K., & Wiegert, T. (2009). Two proteolytic modules are involved in regulated intramembrane proteolysis of *Bacillus subtilis* RsiW. *Molecular Microbiology*, *74*, 1412-1426.

Helmann, J. D. (2002). The extracytoplasmic function (ECF) σ factors. *Advances in Microbial Physiology*, *46*, 47-110.

Helmann, J. D., & Chamberlin, M. J. (1988). Structure and function of bacterial σ factors. *Annual Review of Biochemistry*, *57*, 839-872.

Helmann, J. D., & DeHaseth, P. L. (1999). Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. *Biochemistry*, ;38, 5959-5967.

Hille, F., & Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*, 20150496.

Hizukuri, Y., & Akiyama, Y. (2012). PDZ domains of RseP are not essential for sequential cleavage of RseA or stress-induced σE activation *in vivo*. *Molecular Microbiology*, *86*, 1232-1245.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, *77*, 51-59.

Ho, T. D., & Ellermeier, C. D. (2012). Extracytoplasmic function *σ* factor activation. *Current Opinion in Microbiology, 15*, 182-188.

Hochstrasser, M. L., & Doudna, J. A. (2015). Cutting it close: CRISPR-associated endoribonuclease structure and function. *Trends in Biochemical Sciences*, *40*, 58-66.

Hodgson, D. A. (1983). Light-induced carotenogenesis in Myxococcus xanthus: genetic analysis of the carR region. *Molecular Microbiology*, *7*, 471-488.

Hook-Barnard, I. G., & Hinton, D. M. (2007). Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regulation and Systems Biology*, *1*, 275-293.

Hook-Barnard, I., Johnson, X. B., & Hinton, D. M. (2006). Escherichia coli RNA polymerase recognition of a σ 70- dependent promoter requiring a -35 DNA element and an extended -10 TGn motif. *Journal of Bacteriology*, *188*, 8352-8359.

Høyland-Kroghsbo, N. M., Paczkowski, J., Mukherjee, S., Broniewski, J., Westra, E., Bondy-Denomy, J., & Bassler, B. L. (2016). Quorum sensing controls the *Pseudomonas aeruginosa* CRISPR-Cas adaptive immune system. *Proceedings of the National Academy of Sciences USA*, *114*, 131-135.

Hu, Y., Morichaud, Z., Chen, S., Leonetti, J. P., & Brodolin, K. (2012). *Mycobacterium tuberculosis* RbpA protein is a new type of transcriptional activator that stabilizes the σacontaining RNA polymerase holoenzyme. *Nucleic Acids Research*, *40*, 6547-6557.

Huang, X., Fredrick, K. L., & Helmann, J. D. (1998). Promoter recognition by Bacillus subtilis sigmaW: autoregulation and partial overlap with the σ X regulon. *Journal of Bacteriology*, *180*, 3765-70.

Huang, X., Pinto, D., Fritz, G., & Mascher, T. (2015). Environmental sensing in Actinobacteria: A comprehensive survey on the signaling capacity of this *phylum*. *Journal of Bacteriology*, *197*, 2517-2535.

Hubin, E. A., Fay, A., Xu, C., Bean, J. M., Saecker, R. M., Glickman, M. S., ... Campbell, E. A. (2017). Structure and function of the mycobacterial transcription initiation complex with the essential regulator RbpA. *eLife*, *6*, e22520.

Hubin, E. A., Tabib-Salazar, A., Humphrey, L. J., Flack, J. E., Olinares, P. D. B., Darst, S. A., ... Paget, M. S. B. (2015). Structural, functional, and genetic analyses of the actinobacterial transcription factor RbpA. *Proceedings of the National Academy of Science USA*, *112*, 7171-7176.

lizuka, T., Jojima, Y., Fudou, R., & Yamanaka, S. (1998). Isolation of myxobacteria from the marine environment. *FEMS Microbiology Letters*, *169*, 317-322.

Iniesta, A. A., García-Heras, F., Abellón-Ruiz, J., Gallego-García, A., & Elías-Arnanz, M. (2012). Two systems for conditional gene expression in *Myxococcus xanthus* inducible by isopropyl-β-d-thiogalactopyranoside or vanillate. *Journal of Bacteriology*, *194*, 5875-5885.

Inouye, S. (1990). Cloning and DNA sequence of the gene coding for the major σ factor from *Myxococcus xanthus*. *Journal of Bacteriology*, *172*, 80-85.

Ishino, Y., Krupovic, M., & Forterre, P. (2018). History of CRISPR-Cas from encounter with a mysterious repeated sequence to genome editing technology. *Journal of Bacteriology*, *200*, 00580-17.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, *169*, 5429-5433.

Islam, S. T., & Mignot, T. (2015). The mysterious nature of bacterial surface (gliding) motility: A focal adhesion-based mechanism in *Myxococcus xanthus*. *Seminars in Cell and Developmental Biology*, *46*, 143-154.

Janssen, G. R., Wireman, J. W., & Dworkin, M. (1977). Effect of temperature on the growth of *Myxococcus xanthus*. *Journal of Bacteriology*, *130*, 561-562.

Jogler, C., Waldmann, J., Huang, X., Jogler, M., Glöckner, F. O., & Mascher, T. (2012). Identification of proteins likely to be involved in morphogenesis, cell division, and signal transduction in planctomycetes by comparative genomics. *Journal of Bacteriology*, *194*, 6419-6430.

Jost, M., Fernández-Zapata, J., Polanco, M. C., Ortiz-Guerrero, J. M., Chen, P. Y. T., Kang, G., ... Drennan, C. L. (2015). Structural basis for gene regulation by a B₁₂dependent photoreceptor. *Nature*, *526*, 536-541.

Kahn, D., & Ditta, G. (1991). Modular structure of FixJ: homology of the transcriptional activator domain with the -35 binding domain of sigma factors. *Molecular Microbiology*, *5*, 987-997.

Kaiser, D., & Dworkin, M. (2008). From glycerol to the genome. ASM Press: 257-284.

Kaiser, D. (1979). Social gliding is correlated with the presence of pili in *Myxococcus xanthus*. *Proceedings of the National Academy of Sciences USA*, *76*, 5952-5956.

Kaiser, D. (2004). Signaling in myxobacteria. Annual Review of Microbiology, 58, 75-98.

Kaiser, D. (2008). *Myxococcus*-from single-cell polarity to complex multicellular patterns. *Annual Review of Genetics*, *42*, 109-30.

Kang, J. G., Paget, M. S. B., Seok, Y. J., Hahn, M. Y., Bae, J. B., Hahn, J. S., ... Roe, J. H. (1999). RsrA, an anti-σ factor regulated by redox change. *EMBO Journal*, *18*, 4292-4298.

Karimova, G., Ullmann, A., & Ladant, D. (2000). A bacterial two-hybrid system that exploits a cAMP signaling cascade in *Escherichia coli*. *Methods in Enzymology*, *328*, 59-73.

Kashefi, K., & Hartzell, P. L. (1995). Genetic suppression and phenotypic masking of a *Myxococcus xanthus frzF*- defect. *Molecular Microbiology*, *15*, 483-494.

Kasinsky, H. E., Lewis, J. D., Dacks, J. B., & Ausio, J. (2001). Origin of H1 linker histones. *FASEB Journal*, *15*, 34-42.

Kaur, G., Dutta, D., & Thakur, K. G. (2014). Crystal structure of *Mycobacterium tuberculosis* CarD, an essential RNA polymerase binding protein, reveals a quasidomain-swapped dimeric structural architecture. *Proteins: Structure, Function and Bioinformatics*, *82*, 879-884.

Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G. & Siksnys, V. (2017). A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science*, *357*, 605-609.

Keseler, I. M., & Kaiser, D. (1997). 54, a vital protein for *Myxococcus xanthus. Proceedings of the National Academy of Sciences USA*, 94, 1979-1984.

Kim, K. Y., Park, J. K., & Park, S. (2016). In *Streptomyces coelicolor* SigR, methionine at the -35 element interacting region 4 confers the -31'-adenine base selectivity. *Biochemical and Biophysical Research Communications*, *470*, 257-262.

Klemenčič, M., Novinec, M., & Dolinar, M. (2015). Orthocaspases are proteolytically active prokaryotic caspase homologues: The case of *Microcystis aeruginosa*. *Molecular Microbiology*, 98, 142-150.

Komano, T., Franceschini, T., & Inouye, S. (1987). Identification of a vegetative promoter in *Myxococcus xanthus*. A protein that has homology to histones. *Journal of Molecular Biology*, *196*, 517-524.

Koonin, E. V. (2018). Hunting for treasure chests in microbial defense islands. *Molecular Cell*, 70, 761-762.

Koonin, E. V, Makarova, K. S., & Wolf, Y. I. (2017). Evolutionary genomics of defense systems in archaea and bacteria. *Annual Review of Microbiology*, *71*, 233-261.

Koonin, E. V, Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology*, *37*, 67-78.

Kroos, L. (2007). The *Bacillus* and *Myxococcus* developmental networks and their transcriptional regulators. *Annual Review of Genetics*, *41*, 13-39.

Kroos, L., & Inouye, S. (2008). Transcriptional regulatory mechanisms during *Myxococcus xanthus* development. *Myxobacteria, ASM Press*, 149-168.

Kroos, L., & Kaiser, D. (1987). Expression of many developmentally regulated genes in *Myxococcus* depends on a sequence of cell interactions. *Genes & Development, 1*, 840-854.

Kroos, L., Kuspa, A., & Kaiser, D. (1986). A global analysis of developmentally regulated genes in *Myxococcus xanthus*. *Developmental Biology*, *117*, 252-266.

Kulbachinskiy, A., & Mustaev, A. (2006). Region 3.2 of the σ subunit contributes to the binding of the 3'-initiating nucleotide in the RNA polymerase active center and facilitates promoter clearance during initiation. *Journal of Biological Chemistry*, *281*, 18273-18276.

Kyungyun, C., & Zusman, D. R. (1999). AsgD, a new two-component regulator required for A-signalling and nutrient sensing during early development of *Myxococcus xanthus*. *Molecular Microbiology*, *34*, 268-281.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*, 680-685.

Lamour, V., Rutherford, S. T., Kuznedelov, K., Ramagopal, U. A., Gourse, R. L., Severinov, K., & Darst, S. A. (2008). Crystal Structure of *Escherichia coli* Rnk, a New RNA Polymerase-Interacting Protein. *Journal of Molecular Biology*, 383, 367-379.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., McWilliam, H., ... Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, 23, 2947-2948.

Lawson, C. L., Swigon, D., Murakami, K. S., Darst, S. A., Berman, H. M., & Ebright,
R. H. (2004). Catabolite activator protein: DNA binding and transcription activation. *Current Opinion in Structural Biology*, *14*, 10-20.

Lee, D. J., Minchin, S. D., & Busby, S. J. W. (2012). Activating Transcription in Bacteria.

Annual Review of Microbiology, 66, 125-152.

Lemke, J. J., Durfee, T., & Gourse, R. L. (2009). DksA and ppGpp directly regulate transcription of the *Escherichia coli* flagellar cascade. *Molecular Microbiology*, *74*, 1368-1379.

León, E., Navarro-Avilés, G., Santiveri, C. M., Flores-Flores, C., Rico, M., González, C., ... Padmanabhan, S. (2010). A bacterial antirepressor with SH3 domain topology mimics operator DNA in sequestering the repressor DNA recognition helix. *Nucleic Acids Research*, *38*, 5226-5241.

Lesley, S. A., & Burgess, R. R. (1989). Characterization of the *Escherichia coli* Transcription Factor σ70: Localization of a region Involved in the Interaction with core RNA polymerase. *Biochemistry*, *28*, 7728-7734.

Lewis, D. E. A., & Adhya, S. (2015). Molecular mechanisms of transcription initiation at gal promoters and their multi-level regulation by GaIR, CRP and DNA loop. *Biomolecules*, *5*, 2782-2807.

Li, M., McClure, W. R., & Susskind, M. M. (1997). Changing the mechanism of transcriptional activation by phage lambda repressor. *Proceedings of the National Academy of Sciences USA*, *94*, 3691-3696.

Li, W., Bottrill, A. R., Bibb, M. J., Buttner, M. J., Paget, M. S. B., & Kleanthous, C. (2003). The role of zinc in the disulphide stress-regulated anti-sigma factor RsrA from *Streptomyces coelicolor. Journal of Molecular Biology*, *333*, 461-472.

Li, X., Wang, B., Feng, L., Kang, H., Qi, Y., Wang, J., & Shi, Y. (2009). Cleavage of RseA by RseP requires a carboxyl-terminal hydrophobic amino acid following DegS cleavage. *Proceedings of the National Academy of Sciences USA*, *106*, 14837-14842.

Lin, W., Mandal, S., Degen, D., Liu, Y., Ebright, Y. W., Li, S., ... Ebright, R. H. (2017). Structural basis of *Mycobacterium tuberculosis* transcription and transcription inhibition. *Molecular Cell*, 70, 60-71.

Liu, T., Li, Y., Wang, X., Ye, Q., Li, H., Liang, Y., ... Peng, N. (2015). Transcriptional regulator-mediated activation of adaptation genes triggers CRISPR *de novo* spacer acquisition. *Nucleic Acids Research*, *43*, 1044-1055.

Liu, X., & Matsumura, P. (1995). An alternative sigma factor controls transcription of flagellar class-III operons in *Escherichia coli*: gene sequence, overproduction, purification and characterization. *Gene*, *164*, 81-84.

Lonetto, M. A., Brown, K. L., Rudd li, K. E., & Buttner, M. J. (1994). Analysis of the

Streptomyces coelicolor sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase or factors involved in the regulation of extracytoplasmic functions. *Microbiology*, *91*, 7573-7577.

Lonetto, M. A., Rhodius, V., Lamberg, K., Kiley, P., Busby, S., & Gross, C. A. (1998). Identification of a contact site for different transcription activators in region 4 of the *Escherichia coli* RNA polymerase σ70 subunit. *Journal of Molecular Biology*, *284*, 1353-1365.

Lonetto, M., Gribskov, M., & Gross, C. A. (1992). The σ 70 family: sequence conservation and evolutionary relationships. *Journal of Bacteriology*, *174*, 3843-9.

López-Rubio, J. J., Elías-Arnanz, M., Padmanabhan, S., & Murillo, F. J. (2002). A repressor-antirepressor pair links two loci controlling light-induced carotenogenesis in *Myxococcus xanthus. Journal of Biological Chemistry*, 277, 7262-7270.

López-Rubio, J. J., Padmanabhan, S., Lázaro, J. M., Salas, M., Murillo, F. J., & Elías-Arnanz, M. (2004). Operator design and mechanism for CarA repressor-mediated downregulation of the photoinducible carB operon in *Myxococcus xanthus*. *Journal of Biological Chemistry*, 279, 28945-28953.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. *The Journal of Biological Chemistry*, *193*, 265-275.

Maillard, A. P., Girard, E., Ziani, W., Petit-Hartlein, I., Kahn, R., & Covès, J. (2014).

The crystal structure of the anti- σ factor CnrY in complex with the σ factor CnrH shows a new structural class of anti- σ factors targeting extracytoplasmic function σ factors. *Journal of Molecular Biology*, 426, 2313-2327.

Makarova, K. S., Grishin, N. V, Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct*, *1*, 7.

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., ... Koonin, E. V. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nature Reviews Microbiology*, *13*, 722-736.

Makarova, K. S., Wolf, Y. I., Snir, S., & Koonin, E. V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *Journal of*

Bacteriology.

Malhotra, A., Severinova, E., & Darst, S. A. (1996). Crystal structure of a σ 70 subunit fragment from *E. coli* RNA polymerase. *Cell*, *87*, 127-136.

Marchetti, M., Malinowska, A., Heller, I., & Wuite, G. J. L. (2017). How to switch the motor on: RNA polymerase initiation steps at the single-molecule level. *Protein Science*, *26*, 1303-1313.

Marcos-Torres, F. J., Perez, J., Gomez-Santos, N., Moraleda-Muñoz, A., & Munoz-Dorado, J. (2016). In depth analysis of the mechanism of action of metal-dependent sigma factors: Characterization of CorE2 from *Myxococcus xanthus*. *Nucleic Acids Research*, 44, 5571-5584.

Marr, M. T., & Roberts, J. W. (1997). Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. *Science*, 276, 1258-1260.

Marraffini, L. A., & Sontheimer, E. J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics*, *11*, 181-190.

Martínez-Laborda, A., Balsalobre, J. M., Fontes, M., & Murillo, F. J. (1990). Accumulation of carotenoids in structural and regulatory mutants of the bacterium *Myxococcus xanthus. Molecular & General Genetics : MGG*, 223, 205-10.

Mascher, T., Hachmann, A. B., & Helmann, J. D. (2007). Regulatory overlap and functional redundancy among Bacillus subtilis extracytoplasmic function σ factors. *Journal of Bacteriology*, *189*, 6919-6927.

Mauriello, E. M. F. (2010). Cell polarity/motility in bacteria: Closer to eukaryotes than expected? *EMBO Journal*, 29, 2258-2259.

Medina-Aparicio, L., Rebollar-Flores, J. E., Gallego-Hernández, A. L., Vázquez, A., Olvera, L., Gutiérrez-Ríos, R. M., ... Hernández-Lucas, I. (2011). The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *Journal of Bacteriology*, *193*, 2396-2407.

Meiser, P., Bode, H. B., & Müller, R. (2006). The unique DKxanthene secondary metabolite family from the myxobacterium *Myxococcus xanthus* is required for developmental sporulation. *Proceedings of the National Academy of Sciences USA*, *103*, 19128-33.

Melissa, B. & Echalier, A. (2014). Structure and Function of MPN (Mpr1/Pad1 N-terminal) Domain- Containing Proteins. *Current Protein & Peptide Science*, *15*, 504-517.

Mercier, R., & Mignot, T. (2016). Regulations governing the multicellular lifestyle of *Myxococcus xanthus. Current Opinion in Microbiology*, *34*, 104-110.

Merika, M., & Thanos, D. (2001). Enhanceosomes. *Current Opinion in Genetics and Development*, *11*, 205-208.

Miller, J. H. (1972). Experiments in molecular genetics. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*, 433, 352-355.

Mitchell, J. E., Zheng, D., Busby, S. J. W., & Minchin, S. D. (2003). Identification and analysis of «extended -10» promoters in *Escherichia coli. Nucleic Acids Research, 31*, 4689-4695.

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*, *60*, 174-182.

Mojica, F. J. M., Díez-Villaseñor, C., Soria, E., & Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology*, *36*, 244-246.

Mojica, F. J. M., & Montoliu, L. (2016). On the origin of CRISPR-Cas technology: from prokaryotes to mammals. *Trends in Microbiology*, *24*, 811-820.

Molodtsov, V., Sineva, E., Zhang, L., Huang, X., Cashel, M., Ades, S. E., & Murakami, K. S. (2018). Allosteric effector ppGpp potentiates the inhibition of transcript initiation by DksA. *Molecular Cell*, 69, 828-839.

Moraleda-Muñoz, A., Pérez, J., Fontes, M., Murillo, F. J., & Muñoz-Dorado, J. (2005). Copper induction of carotenoid synthesis in the bacterium *Myxococcus xanthus*. *Molecular Microbiology*, *56*, 1159-1168.

Moreno, A. J., Fontes, M., & Murillo, F. J. (2001). *ihfA* gene of the bacterium *Myxococcus xanthus* and its role in activation of carotenoid genes by blue light. *Journal of Bacteriology*, *183*, 557-569.

Müller-Hill, B. (1998). Some repressors of bacterial transcription. *Current Opinion in Microbiology*, *1*, 145-151.

Muñoz-Dorado, J., Inouye, S., & Inouye, M. (1991). A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus*, a gramnegative bacterium. *Cell*, *67*, 995-1006.

Muñoz-Dorado, J., Marcos-Torres, F. J., García-Bravo, E., Moraleda-Muñoz, A., & Pérez, J. (2016). Myxobacteria: Moving, killing, feeding, and surviving together. *Frontiers*

in Microbiology, 7, 781.

Murakami, K. S. (2013). X-ray crystal structure of *Escherichia coli* RNA polymerase σ70 holoenzyme. *The Journal of Biological Chemistry*, 288, 9126-34.

Murakami, K. S. (2015). Structural biology of bacterial RNA polymerase. *Biomolecules*, *5*, 848-864.

Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., & Darst, S. A. (2002). Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex. *Science*, *296*, 1285-1290.

Nakahigashi, K., Yanagi, H., & Yura, T. (1995). Isolation and sequence analysis of rpoH genes encoding σ 32 homologs from gram negative bacteria: conserved mRNA and protein segments for heat shock regulation. *Nucleic Acids Research*, *23*, 4383-90.

Nan, B., Bandaria, J. N., Moghtaderi, A., Sun, I.-H., Yildiz, A., & Zusman, D. R. (2013). Flagella stator homologs function as motors for myxobacterial gliding motility by moving in helical trajectories. *Proceedings of the National Academy of Sciences USA*, *110*, 1508-1513.

Nan, B., & Zusman, D. R. (2016). Novel mechanisms power bacterial gliding motility. *Molecular Microbiology*, *101*, 186-93.

Nariya, H., & Inouye, S. (2005). Identification of a protein Ser/Thr kinase cascade that regulates essential transcriptional activators in *Myxococcus xanthus* development. *Molecular Microbiology*, *58*, 367-379.

Nariya, H., & Inouye, S. (2006). A protein Ser/Thr kinase cascade negatively regulates the DNA-binding activity of MrpC, a smaller form of which may be necessary for the *Myxococcus xanthus* development. *Molecular Microbiology*, *60*, 1205-1217.

Navarro-Avilés, G., Jiménez, M. A., Pérez-Marín, M. C., González, C., Rico, M., Murillo, F. J., ... Padmanabhan, S. (2007). Structural basis for operator and antirepressor recognition by *Myxococcus xanthus* CarA repressor. *Molecular Microbiology*, 63, 980-994.

Newman, J. D., Anthony, J. R., & Donohue, T. J. (2001). The importance of zincbinding to the function of *Rhodobacter sphaeroides* ChrR as an anti-sigma factor. *Journal of Molecular Biology*, *313*, 485-499.

Nicolás, F. J., Cayuela, M. L., Martínez-Argudo, I. M., Ruiz-Vázquez, R. M., & Murillo,

F. J. (1996). High mobility group I(Y)-like DNA-binding domains on a bacterial transcription factor. *Proceedings of the National Academy of Sciences USA*, *93*, 6881-6885.

Nicolás, F. J., Ruiz-Vázquez, R. M., & Murillo, F. J. (1994). A genetic link between light response and multicellular development in the bacterium *Myxococcus xanthus*. *Genes and Development*, *8*, 2375-2387.

Niewoehner, O., García-Doval, C., Rostol, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., ... Jinek, M. (2017). Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature*, *548*, 543-548.

Nuñez, J. K., Lee, A. S. Y., Engelman, A., & Doudna, J. A. (2015). Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature*, *519*, 193-198.

Ortiz-Guerrero, J. M., Polanco, M. C., Murillo, F. J., Padmanabhan, S., & Elías-

Arnanz, M. (2011). Light-dependent gene regulation by a coenzyme B12-based photoreceptor. *Proceedings of the National Academy of Sciences USA*, *108*, 7565-7570.

Padmanabhan, S., Elías-Arnanz, M., Carpio, E., Aparicio, P., & Murillo, F. J. (2001).

Domain architecture of a High Mobility Group A-type bacterial transcriptional factor. *Journal of Biological Chemistry*, 276, 41566-41575.

Padmanabhan, S., Jost, M., Drennan, C. L., & Elías-Arnanz, M. (2017). A new facet of vitamin B12 : gene regulation by cobalamin-based photoreceptors. *Annual Review of Biochemistry*, *86*, 485-514.

Paget, M. S. B. (2015). Bacterial σ factors and anti-sigma factors: structure, function and distribution. *Biomolecules*, *5*, 1245-1265.

Paget, M. S. B., & Helmann, J. D. (2003). The σ^{70} family of σ factors. *Genome Biology*, *4*, 203.

Patterson, A. G., Chang, J. T., Taylor, C., & Fineran, P. C. (2015). Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Research, 43*, 6038-6048.

Patterson, A. G., Jackson, S. A., Taylor, C., Evans, G. B., Salmond, G. P. C., Przybilski, R., ... Fineran, P. C. (2016). Quorum sensing controls adaptive immunity through the regulation of ,ultiple CRISPR-Cas systems. *Molecular Cell*, *64*, 1102-1108.

Patterson, A. G., Yevstigneyeva, M. S., & Fineran, P. C. (2017). Regulation of CRISPR–Cas adaptive immune systems. *Current Opinion in Microbiology*, 37, 1-7.

Paul, B. J., Barker, M. M., Ross, W., Schneider, D. A., Webb, C., Foster, J. W., & Gourse, R. L. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell*, *118*, 311-322.

Paul, B. J., Berkmen, M. B., & Gourse, R. L. (2005). DksA potentiates direct activation of amino acid promoters by ppGpp. *Proceedings of the National Academy of Sciences USA*, *102*, 7823-7828.

Peñalver-Mellado, M. (2005). Identificación de proteínas de *Myxococcus xanthus* que interaccionan con el factor transcripcional CarD. *Tesis Doctoral, Universidad de Murcia.*

Peñalver-Mellado, M., García-Heras, F., Padmanabhan, S., García-Moreno, D., Murillo, F. J., & Elías-Arnanz, M. (2006). Recruitment of a novel zinc-bound transcriptional factor by a bacterial HMGA-type protein is required for regulating multiple processes in *Myxococcus xanthus*. *Molecular Microbiology*, *61*, 910-926.

Pereira, S. F. F., Goss, L., & Dworkin, J. (2011). Eukaryote-like Serine/Threonine kinases and phosphatases in bacteria. *Microbiology and Molecular Biology Reviews*, *75*, 192-212.

Pérez-Marín, M. C., Padmanabhan, S., Polanco, M. C., Murillo, F. J., & Elías-Arnanz,
M. (2008). Vitamin B₁₂ partners the CarH repressor to downregulate a photoinducible promoter in *Myxococcus xanthus*. *Molecular Microbiology*, 67, 804-819.

Perez, J., Castaneda-Garcia, A., Jenke-Kodama, H., Muller, R., & Munoz-Dorado, J. (2008). Eukaryotic-like protein kinases in the prokaryotes and the myxobacterial kinome. *Proceedings of the National Academy of Sciences USA*, *105*, 15950-15955.

Petushkov, I., Esyunina, D., Mekler, V., Severinov, K., Pupov, D., & Kulbachinskiy, A. (2017). Interplay between σ region 3.2 and secondary channel factors during promoter escape by bacterial RNA polymerase. *Biochemical Journal*, *474*, 4053-4064.

Plagens, A., Richter, H., Charpentier, E., & Randau, L. (2015). DNA and RNA interference mechanisms by CRISPR-Cas surveillance complexes. *FEMS Microbiology Reviews*, *39*, 442-463.

Pompidor, G., Girard, E., Maillard, A. P., Ramella-Pairin, S., Bersch, B., Kahn, R., & Covès, J. (2009). Biostructural analysis of the metal-sensor domain of CnrX from *Cupriavidus metallidurans* CH34. *International Journal of General and Molecular Microbiology*, 96, 141-148.

Pourcel, C., Salvignol, G., & Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology*, *151*, 653-663.

Pyenson, N. C., & Marraffini, L. A. (2017). Type III CRISPR-Cas systems: when DNA cleavage just isn't enough. *Current Opinion in Microbiology*, *37*, 150-154.

Quinones, M., Kimsey, H. H., Ross, W., Gourse, R. L., & Waldor, M. K. (2006). LexA represses CTXphi transcription by blocking access of the alpha C-terminal domain of RNA polymerase to promoter DNA. *The Journal of Biological Chemistry*, *281*, 39407-39412.

Rajagopalan, R., & Kroos, L. (2017). The *dev* operon regulates the timing of sporulation during *Myxococcus xanthus* development. *Journal of Bacteriology*, *199*, 00788-16.

Rajagopalan, R., Wielgoss, S., Lippert, G., Velicer, G. J., & Kroos, L. (2015). devl is an evolutionarily young negative regulator of *Myxococcus xanthus* development. *Journal of Bacteriology*, *197*, 1249-1262.

Rajasekar, K. V, Zdanowski, K., Yan, J., Hopper, J. T. S., Francis, M. L. R., Seepersad, C., ... Kleanthous, C. (2016). The anti- σ factor RsrA responds to oxidative stress by reburying its hydrophobic core. *Nature Communications*, *7*, 12194.

Reichenbach, H. (1999). The ecology of the myxobacteria. *Environmental Microbiology*, *1*, 15-21.

Reichenbach, H., & Höfle, G. (1993). Biologically active secondary metabolites from myxobacteria. *Biotechnology Advances*, *11*, 219-277.

Rhodius, V. A., & Busby, S. J. W. (2000). Interactions between Activating Region 3 of the *Escherichia coli* cyclic AMP receptor protein and region 4 of the RNA polymerase σ 70subunit: Application of suppression genetics. *Journal of Molecular Biology*, *299*, 311-324.

Rhodius, V. A., Segall-Shapiro, T. H., Sharon, B. D., Ghodasara, A., Orlova, E., Tabakh, H., ... Voigt, C. A. (2013). Design of orthogonal genetic switches based on a crosstalk map of σs, anti-σs, and promoters. *Molecular Systems Biology*, *9*, 702.

Richter, C., Chang, J. T., & Fineran, P. C. (2012). Function and regulation of clustered

regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. *Viruses*, *9*, 702.

Risser, D. D., & Callahan, S. M. (2008). HetF and PatA control levels of HetR in *Anabaena sp.* strain PCC 7120. *Journal of Bacteriology*, *190*, 7645-7654.

Rollie, C., Schneider, S., Brinkmann, A. S., Bolt, E. L., & White, M. F. (2015). Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *eLife*, *4*, 08716.

Ross, W., Ernst, A., & Gourse, R. L. (2001). Fine structure of *E. coli* RNA polymerasepromoter interactions: alpha subunit binding to the UP element minor groove. *Genes & Development*, *15*, 491-506.

Ross, W., & Gourse, R. L. (2009). Analysis of RNA polymerase-promoter complex formation. *Methods*, *47*, 13-24.

Ross, W., Sanchez-Vazquez, P., Chen, A. Y., Lee, J. H., Burgos, H. L., & Gourse, R. L. (2016). PpGpp binding to a site at the RNAP-DksA interface accounts for its dramatic effects on transcription initiation during the stringent response. *Molecular Cell*, *62*, 811-823.

Rouillon, C., Athukoralage, J S., Graham, S., Grüschow, S. & White M. F. (2018). Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *eLife*, *7*, e36734.

Ruff, E. F., Thomas Record, M., & Artsimovitch, I. (2015). Initial events in bacterial transcription initiation. *Biomolecules*, *5*, 1035-1062.

Ruiz-Vázquez, R. M., & Murillo, F. J. (1984). Abnormal motility and fruiting behavior of *Myxococcus xanthus* bacteriophage-resistant strains induced by a clear-plaque mutant of bacteriophage Mx8. *Journal of Bacteriology*, *160*, 818-821.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press*.

Sambrook, J., & Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Samson, J. E., Magadán, A. H., Sabri, M., & Moineau, S. (2013). Revenge of the phages: Defeating bacterial defences. *Nature Reviews Microbiology*, *11*, 675-687.

Sánchez-Romero, M. A., Cota, I., & Casadesús, J. (2015). DNA methylation in bacteria: From the methyl group to the methylome. *Current Opinion in Microbiology*, 25,
9-16.

Schwartz, E. C., Shekhtman, A., Dutta, K., Pratt, M. R., Cowburn, D., Darst, S. A., & Muir, T. W. (2008). A full-length Group 1 Bacterial σ factor adopts a compact structure incompatible with DNA binding. *Chemistry and Biology*, *15*, 1091-1103.

Sefcikova, **J.**, **Roth**, **M.**, **Yu**, **G.**, **& Li**, **H.** (2017). Cas6 processes tight and relaxed repeat RNA via multiple mechanisms: a hypothesis. *BioEssays*, *39*.

Severinov, K., Fenyo, D., Severinova, E., Mustaev, A., Chait, B. T., Goldfarb, A., & Darst, S. A. (1994). The σ subunit conserved region 3 is part of «5'-face» of active center of *Escherichia coli* RNA polymerase. *Journal of Biological Chemistry*, 269, 20826-20828.

Sharp, M. M., Chan, C. L., Lu, C. Z., Marr, M. T., Nechaev, S., Merritt, E. W., ... Gross,
C. A. (1999). The interface of σ with core RNA polymerase is extensive, conserved, and functionally specialized. *Genes and Development*, *13*, 3015-3026.

Shimada, T., Fujita, N., Yamamoto, K., & Ishihama, A. (2011). Novel roles of camp receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS One*, *6*, e20081.

Shimada, T., Yamazaki, Y., Tanaka, K., & Ishihama, A. (2014). The whole set of constitutive promoters recognized by RNA polymerase RpoD holoenzyme of *Escherichia coli*. *PLoS One*, *9*, e90447.

Shimkets, L., & Woese, C. R. (1992). A phylogenetic analysis of the myxobacteria: basis for their classification. *Proceedings of the National Academy of Sciences USA*, *89*, 9459-9463.

Shin, M., Kang, S., Hyun, S. J., Fujita, N., Ishihama, A., Valentin-Hansen, P., & Choy,
H. E. (2001). Repression of deoP2 in *Escherichia coli* by CytR: conversion of a transcription activator into a repressor. *EMBO Journal*, 20, 5392-5399.

Shmakov, S., Smargon, A., Scott, D., Cox, D., Pyzocha, N., Yan, W., ... Koonin, E.
V. (2017). Diversity and evolution of class 2 CRISPR-Cas systems. *Nature Reviews Microbiology*, *15*, 169-182.

Shuler, M. F., Tatti, K. M., Wade, K. H., & Moran, C. P. (1995). A single amino acid substitution in σ^{E} affects its ability to bind core RNA polymerase. *Journal of Bacteriology*, *177*, 3687-3694.

Siegele, D. A., Hu, J. C., Walter, W. A., & Gross, C. A. (1989). Altered promoter recognition by mutant forms of the σ 70 subunit of *Escherichia coli* RNA polymerase. *Journal of Molecular Biology*, *206*, 591-603.

Silas, S., Lucas-Elio, P., Jackson, S. A., Aroca-Crevillén, A., Hansen, L. L., Fineran, P. C., ... Sánchez-Amat, A. (2017). Type III CRISPR-Cas systems can provide redundancy to counteract viral escape from type I systems. *eLife*, *6*, e27601.

Silas, S., Mohr, G., Sidote, D. J., Markham, L. M., Sanchez-Amat, A., Bhaya, D., ... Fire, A. Z. (2016). Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science*, *351*, 351, aad4234.

Sineva, E., Savkina, M., & Ades, S. E. (2017). Themes and variations in gene regulation by extracytoplasmic function (ECF) σ factors. *Current Opinion in Microbiology*, *36*, 128-137.

Singh, B. N. (1947). Number and lytic action on bacteria. *Journal of General Microbiology*, *1*, 1-10.

Skerker, J. M., & Berg, H. C. (2001). Direct observation of extension and retraction of type IV pili. *Proceedings of the National Academy of Sciences USA*, *98*, 6901-6904.

Sorek, R., Kunin, V., & Hugenholtz, P. (2008). CRISPR - A widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology*, *6*, 181-186.

Souza, B. M., Castro, T. L. D. P., Carvalho, R. D. D. O., Seyffert, N., Silva, A., Miyoshi, A., & Azevedo, V. (2014). σ-ECF factors of gram-positive bacteria: A focus on *Bacillus subtilis* and the CMNR group. *Virulence*, *5*, 587-600.

Sperandio, V., Mellies, J. L., Delahay, R. M., Frankel, G., Adam Crawford, J., Nguyen, W., & Kaper, J. B. (2000). Activation of enteropathogenic *Escherichia coli* (EPEC) LEE2 and LEE3 operons by Ler. *Molecular Microbiology*, *38*, 781-793.

Spröer, C., Reichenbach, H., & Stackebrandt, E. (1999). The correlation between morphological and phylogenetic classification of myxobacteria. *International Journal of Systematic Bacteriology*, *49*, 1255-1262.

Srivastava, D. B., Leon, K., Osmundson, J., Garner, A. L., Weiss, L. A., Westblade,
L. F., ... Campbell, E. A. (2013). Structure and function of CarD, an essential mycobacterial transcription factor. *Proceedings of the National Academy of Sciences USA*, *110*, 12619-12624.

Stallings, C. L., Stephanou, N. C., Chu, L., Hochschild, A., Nickels, B. E., & Glickman, M. S. (2009). CarD is an essential regulator of rRNA transcription required for *Mycobacterium tuberculosis* persistence. *Cell*, *138*, 146-159.

Staroń, A., Sofia, H. J., Dietrich, S., Ulrich, L. E., Liesegang, H., & Mascher, T.

(2009). The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. *Molecular Microbiology*, 74, 557-581.

Steele, D. L., El-Kabbani, O., Dunten, P., Windsor, L. J., Kammlott, R. U., Crowther, R. L., ... Birktoft, J. J. (2000). Expression, characterization and structure determination of an active site mutant (Glu202-Gln) of mini-stromelysin-1. *Protein Engineering*, *13*, 397-405.

Stern, A., Keren, L., Wurtzel, O., Amitai, G., & Sorek, R. (2010). Self-targeting by CRISPR: Gene regulation or autoimmunity? *Trends in Genetics*, *26*, 335-340.

Sun, H., Zusman, D. R., & Shi, W. (2000). Type IV pilus of *Myxococcus xanthus* is a motility apparatus controlled by the frz chemosensory system. *Current Biology*, *10*, 1143-1146.

Swarts, D. C., Hegge, J. W., Hinojo, I., Shiimori, M., Ellis, M. A., Dumrongkulraksa, J., ... Van Der Oost, J. (2015). Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Research*, *43*, 5120-5129.

Tagami, H., & Aiba, H. (1999). An inactive open complex mediated by an UP element

at Escherichia coli promoters. Proceedings of the National Academy of Sciences USA, 96, 7202-7207.

Thanos, D., & Maniatis, T. (1992). The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene. *Cell*, *71*, 777-789.

Trepreau, J., De Rosny, E., Duboc, C., Sarret, G., Petit-Hartlein, I., Maillard, A. P., ... Covés, J. (2011). Spectroscopic characterization of the metal-binding sites in the periplasmic metal-sensor domain of CnrX from *Cupriavidus metallidurans* CH34. *Biochemistry*, *50*, 9036-9045.

Trepreau, J., Girard, E., Maillard, A. P., De Rosny, E., Petit-Hartlein, I., Kahn, R., & Covès, J. (2011). Structural basis for metal sensing by CnrX. *Journal of Molecular Biology*, 408, 766-779.

Towbin, H., Staehelin, T., Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America,*

76,4350-4354.

Ueki, T., & Inouye, S. (1998). A new sigma factor, SigD, essential for stationary phase is also required for multicellular differentiation in *Myxococcus xanthus*. *Genes Cell*, *3*, 371-385.

Ueki, T., & Inouye, S. (2001). SigB, SigC, and SigE from *Myxococcus xanthus* homologous to σ 32 are not required for heat shock response but for multicellular differentiation. *Journal of Molecular Microbiology and Biotechnology*, *3*, 287-293.

Ueki, T., Xu, C. Y., & Inouye, S. (2005). SigF, a new sigma factor required for a motility system of *Myxococcus xanthus*. *Journal of Bacteriology*, *187*, 8537-8541.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., & Leunissen, J.A. (2007). Primer3Plus, an enchanced web interface to Primer3. *Nucleic Acids Res, 35*, W71-74.

Van der Woude, M. W., & Henderson, I. R. (2008). Regulation and Function of Ag43 (Flu). *Annual Review of Microbiology*, 62, 153-169.

Van Orden, M. J., Klein, P., Babu, K., Najar, F. Z., & Rajan, R. (2017). Conserved DNA motifs in the type II-A CRISPR leader region. *PeerJ*, *5*, 3161.

Viswanathan, P., Murphy, K., Julien, B., Garza, A. G., & Kroos, L. (2007). Regulation of dev, an operon that includes genes essential for *Myxococcus xanthus* development and CRISPR-associated genes and repeats. *Journal of Bacteriology*, *189*, 3738-3750.

Waldburger, C., Gardella, T., Wong, R., & Susskind, M. M. (1990). Changes in conserved region 2 of *Escherichia coli* σ 70 affecting promoter recognition. *Journal of Molecular Biology*, *215*, 267-76.

Wall, D., & Kaiser, D. (1999). Type IV pili and cell motility. *Molecular Microbiology*, 32, 01-10.

Wall, D., Wu, S. S., & Kaiser, D. (1998). Contact stimulation of Tgl and type IV pili in *Myxococcus xanthus*. *Journal of Bacteriology*, *180*, 759-761.

Wallace, R. A., Black, W. P., Yang, X., & Yang, Z. (2014). A CRISPR with roles in

Myxococcus xanthus development and exopolysaccharide production. *Journal of Bacteriology*, *196*, 4036-4043.

Waltersperger, S., Widmer, C., Wang, M., & Baumann, U. (2010). Crystal structure of archaemetzincin AmzA from *Methanopyrus kandleri* at 1.5 Å resolution. *Proteins: Structure, Function and Bioinformatics*, 78, 2720-2723.

Ward, M. J., Lew, H., Treuner-Lange, A., & Zusman, D. R. (1998). Regulation of motility behavior in *Myxococcus xanthus* may require an extracytoplasmic-function σ factor. *Journal of Bacteriology*, *180*, 5668-5675.

Wecke, T., Halang, P., Staroń, A., Dufour, Y. S., Donohue, T. J., & Mascher, T. (2012). Extracytoplasmic function σ factors of the widely distributed group ECF41 contain a fused regulatory domain. *MicrobiologyOpen*, *1*, 194-213.

Weiss, L. A., Harrison, P. G., Nickels, B. E., Glickman, M. S., Campbell, E. A., Darst, S. A., & Stallings, C. L. (2012). Interaction of CarD with RNA polymerase mediates *Mycobacterium tuberculosis* viability, rifampin resistance, and pathogenesis. *Journal of Bacteriology*, *194*, 5621-5631.

Westblade, L. F., Campbell, E. A., Pukhrambam, C., Padovan, J. C., Nickels, B. E., Lamour, V., & Darst, S. A. (2010). Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction. *Nucleic Acids Research*, *38*, 8357-8369.

Westra, E. R., Pul, Ü., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., ... Brouns, S. J. J. (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Molecular Microbiology*, 77, 1380-1393.

Wielgoss, S., Didelot, X., Chaudhuri, R. R., Liu, X., Weedall, G. D., Velicer, G. J., & Vos, M. (2016). A barrier to homologous recombination between sympatric strains of the cooperative soil bacterium *Myxococcus xanthus*. *ISME Journal*, *10*, 2468-2477.

Wolgemuth, C., Hoiczyk, E., Kaiser, D., & Oster, G. (2002). How myxobacteria glide. *Current Biology*, *12*, 369-377.

Yanisch-Perron, C., Vieira, J., & Messing, J. (1985). Improved M13 phage cloning vectors and host strain: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene, 33*, 103-119.

Yhang, Z., & Higgs, P. I. (2014). Myxobacteria. Genomics, cellular and molecular biology. *Caister Academic Press*.

Yie, J., Liang, S., Merika, M., & Thanos, D. (1997). Intra- and intermolecular cooperative binding of high-mobility-group protein I(Y) to the beta-interferon promoter. *Molecular and Cellular Biology*, *17*, 3649-3662.

Yu, R., & Kaiser, D. (2007). Gliding motility and polarized slime secretion. *Molecular Microbiology*, 63, 454-467.

Yuzenkova, Y., Tadigotla, V. R., Severinov, K., & Zenkin, N. (2011). A new basal promoter element recognized by RNA polymerase core enzyme. *EMBO Journal, 30*, 3766-3775.

Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M.,

... Mustaev, A. (1996). Mapping of catalytic residues in the RNA polymerase active center. *Science*, 273, 107-109.

Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., & Darst, S. A. (1999). Crystal structure of thermus aquaticus core RNA polymerase at 3.3 Å resolution. *Cell*, 98, 811-824.

Zhang, N., & Buck, M. (2015). A perspective on the enhancer dependent bacterial RNA polymerase. *Biomolecules*, *5*, 1012-1019.

Zhang, Y., Feng, Y., Chatterjee, S., Tuske, S., Ho, M. X., Arnold, E., & Ebright, R. H. (2012). Structural basis of transcription initiation. *Science*, *338*, 1076-1080.

Zheleznova Heldwein, E. E., & Brennan, R. G. (2001). Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature*, *409*, 378-382.

Zheng, D., Constantinidou, C., Hobman, J. L., & Minchin, S. D. (2004). Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Research*, *32*, 5874-5893.

Zuber, P., Healy, J., Carter, H. L., Cutting, S., Moran, C. P., & Losick, R. (1989). Mutation changing the specificity of an RNA polymerase sigma factor. *Journal of Molecular Biology*, *206*, 605-614.