

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I



CARACTERIZACIÓN MOLECULAR DE LA DEGRADACIÓN
ANAERÓBICA DE TOLUENO Y *m*-XILENO EN *Azoarcus* sp. CIB

TESIS DOCTORAL DE:

BLAS BLÁZQUEZ CASTIÑEIRA

BAJO LA DIRECCIÓN DE:

**EDUARDO DÍAZ
MANUEL CARMONA**

Madrid

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I



**CARACTERIZACIÓN MOLECULAR DE LA DEGRADACIÓN
ANAERÓBICA DE TOLUENO Y *m*-XILENO EN *Azoarcus* sp. CIB**

TESIS DOCTORAL

BLAS BLÁQUEZ CASTIÑEIRA

Madrid, 2009

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I



**CARACTERIZACIÓN MOLECULAR DE LA DEGRADACIÓN
ANAERÓBICA DE TOLUENO Y *m*-XILENO EN *Azoarcus* sp. CIB**

TESIS DOCTORAL
BLAS BLÁQUEZ CASTIÑEIRA

DIRECTORES:

Dr. EDUARDO DÍAZ

Dr. MANUEL CARMONA

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
CENTRO DE INVESTIGACIONES BIOLÓGICAS

Madrid, 2009

La biología molecular ha mostrado que hasta los más sencillos de todos los sistemas vivos en la Tierra hoy, las células bacterianas, son tremendamente complejas.

Michael Denton (1943-).
Wikiquote.

A mi familia

Agradecimientos

Quisiera expresar mi agradecimiento a todas las personas que han colaborado en la realización de esta Tesis.

ÍNDICE

ABREVIATURAS	v
I. INTRODUCCIÓN	1
1. Los compuestos aromáticos.	1
2. Rutas de degradación de compuestos aromáticos.	2
2.1 Rutas catabólicas aeróbicas de compuestos aromáticos.	3
2.1.1. Rutas aeróbicas clásicas.	4
2.1.2. Rutas aeróbicas híbridas.	5
2.2 Rutas catabólicas anaeróbicas de compuestos aromáticos.	6
2.2.1. La ruta central del benzoil-CoA.	8
2.2.1.1. Ruta alta del benzoil-CoA.	10
2.2.1.2. Ruta baja del benzoil-CoA.	11
3. Regulación de las rutas de degradación de compuestos aromáticos.	12
3.1. Regulación específica.	13
3.2. Regulación global o sobreimpuesta.	15
4. La degradación anaeróbica del tolueno y xilenos.	16
4.1. Ruta de la degradación anaeróbica del tolueno.	17
4.2. Ruta de la degradación anaeróbica de los xilenos.	19
II. OBJETIVOS	21
III. MATERIALES Y MÉTODOS	23
1. Cepas bacterianas.	23
2. Plásmidos	23
3. Medios y condiciones de cultivo.	25
4. Ensayos de quimiotaxis.	27
4.1. Ensayo de quimiotaxis a tiempos cortos (<i>Agarose Plug Assay</i>).	27
4.2. Ensayo de quimiotaxis en placa de cultivo.	28
5. Ensayo de formación de biopelículas en placas multipocillo.	28
6. Técnicas de manipulación de ADN.	29
6.1. Aislamiento del ADN.	29
6.2. Reacción de amplificación en cadena con ADN polimerasa termorresistente (PCR).	29
6.3. Búsqueda de genes en la genoteca fágica de <i>Azoarcus</i> sp. CIB.	33

6.4. Secuenciación de ADN.	33
7. Extracción de ARN y ensayos de retrotranscripción-PCR (RT-PCR).	33
8. Determinación del sitio de inicio de la transcripción.	34
9. Procedimientos de transferencia genética.	35
9.1. Transformación y electroporación.	35
9.2. Transferencia de plásmidos por conjugación.	35
10. Disrupción insercional de genes cromosómicos de <i>Azoarcus</i> sp. CIB	36
11. Técnicas de manipulación de proteínas.	36
11.1. Obtención de los extractos proteicos.	36
11.2. Electroforesis en geles de poliacrilamida-SDS (SDS-PAGE).	37
11.3. Modelado de la proteína glutaril-CoA deshidrogenasa de <i>Azoarcus</i> sp. CIB.	37
12. Análisis de los datos de secuencia.	37
13. Ensayos enzimáticos.	38
13.1. Ensayo de actividad β -galactosidasa.	38
13.2. Ensayo de actividad glutaril-CoA deshidrogenasa.	39
IV. RESULTADOS	41
1. Estudio de los determinantes genéticos de la ruta periférica para el catabolismo anaeróbico de tolueno y <i>m</i> -xileno en <i>Azoarcus</i> sp. CIB.	41
1.1. Crecimiento anaeróbico de <i>Azoarcus</i> sp. CIB en tolueno y <i>m</i> -xileno.	41
1.2. Identificación del cluster génico responsable de la ruta periférica para el catabolismo anaeróbico del tolueno en <i>Azoarcus</i> sp. CIB	42
1.3. Confirmación de la implicación del <i>cluster bss/bbs</i> de <i>Azoarcus</i> sp. CIB en la degradación anaeróbica de tolueno.	45
1.4. Análisis de la secuencia del <i>cluster bss/bbs</i> de <i>Azoarcus</i> sp. CIB.	46
1.5. El gen <i>tol</i> .	48
1.6. El cluster <i>bss/bbs</i> está implicado en el catabolismo anaeróbico de <i>m</i> -xileno en <i>Azoarcus</i> sp. CIB.	53
1.7. Análisis transcripcional del cluster de degradación de tolueno y <i>m</i> -xileno en <i>Azoarcus</i> sp. CIB.	54
1.7.1. Análisis de las regiones promotoras <i>PtdiS</i> , <i>PbssD</i> y <i>PbbsA</i> .	55
1.7.2. Estudio de la expresión del cluster génico <i>bss/bbs</i> en <i>Azoarcus</i> sp. CIB.	56

1.7.3. Identificación de una cuarta unidad transcripcional en el <i>cluster bss/bbs</i> .	58
1.8. Estudio de la regulación transcripcional de los operones catabólicos <i>bss</i> y <i>bbs</i> .	59
1.8.1. Clonación y actividad promotora de la región <i>PbbsA</i> en <i>E. coli</i> .	60
1.8.2. Actividad de los promotores <i>PbbsA</i> y <i>PbssD</i> en <i>Azoarcus</i> sp. CIB.	62
1.8.3. TdiR es un activador transcripcional de los genes <i>bss</i> y <i>bbs</i> .	64
1.8.4. Caracterización de los inductores del operón catabólico <i>bbs</i> .	65
1.8.5. Caracterización de los inductores del operón catabólico <i>bss</i> .	67
1.9. Regulación sobreimpuesta de los genes <i>bbs</i> en <i>Azoarcus</i> sp. CIB.	68
1.9.1. Represión catabólica de los genes <i>bbs</i> .	68
1.9.2. Expresión de los genes <i>bbs</i> en presencia de oxígeno.	69
2. Caracterización y regulación del gen <i>gcdH</i> implicado en la ruta baja de degradación anaeróbica de tolueno en <i>Azoarcus</i> sp. CIB.	70
2.1. Estudio de la inducción de la actividad glutaril-CoA deshidrogenasa (GcdH) de <i>Azoarcus</i> sp. CIB.	70
2.2. Identificación del gen <i>gcdH</i> de <i>Azoarcus</i> sp. CIB.	71
2.3. Clonación y caracterización del gen <i>gcdH</i> de <i>Azoarcus</i> sp. CIB	73
2.4 Estudio de la regulación del gen <i>gcdH</i> de <i>Azoarcus</i> sp. CIB.	77
2.4.1. Estudio de la expresión del gen <i>gcdH</i> en <i>Azoarcus</i> sp. CIB	77
2.4.2. El gen <i>gcdR</i> codifica un activador transcripcional del gen <i>gcdH</i> .	78
2.4.3. Caracterización del sistema regulador GcdR/Pg en <i>E. coli</i> .	79
2.4.4 Determinación del sitio de inicio de la transcripción en el promotor Pg.	82
2.5. Clonación y regulación de otros genes <i>gcdH</i> bacterianos: el gen <i>gcdH</i> de <i>P. putida</i> .	84
V. ANEXO I	87
VI. ANEXO II	99

VII. DISCUSIÓN.	101
1. Identificación y caracterización genética del cluster responsable de la ruta periférica de degradación anaeróbica de tolueno y <i>m</i> -xileno en <i>Azoarcus</i> sp. CIB.	103
2. Análisis del gen <i>tol</i> .	145
3. Genómica comparada y relaciones evolutivas de los clusters <i>bss/bbs</i> en diferentes microorganismos.	148
4. Regulación del cluster de degradación anaeróbica de tolueno y <i>m</i> -xileno en <i>Azoarcus</i> sp. CIB.	151
5. Caracterización del gen <i>gcdH</i> de la ruta baja de degradación anaeróbica de tolueno en <i>Azoarcus</i> sp. CIB	151
6. Caracterización de la regulación del gen <i>gcdH</i> .	152
VII. CONCLUSIONES.	139
IX. BIBLIOGRAFÍA.	141

ABREVIATURAS

A_{λ}	Absorbancia medida a λ nm
ADN	Ácido desoxirribonucleico
ADNc	ADN complementario
ADP	Adenosina difosfato
cAMP	Adenosín monofosfato cíclico
ARN	Ácido ribonucleico
ARNasa	Ribonucleasa
ARNm	ARN mensajero
ARNP	ARN polimerasa
Ap^r	Resistencia a ampicilina
ATP	Adenosina trifosfato
BSA	Seroalbúmina bovina
CMI	Concentración mínima inhibitoria
Cm^r	Resistencia a cloranfenicol
CoA	Coenzima A
C-terminal	Carboxilo terminal
Da	Dalton
DGC	Dominio diguanilato ciclasa
DH	Deshidrogenasa
di-GMPc	Guanosina difosfato dicíclico
DMSO	Dimetilsulfóxido
dNTP	Desoxinucleótido trifosfato
ϵ	Coefficiente de extinción molar
EAL	Dominio fosfodiesterasa de di-GMPc
EDTA	Ácido etilendiaminotetracético
FAD	Dinucleótido de flavina y adenina oxidado
FADH₂	Dinucleótido de flavina y adenina reducido
FMN	Mononucleótido de flavina
Gm^r	Resistencia a gentamicina
GTP	Guanosina trifosfato
HEPES	Ácido 4-(2-hidroxietil)-1-piperacinetanosulfónico
HK	Dominio histidín-quinasa de proteínas sensoras

HMN	2,2,4,4,6,8,8-heptametilnonano
HTH	Dominio hélice-giro-hélice
IPTG	Isopropil- β -D-tiogalato piranosido
kb	1000 pares de bases
kDa	1000 dalton
Km^r	Resistencia a kanamicina
LB	Medio de cultivo Luria-Bertani
M63	Medio mínimo M63
MC	Medio mínimo MC
MCS	Sitio de clonación múltiple
NAD	Nicotinamina-adenina-dinucleótido oxidado
NADH	Nicotinamina-adenina-dinucleótido reducido
NADPH	Fosfato de Nicotinamina-adenina-dinucleótido reducido
NCBI	<i>National Center for Biotechnology Information</i>
nt	nucleótido(s)
N-terminal	Amino terminal
ORF	Marco abierto de lectura
PAGE	Electroforesis en gel de poliacrilamida
PAS	Dominio sensor Per-Arnt-Sim
pb	Pares de bases
p/v	relación peso/volumen
PCR	Reacción de amplificación en cadena con ADN polimerasa termorresistente
PDB	<i>Protein Data Bank</i>
PEG	Polietilenglicol
RBS	Secuencia de unión a ribosoma
REC	Dominio receptor de respuesta (<i>Receiver</i>) en sistemas reguladores de dos componentes
REP	Secuencia palindrómica repetitiva extragénica
Rf^r	Resistencia a rifampicina
RNAP	ARN polimerasa
RNAsa	Ribonucleasa
RT-PCR	Reacción de retrotranscripción acoplada a PCR
σ	Subunidad sigma de la RNA polimerasa

Sm^r	Resistencia a estreptomicina
SDS	Dodecilsulfato sódico
Tris	Trihidroximetilaminometano
U	Unidad de actividad enzimática
v/v	relación volumen/volumen
X-gal	5-Bromo-4-cloro-3-indolil- β -D-galactopiranosido

I. INTRODUCCIÓN

1. Los compuestos aromáticos.

Los compuestos aromáticos, que incluyen estructuras homocíclicas (derivadas del benceno) o heterocíclicas (N-; S-; O-heterociclos), constituyen, después de los carbohidratos, la clase de compuestos orgánicos más abundante en la naturaleza (Adler, 1977). Así el 25% de la biomasa de la Tierra está constituida por lignina, polímero de compuestos aromáticos. Otras fuentes naturales de compuestos aromáticos son los flavonoides, taninos, aminoácidos aromáticos y los combustibles fósiles (Gibson y Harwood, 2002). Un buen número de compuestos aromáticos son producidos en grandes cantidades para uso industrial, e. g., fabricación de disolventes, pesticidas, plásticos, fibras sintéticas, etc (Swoboda-Colberg, 1995). Durante el último siglo se ha producido un aumento considerable en la liberación de compuestos aromáticos a la biosfera debido a esta actividad industrial (OCDE, 1994), teniendo algunos de ellos estructuras o sustituyentes que no se encuentran en la naturaleza de forma habitual (compuestos xenobióticos). Estos compuestos xenobióticos no son degradados, o lo hacen muy lentamente, por lo que constituyen una importante fuente de contaminación ambiental (Pieper y Reineke, 2000). Algunos ejemplos de estos compuestos xenobióticos son, por ejemplo, los derivados polihalogenados de las dioxinas y los bifenilos policlorados (PCBs), los compuestos nitroaromáticos y los derivados fenólicos halogenados.

Dada la gran abundancia de los compuestos aromáticos, su reciclaje es vital para el mantenimiento del ciclo del carbono en el planeta (Harwood y Parales, 1996). Pese a la gran estabilidad termodinámica que le otorga la configuración electrónica del anillo bencénico a estas moléculas, las bacterias han desarrollado estrategias para la mineralización de la mayoría de estos compuestos presentes en la naturaleza (Gibson y Harwood, 2002; Vaillancourt y col., 2004; Díaz, 2004). Estas estrategias de mineralización también pueden ser usadas por las bacterias para la degradación, al menos parcial, de los compuestos xenobióticos que comparten estructuras similares a los compuestos de origen natural (Díaz, 2004; McLeod y Eltis, 2008). La aplicación del enorme potencial catabólico de las bacterias frente a los compuestos aromáticos ha permitido el desarrollo de líneas de investigación que tienen también gran interés socio-económico, tales como la biorremediación, biotransformación, desarrollo de biosensores, etc (Díaz, 2004).

2. Rutas de degradación de compuestos aromáticos.

Aunque la mineralización de los compuestos aromáticos supone un reto importante para los microorganismos, dada la estabilidad de dichos compuestos, los microorganismos han desarrollado soluciones que implican una gran variedad de “rutas periféricas” encargadas de activar los diversos sustratos aromáticos a un número limitado de intermediarios centrales, los cuales posteriormente perderán su carácter aromático y se transformarán en compuestos del metabolismo primario de la célula a través de las denominadas “rutas centrales” (Heider y Fuchs, 1997; Heider y col., 1998; Harwood y col., 1999; Pieper y col., 2004). A esta estrategia de canalizar un gran número de rutas periféricas distintas hacia unas pocas rutas centrales se le ha denominado “embudo catabólico” (Harayama y Timmis, 1989).

El catabolismo de los compuestos aromáticos implica dos procesos secuenciales: 1) modificación de la densidad electrónica del anillo aromático, y 2) ruptura del anillo. Estos procesos suponen la aplicación de reacciones de oxidación o reducción dependiendo de si el metabolismo se lleva a cabo en condiciones aeróbicas o anaeróbicas, respectivamente (Heider y Fuchs, 1997; Díaz, 2004). Dependiendo del aceptor final de los electrones, los compuestos aromáticos están sometidos a dos tipos básicos de catabolismo. Si el aceptor es una molécula de oxígeno, se trata de una degradación aeróbica. Si el aceptor final de los electrones es otra molécula distinta al oxígeno, la degradación es de tipo anaeróbica. Estos catabolismos presentan características diferenciales, así en el metabolismo aeróbico generalmente se produce la desestabilización del anillo aromático a través de reacciones de oxidación que introducen grupos hidroxilo, mientras que en el catabolismo anaeróbico la desestabilización del anillo aromático se lleva a cabo generalmente mediante la formación de un derivado tioesterificado de coenzima A (CoA) que favorece las posteriores reacciones de reducción del anillo.

La degradación aeróbica de compuestos aromáticos es el mecanismo de mineralización mejor estudiado, en ella la molécula de oxígeno no sólo constituye el aceptor final de electrones sino que también es el reactivo necesario para la actuación de las oxigenasas que participan en la activación y ruptura del anillo aromático. Los diferentes sustratos son generalmente convertidos en derivados mono-, di-, o tri-hidroxilados mediante la incorporación de átomos de oxígeno, lo que permite la activación termodinámica del anillo aromático. En una siguiente etapa, estos intermediarios hidroxilados sufren la apertura de la molécula, catalizada por dioxigenasas de apertura del

anillo, y la consiguiente pérdida de aromaticidad. En las rutas aeróbicas híbridas, la activación del anillo aromático implica la tioesterificación con CoA y la apertura del anillo no conlleva la participación de dioxigenasas de ruptura. Finalmente, los productos de apertura son fácilmente oxidados hasta intermediarios del metabolismo central de la célula (van der Meer, 1997; Díaz y col., 2001; Jiménez y col., 2002).

En la degradación anaeróbica de compuestos aromáticos, el metabolismo bacteriano requiere de otros aceptores finales de electrones diferentes al oxígeno. Existen numerosos ecosistemas en los que el aporte de oxígeno se encuentra limitado como ocurre en los suelos con difícil drenaje, las aguas estancadas y subterráneas, las aguas que tienen una elevada actividad bacteriana, los sedimentos marinos, el tracto digestivo de animales, etc. (Widdel y Rabus, 2001). En estos ambientes anaeróbicos la acumulación de compuestos aromáticos es muy abundante y su degradación por parte de los microorganismos es un proceso de especial importancia. Los microorganismos en estas condiciones emplean como aceptores finales de electrones moléculas inorgánicas tales como el NO_3^- o NO_2^- que son reducidos hasta N_2 (desnitrificación) (Coates y col., 2001b), el Fe^{3+} que es reducido a Fe^{2+} (reducción de hierro) (Lack y col., 2002), el SO_4^{2-} que es reducido a SO_2^{2-} o S_2 (reducción de sulfato) (Morasch y col., 2004; Peters y col., 2004) o el CO_2 que es reducido a metano (CH_4) (metanogénesis) (Spormann y Widdel, 2000). Otros compuestos que también pueden actuar como aceptores finales de electrones son el Mn^{4+} , ClO_4^- , Cr^{4+} , V^{4+} , óxido de trimetilamina, dimetilsulfóxido, fumarato, etc. (Gibson y Harwood, 2002; Coates y col., 2002; Lovley, 2003; Chakraborty y Coates, 2004). Algunos organismos fermentadores también pueden degradar compuestos aromáticos, aunque generalmente se acoplan a otros organismos metanogénicos o reductores de sulfato para hacer el proceso más eficiente (Elshahed y col., 2001). Finalmente, algunos organismos fototrofos son capaces de utilizar anaeróbicamente compuestos aromáticos como única fuente de carbono, utilizando la luz solar como fuente de energía (Gibson y Harwood, 2002).

2.1 Rutas catabólicas aeróbicas de compuestos aromáticos.

Las rutas de catabolismo aeróbico de compuestos aromáticos han sido ampliamente estudiadas en multitud de bacterias así como en hongos (revisiones en van der Meer y col., 1992; Harwood y Parales, 1996; Pieper y col., 2004).

2.1.1. Rutas aeróbicas clásicas.

La estrategia clásica de mineralización aeróbica del anillo aromático se fundamenta en la desestabilización del mismo mediante rutas periféricas que implican la incorporación de grupos hidroxilo. Con esta adición de grupos hidroxilo la estructura electrónica del anillo aromático se ve aumentada y es más susceptible de ser modificada por oxigenasas de ruptura que catalizan su conversión a un compuesto alifático, el cual será oxidado progresivamente hasta obtener intermedios del ciclo de Krebs. La incorporación de los grupos hidroxilo en el anillo bencénico lo realizan las oxigenasas, monooxigenasas y/o dioxigenasas según incorporen uno o dos grupos hidroxilo (Gibson y Parales, 2000), generándose así derivados aromáticos mono-, di-, o tri-hidroxilados (e. g. catecol, protocatecuato, homoprotocatecuato, gentisato, homogentisato, etc.) (Fig. 1). Las enzimas de las “rutas periféricas” son más o menos específicas de sus sustratos y generalmente se inducen por dichos sustratos (Heider y Fuchs, 1997). Otra característica frecuente es que un mismo compuesto aromático puede degradarse a través de distintas rutas catabólicas como ocurre, por ejemplo, con el tolueno (Fig. 1).

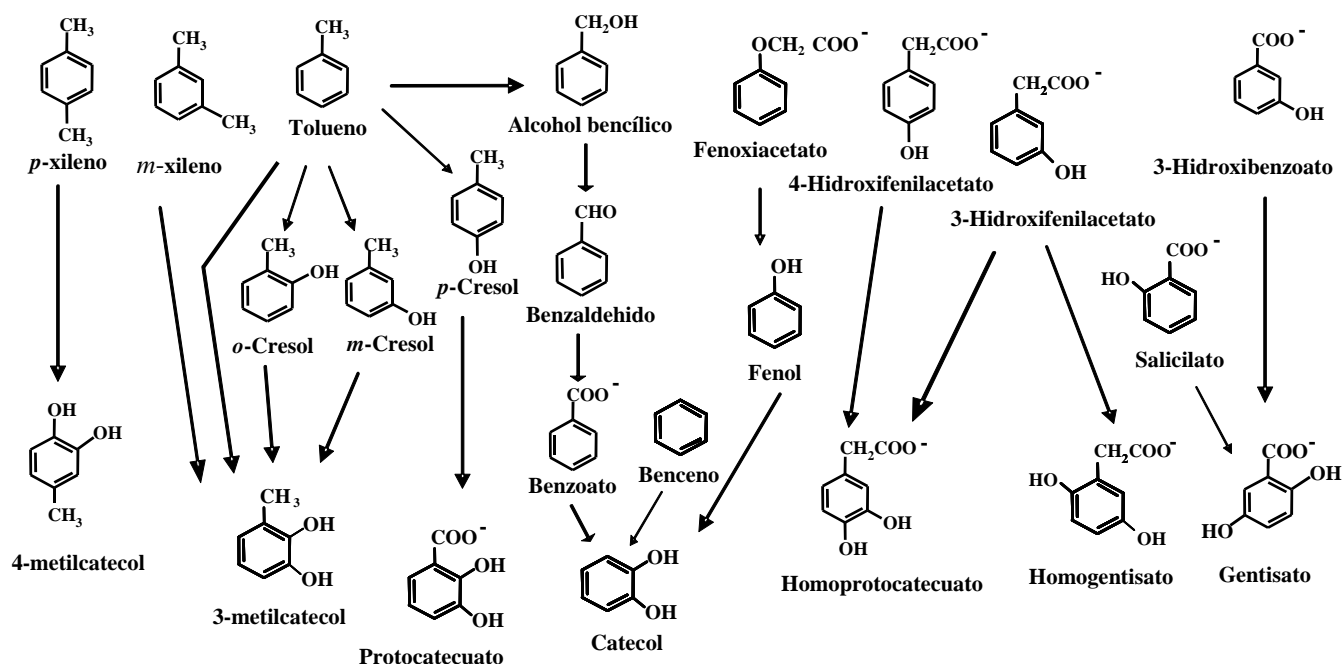


Figura 1. Esquema de algunas de las rutas periféricas e intermediarios centrales del catabolismo aeróbico bacteriano de compuestos aromáticos. Se muestran algunos de los intermediarios centrales dihidroxilados más frecuentes: el catecol y sus derivados (3-metilcatecol, 4-metilcatecol, protocatecuato y homoprotocatecuato), el gentisato y el homogentisato. El tolueno puede degradarse aeróbicamente por diversas rutas catabólicas que se indican (figura modificada de Bartolomé-Martín y col., 2006)

Los intermediarios centrales sufren, a través de las “rutas centrales”, la apertura del anillo aromático mediante dioxigenasas de ruptura de anillo (Harayama y Reikik, 1989; Harayama y Timmis, 1992; van der Meer y col., 1992; Harwood y Parales, 1996). La ruptura del anillo se puede producir entre los dos grupos hidroxilo, denominada ruptura *orto* y catalizada por las intradiol dioxigenasas (dependientes de Fe_{III}), o puede producirse por una ruptura adyacente a uno de los grupos hidroxilo, denominada en este caso ruptura *meta* y catalizada por las extradiol dioxigenasas (dependientes de Fe_{II}). Se han descrito también compuestos monohidroxilados sustrato de las dioxigenasas de ruptura, como el 2-aminofenol (Takenaka y col., 1997) y el 1-hidroxi-2-naftoato (Iwabuchi y Harayama, 1998), así como trihidroxilados, como el hidroxiquinol (Armengaud y col., 1999) y el galato (Nogales y col., 2005).

2.1.2. Rutas aeróbicas híbridas.

En el catabolismo aeróbico de compuestos aromáticos se han descrito algunas rutas que combinan características del catabolismo aeróbico y del anaeróbico, y se denominan rutas híbridas. En estas rutas los sustratos son transformados en derivados de CoA mediante una tioesterificación mediada por enzimas denominadas CoA ligasas. Este proceso de tioesterificación es una característica típica de las rutas anaeróbicas de degradación de compuestos aromáticos, ya que facilita la desestabilización del anillo aumentando su densidad electrónica (Heider y Fuchs, 1997). Para que se lleve a cabo el proceso de tioesterificación mediante la actividad de las enzimas CoA ligasas es necesario CoA, Mg²⁺ y ATP (Boll y Fuchs, 1995). La apertura del anillo aromático y pérdida de aromaticidad del derivado CoA se lleva a cabo mediante reacciones de oxigenación u oxigenación/reducción y posterior actuación de enzimas que atacan el compuesto alicíclico generado (Gescher y col., 2005).

La participación de enzimas CoA ligasas en este tipo de catabolismo se ha descrito para la degradación del ácido fenilacético en distintas bacterias (Luengo y col., 2001; Bartolomé-Martín y col., 2004); para la degradación del 2-aminobenzoato en *Azoarcus evansii* (Schühle y col., 2001); y para la degradación del benzoato en *Thauera aromatica* (Schühle y col., 2003), *Burkholderia xenovorans* LB400 (Denef y col., 2004) y *A. evansii*, siendo esta última ruta aeróbica híbrida la mejor estudiada hasta la fecha (Fig. 2) (Gescher y col., 2005; Gescher y col., 2006). También han sido descritas rutas híbridas en el metabolismo aeróbico de compuestos aromáticos heterocíclicos, como en la degradación

del ácido isonicotínico por *Mycobacterium* sp. INA1 (Kretzer y col., 1993) y en la degradación de los ácidos 2-furano y 2-tiofeno carboxilato en algunas bacterias (Koenig y Andreesen, 1990; Cripps, 1973).

El uso de estas rutas híbridas puede representar una ventaja en los microorganismos anaerobios facultativos que vivan en ambientes en los que la concentración de oxígeno sufra importantes variaciones, ya que el intermediario central derivado de CoA formado sería sustrato tanto para las rutas aeróbicas como para las anaeróbicas (Niemetz y col., 1995).

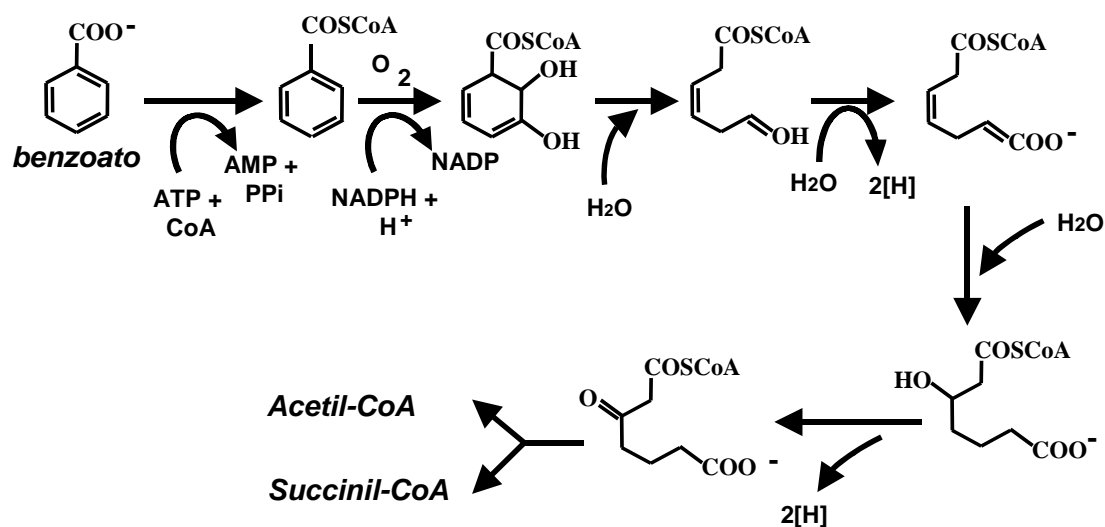


Figura 2. Esquema de la degradación aeróbica de benzoato en la cepa *Azoarcus Evansii*. Se muestran los distintos intermediarios que se forman en la degradación aeróbica de benzoato hasta acetil-CoA y succinil-CoA en la bacteria *A. Evansii*. Figura modificada de Gescher y col., 2005.

2.2 Rutas catabólicas anaeróbicas de compuestos aromáticos.

A diferencia de la degradación aeróbica, los procesos relacionados con el catabolismo anaeróbico de compuestos aromáticos sólo han sido estudiados en las últimas décadas y todavía se desconocen muchos aspectos de la bioquímica y, especialmente, de la genética que controla este tipo de catabolismo. En gran parte, esto es debido a la gran dificultad que conlleva el aislamiento y manipulación de cultivos bacterianos puros capaces de utilizar anaeróticamente compuestos aromáticos como única fuente de carbono. Hasta la fecha los estudios sobre el catabolismo anaeróbico de compuestos aromáticos se han realizado en unos pocos microorganismos tales como las α -proteobacterias *Rhodopseudomonas palustris* (Harwood y col., 1999; Eglund y col., 2001), y *Magnetospirillum* spp. (Shinoda y col., 2005); las β -proteobacterias *Thauera aromatica*

(Leuthner y Heider, 2000; Philipp y Schink, 2000), y *Azoarcus* spp. (Anders y col., 1995; Hirsch y col., 1998; Rabus, 2005; López-Barragán y col., 2004a); las δ -proteobacterias reductoras de hierro (*Geobacter* spp. (Coates y col., 2001a)) y de azufre (*Desulfococcus multivorans* (Peters y col., 2004)) o la bacteria fermentadora *Syntrophus aciditrophicus* (Elshahed y col., 2001). No obstante, los estudios de metagenómica revelan que la capacidad de degradar anaeróbicamente compuestos aromáticos está mucho más extendida en los ecosistemas naturales de lo que inicialmente se creía, lo que refuerza su papel biológico esencial en el ciclo del carbono y permite augurar el descubrimiento de numerosas nuevas rutas (genes y enzimas) en un futuro próximo (Foght, 2008).

Las rutas periféricas de degradación anaeróbica catalizan la conversión de numerosos compuestos monoaromáticos a unos pocos intermediarios centrales como son el resorcinol (1,3-dihidroxibenceno), 6-hidroxinicotinato, floroglucinol (1,3,5-trihidroxibenceno), hidroxihidroquinona (1,2,4-trihidroxibenceno), 2-aminobenzoil-CoA, 3-hidroxibenzoil-CoA, 3-metilbenzoil-CoA y el benzoil-CoA (Heider y Fuchs, 1997; Harwood y col., 1999; Gibson y Harwood, 2002; Alhapel y col., 2006; Fuchs, 2008) (Fig. 3). Un mismo compuesto aromático puede ser metabolizado por diferentes rutas catabólicas en función de la cepa bacteriana y del tipo de respiración que esté utilizando (Schink y col., 2000). Entre los intermediarios centrales, el más común en la degradación anaeróbica es el benzoil-CoA (Harwood y col., 1999), ya que en él convergen diferentes rutas periféricas a través de las cuales son degradados un gran variedad de compuestos tales como el etilbenceno, bencilalcohol, *p*-cresol, tolueno, fenol, fenilacetato, fenilpropionato, fenilalanina, vanilina, anilina, benzoato, y sus derivados hidroxilados o aminados, así como moléculas aromáticas halogenadas (Heider y Fuchs, 1997; Harwood y col., 1999; Schink y col., 2000). Los compuestos floroglucinol, resorcinol, 6-hidroxinicotinato e hidroxihidroquinona, presentan menor aromaticidad que el benzoil-CoA, lo que permite que no se requiera la activación mediada por CoA, siendo la reducción de su anillo aromático independiente de ATP y producida por enzimas que utilizan directamente NAD(P)H, en el caso del floroglucinol y la hidroxiquinona (Haddock y Ferry, 1989; Schink y col., 2000), o ferredoxinas, en el caso del resorcinol (Kluge y col., 1990) y 6-hidroxinicotinato (Alhapel y col., 2006).

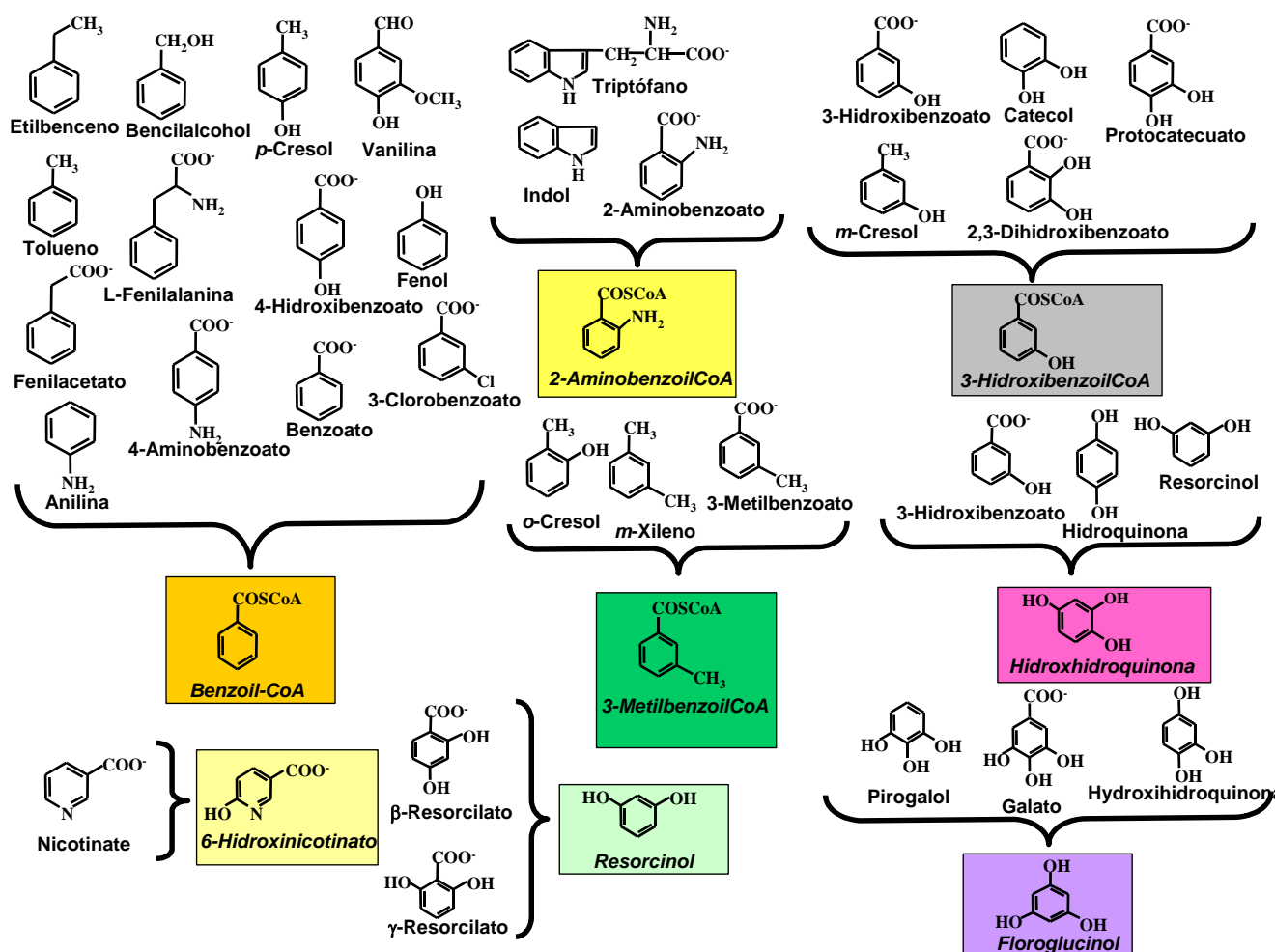


Figura 3. Esquema del catabolismo anaeróbico bacteriano de compuestos monoaromáticos. En diferentes colores se muestran los intermediarios centrales del catabolismo anaeróbico: benzoil-CoA, 2-aminobenzoil-CoA, 3-hidroxibenzoil-CoA, 3-metilbenzoil-CoA, hidroxihidroquinona, 6-hidroxinicotinato, resorcinol y floroglucinol.

2.2.1. La ruta central del benzoil-CoA.

Al ser el benzoil-CoA el intermediario más frecuente en la degradación anaeróbica de compuestos aromáticos se han realizado multitud de estudios sobre su ruta de degradación. Los genes responsables de la ruta fueron descritos por primera vez en la α -proteobacteria *R. palustris* (Egland y col., 1997). Actualmente, la bioquímica y genética de la ruta ha sido estudiada tanto en la bacteria *R. palustris* como en las β -proteobacterias *T. aromatica* y *Azoarcus* spp. (Breese y col., 1998; Harwood y col., 1999; López-Barragán y col., 2004a). Recientemente, se han identificado algunos de los genes responsables de la ruta central del benzoil-CoA en otras bacterias anaerobias facultativas (e.g. *Magnetospirillum* spp.) (Shinoda y col., 2005) así como en organismos anaerobios

estrictos tales como *G. metallireducens* (reductor de hierro) y, parcialmente, en *Syntrophus aciditrophicus* (fermentador) (Wischgoll y col., 2005; Peters y col., 2007).

Dado que el benzoato se incorpora directamente a la ruta central del benzoil-CoA a través de una activación mediada por la benzoato-CoA ligasa (Fig. 4), este compuesto se ha utilizado como sustrato modelo para todos los estudios de caracterización de dicha ruta central (Carmona y col., 2009).

La ruta central del benzoil-CoA se estructura en dos segmentos metabólicos que reducen y convierten el benzoil-CoA en un compuesto alifático (ruta alta del benzoil-CoA) y transforman este último mediante β -oxidación en acetil-CoA (ruta baja del benzoil-CoA) (Fig. 4). Mientras que la ruta alta es específica del catabolismo de compuestos aromáticos, la ruta baja es común para el catabolismo de compuestos aromáticos y otros compuestos tales como los ácidos dicarboxílicos de número impar de átomos de carbono (Harrison y Harwood, 2005).

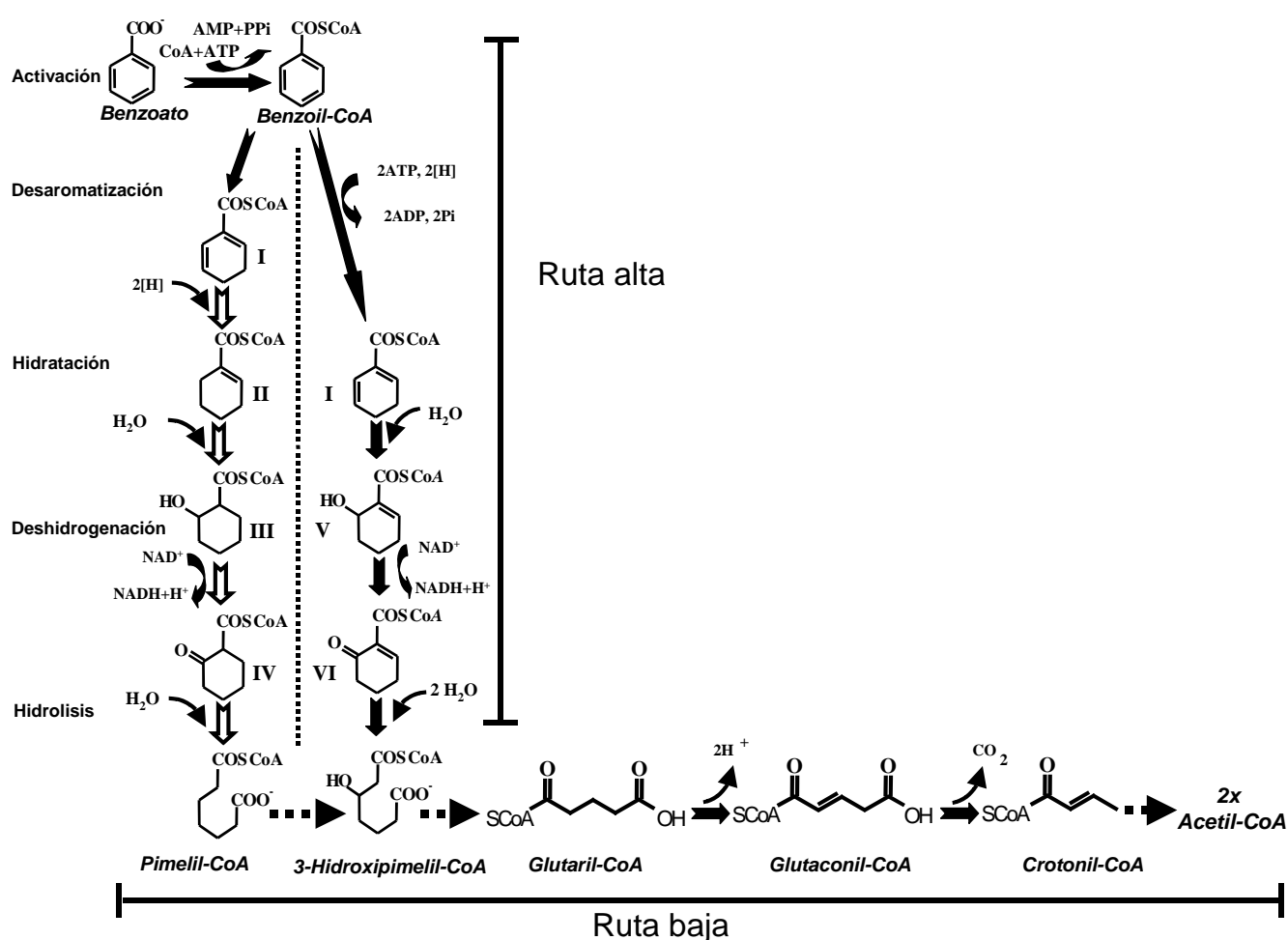


Figura 4. Esquema del catabolismo anaeróbico del benzoato en bacterias anaerobias facultativas . Se indican las reacciones enzimáticas individuales con flechas continuas y con flechas discontinuas se indican varias reacciones enzimáticas consecutivas. Con flechas blancas y negras se indican las reacciones de la ruta alta en *R. Palustris* y en el resto de los anaerobios facultativos, respectivamente. Los intermediarios mostrados son: I, ciclohex-1,5-dieno-1-carbonil-CoA; II, ciclohex-1-eno-1-carbonil-CoA; III, 2-hidroxi-ciclohexano-1-carbonil-CoA; IV, 2-cetociclohexano-1-carbonil-CoA; V, 6-hidroxiciclohex-1-eno-1-carbonil-CoA; VI, 6-cetociclohex-1-eno-1-carbonil-CoA.

2.2.1.1. Ruta alta del benzoil-CoA.

El primer paso de la ruta central de degradación de benzoil-CoA en las cepas *T. aromatica* y *Azoarcus* spp. consiste en la reducción enzimática del benzoil-CoA hasta el intermediario alicíclico, ciclohex-1,5-dieno-1-carbonil-CoA. La actividad enzimática responsable de esta reacción es la benzoil-CoA reductasa y requiere del aporte de ATP para poder vencer la barrera energética de la reducción del anillo aromático (Boll, 2005). El producto de la reducción del anillo aromático es posteriormente hidratado por una dienoil-CoA hidratasa específica que forma 6-hidroxiciclohex-1-eno-1-carbonil-CoA (Laempe y col., 1998). Posteriormente, este intermediario sufre una oxidación dependiente de NAD^+ catalizada por una β -hidroxiacil-CoA deshidrogenasa, generando el intermediario 6-cetociclohex-1-eno-1-carbonil-CoA (Laempe y col., 1999). Finalmente, este intermediario sufre la adición de agua y la ruptura de un enlace C-C generando el compuesto alifático 3-hidroxi pimelil-CoA, en una reacción catalizada por una enzima ácido alicíclico-CoA hidrolasa (Laempe y col., 1999; Harwood y col., 1999) (Fig. 4). Una ruta alta similar se ha propuesto para *Magnetospirillum* spp. (López-Barragán y col., 2004b; Shinoda y col., 2005).

En *R. palustris*, la naturaleza del producto final de la reducción enzimática del anillo aromático difiere de la descrita en *T. aromatica* o *Azoarcus* spp. Así, en este organismo el benzoil-CoA sufre dos reducciones consecutivas del anillo aromático, generándose el derivado monoeno, ciclohex-1-eno-1-carbonil-CoA (Gibson y Gibson, 1992; Harwood y col., 1999). La reacción que sigue a la reducción del anillo aromático en *R. palustris* es catalizada por una enoil-CoA hidratasa específica que convierte el ciclohex-1-eno-1-carbonil-CoA en 2-hidroxi-ciclohexano-1-carbonil-CoA, el cual sufre una oxidación catalizada por una alcohol deshidrogenasa dependiente de NAD^+ dando lugar a 2-cetociclohexano-1-carbonil-CoA (Pelletier y Harwood, 2000). Este intermediario es convertido en el compuesto alifático pimelil-CoA mediante la hidrólisis del anillo aromático catalizada por una ácido alicíclico-CoA hidrolasa (Pelletier y Harwood, 1998) (Fig. 4). El intermediario pimelil-CoA es posteriormente oxidado a 3-hidroxi pimelil-CoA (Harwood y col., 1999) y la ruta converge con la descrita en *T. aromatica* y *Azoarcus* spp. En organismos anaerobios estrictos tales como *G. metallireducens*, *S. aciditrophicus*, *D. multivorans*, la reducción del benzoil-CoA conlleva una reductasa completamente distinta a la dependiente de ATP que se ha caracterizado en anaerobios facultativos. Dicha reductasa todavía no se ha caracterizado bioquímicamente y el mecanismo de la reducción

parece involucrar un complejo sistema enzimático que utiliza el potencial de membrana como fuente energética (Peters y col., 2007). No obstante, el producto de la reducción del benzoil-CoA en organismos anaerobios estrictos es el mismo (ciclohex-1,5-dieno-1-carbonil-CoA) que el observado en organismos desnitrificantes (Fig. 4), y los genes responsables de la conversión de este en 3-hidroxi-pimelil-CoA también muestran una significativa similitud con los descritos en organismos desnitrificantes (Peters y col., 2007).

2.2.1.2. Ruta baja del benzoil-CoA.

La degradación del compuesto alifático C₇-dicarboxil-CoA se denomina “ruta baja”, y es llevada a cabo por reacciones de β -oxidación, típicas del metabolismo de ácidos grasos dicarboxílicos, y descarboxilación para generar crotonil-CoA, acetil-CoA y CO₂. El crotonil-CoA es convertido en dos moléculas de acetil-CoA a través de la ruta clásica de β -oxidación de ácidos grasos de cadena corta (Gallus y Schink, 1994) (Fig. 4). Esta ruta baja, además de ser esencial para el metabolismo de muchos compuestos aromáticos, es utilizada por compuestos alicíclicos (e. g. ciclohexanocarboxilato) y algunos ácidos dicarboxílicos (glutarato, pimelato) (Gibson y Harwood, 2002).

Hasta la fecha, se ha descrito un *cluster* génico (*pim* FABCDE) en la bacteria *R. palustris*, que contendría todos los genes necesarios para la β -oxidación de ácidos dicarboxílicos de cadena impar hasta glutaril-CoA (Harrison y Harwood, 2005). Estos genes se inducen cuando la cepa está creciendo en benzoato o pimelato anaeróbicamente. Mutaciones en los genes *pim* no impiden el crecimiento anaeróbico de los mutantes en los compuestos aromáticos, si bien éstos muestran un mayor tiempo de duplicación que la cepa silvestre. Este resultado sugiere que en *R. palustris* existen otros genes distintos a los *pim* que intervienen también en la β -oxidación de ácidos dicarboxílicos y, por ende, en la ruta baja del benzoil-CoA (Harrison y Harwood, 2005; ver Berkmoes y col., 2006). En organismos anaerobios estrictos como *G. metallireducens* la ruta baja del benzoil-CoA parece involucrar sistemas enzimáticos más complejos, y todavía sin caracterizar, que los presentes en anaerobios facultativos (Wischgoll y col., 2005).

En la mayoría de los organismos, el intermediario glutaril-CoA es oxidado y descarboxilado a crotonil-CoA por una glutaril-CoA deshidrogenasa (Fig. 4). La enzima cataliza inicialmente una reacción de deshidrogenación del glutaril-CoA a glutaconil-CoA, seguida de una reacción de descarboxilación y conversión del glutaconil-CoA en crotonil-

CoA (Gomes y col., 1981; Fu y col., 2004) (Fig. 4). Aunque la actividad glutaril-CoA deshidrogenasa ha sido detectada en células creciendo anaeróbicamente en compuestos aromáticos tales como benzoato (Blakley, 1978; Härtel y col., 1993; Heider y col., 1998; Elshahed y col., 2001), no se conoce el determinante genético bacteriano que codifica dicha actividad.

En las bacterias fermentadoras, el glutaril-CoA es oxidado por una glutaril-CoA deshidrogenasa dependiente de NAD^+ y el glutaconil-CoA es descarboxilado por la acción de una glutaconil-CoA descarboxilasa multicomponente anclada a la membrana y que acopla la descarboxilación con el transporte de sodio a través de la membrana y la síntesis de ATP (Dimroth y Schink, 1998). Esta conversión de glutaril-CoA a crotonil-CoA acoplada a la formación de ATP constituye una estrategia de obtener energía en aquellos organismos que poseen un rendimiento energético muy limitado, como es el caso de los organismos fermentadores (Schöcke y Schink, 1999; Elshahed y McInerney, 2001)

3. Regulación de las rutas de degradación de compuestos aromáticos.

Los compuestos aromáticos no suelen ser una fuente de carbono preferente en las bacterias, por lo que las rutas de degradación de dichos compuestos suelen estar sometidas a un estricto control. Este control debe asegurar una respuesta adecuada y rápida a la presencia de un determinado sustrato. Aunque la regulación puede ser a diferentes niveles (transcripcional, traduccional, postraduccional), la regulación transcripcional ha sido la más estudiada en el catabolismo de compuestos aromáticos (Carmona y col., 2008). El control transcripcional se ejerce a dos niveles: 1) una regulación específica de cada ruta y mediada por reguladores generalmente codificados en el propio “cluster” catabólico, y 2) una regulación sobreimpuesta a la anterior y que responde a mecanismos de regulación globales que operan en función del estado energético de la célula o de la presencia de otra fuente de carbono preferencial para la bacteria (Fig. 5) (Díaz y Prieto, 2000; Cases y de Lorenzo, 2001; Rojo, 2001; Shingler, 2003; Prieto y col., 2004; Tropel y van der Meer, 2004; Cases y de Lorenzo, 2005a).

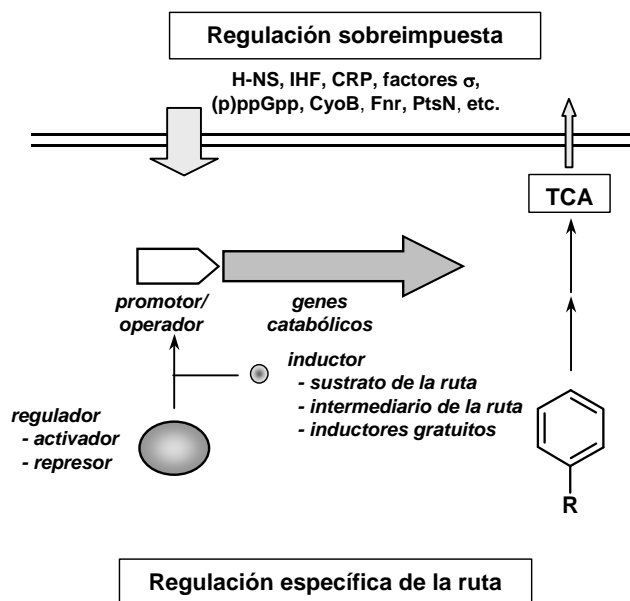


Figura 5. Esquema de los dos niveles de regulación transcripcional del catabolismo de compuestos aromáticos en bacterias. En la degradación de un compuesto aromático, la interacción regulador-promotor es el mecanismo de regulación específico en presencia de la molécula inductora de la ruta. Un segundo nivel de control está constituido por la regulación global o sobreimpuesta, en la que participan reguladores y factores transcripcionales como H-NS (*histone-like nucleoid structuring*), IHF (*Integration Host Factor*), CRP (*cAMP Receptor Protein*), Fnr (*Fumarate nitrate reductase regulator*), factores alternativos de la ARN polimerasa (ARNP), (p)ppGpp (guanosina tetra o pentafofato), CyoB (Citocromo *o* ubiquinol oxidasa), PtsN (proteína IIA^{Ntr} del sistema fosfoenolpiruvato:azúcar fosfotransferasa), etc. Figura modificada de Díaz y Prieto (2000).

3.1. Regulación específica.

Dentro del catabolismo de compuestos aromáticos en bacterias existe una gran variedad de reguladores específicos que pertenecen a distintas familias de proteínas reguladoras (Tabla 1). Estas proteínas reguladoras, y los promotores controlados por ellas, son los elementos clave que dirigen la regulación específica de la transcripción de los operones catabólicos, de forma que cuando el sustrato catabólico (o un intermediario de la ruta) está presente se activa la transcripción de estos operones (Tropel y van der Meer, 2004; Carmona y col., 2008).

El mecanismo de control de los reguladores implicados en el catabolismo de aromáticos se basa en la interacción de la proteína, a través de un dominio hélice-giro-hélice (HTH) de interacción al ADN, con la región operadora del promotor regulado y, a través de otro dominio proteico, con el correspondiente inductor. Las moléculas inductoras suelen ser los propios sustratos de la ruta, o bien alguno de los metabolitos (aromáticos o no aromáticos) que se generan en la degradación. A veces, análogos estructurales del sustrato pueden actuar también como inductores de la ruta, aunque no constituyan un sustrato para las proteínas catabólicas correspondientes (inductores gratuitos) (Díaz y Prieto, 2000; Tropel y van der Meer, 2004).

Tabla 1. Reguladores de rutas de degradación anaeróbica de compuestos aromáticos en bacterias. Se indica la familia a la que pertenecen, en que microorganismo han sido descritos, sustrato de la ruta que regulan, su función y su N° de acceso en el GenBank.

Proteína reguladora	Familia	Microorganismo	Sustrato de la ruta regulada	Función	N° Acceso
BzdR	BzdR	<i>Azoarcus</i> sp. CIB	Benzoato	Represor	AAQ08805
BadR	MarR	<i>R. palustris</i> CGA009	Benzoato	Activador	AAC23923
ORF11	XylR/DmpR	<i>T. aromatica</i> K172	Fenol	Desconocida	CAC12685
Gmet1542	XylR/DmpR	<i>G. metallireducens</i> GS-15	Tolueno	Desconocida	ABB31776
Gmet1520	TetR	<i>G. metallireducens</i> GS-15	Tolueno	Desconocida	ABB31754
TutC1/B1	TCRS (<i>two-component regulatory system</i>)	<i>T. aromatica</i> T1	Tolueno	Posible activador	AAD12187/ AAD12186
TdiS/R	TCRS (<i>two-component regulatory system</i>)	<i>Azoarcus</i> sp. EbN1	Tolueno	Posible activador	CAI07156/ CAI07155
Tcs2/Tcr2	TCRS (<i>two-component regulatory system</i>)	<i>Azoarcus</i> sp. EbN1	Etilbenceno	Desconocida	CAI07438/ CAI07439
EbA715	TetR	<i>Azoarcus</i> sp. EbN1	3-Hidroxibenzoato	Desconocida	CAI06487
HbaR	FNR/CRP	<i>R. palustris</i> CGA009	4-Hidroxibenzoato	Activador	CAE26117
AcpR	FNR/CRP	<i>Azoarcus</i> sp. CIB	Benzoato	Activador	AAY81959
AadR	FNR/CRP	<i>R. palustris</i> CGA009	Benzoato/ 4-hidroxibenzoato	Activador	CAE29675
NicR	LysR	<i>Eubacterium barkeri</i>	Nicotinato	Desconocida	ABC88392

Las proteínas reguladoras pueden actuar como represores o activadores, y su gran diversidad estructural sugiere que han evolucionado independientemente. Por su parte, los promotores de los genes catabólicos poseen un cierto nivel de inespecificidad con respecto a las señales a las que responden, eso les permite evolucionar y poder ser reclutados para controlar nuevas rutas. Otra característica de la regulación transcripcional del catabolismo de compuestos aromáticos es la existencia de cascadas de regulación con diferentes puntos de control, de manera que los genes se inducen selectivamente y se evita el gasto energético que supone la expresión de rutas periféricas innecesarias (Díaz y Prieto, 2000, Cases y de Lorenzo, 2001; Cases y de Lorenzo, 2005a y b). Si bien se han caracterizado un gran número de reguladores del catabolismo aeróbico de compuestos aromáticos, la regulación transcripcional de rutas anaeróbicas ha sido mucho menos estudiada (Carmona y col., 2008). El gen *aadR* de *R. palustris* (Dispensa y col., 1992) fue el primer gen regulador que se describió relacionado con el catabolismo anaeróbico de compuestos aromáticos, estando involucrado en la regulación de los genes *bad* y *hba* responsables del catabolismo de benzoato y 4-hidroxibenzoato.

Los reguladores transcripcionales específicos mejor estudiados hasta la fecha son las proteínas BadR y BzdR que controlan las rutas anaeróbicas para la degradación de benzoato en *R. palustris* y *Azoarcus* sp. CIB (England y Harwood, 1999; Barragán y col., 2005) (Tabla 1). Se han descrito también sistemas reguladores de dos componentes

formados por una proteína histidín-quinasa sensora (Mascher y col., 2006) y una proteína reguladora, con un dominio N-terminal de fosforilación y un dominio C-terminal HTH de unión a ADN. En catabolismo anaeróbico de compuestos aromáticos se han descrito al menos tres posibles sistemas de dos componentes para la degradación de hidrocarburos (Leuthner y Heider, 1998) (Tabla 1).

3.2. Regulación global o sobreimpuesta.

En un segundo nivel de control transcripcional, los promotores catabólicos se encuentran sujetos a una regulación basada en el estado fisiológico de la bacteria, en respuesta a condiciones ambientales generales y no sólo a la presencia de determinados inductores relacionados con la ruta de degradación (Cases y de Lorenzo, 2005a). Esta regulación sobreimpuesta está mediada por reguladores globales tales como factores σ alternativos de la RNAP, las proteínas CRP (*cAMP Receptor Protein*), IHF (*Integration Host Factor*), Fis (*Factor for Inversion Stimulation*), FNR (*Fumarate and Nitrate Reduction*) o H-NS (*Histone like Nucleoid-Structuring Protein*) (Cases y de Lorenzo, 2001; Browning y Busby, 2004), o por otros factores reguladores, tales como el (p)ppGpp (Srivatsan y Wang, 2008). Algunos sistemas de regulación sobreimpuesta implican proteínas de la cadena de transporte de electrones, como CyoB o el sistema PTS, como PtsN (Petruschka y col., 2001; Pflüger y de Lorenzo, 2008).

Uno de los tipos de regulación global más importantes es la represión catabólica, fenómeno común en la regulación de las rutas catabólicas de compuestos aromáticos (Shingler, 2003) y que hace referencia tanto al proceso regulador que permite a las células usar preferentemente una fuente de carbono determinada frente a otras cuando hay una mezcla de ellas en el medio (Magasanik, 1970; Ullmann, 1996), como a una represión parcial que provoca un descenso en la transcripción a partir de un promotor en presencia de una fuente de carbono adicional (Shingler, 2003). Las estrategias de regulación involucradas en los procesos de represión catabólica pueden implicar al sistema de transporte del sustrato en cuestión o la expresión de las rutas catabólicas de las fuentes de carbono no preferentes (Stülke y Hillen, 1999). La represión catabólica suele estar mediada por la proteína reguladora CRP en enterobacterias, si bien la base molecular de esta regulación en otros organismos todavía no se conoce en detalle (Ferrández y col., 2000; Galán y col., 2001). Proteínas como CyoB, PtsN, y BphQ se han visto implicadas en procesos de represión catabólica de rutas de degradación de compuestos aromáticos en

bacterias de los géneros *Pseudomonas* y *Acidovorax* (Petruschka y col., 2001; Ohtsubo y col., 2006; Pflüger y de Lorenzo, 2008). La proteína Crc de *Pseudomonas* está implicada, mediante un mecanismo de control postranscripcional, en la represión que ejercen los aminoácidos sobre la expresión de rutas catabólicas de compuestos aromáticos (Morales y col., 2004; Moreno y Rojo, 2008).

Al igual que ocurre con la regulación específica, la regulación sobreimpuesta de rutas de catabolismo anaeróbico de compuestos aromáticos ha sido poco estudiada y se reduce a la regulación global dependiente de oxígeno en organismos anaeróbicos. Esta regulación está mediada por proteínas de la familia FNR/CRP. Un ejemplo de este tipo de regulación es el que se da en la degradación anaeróbica de benzoato en las cepas *R. palustris* y *Azoarcus* sp. CIB y en el que las proteínas AadR y AcpR, respectivamente, actúan como activadores transcripcionales permitiendo la expresión de los genes catabólicos solamente en ausencia de oxígeno (Tabla1) (Egland y Harwood, 1999; Durante-Rodríguez y col., 2006).

4. La degradación anaeróbica del tolueno y los xilenos.

El tolueno y los xilenos (*o*-, *m*- y *p*-xileno) son compuestos aromáticos, hidrofóbicos y extremadamente volátiles que se emplea en las industrias petroquímicas y de procesamiento de polímeros. Estos compuestos forman parte del grupo de los denominados compuestos BTEX (junto con el benceno y etilbenceno) y son importantes contaminantes que se encuentran presentes en acuíferos, sedimentos procedentes de filtraciones, vertidos de petróleo o productos derivados de la gasolina (Lin y col., 2002). En el campo de la degradación anaeróbica de los hidrocarburos aromáticos, la degradación de los compuestos BTEX es la mejor estudiada. Por otro lado, dentro de los compuestos BTEX, la degradación anaeróbica de tolueno es la que se ha caracterizado mejor en un buen número de bacterias tanto anaerobias facultativas (e.g. *Thauera* spp., *Azoarcus* spp., *Magnetospirillum* spp.) como estrictas (*Geobacter* spp.). Así, el estudio de la activación anaeróbica del tolueno ha permitido descubrir una nueva estrategia general de activación anaeróbica de hidrocarburos aromáticos basada en la formación de succinil-derivados que ya son susceptibles de un ataque enzimático mediante reacciones similares a las de β -oxidación (Fuchs, 2008, Foght, 2008).

4.1. Ruta de la degradación anaeróbica del tolueno.

Aunque el primer cultivo puro con capacidad de degradar tolueno anaeróbicamente fue la cepa reductora de hierro, *Geobacter metallireducens* GS-15 (Lovley y Phillips., 1989), la mayoría de los estudios realizados sobre el catabolismo anaeróbico del tolueno se han realizado en bacterias desnitrificantes tales como *T. aromatica* K172, *T. aromatica* T1, *Azoarcus* sp. T y *Azoarcus* sp. EbN1 (Chakraborty y Coates, 2004).

Los genes implicados en la degradación anaeróbica del tolueno se describieron inicialmente en la bacteria *T. aromatica* K172 y fueron denominados *bss* (*benzylsuccinate synthase*) los que codifican la enzima bencilsuccinato sintasa responsable de la activación del tolueno a bencilsuccinato (Leuthner y col., 1998; Coschigano y col., 1998), y *bbs* (*β-oxidation of benzylsuccinate*) los que codifican las enzimas implicadas en la transformación del bencilsuccinato en benzoil-CoA (Leuthner y Heider, 2000).

El primer paso en el catabolismo anaeróbico del tolueno es la activación de éste mediante su adición, a través de su grupo metilo, al ácido fumárico para formar (*R*)-bencilsuccinato (Fig. 6). Esta reacción está catalizada por la bencilsuccinato sintasa, enzima heterohexamérica (subunidades α_2 , β_2 y γ_2) codificada por los genes *bssC* (subunidad γ), *bssA* (subunidad α), *bssB* (subunidad β) (denominados *tutFDG* en *T. aromatica* T1). La inactivación de cualquiera de estas subunidades impide que la enzima sea funcional (Achong y col., 2001). La reacción es estrictamente anaeróbica ya que en presencia de oxígeno la subunidad α se fragmenta con la consecuente inactivación irreversible de la enzima (Leuthner y col., 1998). El mecanismo catalítico de la enzima ha sido ampliamente estudiado y se sabe que implica un radical glicil (Bhandare y col., 2006; Li y col., 2009). Para que la enzima bencilsuccinato sintasa pueda catalizar la reacción de adición del tolueno al fumarato necesita ser activada por la proteína BssD (codificada por el gen *bssD*; *tutE* en *T. aromatica* T1), que introduce el radical catalítico glicil en la subunidad α a través de una reacción dependiente de *S*-adenosilmetionina (Selmer y col., 2005). La proteína BssE (denominada TutH en *T. aromatica* T1) presenta un dominio tipo Walker de unión a ATP/GTP que es similar al que presentan chaperonas necesarias para el ensamblaje de complejos proteicos, lo cual sugiere que esta proteína es necesaria para el ensamblaje o activación de la bencilsuccinato sintasa (Boll y col., 2002; Hermuth y col., 2002).

La degradación del (*R*)-bencilsuccinato hasta benzoil-CoA y succinil-CoA se lleva a cabo por un proceso similar a una β -oxidación. El primer paso de la β -oxidación lo

realiza la succinil-CoA:(*R*)-bencilsuccinato CoA-transferasa, proteína formada por dos subunidades codificadas por los genes *bbsE* y *bbsF*, que activa el bencilsuccinato mediante su tioesterificación con CoA formando bencilsuccinil-CoA (Leutwein y Heider, 2001). La segunda enzima de la β -oxidación es una bencilsuccinil-CoA deshidrogenasa, codificada por el gen *bbsG*, que genera (*E*)-fenil-itaconil-CoA (Leutwein y Heider, 2002). El resto de las reacciones hasta benzoil-CoA se han propuesto pero todavía no se han caracterizado ni bioquímicamente ni genéticamente. Así, la tercera enzima implicada sería una presunta hidroxiaxil-CoA hidratasa codificada por el gen *bbsH*, encargada de transformar el (*E*)-fenil-itaconil-CoA en 2-[hidroxi(fenil)metil]-succinil-CoA (Leuthner y Heider, 2000). La siguiente reacción la catalizará una hidroxiaxil-CoA deshidrogenasa que o bien está formada por dos subunidades o bien son dos isoenzimas codificadas por los genes *bbsCD*, encargada de transformar el 2-[hidroxi(fenil)metil]-succinil-CoA en benzoil-succinil-CoA (Leuthner y Heider, 2000). En la última reacción de β -oxidación, una benzoil-succinil-CoA tiolasa presuntamente codificada por los genes *bbsAB*, sería la responsable de producir benzoil-CoA y succinil-CoA (Kühner y col., 2005).

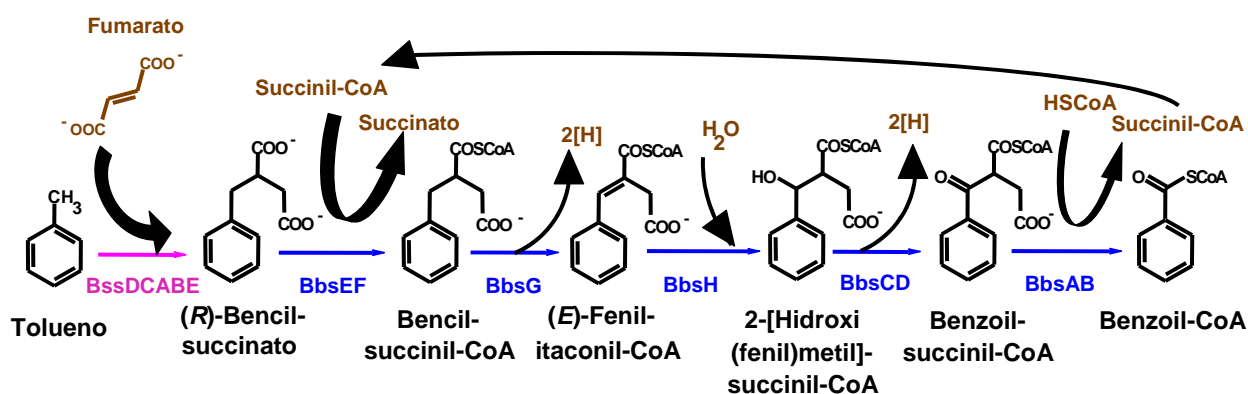


Figura 6. Esquema propuesto para la ruta periférica de degradación anaeróbica de tolueno. Junto a cada reacción enzimática de la ruta se indican los productos genéticos implicados. En color morado se muestra la bencilsuccinato sintasa (y las proteínas activadoras BssDE) implicada en la adición del grupo metilo del tolueno a una molécula de fumarato, mientras que en color azul se muestran las posibles enzimas implicadas en la β -oxidación del bencilsuccinato formado hasta benzoil-CoA. Figura modificada de Kühner y col., 2005.

Los genes responsables de la degradación anaeróbica de tolueno en bacterias desnitrificantes (*Thauera* spp., *Azoarcus* spp., *Magnetospirillum* spp.) están organizados en un *cluster* que incluye tanto a los genes catabólicos *bss* (*tut* en *T. aromatica* T1) y *bbs*, como a los genes reguladores *tdiSR* que codifican un presunto sistema regulador de dos componentes de la subfamilia TodS/TodT (Tabla 1) (Kube y col., 2004).

Entre los organismos anaerobios estrictos y degradadores de tolueno se conoce el *cluster* génico implicado en dicha ruta en *G. metallireducens* GS-15, el cual también

contiene homólogos de los genes *bbs* y *bss* caracterizados en bacterias anaerobias facultativas si bien los presuntos genes reguladores (de las familias XylR y TetR) difieren de los genes *tdiSR* (Kühner y col., 2005; Butler y col., 2007). Hasta la fecha, los únicos *clusters* de degradación anaeróbica de tolueno secuenciados completamente pertenecen a las cepas *Azoarcus* sp. EbN1 y *G. metallireducens* GS-15.

En *Azoarcus* sp. EbN1 estudios de proteómica han revelado la posible participación en el catabolismo anaeróbico del tolueno de otras proteínas cuyos genes codificantes están localizados en la proximidad del *cluster bss-bbs*, y que podrían estar implicadas en la extrusión del tolueno para reducir los efectos tóxicos del hidrocarburo sobre la célula (Kühner y col., 2005; Trautwein y col., 2008).

Los genes *bss* constituyen un operón cuya inducción se produce cuando las células se cultivan anaeróbicamente en presencia de tolueno, y se ha determinado el sitio de inicio de la transcripción en *T. aromatica* K172 y *Azoarcus* sp. T (Achong y col., 2001; Hermuth y col., 2002), sugiriendo la existencia de promotores dependientes de la subunidad σ^{70} de la ARNP (Achong y col., 2001). En *T. aromatica* T1, los genes *tut* (*bss*) parecen organizarse en dos operones diferentes, *tutE* (*bssD*) y *tutFDGH* (*bssCABE*) (Coschigano, 2000).

En *T. aromatica* K172 se ha descrito, además la existencia de un control postranscripcional para los genes *bss*, el cual implicaría un procesamiento del ARNm y la formación de dos transcritos (*bssDCABE* y *bssCABE*) de corta y larga vida media, respectivamente. Este control permitiría que la bencilsuccinato sintasa (BssCAB) y su activador (BssD) puedan ser inducidos de forma coordinada pero luego los niveles proteicos se ajusten de forma que se pueda explicar la significativa menor abundancia observada para la proteína BssD (Hermuth y col., 2002).

4.2. Ruta de la degradación anaeróbica de los xilenos.

A diferencia del tolueno, la degradación anaeróbica de los xilenos ha sido mucho menos estudiada, siendo el *m*-xileno el isómero cuya degradación está mejor caracterizada.

La degradación anaeróbica del *m*-xileno es similar a la del tolueno, ya que la primera reacción es la adición al grupo metilo del *m*-xileno del ácido fumárico para formar 3-metilbencilsuccinato (Krieger y col., 1999). Esta reacción está también catalizada por la enzima bencilsuccinato sintasa (BssABC) en *Azoarcus* sp. T (Achong y col., 2001;

Verfürth y col., 2004) lo que está de acuerdo con el hecho de que los genes *bss* se induzcan en células cultivadas en *m*-xileno (Achong y col., 2001). Las siguientes reacciones de β -oxidación del 3-metilbencilsuccinato hasta 3-metilbenzoil-CoA no se han caracterizado hasta la fecha y se ha propuesto la existencia de isoenzimas Bbs distintas a las implicadas en la degradación de tolueno (Krieger y col., 1999; Verfürth y col., 2004).

Aunque se desconoce la genética y bioquímica de la degradación anaeróbica del *o*- y *p*-xileno, algunos estudios apuntan a que la degradación comienza con la adición de un grupo fumarato y una posterior β -oxidación similar a la que ocurre en la degradación del *m*-xileno (Morasch y Meckenstock, 2005). La degradación anaeróbica del *o*-xileno puede ser llevada a cabo por diferentes bacterias en condiciones de reducción de nitrato o sulfato (Rabus y Widdel, 1995; Morasch y col., 2004), en cambio hasta la fecha no ha sido posible aislar un cultivo puro capaz de degradar anaeróbicamente el *p*-xileno (Chakraborty y Coates, 2004).

Ya que en los ambientes anóxicos contaminados por compuestos aromáticos son muy frecuentes las bacterias del género *Azoarcus* (Lin y col., 2002), y teniendo en cuenta que nuestro grupo ha caracterizado por primera vez la ruta alta del benzoil-CoA en una bacteria del género *Azoarcus* que degrada anaeróbicamente tolueno y *m*-xileno, *Azoarcus* sp. CIB (López-Barragán y col., 2004a), y que el mecanismo de su control transcripcional es actualmente el mejor estudiado con regulación de rutas anaeróbicas (Barragán y col., 2005; Durante-Rodríguez y col., 2006; Durante-Rodríguez y col., 2008), la cepa *Azoarcus* sp. CIB se presenta como un sistema modelo ideal para ampliar los estudios sobre la genética y regulación de las rutas periféricas y centrales de degradación de compuestos aromáticos, concretamente de aquellas implicadas en el catabolismo anaeróbico del tolueno y *m*-xileno.

II. OBJETIVOS

Como se ha comentado en la Introducción de esta memoria, la bacteria *Azoarcus* sp. CIB se ha convertido en un microorganismo modelo para el estudio del catabolismo anaeróbico de compuestos aromáticos, especialmente en los que concierne a la regulación de la expresión de los genes implicados en este catabolismo.

A la versatilidad metabólica de la cepa CIB frente a los compuestos aromáticos, y a su capacidad de ser manipulada genéticamente, se suma la reciente secuenciación completa de su genoma, lo que facilita en gran medida los estudios genéticos y los abordajes globales en este organismo.

Por otro lado, al comiendo de esta Tesis no existía un estudio completo de la regulación transcripcional de los genes implicados en el catabolismo anaeróbico periférico del tolueno y los determinantes genéticos de la degradación anaeróbica del *m*-xileno tampoco habían sido identificados claramente en ninguna bacteria. Asimismo, varios de los genes responsables de las etapas finales del catabolismo anaeróbico de estos hidrocarburos aromáticos tampoco habían sido caracterizados.

Por lo expuesto anteriormente, en este trabajo se planteó el estudio y caracterización de los genes responsables de la ruta periférica de degradación anaeróbica de tolueno y *m*-xileno en la bacteria *Azoarcus* sp. CIB, así como el estudio del gen responsable de una de las etapas clave de la ruta baja del benzoil-CoA, la mediada por la enzima glutaril-CoA deshidrogenasa.

Los objetivos básicos de esta tesis han sido:

1. Identificar, clonar y caracterizar los genes implicados en la ruta periférica de la degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB.
2. Estudiar la organización y regulación transcripcional de los operones implicados en la degradación periférica de tolueno y *m*-xileno.
3. Identificar, clonar y caracterizar el gen *gcdH* que codifica la glutaril-CoA deshidrogenasa de la ruta baja del benzoil-CoA en *Azoarcus* sp. CIB.
4. Estudiar la regulación transcripcional del gen *gcdH*.

III. MATERIALES Y MÉTODOS

1. Cepas bacterianas.

Las cepas de bacterias que se han utilizado en este trabajo se detallan a continuación en la tabla 2 junto con sus genotipos y características más relevantes.

Tabla 2. Cepas bacterianas empleadas en este trabajo.

Cepa	Genotipo/fenotipo relevante	Referencia
<i>E. coli</i>		
DH5 α	(ϕ 80 <i>lacZ</i> Δ M15) <i>endA1 recA1 hsdR17</i> ($r_K^- m_K^-$) <i>supE44</i> Δ <i>lacU169</i>	Sambrook y Russell, 2001
DH10B	F ⁺ , <i>mcrA</i> Δ (<i>mrr hsdRMS-mcrBC</i>) ϕ 80 <i>dlac</i> Δ M15 Δ <i>lacX74 deoR recA1 araD139</i> Δ (<i>ara-leu</i>)7697 <i>galU galK</i> λ <i>rpsL endA1 nupG</i>	Invitrogen
S17-1 λ <i>pir</i>	Tp ^f Sm ^r <i>recA thi hsdRM</i> ⁺ RP4::2-Tc::Mu::Km Tn7 λ <i>pir</i> lisogénico	de Lorenzo y Timmis, 1994
MC4100	<i>araD139</i> Δ (<i>argF-lac</i>)U169 <i>rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR</i>	Casadaban, 1976
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proA</i> ⁺ <i>B</i> ⁺ <i>lacI</i> ^f Δ (<i>lacZ</i>)M15 Tn10 (Tc ^r)]	Sambrook y Russell, 2001
TRMG	MG1655 <i>csrA::km</i> ^R	Romeo y col., 1993
<i>Azoarcus</i> sp. CIB		
<i>Azoarcus</i> sp.CIB <i>dbzdN</i>	Cepa parental degradadora de tolueno y <i>m</i> -xileno	López-Barragán y col. 2004
<i>Azoarcus</i> sp.CIB <i>dtiR</i>	Derivado de la cepa CIB con disrupción insercional del gen <i>bzdN</i>	López-Barragán y col. 2004
<i>Azoarcus</i> sp.CIB <i>dtiS</i>	Derivado de la cepa CIB con disrupción insercional del gen <i>tdiR</i>	Este trabajo
<i>Azoarcus</i> sp.CIB <i>dbssD</i>	Derivado de la cepa CIB con disrupción insercional del gen <i>tdiS</i>	Este trabajo
<i>Azoarcus</i> sp.CIB <i>dtol</i>	Derivado de la cepa CIB con disrupción insercional del gen <i>bssD</i>	Este trabajo
<i>Azoarcus</i> sp.CIB <i>dbbsB</i>	Derivado de la cepa CIB con disrupción insercional del gen <i>tol</i>	Este trabajo
<i>Azoarcus</i> sp.CIB <i>dgcdH</i>	Derivado de la cepa CIB con disrupción insercional del gen <i>bbsB</i>	Este trabajo
<i>Azoarcus</i> sp.CIB <i>dgcdR</i>	Derivado de la cepa CIB con disrupción insercional del gen <i>gcdH</i>	Este trabajo
<i>Pseudomonas putida</i> KT2440	<i>P. putida</i> mt-2 curada del plásmido TOL	Franklin y col., 1981

2. Plásmidos

Los plásmidos utilizados en este trabajo se describen en la Tabla 3 donde se muestra el nombre del plásmido, las características más relevantes y la referencia.

Tabla 3. Plásmidos utilizados en este trabajo.

Plásmido	Descripción	Referencia
pGEM-T Easy	Ap ^r , oriColE1, lacZα, vector de clonación para fragmentos de PCR	Promega
pK18mob	Km ^r , oriColE1 Mob ⁺ , lacZα, vector suicida para la construcción de mutantes por recombinación homóloga.	Schäfer y col., 1994
pK18mobtdiS	Km ^r , derivado de pK18mob que contiene un fragmento interno EcoRI (623 pb) del gen <i>tdiS</i>	Este trabajo
pK18mobtdiR	Km ^r , derivado de pK18mob que contiene un fragmento interno XbaI/SalI (400 pb) del gen <i>tdiR</i>	Este trabajo
pK18mobbssD	Km ^r , derivado de pK18mob que contiene un fragmento interno PstI/BamHI (593 pb) del gen <i>bssD</i>	Este trabajo
pK18mobtol	Km ^r , derivado de pK18mob que contiene un fragmento interno XbaI/SmaI (533 pb) del gen <i>tol</i>	Este trabajo
pK18mobbbsB	Km ^r , derivado de pK18mob que contiene un fragmento interno PstI/BamHI (512 pb) del gen <i>bbsB</i>	Este trabajo
pK18mobgcdH	Km ^r , derivado de pK18mob que contiene un fragmento interno EcoRI/PstI (578 pb) del gen <i>gcdH</i>	Este trabajo
pK18mobgcdR	Km ^r , derivado de pK18mob que contiene un fragmento interno XbaI/PstI (497pb) del gen <i>gcdR</i>	Este trabajo
pSJ3	Ap ^r , oriColE1, lacZ, vector de búsqueda de promotores mediante fusiones traduccionales al gen <i>lacZ</i> , flanqueado por dianas de restricción NotI	Ferrández y col., 1998
pSJPbbsA	Ap ^r , vector pSJ3 que contiene un fragmento de 351 pb XbaI/PstI con la fusión traduccional <i>PbbsA::lacZ</i> .	Este trabajo
pSJ3Pg	Ap ^r , derivado del vector pSJ3 que contiene un fragmento de 303 pb XbaI/KpnI con la fusión traduccional <i>Pg::lacZ</i>	Este trabajo
pBBR1MCS-5	Gm ^r , <i>ori</i> pBBR1, Mob ⁺ , lacZα, vector de clonación de amplio espectro de huésped	Kovach y col., 1995
pJN2133	Gm ^r , derivado de pJN105 que contiene el gen <i>PA2133</i> de <i>P. aeruginosa</i> PAO1	Hickman y col., 2005
pBBR5T	Gm ^r , derivado de pBBR1MCS-5 que contiene un terminador transcripcional del fago T7 en posición 5' del gen <i>lacZ</i> .	J.I.Jiménez
pBBRPbbsA	Gm ^r , derivado de pBBR5T que contiene un fragmento de 3.5 kb EcoRI/HindIII con la fusión traduccional <i>PbbsA::lacZ</i>	Este trabajo
pBBRPbssD	Gm ^r , derivado de pBBRPbbsA que contiene la fusión traduccional <i>PbssD::lacZ</i>	Este trabajo
pBBRPg	Gm ^r , derivado de pBBR5T que contiene un fragmento de 3.5-kb EcoRI/HindIII con la fusión traduccional <i>Pg::lacZ</i>	Este trabajo
pVTR-B	Cm ^r , <i>ori</i> pSC101 lacI ^q , <i>Ptr_c</i> , vector de clonación de bajo número de copias	Pérez-Martin y de Lorenzo, 1996
pVTRtdiRS	Cm ^r , derivado de pVTR-B que contiene clonados los genes <i>tdiS</i> y <i>tdiR</i> bajo el control de <i>Ptr_c</i> .	Este trabajo
pIZ1016	Gm ^r , vector de amplio rango de huésped derivado de pBBR1MCS-5, que permite la expresión de genes clonados bajo el control del sistema <i>lacI^q/Ptac</i>	Moreno-Ruiz y col., 2003
pIZToIR	Gm ^r , derivado de pIZ1016 que contiene clonado el gen <i>tol</i> como un fragmento de 3,1 kb HindIII/XbaI	Este trabajo
pGCDH	Gm ^r , derivado de pIZ1016 que contiene clonado el gen <i>gcdH</i> como un fragmento de 1,3 kb XbaI/ClaI	Este trabajo
pCK01	Cm ^r , <i>ori</i> pSC101, vector de clonación de bajo número de copias con <i>polylinker</i> de pUC18 flanqueado por dianas NotI	Fernández y col., 1995
pGCDR	Cm ^r , derivado de pCK01 que contiene clonado el gen <i>gcdR</i> bajo el control de <i>Plac</i> como un fragmento de 857 pb KpnI/XbaI	Este trabajo
pGCDHKT	Gm ^r , derivado de pBBR1MCS-5 que contiene clonado el gen <i>gcdH</i> de <i>P. putida</i> KT2440 con su propio promotor como un fragmento de 1,4 kb ClaI/XbaI,	Este trabajo

3. Medios y condiciones de cultivo.

El medio rico utilizado para cultivar las células de *E. coli* y *P. putida* fue el medio Luria-Bertani (LB) (Sambrook y Russell, 2001). Para obtener medio LB sólido se suplementó con Bacto Agar (Pronadisa) al 1,5 % (p/v). Cuando fue necesario, se añadió 5-bromo-4-cloro-3-indolil- β -D-galactopiranosido (X-Gal), a una concentración de 0,08 mM, así como isopropil-1-tio- β -galactopiranosido (IPTG) a una concentración de 1 mM a 5 mM según las necesidades experimentales.

Los medios mínimos que se utilizaron en este trabajo fueron el medio M63 (Miller, 1972) suplementado con vitaminas y aminoácidos a las concentraciones recomendadas en Gerhardt y col., 1994 para cultivar las células de *E. coli*, y el medio mínimo MC (López-Barragán y col., 2004a) para cultivar las células de *Azoarcus* sp. CIB.

El medio MC se basa en la mezcla de un medio basal pH 7,5 denominado MA, y suplementado con elementos traza y vitaminas. El medio MA está compuesto por:

KH ₂ PO ₄	0,33 g
Na ₂ HPO ₄	1,2 g
NH ₄ Cl	0,11 g
MgSO ₄ ·7H ₂ O	0,1 g
CaCl ₂	0,04 g
H ₂ O destilada	Hasta 1 l

Este medio basal fue suplementado con una solución de elementos traza (pH 6.5), cuya composición (1000x) se detalla a continuación:

Ácido nitrilotriacético (NTA)	1,5 g	ZnSO ₄ ·7H ₂ O	0,18 g
MgSO ₄ ·7H ₂ O	3,0 g	CuSO ₄ ·5H ₂ O	0,01 g
MnSO ₄ ·2H ₂ O	0,5 g	KAl(SO ₄) ₂ ·12H ₂ O	0,02 g
NaCl	1,0 g	H ₃ BO ₃	0,01 g
FeSO ₄ ·7H ₂ O	0,1 g	Na ₂ MoO ₄ ·2H ₂ O	0,01 g
CoSO ₄ ·7H ₂ O	0,18 g	NiCl ₂ ·6H ₂ O	0,025 g
NaSeO ₃ ·5H ₂ O	0,3 g	H ₂ O destilada	Hasta 1 l

Y con una solución de vitaminas cuya composición (1000x) es la siguiente:

Biotina	20 mg	Ácido fólico	20 mg
Piridoxina-HCl	10 mg	Tiamina.HCl·2H ₂ O	50 mg
Riboflavina	50 mg	Ácido nicotínico	50 mg
D-Pantotenato cálcico	50 mg	Vitamina B ₁₂	50 mg
Ácido <i>p</i> -aminobenzoico	50 mg	H ₂ O destilada	Hasta 1 l

Los medios M63 y MC fueron suplementados con la fuente de carbono requerida. Las soluciones concentradas de las diferentes fuentes de carbono fueron esterilizadas por calor húmedo en autoclave a 120°C y 1 atmósfera de presión, o mediante filtración utilizando filtros estériles (Millipore) de 0,2 µm de diámetro, y añadidas al medio de cultivo de forma aseptica. Para el empleo de los diferentes hidrocarburos aromáticos tales como el tolueno, *m*-xileno, estireno, etc., estos se añadieron a una concentración de 250 mM en 2 ml de una fase orgánica de 2,2,4,4,6,8,8-heptametilnonano (HMN) que fue añadida al medio de cultivo (15 ml). Atendiendo al log Po/w del tolueno y *m*-xileno, 2,8 y 3,2, respectivamente, la concentración de estos dos compuestos en la fase acuosa cuando se añaden a una concentración de 250 mM en HMN es de 0,4 mM y 0,15 mM, respectivamente.

El medio de cultivo CFA (*Colony Forming Antigen*) (Jackson y col., 2002) fue empleado para la formación de biopelículas en *E. coli*, y contiene 10 g de casaaminoácidos, 1,5 g de extracto de levadura, 50 mg de MgSO₄ y 5 mg de MnCl₂ por litro, pH 7,4.

Para obtener condiciones de anaerobiosis, el medio de cultivo así como las soluciones de elementos traza, KNO₃ (aceptor final de electrones), sulfuro ferroso y todos los compuestos que fueron utilizados como fuente de carbono, se dispensaron en frascos de vidrio, los cuales se taparon con tapones de goma y se sellaron con arandelas de aluminio. A través del tapón de goma se inyectó una aguja conectada a un depósito de nitrógeno (Alpha Gas), permitiendo pasar un flujo de este gas al medio líquido durante varios minutos con el objeto de eliminar el oxígeno. Finalmente, todos los frascos fueron autoclavados. Para obtener una solución anaeróbica y estéril de vitaminas (las cuales son sensibles al calor) se realizó la misma operación descrita anteriormente pero con frascos de vidrio vacíos en los que, después de ser esterilizados por autoclave, se inyectó la solución de vitaminas mediante una jeringa conectada a un filtro estéril Millipore (0,2 µm). Una vez preparadas estas soluciones, fueron añadidas al medio basal MA para obtener así el medio MC suplementado con las correspondientes fuentes de carbono. Tanto la filtración de las soluciones de vitaminas como la inoculación de los medios de cultivo con las distintas cepas bacterianas, se realizaron en una cámara de anaerobiosis (Forma Scientific).

Para los crecimientos aeróbicos, las cepas de *E. coli* fueron cultivadas a 37°C mientras que *Azoarcus* sp. CIB fue cultivado a 30°C, ambas bacterias con una agitación de 250 revoluciones por minuto. El crecimiento anaeróbico de *Azoarcus* y *E. coli* se

realizó sin agitación y a una temperatura de 30°C (37°C para *E. coli*) utilizando KNO₃ 10 mM como aceptor final de electrones y sulfuro ferroso (0,044 mg/ml) para favorecer el ambiente reductor. El crecimiento en medio líquido de los cultivos fue monitorizado midiendo la absorbancia a 600 nm (A_{600}) empleando un espectrofotómetro Shimadzu UV-260. La morfología celular se analizó con un microscopio de contraste de fase Nikon OPTIPHOT-2. La desnitrificación fue monitorizada midiendo los niveles de NO₃ y NO₂ con el *Test Nitrates* (Merck).

Los antibióticos se prepararon en soluciones (1000x) en agua excepto el cloranfenicol, que se preparó en etanol al 100% y la rifampicina, que se disolvió en dimetilsulfóxido (DMSO) 100%. Las soluciones preparadas se esterilizaron por filtración y se conservaron a -20°C. Los antibióticos se utilizaron a las concentraciones finales que se indican a continuación: ampicilina (Ap, 100 µg/ml), cloranfenicol (Cm, 34 µg/ml), kanamicina (Km, 50 µg/ml), rifampicina (Rf, 50 µg/ml) y gentamicina (Gm, 10 µg/ml).

Durante cortos períodos de tiempo (menos de un mes) *E. coli* se conservó a 4°C en placas de medio LB o de medio mínimo. Los cultivos líquidos anaeróbicos de *Azoarcus* sp. CIB se conservaron adecuadamente durante un periodo aproximado de un mes a 4°C en frascos de vidrio. Para su conservación a largo plazo, las bacterias se congelaron en el medio de cultivo correspondiente con glicerol al 15% (v/v) y se mantuvieron a -80°C.

4. Ensayos de quimiotaxis.

4.1. Ensayo de quimiotaxis a tiempos cortos (*Agarose Plug Assay*).

Para el ensayo de quimiotaxis a tiempos cortos se realizó el protocolo de Parales y Harwood, 2002 con modificaciones. Los cultivos celulares de *Azoarcus* sp. CIB y *Azoarcus* sp. CIBdtol se realizaron anaeróticamente en medio mínimo MC con piruvato 0,2% como fuente de carbono, hasta alcanzar la fase exponencial. Se tomaron 150 µl del cultivo y en condiciones anaeróbicas se depositaron en un portaobjetos que contiene una gota (10 µl) de agarosa *low melting* (Ambion) al 2% en solución MA 1x y tolueno 1mM, entre dos topes de una mezcla de parafina-baselina (9:1 v/v) y el correspondiente cubreobjetos (Fig. 7). Para conservar la anaerobiosis, se sellaron los bordes con laca de

uñas antes de sacarlo fuera de la cámara de anaerobiosis. La preparación se observó al microscopio de contraste de fases y se tomaron fotografías del borde de la gota a los 5, 15 y 25 minutos.

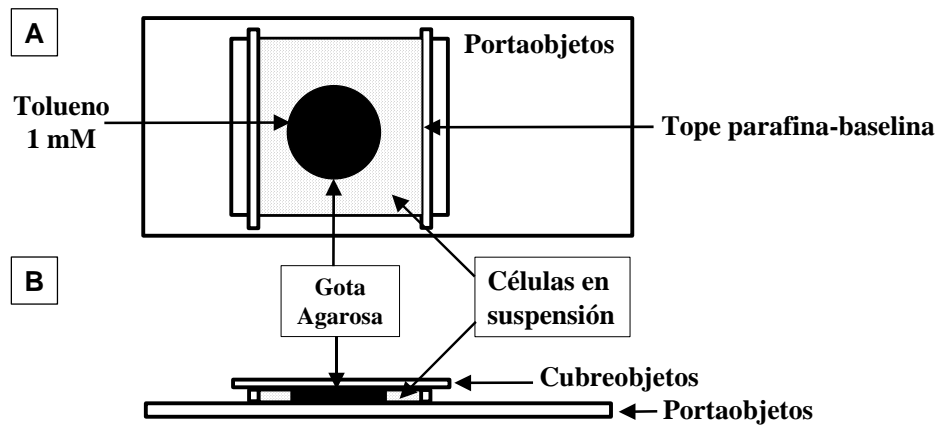


Figura 7. Esquema de la preparación microscópica utilizada para el ensayo de quimiotaxis (Agarose Plug Assay). A) En un portaobjetos se añade una gota de agarosa *low melting* al 2% en solución MA 1X y con tolueno 1mM. Las células se depositan en el espacio localizado entre la gota de agarosa y los tope de parafina en el portaobjetos. Se sellan los bordes con laca para conservar la anaerobiosis. B) Sección transversal de la preparación.

4.2. Ensayo de quimiotaxis en placa de cultivo.

Para ensayar quimiotaxis en medio sólido se utilizan células de las cepas *Azoarcus* sp. CIB, *Azoarcus* sp. CIBdtol y *Azoarcus* sp. CIBdtol (pIZTolR) cultivadas en piruvato 0,2% hasta alcanzar la fase exponencial. Dichas células se inoculan con un palillo en placas anaerobias con agar al 0,3%. Se utilizó medio mínimo MC suplementado con nitrato 10 mM, IPTG 5mM, y, como fuente de carbono, una mezcla de piruvato 0,02% más tolueno 1mM. Se midió el diámetro del halo de crecimiento a los 17 días de incubación en condiciones anaeróbicas.

5. Ensayo de formación de biopelículas en placas multipocillo.

Para el ensayo de formación de biopelículas, se cultivaron anaeróticamente células durante 12 h en medio CFA con los antibióticos correspondientes. Se realizaron diluciones de estos cultivos hasta alcanzar una A_{600} de 0,04 en un volumen de 200 μ l de medio CFA en placas de poliestireno estériles de 96 pocillos, incubándose 24 h a 30°C en la cámara de anaerobios. Posteriormente, se añadieran a cada pocillo 50 μ l de cristal violeta 1%, incubándose durante 15 min a temperatura ambiente para después lavar con agua destilada 3 veces dejando secar al aire los pocillos. Para la cuantificación de la

biopelícula, se solubilizaron las células pegadas a las paredes y la base del pocillo con etanol, y se midió su absorbancia a 595 nm. Al valor de A_{595} se le restó la absorbancia basal debida al medio CFA sin células.

6. Técnicas de manipulación de ADN.

La manipulación del ADN así como otras técnicas de biología molecular fueron aplicadas esencialmente como ha sido descrito por Sambrook y Rusell (2001). Las endonucleasas de restricción fueron suministradas por Amersham Biosciences y New England Biolabs. La enzima T4 ADN ligasa fue proporcionada por USB (Amersham), la ADN polimerasa I y la polimerasa Pfu fueron suministradas por Biotools B&M Labs, S.A. La transcriptasa reversa AMV1 y la polinucleótido quinasa del fago T4 fueron proporcionadas por Promega. Todas las enzimas se emplearon atendiendo a las especificaciones de las diferentes casas comerciales.

6.1. Aislamiento del ADN.

Los fragmentos de ADN se purificaron empleando geles de agarosa, mediante el *kit GeneClean* (BIO 101) o el *High Pure PCR Product Purification Kit* (Roche). La purificación de ADN plasmídico se realizó mediante el uso de *High Pure™ Plasmid Purification Kit* (Roche). El ADN cromosómico se extrajo utilizando el *Genra Puregene Yeast/Bact. Kit* de Quiagen.

6.2. Reacción de amplificación en cadena con ADN polimerasa termorresistente (PCR).

Las mezclas de reacción contenían $MgCl_2$ 1.5 mM, dNTPs 0,2 mM, ADN polimerasa I o Pfu polimerasa 1 U, el ADN molde y la correspondiente pareja de oligonucleótidos.

La amplificación del ADN se realizó en un equipo *Mastercycler Gradient* (Eppendorf). Los distintos oligonucleótidos empleados en las reacciones de PCR se indican en la Tabla 4, en la que se detallan la secuencia de nucleótidos y la finalidad del producto de PCR. Las dianas de restricción que se han diseñado en algunos de los oligonucleótidos se marcan mediante subrayados. Algunos de los oligonucleótidos

fueron sintetizados empleando un equipo Oligo-1000M (Beckman Instruments, Inc.) en el servicio de Química de Proteínas del Centro de Investigaciones Biológicas, si bien la mayoría fueron adquiridos a Sigma-Genosys.

Tabla 4. Oligonucleótidos sintéticos utilizados en este trabajo. Se detalla la secuencia y finalidad de los oligonucleótidos empleados. Las secuencias subrayadas se corresponden con las dianas de restricción creadas (indicadas entre paréntesis).

Oligonucleótido	Secuencia 5'-3' (diana de restricción)	Descripción
bssA5'	ATGAGCGATGTACAGACGTTGGAATACAAGG GG	Oligonucleótido de <i>Azoarcus</i> sp. T empleado para amplificar un fragmento del gen <i>bssA</i> de <i>Azoarcus</i> sp. CIB.
bssA3'	CCCAGGCGATGACCGCTTTGCAGGCG	Oligonucleótido de <i>Azoarcus</i> sp. T empleado para amplificar un fragmento del gen <i>bssA</i> de <i>Azoarcus</i> sp. CIB.
bbsA5'	ATGTCAGAGCAAAAACCGAAC	Oligonucleótido de <i>T. aromatica</i> K172 empleado para amplificar un fragmento del gen <i>bbsA</i> de <i>Azoarcus</i> sp. CIB.
bbsA3'	GATCTTGCGGAACCTGTAGCTC	Oligonucleótido de <i>T. aromatica</i> K172 empleado para amplificar un fragmento del gen <i>bbsA</i> de <i>Azoarcus</i> sp. CIB.
tdiRS3'	A <u>ACTGCAGC</u> CCTGACTTCGAGCATCACCGCGC (<i>Pst</i> I)	Empleado para amplificar los genes <i>tdiS</i> y <i>tdiR</i>
tdiRS5'	G <u>CTCTAGAT</u> GACCTAAGGAGGTAATCCATGAA GAATGACGTGGTCGCGG (<i>Xba</i> I)	Empleado para amplificar los genes <i>tdiS</i> y <i>tdiR</i>
tdiRint5'	CACGGTTTTGGGGCTGATCC	Empleado para amplificar un fragmento interno del gen <i>tdiR</i> y la región intergénica <i>tdiR-tdiS</i>
tdiR131.5'	G <u>CTCTAGAC</u> GACGCGGATATTCAGGTTCTTA CGG (<i>Xba</i> I)	Empleado para amplificar un fragmento interno del gen <i>tdiR</i>
tdiS.F3'	GTCGAGTCGCACGGCGGGCAGC	Empleado para amplificar la región intergénica entre <i>tdiR</i> y <i>tdiS</i>
tdiS3'	GATCTCCGTTCCGTCGTGCAGG	Empleado para amplificar un fragmento interno del gen <i>tdiS</i>
tdiS5'	CATCTCCGAAACCGCACGATCAAGC	Empleado para amplificar un fragmento interno del gen <i>tdiS</i>
ptdiS5'	GG <u>ACTAGT</u> CACCGCGACCACGTCATTCTTCAT G (<i>Spe</i> I)	Empleado para amplificar el promotor <i>PbssD</i>
pbssD3'	CGGGATCCCTCATGACAGCCTCTGCGGTTTCA TGC (<i>Bam</i> HI)	Empleado para amplificar el promotor <i>PbssD</i>
bssD5'	A <u>ACTGCAGG</u> CCTGACGGACGCGCACACC (<i>Pst</i> I)	Empleado para amplificar un fragmento interno del gen <i>bssD</i>
3T7.3	CATGCAGCTGGAGGGTGTGCG	Empleado para amplificar la región intergénica entre <i>bssD</i> y <i>bssA</i>
bssD3'	CGGGATCCGCAATCGTTACGTTCTTGATGCC TGC (<i>Bam</i> HI)	Empleado para amplificar un fragmento interno de <i>bssD</i>
bssA5'new	CGCTCAATTTACACCTGAAGATC	Empleado para amplificar un fragmento interno del gen <i>bssA</i>
bssAint3'	GATGCTCGTGCAACTCGTCGGATGG	Empleado para amplificar la región intergénica entre <i>bssD</i> y <i>bssA</i>
bssA3'new	CAAGGACGGTTTCGTAGGGCGGTAC	Empleado para amplificar un fragmento interno del gen <i>bssA</i>
3T7.7	GACTTCAGCGCGTCCGATCTCG	Empleado para amplificar la región intergénica entre <i>bssA</i> y <i>bssE</i>

Tabla 4. Oligonucleótidos (continuación).

Oligonucleótido	Secuencia 5'-3' (diana de restricción)	Descripción
bssEint3'	GGGGTCCACTTCCGGGATGAACAAGCC	Empleado para amplificar la región intergénica entre <i>bssA</i> y <i>bssE</i>
3T7.9	CGAATACGTTTCAGTACATTGCCGC	Empleado para amplificar la región intergénica entre <i>bssE</i> y <i>bssF</i>
bssFint3'	GGACGCGCAGCCTGAGCATCTCG	Empleado para amplificar la región intergénica entre <i>bssE</i> y <i>bssF</i>
bssGint5'	CGCCGCCAATCGCAATCTCGC	Empleado para amplificar la región intergénica entre <i>bssG</i> y <i>bssI</i>
3T3.14	CGTGACGACAAGCCCGCCATC	Empleado para amplificar la región intergénica entre <i>bssG</i> y <i>bssI</i>
c2A200int5'	CTCGTCTACCTGCTGAACCGCCGC	Empleado para amplificar la región intergénica entre <i>bssJ</i> y <i>bssK</i>
3T3.12	CGGCGCCGGGCATCTTCAGC	Empleado para amplificar la región intergénica entre <i>bssJ</i> y <i>bssK</i>
gtolSR3'	GCTCTAGAGGGGCCATACGGGAACCGCGC (<i>Xba</i> I)	Empleado para amplificar el gen <i>tol</i>
tolSR2244.3'	GACCGGGAAAGATCGCGAGGCC	Empleado para amplificar un fragmento interno de <i>tol</i>
3T3.6	GCCGAATCGGTGCAGGACTATG	Empleado para amplificar un fragmento interno del gen <i>tol</i>
TutC3'	GCTCTAGATCTTTTCGGAATCCGCTGCGGC (<i>Xba</i> I)	Empleado para amplificar un fragmento interno del gen <i>tol</i>
3T3.4	CGGGCAAGGAGCGGGTCGTC	Empleado para amplificar un fragmento interno del gen <i>tol</i>
tolSRint3'	GCCGTCGAGGATGATGCGTGCC	Empleado para amplificar la región intergénica entre <i>tol</i> y <i>bbsI</i>
gtolSR5'	CCCAAGCTTTGACCTAAGGAGGTAAGCGATGT CGACGTTCGATCGTCAGTTATTCTCC (<i>Hind</i> III)	Empleado para amplificar el gen <i>tol</i>
3T3.3	CGAGCACATGGCGGGGCTGTAC	Empleado para amplificar la región intergénica entre <i>tol</i> y <i>bbsI</i>
bbsIint3'	CGGCGTTCGGCGTCCACCACC	Empleado para amplificar la región intergénica entre <i>bbsI</i> y <i>bbsH</i>
bbsH3'	TCAGGCGCCGCGGAATTCCG	Empleado para amplificar un producto de PCR de 5,9 kb
3T3.2	CGGGCTTGTCAGCGACGTGG	Empleado para amplificar la región intergénica entre <i>bbsI</i> y <i>bbsH</i>
bbsH333.3'	CATCTCGAGCCCGCCACCCACC	Empleado para amplificar la región intergénica entre <i>bbsH</i> y <i>bbsG</i>
bbsG.F5'	GCTCACGCGTGTCCCGACCG	Empleado para amplificar la región intergénica entre <i>bbsH</i> y <i>bbsG</i>
bbsGint3'	GTGAACCGCCGGGCCACCTCC	Empleado para amplificar la región intergénica entre <i>bbsG</i> y <i>bbsF</i>
bbsFint5'	GGAAATGAAGACCGCTGCGCCG	Empleado para amplificar la región intergénica entre <i>bbsG</i> y <i>bbsF</i>
BbsC.F5'	GTGAGCACGGTCGTCATGGGCG	Empleado para amplificar un producto de PCR de 5,9 kb
bbsB3'	AAGGATCCCCCGTAACGTGCGTCATCGCCAC (<i>Bam</i> HI)	Empleado para amplificar un fragmento interno del gen <i>bbsB</i>
bbsB5'	CCGAGGATCTGGCGATGATCACCG	Empleado para amplificar un fragmento interno del gen <i>bbsB</i>
bbsA331.3'	GGTTGTCGCGAACGGGCCCC	Empleado para amplificar un fragmento interno del gen <i>bbsA</i>
bbsA5'new	ATCTCAAGGGGTATCGCTGCAAGG	Empleado para amplificar un fragmento interno del gen <i>bbsA</i>
pbbsA3'	AACTGCAGGACATGACGCCTCCGCAGCATTTG (<i>Pst</i> I)	Empleado para amplificar el promotor <i>PbbsA</i> .

Tabla 4. Oligonucleótidos (continuación).

Oligonucleótido	Secuencia 5'-3' (diana de restricción)	Descripción
pbbsA5'	CCTCTAGACGACCGGACAGCTCAAGGCGC (<i>Xba</i> I)	Empleado para amplificar el promotor <i>PbbsA</i>
3GcdH	GCTCTAGAGGAAGAAGGAACCCCGGTGGCC (<i>Xba</i> I)	Empleado para amplificar el gen <i>gcdH</i>
5GcdH	CGATCGATTGACCTAAGGAGGTAACATGG CAACGAACCGCGTGACTTTC (<i>Cl</i> aI)	Empleado para amplificar el gen <i>gcdH</i>
3'GcdEbN1	CTCGTCGGAGATGCCATTGC	Empleado para amplificar un fragmento interno de 700 pb del gen <i>gcdH</i>
5'GcdEbN1	GTCGTCGCTGGTGATGGTGC	Empleado para amplificar un fragmento interno de 700 pb del gen <i>gcdH</i>
Gcd3'	GCGGATGTCGAGCGACTTGCCG	Empleado para amplificar un fragmento interno de 650 pb del gen <i>gcdH</i>
Gcd5'	CCTGCCAAGCTCGCCACCGG	Empleado para amplificar un fragmento interno de 650 pb del gen <i>gcdH</i>
Pgc3'	GCTCTAGAGCCATGTTGAATCTCCTCAATGCG GC (<i>Xba</i> I)	Empleado para amplificar el promotor <i>Pg</i>
Pgl5'	GGGGTACCCATCGATGCCTCCATAAAGTGC (<i>Kpn</i> I)	Empleado para amplificar el promotor <i>Pg</i>
gLysR5'	GGGGGTACCCTGACCTAAGGAGGTAATCGAT GCGCCGCAAGATCCCC (<i>Kpn</i> I)	Empleado para amplificar el gen <i>gcdR</i>
GcdR5'	GCTCTAGAGCTTACGCGCGCGGCCGAGG (<i>Xba</i> I)	Empleado para amplificar un fragmento interno del gen <i>gcdR</i>
LysRint3'	GTTCTCGGCATCAGGAAATG	Empleado para amplificar un fragmento interno del gen <i>gcdR</i>
gLysR3'	GCTCTAGACCGCCGCGCTCACTCCAGCCCG (<i>Xba</i> I)	Empleado para amplificar el gen <i>gcdR</i>
gcdRHis3'	CCCAAGCTTCCGCGCGCTCACTCCAGCCCG (<i>Hind</i> III)	Empleado para analizar la inserción del gen <i>gcdR</i>
Ppgcd5'	CCATCGATCAGGGCGGCGGTACTGGGGATC (<i>Cl</i> aI)	Empleado para amplificar el gen <i>gcdH</i> de <i>P. putida</i>
Ppgcd3'	GCTCTAGAGCAGATGTGATAGCGGCCCATGG (<i>Xba</i> I)	Empleado para amplificar el gen <i>gcdH</i> de <i>P. putida</i>
F24	CGCCAGGGTTTTCCAGTCACGAC	Empleado para analizar las inserciones generadas mediante los derivados del plásmido pK18 <i>mob</i> (Sambrook y Rusell, 2004)
R24	AGCGGATAACAATTTACACAGGA	Empleado para analizar las inserciones generadas mediante los derivados del plásmido pK18 <i>mob</i> (Sambrook y Rusell, 2004)
Lac57	CGATTAAGTTGGGTAACGCCAGGG	Oligonucleótido que hibrida en el gen <i>lacZ</i> , empleado para la identificación del +1 del promotor <i>Pg</i>
λLeft	CTGTGATGCCATGGTGTCCGACTTATGCC	Secuenciación del extremo 5' del fragmento de ADN clonado en el fago λDASH II
λRight	GTCCAACCCAGATAACGATCATATACATGGTCTC	Secuenciación del extremo 3' del fragmento de ADN clonado en el fago λDASH II

6.3. Búsqueda de genes en la genoteca fágica de *Azoarcus* sp. CIB.

La genoteca fágica de *Azoarcus* sp. CIB en el fago λ DASH II (López-Barragán y col., 2004a) se utilizó para la infección de *E. coli* XL1-Blue, como se describe en Sambrook y Rusell (2001). Para la búsqueda de genes en la genoteca, se utilizaron como sondas fragmentos de dichos genes obtenidos por PCR del ADN de la cepa CIB con los oligonucleótidos correspondientes para cada caso (Tabla 4). Estos fragmentos de ADN fueron marcados con digoxigenina por la técnica de *random primer* (Roche) y utilizados para la hibridación con las placas de lisis fágicas que fueron detectadas con el *DIG luminescence detection* kit (Roche). El título de los fagos seleccionados fue escalado y su ADN purificado como se describe en Sambrook y Rusell (2001), para su posterior análisis y secuenciación.

6.4. Secuenciación de ADN.

La secuenciación del ADN se llevó a cabo en el servicio de Secuenciación Automática (Secugen) del Centro de Investigaciones Biológicas utilizando un secuenciador automático modelo ABI Prism 3730 (Applied Biosystems). Para la reacción de secuenciación se utilizó el “Dye Terminator Cycle Sequencing Ready Reaction Kit” de Applied Biosystems, y la ADN polimerasa AmpliTaq FS, de acuerdo con las recomendaciones del proveedor. Las reacciones de amplificación se llevaron a cabo mediante la técnica de PCR con un termociclador “Gene Amp PCR System 2400” de Perkin-Elmer.

En el caso concreto de los experimentos de extensión por cebador (*primer extension*), la secuenciación del ADN patrón se llevó a cabo según el método de terminación de la polimerización por dideoxinucleótidos (Sanger y col., 1977) utilizando el *T7 Sequencing™ Kit* (Amersham Pharmacia Biotech), y [α - 32 P]dCTP ($14,8 \times 10^{12}$ Bq/mmol, $3,7 \times 10^5$ Bq/ μ l) (Amersham Biosciences) como deoxinucleósido trifosfato marcado radiactivamente, siguiendo las recomendaciones de los fabricantes.

7. Extracción de ARN y ensayos de retrotranscripción-PCR (RT-PCR).

Para la extracción de ARN se cultivaron anaerómicamente 80 ml de células de *Azoarcus* sp. CIB y sus derivados en medio mínimo MC empleando las fuentes de carbono que se detallan en cada caso. Los cultivos fueron recogidos a la mitad de su

fase exponencial o en la fase estacionaria y se extrajo su ARN siguiendo las indicaciones del kit RNeasy (Qiagen). A continuación, el ARN fue procesado para la eliminación de ADN contaminante mediante el tratamiento con el *DNAse and Removal treatment kit* (Ambion) siguiendo las especificaciones del fabricante.

Para la obtención del ADN complementario ADNc se utilizó 1 µg de ARN purificado y 10 U de la transcriptasa reversa AMV (*Avian Myeloblastosis Virus reverse transcriptase*; Promega), en un volumen final de 20 µl, y se incubó con la mezcla de hexámeros durante 1 h a 42 °C. Se emplearon 2 µl de esta reacción (ADNc) como molde para la reacción posterior de PCR. El ADNc se amplificó utilizando 1 U de la ADN polimerasa I (Biotools) y la pareja de oligonucleótidos correspondiente para cada caso. Para cada una de las reacciones de RT-PCR se incluyó una reacción control en la que no se empleó la transcriptasa reversa, y en la que no se obtuvo ninguna banda de amplificación, lo que indica que las preparaciones de ARN no contenían ADN contaminante.

8. Determinación del sitio de inicio de la transcripción.

La determinación del sitio de iniciación de la transcripción de la región promotora *Pg* se llevó a cabo empleando células de *E. coli* MC4100 que contenían los plásmidos pSJ3Pg y pGCDR, las cuales fueron cultivadas aeróbicamente a 37 °C en medio M63 suplementado con glicerol 30 mM como fuente de carbono hasta alcanzar una A_{600} de 0,5, añadiéndose glutarato 1 mM como inductor y recogiendo las células a las 2 h de inducción. Las células fueron lavadas, resuspendidas en tampón TE con lisozima (50 mg/ml; Sigma) y posteriormente lisadas mediante varios ciclos de congelación-descongelación. El ARN total fue extraído mediante el uso del *RNeasy Mini kit* (Qiagen) de acuerdo con las instrucciones del fabricante.

Se procedió al análisis del sitio de inicio de la transcripción (+1) mediante la técnica de extensión por cebador (*primer extension*). La extensión se realiza con la transcriptasa reversa y un cebador que hibrida con el ARNm en una posición cercana al supuesto sitio de iniciación de la transcripción. Para la identificación de la posición +1 del promotor *Pg* se empleó el oligonucleótido Lac57 (Tabla 4), que hibrida en el gen *lacZ* de la fusión traduccional *Pg::lacZ*, el cual fue marcado previamente en su extremo 5' con la polinucleótido quinasa del fago T4, en presencia de [γ - 32 P]ATP (3000 Ci/mmol; Amersham Biosciences). Las reacciones de extensión por cebador se llevaron

a cabo empleando 20 µg de ARN total en un volumen máximo de 6 µl y en presencia de la *Super ScriptTM II RNase H Reverse Transcriptase*, según el método descrito por Martín y col. (1995). Para determinar la longitud de los productos obtenidos en la reacción de extensión, se realizó la secuenciación del plásmido pSJ3Pg utilizando el mismo oligonucleótido Lac57, como se describe en el apartado 6.4 de esta sección. Los productos obtenidos en las reacciones fueron analizados por electroforesis en geles de poliacrilamida 6%-urea 8M, detectándose por autorradiografía (se utilizaron películas HyperfilmTM-MP (Amersham Biosciences) con pantallas amplificadoras de la señal (Cronex Lightning Plus DuPont).

9. Procedimientos de transferencia genética.

9.1. Transformación y electroporación.

Las células de *E. coli* fueron modificadas genéticamente por transformación tras hacerlas competentes mediante el método de RbCl (Sambrook y Rusell, 2001), o bien mediante electroporación (Wirth y col., 1989), empleando un equipo Gene Pulser/Pulse Controller (Bio-Rad) utilizando las siguientes condiciones: 2.5 kV, 25 µF, 200 Ω.

9.2. Transferencia de plásmidos por conjugación.

Los plásmidos se movilaron a *Azoarcus* sp. CIB y sus derivados por conjugación biparental utilizando la cepa (*E. coli* S17-1λ*pir*) como donador, según el protocolo establecido por de Lorenzo y Timmis (1994) con algunas modificaciones. Se utilizó una cantidad de células de la cepa donadora *E. coli* S17-1λ*pir*, correspondiente a una A_{600} de 5, y de la cepa receptora (*Azoarcus*) cultivada en medio mínimo MC con succinato 0,2 % se utilizó una cantidad de células correspondiente a una A_{600} de 35. Los exconjugantes fueron seleccionados aeróbicamente en placas de medio mínimo MC conteniendo citrato 0,2% (contraselección de la cepa donadora) y el correspondiente antibiótico del plásmido y, en su caso, el de la cepa receptora (cuando se utilizaran derivados de *Azoarcus* sp. CIB).

10. Disrupción insercional de genes cromosómicos de *Azoarcus* sp. CIB.

Para la construcción de las cepas mutantes *Azoarcus* sp. CIBdt*diR*, *Azoarcus* sp. CIBdt*diS*, *Azoarcus* sp. CIBdb*ssD*, *Azoarcus* sp. CIBdt*ol*, *Azoarcus* sp. CIBdb*bbsB*, *Azoarcus* sp. CIBdg*cdH* y *Azoarcus* sp. CIBdg*cdR* (Tabla 2), se procedió a la disrupción insercional de los genes *tdiR*, *tdiS*, *bssD*, *tol*, *bbsB*, *gcdH* y *gcdR*, respectivamente, por recombinación homóloga. Para ello, se amplificaron fragmentos internos de cada uno de los genes citados, y se clonaron en el lugar de clonación múltiple de pK18*mob*, un plásmido movilizable que sólo replica en *E. coli* pero no en otras bacterias tales como *Azoarcus*. Las construcciones resultantes pK18*mobtdiS*, pK18*mobtdiR*, pK18*mobbssD*, pK18*mobtol*, pK18*mobbbsB*, pK18*mobgcdH* y pK18*mobgcdR*, respectivamente (Tabla 3), fueron transferidas desde la cepa donadora, *E. coli* S17-1 λ *pir*, a la cepa receptora, *Azoarcus* sp. CIB, mediante conjugación biparental. Los transconjugantes obtenidos son el resultado de una recombinación homóloga entre el gen nativo cromosómico y el correspondiente fragmento interno clonado en el vector suicida, con la consiguiente disrupción génica por integración plasmídica. Dichos transconjugates fueron seleccionados en medio mínimo MC suplementado con kanamicina (50 μ g/ml) y con citrato 0,2% (p/v) como fuente de carbono para la contraselección de la cepa donadora. La correcta integración de los plásmidos suicidas en los genes descritos se comprobó mediante amplificación por PCR usando un oligonucleótido del plásmido (F24 o R24; Tabla 4) y un oligonucleótido externo al fragmento interno usado para realizar la recombinación homóloga.

11. Técnicas de manipulación de proteínas.

11.1. Obtención de los extractos proteicos.

Para la obtención de extractos crudos de *E. coli* y *Azoarcus* sp. CIB, se cultivaron las células en medio LB o medio mínimo MC con la fuente de carbono apropiada en cada caso, hasta alcanzar la mitad de la fase exponencial. En los casos en los que fue necesario, se añadió IPTG 1 mM y la inducción se mantuvo durante 2 h. Los cultivos se concentraron 20 veces, tras centrifugar y resuspender en el tampón Tris-HCl 100 mM pH 7,5, y se lisaron empleando la prensa de French (Aminco Corp.) operada a una presión de 20.000 psi. Las células rotas se centrifugaron a 4°C en un rotor SS34

(Sorvall) a 14.000 rpm durante 15 min, recogándose la fracción soluble en el sobrenadante (extracto crudo). La concentración de proteínas en el extracto crudo se determinó mediante el método de Bradford (1976) con seroalbúmina bovina como estándar.

11.2. Electroforesis en geles de poliacrilamida-SDS (SDS-PAGE).

Las electroforesis analíticas de proteínas se realizaron en todos los casos en condiciones desnaturalizantes en presencia de dodecilsulfato sódico (SDS). La técnica utilizada fue la descrita por Laemmli (1970) empleando geles de poliacrilamida en placa (PAGE) a una concentración del 12.5%. Las muestras se hirvieron durante 5 min en presencia del tampón de ruptura (Tris-HCl 6,25 mM pH 6,8; SDS 2%, β -mercaptoetanol 5%, glicerol 10% y azul de bromofenol 0,005%). Las electroforesis se realizaron a temperatura ambiente y a 50 mA (corriente constante), utilizando un electrolito que contenía Tris-HCl 25 mM pH 8,8, glicina 192 mM y SDS 0,1%. Para la visualización de las proteínas en los geles, éstos se tiñeron con azul brillante de Coomassie R-250, según se describe en Swank y Munkres (1971). Las proteínas empleadas como marcadores de tamaño molecular (fosforilasa B (97,4 kDa), seroalbúmina bovina (66,2 kDa), ovoalbúmina (45 kDa), anhidrasa carbónica (31 kDa), e inhibidor de tripsina (21,5 kDa)) se adquirieron de Bio-Rad.

11.3. Modelado de la proteína glutaril-CoA deshidrogenasa de *Azoarcus* sp. CIB.

La estructura tridimensional de la proteína glutaril-CoA deshidrogenasa (GcdH) de *Azoarcus* sp. CIB se ha modelado usando el programa *Loopp program* (Teodorescu y col., 2004). El molde utilizado fue la proteína glutaril-CoA deshidrogenasa humana (PDB 55669741).

12. Análisis de los datos de secuencia.

Los análisis de secuencias de nucleótidos se realizaron en el servidor BIOSUPPORT (<http://bioinfo.hku.hk/package.html>). Las búsquedas de marcos abiertos de lectura (ORFs) se realizaron con el programa ORF FINDER del servidor del *National Center for Biotechnology Information* (NCBI)

(<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Las secuencias de aminoácidos de las distintas ORF se compararon con las presentes en las bases de datos de genomas secuenciados o en vías de secuenciación empleando el algoritmo TBLASTN (Altschul y col., 1990) del servidor del NCBI (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). Las búsquedas de similitudes de secuencia se realizaron utilizando el programa BLAST en el servidor del NCBI (<http://www.ncbi.nlm.nih.gov/blast/>). Los alineamientos de secuencias de nucleótidos y los alineamientos múltiples de proteínas se realizaron con los programas DIALIGN (Morgenstern y col., 1996) y CLUSTALW (Thompson y col., 1994), respectivamente, en el servidor de BIOSUPPORT. Las masas moleculares de las proteínas se calcularon con el programa Compute pI/Mw del servidor ExPASy (http://www.expasy.org/tools/pi_tool.html), mientras que las predicciones de estructura secundaria se realizaron con el programa SCRATCH (<http://www.igb.uci.edu/tools/scratch/>; Cheng y col., 2005). La búsqueda de dominios conservados se realizó con el programa CD-SEARCH (Marchler-Bauer y Bryant, 2004) del NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Los árboles filogenéticos se han visualizado con el programa *Tree View program*, versión 1.6.6 (Page, 1996).

13. Ensayos enzimáticos.

13.1. Ensayo de actividad β -galactosidasa.

Para el ensayo de actividad β -galactosidasa de las fusiones traduccionales *PbssD::lacZ* y *PbbsA::lacZ* en la cepa *Azoarcus* sp. CIB y sus derivados, las células fueron cultivadas en medio mínimo MC suplementado con piruvato 0,2% como única fuente de carbono (condiciones de no inducción), y con una mezcla de piruvato 0,2% y tolueno 250mM ó *m*-xileno 250 mM en fase HMN (condiciones de inducción). Los cultivos se incubaron anaeróbicamente a 30°C hasta alcanzar la fase estacionaria de crecimiento.

Para el ensayo de actividad β -galactosidasa de las células de *E. coli* MC4100 portadoras de la fusión traduccional *Pg::lacZ*, éstas fueron cultivadas en medio mínimo M63 con glicerol 30 mM como fuente de carbono y en presencia de distintos

compuestos según se detalla en cada caso. Los cultivos se incubaron a 37°C y fueron recogidos cuando alcanzaron una A_{600} de 0,5.

La actividad β -galactosidasa se analizó permeabilizando las células, y las unidades de actividad enzimática (Unidades Miller) se determinaron según el método descrito por Miller (1972).

13.2. Ensayo de actividad glutaril-CoA deshidrogenasa.

La actividad glutaril-CoA deshidrogenasa de los extractos proteicos se valoró espectrofotométricamente a 300 nm y 30 °C, utilizando como aceptor artificial de electrones el catión ferrocenio tal y como describen Härtel y col., 1993. La mezcla de reacción (600 μ l) contenía 10 μ l de extracto proteico de células de *E. coli* ó 100 μ l de extracto proteico de células de *Azoarcus* en un tampón Tris-HCl 100 mM pH 7.5, que contenía hexafluorofosfato de ferrocenio 0,2 mM, FAD 0,01 μ M y glutaril-CoA 0,2 mM. El coeficiente de extinción molar (ϵ) para el glutaril-CoA a 300 nm es de 4300 $M^{-1}cm^{-1}$ (Lehman y col., 1990).

IV. RESULTADOS

1. ESTUDIO DE LOS DETERMINANTES GENÉTICOS DE LA RUTA PERIFÉRICA PARA EL CATABOLISMO ANAERÓBICO DE TOLUENO Y *m*-XILENO EN *Azoarcus* sp. CIB.

En trabajos anteriores realizados en nuestro laboratorio se describió la capacidad de *Azoarcus* sp. CIB de crecer anaeróbicamente utilizando tolueno y *m*-xileno como únicas fuente de carbono y energía (López-Barragán y col., 2004a). Sin embargo no se estableció ni la cinética de crecimiento, ni la ruta bioquímica que este catabolismo cursaba, ni los genes que codificaban las enzimas implicadas en este proceso en *Azoarcus* sp. CIB.

1.1. Crecimiento anaeróbico de *Azoarcus* sp. CIB en tolueno y *m*-xileno.

En este trabajo, en primer lugar, procedimos a estudiar la cinética de crecimiento anaeróbico de *Azoarcus* sp. CIB en tolueno y *m*-xileno. Para ello, células de *Azoarcus* sp. CIB fueron cultivadas en medio mínimo MC suplementado con KNO₃ 10 mM como aceptor final de electrones y tolueno 250 mM (0,4 mM en fase acuosa) o *m*-xileno 250 mM (0,15 mM en fase acuosa) añadidos a una fase de HMN (ver apartado 3 de Materiales y Métodos) como únicas fuente de carbono y energía. En la figura 8 puede observarse que *Azoarcus* sp. CIB alcanza la fase estacionaria de crecimiento a las 72 h y posee un tiempo de duplicación de aproximadamente 12 h en tolueno y de 24 h en *m*-xileno. Además, el crecimiento máximo de *Azoarcus* sp. CIB en *m*-xileno es inferior al observado en tolueno.

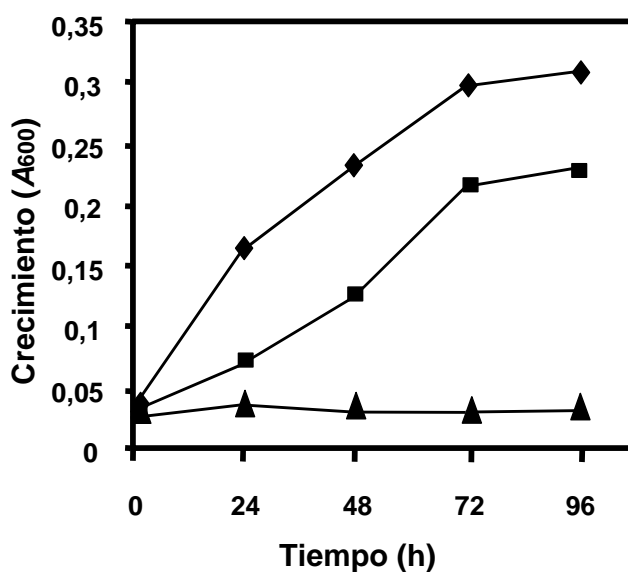


Figura 8. Crecimiento anaeróbico de *Azoarcus* sp. CIB utilizando tolueno o *m*-xileno como únicas fuentes de carbono. Se valoró el incremento de absorbancia a 600 nm (A_{600}) de *Azoarcus* sp. CIB cultivado anaeróbicamente en medio mínimo MC con KNO₃ 10 mM como aceptor final de electrones, y con tolueno 250 mM (rombos) o *m*-xileno 250 mM (cuadrados) como únicas fuentes de carbono suministrados en HMN. Los triángulos representan células control cultivadas sin fuente de carbono.

Para determinar la toxicidad del tolueno y del *m*-xileno en un sistema de dos fases se realizaron cultivos anaeróbicos de *Azoarcus* sp. CIB creciendo en diferentes concentraciones de estos hidrocarburos (Fig. 9), siendo las concentraciones mínimas inhibitorias de 300 mM para el *m*-xileno (aproximadamente 0,19 mM en fase acuosa) y 500-600 mM para el tolueno (aproximadamente 0,8-0,9 mM en fase acuosa), respectivamente. Estos resultados indican la mayor toxicidad del *m*-xileno y están de acuerdo con el crecimiento más lento observado para la cepa CIB en *m*-xileno (Fig. 8)

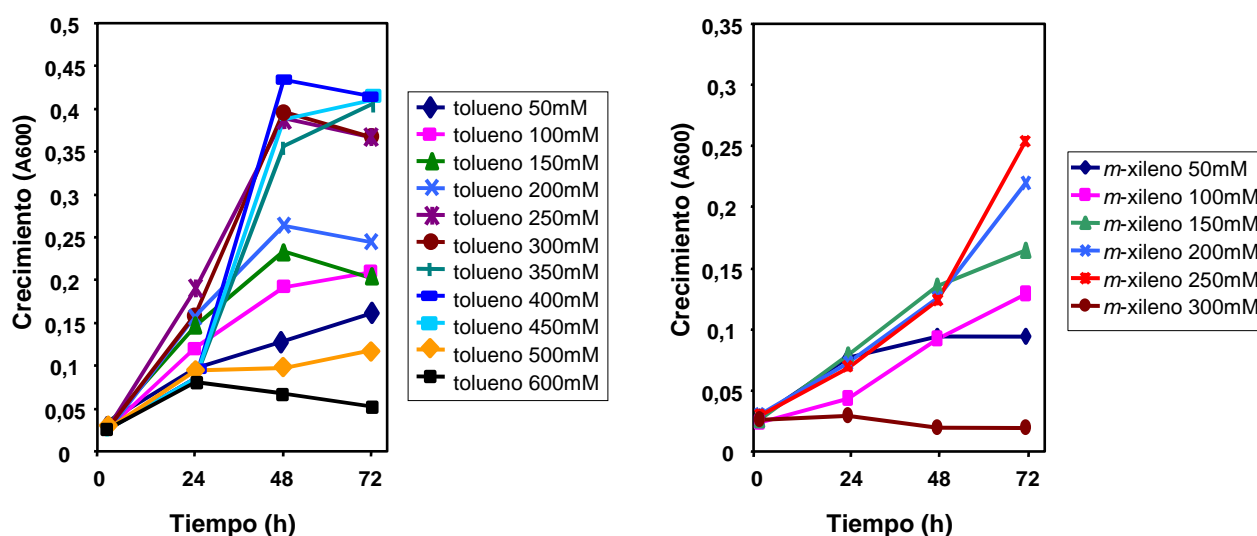


Figura 9. Crecimiento anaeróbico de *Azoarcus* sp. CIB utilizando tolueno o *m*-xileno como únicas fuentes de carbono a diferentes concentraciones. Se valoró el incremento de absorbancia a 600 nm (A_{600}) de *Azoarcus* sp. CIB cultivado anaeróbicamente en medio mínimo MC con KNO_3 10 mM como aceptor final de electrones, y con diferentes concentraciones de tolueno o *m*-xileno suministrados en HMN como únicas fuentes de carbono.

1.2. Identificación del *cluster* génico responsable de la ruta periférica para el catabolismo anaeróbico del tolueno en *Azoarcus* sp. CIB

Tras analizar el crecimiento anaeróbico de *Azoarcus* sp. CIB en tolueno y *m*-xileno se procedió a identificar los genes involucrados en este proceso. El hecho de que la cepa *Azoarcus* sp. CIBdbzdN (tabla 2), cepa mutante en el *cluster* *bzd* responsable de la degradación anaeróbica de benzoato/benzoil-CoA y, por lo tanto, incapaz de crecer en benzoato como única fuente de carbono (López-Barragán y col., 2004a), no fuese capaz de emplear tolueno como única fuente de carbono y energía, sugería que el catabolismo anaeróbico de éste implica una ruta periférica que lo convierte en benzoil-CoA, de forma similar a la ruta *bss* y *bbs* ya caracterizada en distintos organismos anaeróbicos

degradadores de tolueno (ver capítulo 4.1 de la Introducción). Para probar en *Azoarcus* sp. CIB si existen genes ortólogos a los *bss/bbs* que podrían estar implicados en el catabolismo anaeróbico del tolueno, se utilizó la información de secuencia disponible sobre los genes *bss/bbs* de organismos desnitrificantes próximos filogenéticamente a la cepa CIB. Se emplearon las secuencias del gen *bssA* de *Azoarcus* sp. T (Achong y col., 2001) y del gen *bbsA* de *T. aromatica* K172 (Leuthner y Heider, 2000), para diseñar dos parejas de oligonucleótidos (tabla 4) que se usaron para amplificar por PCR sendos fragmentos de 640 pb (*bssA*) y 400 pb (*bbsA*) empleando como molde el ADN de *Azoarcus* sp. CIB. Los fragmentos obtenidos fueron marcados con digoxigenina y usados como sondas para la búsqueda de los genes responsables del catabolismo anaeróbico del tolueno en una genoteca fágica de ADN de *Azoarcus* sp. CIB construida en el fago λ DASH II de *E. coli* (López-Barragán y col., 2004a) (ver detalles en el apartado 6.3 de Materiales y Métodos).

Con la sonda *bssA* se identificaron dos clones positivos denominados λ bssA2 y λ bssA3 (Fig. 10). Mediante amplificación por PCR se confirmó la presencia del gen *bssA* en dichos fagos. Se realizó posteriormente la secuenciación de los extremos del clon λ bssA2 empleando los oligonucleótidos λ Left y λ Right (Tabla 4), que flanquean el sitio de clonación múltiple del fago λ DASH II. La secuencia del extremo 5' mostró elevada similitud con un fragmento del gen *tdiS* de *Azoarcus* sp. T, uno de los presuntos genes reguladores del catabolismo anaeróbico del tolueno en esta bacteria (Achong y col., 2001). Por el contrario, la secuencia del extremo 3' no mostró similitud con ningún gen relacionado con el catabolismo anaeróbico del tolueno en bacterias, lo que sugería que el fago λ bssA2 contenía un ADN quimérico generado durante la construcción de la genoteca. A partir de la secuencia que posee similitud con el gen *tdiS* de *Azoarcus* sp. T, se diseñaron nuevos oligonucleótidos, denominados *tdiS3'* y *tdiS5'* (Tabla 4), que amplificaron un fragmento de ADN (622 pb) del genoma de la cepa CIB, el cual fue posteriormente marcado con digoxigenina y utilizado como sonda (*tdiS*) con el fin de rastrear de nuevo fagos portadores del cluster *bss* en la genoteca de *Azoarcus* sp. CIB. Usando esta estrategia, se identificó el clon λ tdiS3 (Fig. 10). La secuencia del extremo 5' de este clon mostraba alta identidad con el gen *tdiR* de *Azoarcus* sp. T, mientras que en el extremo 3' no presentaba similitud con ningún gen relacionado con el catabolismo anaeróbico del tolueno, sugiriendo de nuevo un fago λ tdiS3 quimérico.

Por otro lado, la secuenciación del clon λ bssA3 mostró la existencia de un fragmento de 17.010 pb cuyo extremo 5' presenta similitud con el gen *bssD*, mientras que su extremo 3' presentaba similitud con el gen *bbsH* de *Azoarcus* sp. T (Fig. 10).

Empleando la sonda *bbsA*, se identificó un clon positivo (λ bbsA4) portador de un fragmento de 3,7 kb que contenía los ortólogos de los genes *bbsC*, *bbsB* y *bbsA* de *Azoarcus* sp. EbN1 (Fig. 10).

Cuando se comparó la organización génica del cluster *bss/bbs* de *Azoarcus* sp. CIB con la de los *clusters* correspondientes de *T. aromatica* K172 (Leuthner y Heider, 2000) y *Azoarcus* sp. EbN1 (Kube y col., 2004) se dedujo que entre el gen *bbsH* del clon λ bssA3 y el gen *bbsC* del clon λ bbsA4 podría localizarse un fragmento de ADN de más de 5 kb donde estarían situados los genes *bbsD*, *bbsE*, *bbsF* y *bbsG*. Para confirmar esta hipótesis, y usando como molde el ADN genómico de *Azoarcus* sp. CIB y los oligonucleótidos *bbsH*3' y *bbsC*.F5' (Tabla 4), se amplificó por PCR un fragmento de 5,9 kb que, como cabría esperar, contenía los ortólogos de los genes *bbsD*, *bbsE*, *bbsF* y *bbsG* en *Azoarcus* sp. CIB (Fig. 10).

La secuenciación de los fragmentos de ADN clonados en los fagos λ tdiS3, λ bssA2, λ bssA3 y λ bbsA4, así como del producto de PCR, generó una secuencia de 29,3 kb que se depositó en el GenBank con el número de acceso DQ988527 (ver Fig. 51 del anexo I). Un análisis inicial de dicha secuencia utilizando el programa ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) reveló la existencia de 24 marcos abiertos de lectura (ORFs) (Fig. 10).

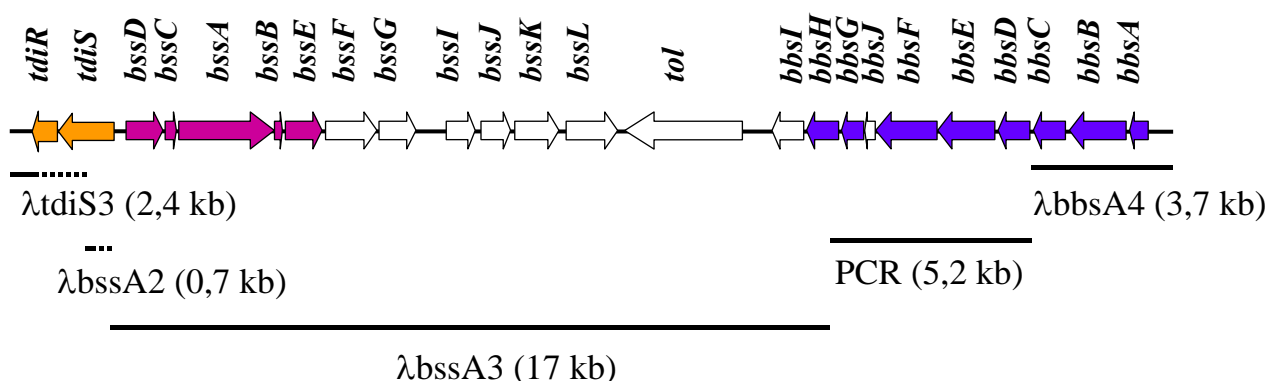


Figura 10. Organización de los genes responsables de la ruta periférica de catabolismo anaeróbico de tolueno en *Azoarcus* sp. CIB. Se representan en color naranja los posibles genes reguladores, *tdiR/tdiS*, en color morado los genes que codifican las posibles subunidades y proteínas activadoras de la bencilsuccinato sintasa, y en color azul se muestran los genes presuntamente implicados en la β -oxidación del bencilsuccinato. En blanco se representan los genes con función desconocida. Las líneas discontinuas representan los clones λ tdiS3 y λ bssA2 que contienen fragmentos de ADN parcialmente secuenciados. Las líneas continuas representan los clones λ bssA3, λ bbsA4 y el producto de PCR de 5,2 kb, que han sido secuenciados en su totalidad.

1.3. Confirmación de la implicación del *cluster bss/bbs* de *Azoarcus* sp. CIB en la degradación anaeróbica de tolueno.

Con el fin de confirmar experimentalmente la implicación del *cluster* génico descrito en el apartado 1.2 en la degradación anaeróbica de tolueno, se construyeron mutantes de *Azoarcus* sp. CIB en los genes *tdiR*, *tdiS*, *bssD* y *bbsB*. Para la disrupción insercional de los genes *tdiR*, *tdiS*, *bssD* y *bbsB* en *Azoarcus* sp. CIB se procedió, en primer lugar, a la subclonación de los fragmentos internos de dichos genes en el vector suicida pK18*mob*, generándose los plásmidos pK18*mobtdiR*, pK18*mobtdiS*, pK18*mobbssD* y pK18*mobbbsB*, respectivamente (Tabla 3, ver apartado 10 de Materiales y Métodos). Dichos plásmidos se transfirieron a *Azoarcus* sp. CIB mediante conjugación biparental seleccionándose en medio mínimo MC con citrato y kanamicina aquellos exconjugantes que poseían integrado el plásmido suicida en el gen *tdiR* (*Azoarcus* sp. CIB*tdiR*); *tdiS* (*Azoarcus* sp. CIB*tdiS*); *bssD* (*Azoarcus* sp. CIB*dbssD*) o en el gen *bbsB* (*Azoarcus* sp. CIB*dbbsB*) (ver apartado 9.2 de Materiales y Métodos) (Fig. 11).

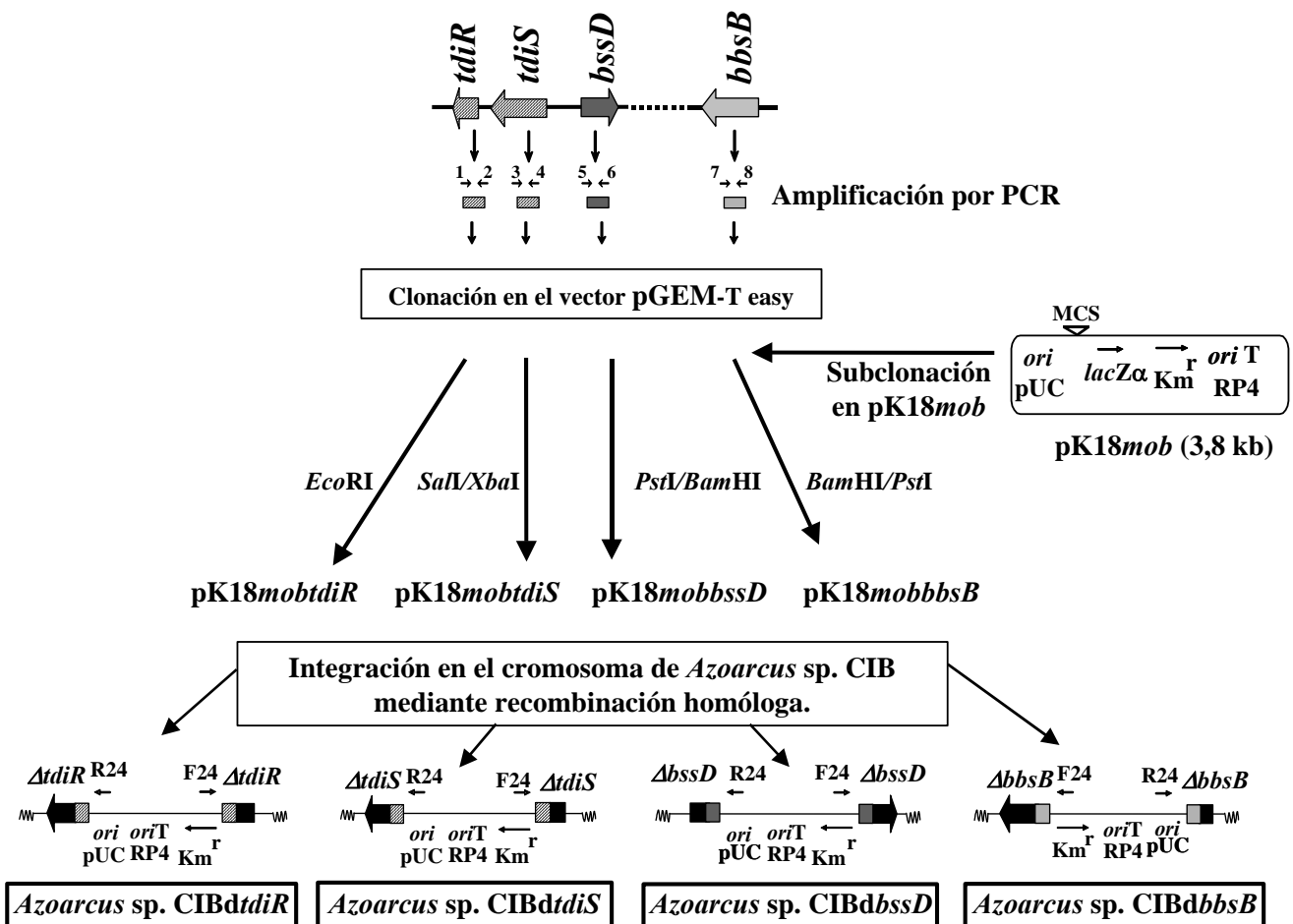


Figura 11. Estrategia seguida para la construcción de los mutantes *Azoarcus* sp. CIB*tdiR*, *Azoarcus* sp. CIB*tdiS*, *Azoarcus* sp. CIB*bssD* y *Azoarcus* sp. CIB*bbsB*. En la parte superior de la figura se esquematizan en rayado los genes *tdiSR*, en gris oscuro el gen *bssD*, y en gris claro el gen *bbsB*. Las flechas con números representan los oligonucleótidos *tdiRint5'* (1), *tdiR131.5'* (2), *tdiS3'* (3), *tdiS5'* (4), *bssD5'* (5), *bssD3'* (6), *bbsB3'* (7) y *bbsB5'* (8), respectivamente (Tabla 4) usados para la amplificación por PCR de los fragmentos internos de los genes *tdiR*, *tdiS*, *bssD* y *bbsB*, respectivamente. Los oligonucleótidos F24 /R24 (Tabla 4) se emplearon para comprobar por PCR la correcta disrupción de los genes *tdiS*, *tdiR*, *bssD* y *bbsB* en el cromosoma de los mutantes de *Azoarcus* sp. CIB. Se indican los orígenes de replicación (*ori*pUC) y de transferencia (*ori*TRP4) de los plásmidos derivados de pK18*mob*. Km^r, representa el gen que codifica la resistencia a kanamicina. *lacZα*, secuencia del gen *lacZα* que codifica al fragmento α de la β-galactosidasa de *E. coli*. MCS, secuencia diana de clonación múltiple (*polylinker*) del plásmido pUC18. Δ, indica gen truncado.

1.4 Análisis de la secuencia del *cluster bss/bbs* de *Azoarcus* sp. CIB.

El análisis detallado mediante comparación de secuencias de cada uno de los 24 ORFs en el *cluster bss/bbs* de *Azoarcus* sp. CIB, reveló, en la mayoría de los casos, una elevada similitud con proteínas implicadas en la degradación anaeróbica de tolueno en distintas cepas bacterianas de los generos *Thauera* y *Azoarcus* (Tabla 5), lo cual estaba de acuerdo con el hecho de que dicho *cluster* contenía los genes responsables de la ruta periférica de degradación anaeróbica de tolueno en *Azoarcus* sp. CIB.

Dada la significativa similitud de secuencia entre los productos génicos del *cluster bss/bbs* y las proteínas implicadas en la degradación anaeróbica del tolueno en *Azoarcus* sp. T, *Azoarcus* sp. EbN1 y *T. aromatica* K172, se ha podido predecir la función para un gran número de los genes *bss/bbs* (tabla 5), así como asumir una ruta bioquímica para el catabolismo anaeróbico periférico del tolueno en *Azoarcus* sp. CIB similar a la ya descrita en otras bacterias (Fig. 12).

Los genes localizados en la posición 5' del *cluster*, genes *tdiS* y *tdiR*, codifican proteínas que guardan una elevada identidad con un presunto sistema regulador de la transcripción de dos componentes descrito en otros *clusters* de degradación anaeróbica de tolueno (Achong y col., 2001). El gen *tdiS* codifica la proteína TdiS, una posible histidín-quinasa sensora de tolueno. La proteína TdiS presenta dos dominios PAS consecutivos en su secuencia, estos dominios están frecuentemente relacionados con la detección de señales ambientales, tales como niveles de oxígeno, potencial rédox, presencia de compuestos aromáticos, etc (Leuthner y Heider, 1998; Mascher y col., 2006), y un dominio histidín-quinasa (HK), relacionado con la transmisión de señal al componente regulador mediante fosforilación (Billwes y col., 1999; Mascher y col., 2006) (Fig. 12). El gen *tdiR* codifica el hipotético regulador transcripcional del sistema

de dos componentes. La proteína TdiR presenta en posición N-terminal un dominio "receiver" (REC) de recepción de la señal de fosforilación desde la proteína sensora (Pao y Saier, 1995; Mascher y col., 2006), mientras que su región C-terminal contiene un dominio HTH de unión a ADN similar al que presentan las proteínas de la familia FixJ/NarL (Ducros y col., 2001) (Fig. 12).

Tabla 5. Análisis del cluster génico responsable de la ruta periférica de degradación anaeróbica de tolueno en *Azoarcus* sp. CIB. En la tabla se detallan los genes del cluster de *Azoarcus* sp. CIB, así como su contenido G+C, la distancia que separa cada gen del siguiente (D) expresada en pares de bases (pb), la longitud en aminoácidos (aa) y la masa molecular (kDa) de los productos génicos, y la hipotética función que se ha asignado a cada uno de los ellos. En los productos génicos relacionados se detalla el nombre de la proteína de mayor similitud de secuencia, la longitud en aminoácidos (aa), la posible función, el grado de identidad que presenta con la proteína homóloga de *Azoarcus* sp. CIB y su N° de acceso en el GenBank. En la figura 51 del Anexo I se detalla la secuencia completa del cluster de degradación anaeróbica del tolueno en *Azoarcus* sp. CIB, así como los oligonucleótidos utilizados para su secuenciación y manipulación génica, y algunas dianas de enzimas de restricción más relevantes.

Gen	G+C %	D (pb)	Producto génico aa/kDa	Hipotética función del producto génico	Productos génicos relacionados				
					Nombre/aa	Función	Organismo	Id %	N° Acceso
<i>tdiR</i>	56	869	218/24,2	Regulador transcripcional	TdiR/218	Posible regulador transcripcional	<i>Azoarcus</i> sp. T	100	AAK50368
<i>tdiS</i>	60,8	69	545/65,5	Histidín-quinasa sensora	TdiS/545	Posible histidín-quinasa sensora	<i>Azoarcus</i> sp. T	100	AAK50369
<i>bssD</i>	56	18	332/37,5	Activador de la bencilsuccinato sintasa	BssD/332	Activador de la bencilsuccinato sintasa	<i>Azoarcus</i> sp. T	99	AAK50370
<i>bssC</i>	48,6	23	60/7	Subunidad γ bencilsuccinato sintasa	BssC/60	Subunidad γ bencilsuccinato sintasa	<i>Azoarcus</i> sp. T	100	AAK50371
<i>bssA</i>	59,5	76	864/97,4	Subunidad α bencilsuccinato sintasa	BssA/864	Subunidad α bencilsuccinato sintasa	<i>Azoarcus</i> sp. T	99	AAK50372
<i>bssB</i>	61,5	47	76/8,8	Subunidad β bencilsuccinato sintasa	BssB/76	Subunidad β bencilsuccinato sintasa	<i>Azoarcus</i> sp. T	100	AAK50373
<i>bssE</i>	56,8	41	286/31,8	Chaperona	BssE/286	Posible chaperona	<i>Azoarcus</i> sp. T	80	YP_158062
<i>bssF</i>	58,8	0	569/64,6	Desconocido	BssF/567	Desconocido	<i>T. aromatica</i> K172	59	CAD88802
<i>bssG</i>	58,3	839	190/21,7	Desconocido	BssG /190	Desconocido	<i>Azoarcus</i> sp. EbN1	63	YP_158064
<i>bssI</i>	66,5	0	194/20,9	Desconocido	C2B001/194	Desconocido	<i>Azoarcus</i> sp. EbN1	83	YP_158066
<i>bssJ</i>	68,1	23	307/34,4	Desconocido	C2A200/307	Desconocido	<i>Azoarcus</i> sp. EbN1	83	YP_158067
<i>bssK</i>	64,3	0	314/35,1	Desconocido	C2A203/315	Azúcar deshidratasa	<i>Azoarcus</i> sp. EbN1	83	YP_158068
<i>bssL</i>	71,9	68	518/54,1	Desconocido	C2A204/520	Desconocido	<i>Azoarcus</i> sp. EbN1	75	YP_158069
<i>tol</i>	66,1	68	1049/115,2	Desconocido	Azo1535/799	Desconocido	<i>Azoarcus</i> sp. BH72	47	YP_933039
<i>bbsI</i>	68,6	231	209/23,5	Desconocido	BbsI/209	Desconocido	<i>T. aromatica</i> K172	80	AAF89844
<i>bbsH</i>	66,7	62	256/27,5	Fenilitaconil-CoA hidratasa	BbsH/256	Posible fenilitaconil-CoA hidratasa	<i>Azoarcus</i> sp. T	94	AAF89843
<i>bbsG</i>	62,5	35	406/45,6	(R)-bencilsuccinil-CoA deshidrogenasa	BbsG/406	(R)-bencilsuccinil-CoA deshidrogenasa	<i>T. aromatica</i> K172	92	Q9KJE8
<i>bbsJ</i>	62,6	0	94/10	Desconocido	BbsJ/94	Desconocido	<i>Azoarcus</i> sp. T	64	YP_158073
<i>bbsF</i>	64,6	0	409/45,1	Subunidad de la succinil-CoA:bencilsuccinato CoA transferasa	BbsF/409	Subunidad de la succinil-CoA:bencilsuccinato CoA transferasa	<i>T. aromatica</i> K172	89	AAF89841
<i>bbsE</i>	64,6	0	410/45,4	Subunidad de la succinil-CoA:bencilsuccinato CoA transferasa	BbsE/410	Subunidad de la succinil-CoA:bencilsuccinato CoA transferasa	<i>T. aromatica</i> K172	87	AAF89840
<i>bbsD</i>	65,5	42	248/26,3	2-[Hidroxi(fenil)metil] succinil-CoA deshidrogenasa	BbsD/248	Posible 2-[Hidroxi(fenil)metil] succinil-CoA deshidrogenasa	<i>Azoarcus</i> sp. EbN1	91	AAF89839
<i>bbsC</i>	66,4	14	250/25,6	2-[Hidroxi(fenil)metil] succinil-CoA deshidrogenasa	BbsC/250	Posible 2-[Hidroxi(fenil)metil] succinil-CoA deshidrogenasa	<i>Azoarcus</i> sp. EbN1	86	YP_158077
<i>bbsB</i>	62,9	13	389/41,3	Benzoilsuccinil-CoA tiolasa	BbsB/389	Posible benzoilsuccinil-CoA tiolasa	<i>Azoarcus</i> sp. EbN1	96	YP_158078
<i>bbsA</i>	60	20	144/16,2	Benzoilsuccinil-CoA tiolasa	BbsA/148	Posible benzoilsuccinil-CoA tiolasa	<i>Azoarcus</i> sp. EbN1	94	YP_158079

Los genes *bssF*, *bssG*, *bssI*, *bssJ*, *bssK*, *bssL*, *bbsI*, y *bbsJ* codifican proteínas de función desconocida en los *clusters* de degradación anaeróbica de tolueno en otras bacterias. De todos los genes presentes en el cluster *bss/bbs* de *Azoarcus* sp. CIB, el único que no se ha identificado en ninguno de los *clusters* equivalentes descritos hasta la fecha en microorganismos degradadores de tolueno es el gen *tol*. Por ello, se procedió a un análisis más detallado de la posible función de este gen en la ruta periférica de la degradación anaeróbica de tolueno en la cepa CIB.

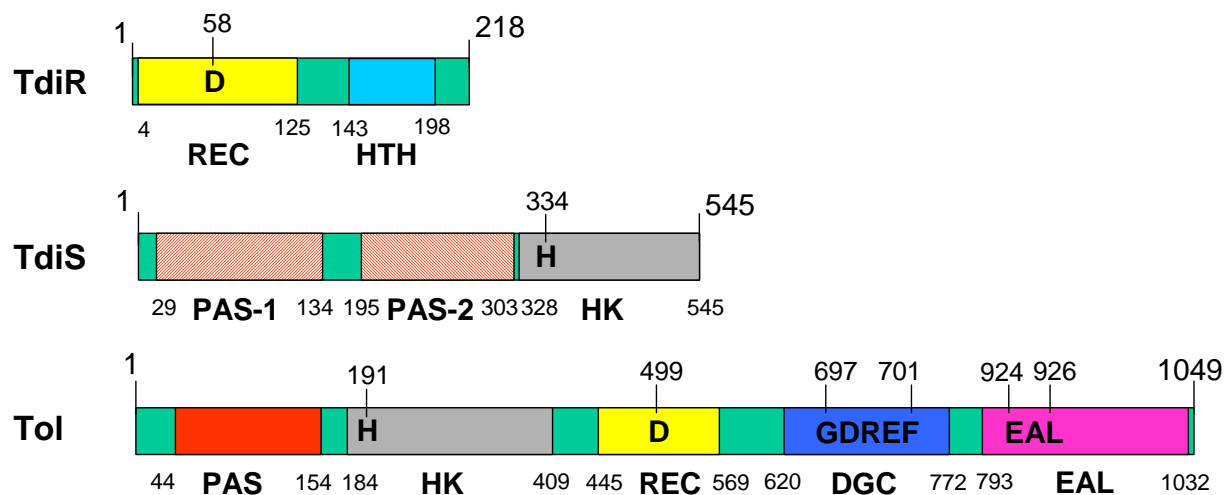


Figura 12. Esquema de la arquitectura modular de las proteínas TdiR, TdiS y Tol de *Azoarcus* sp. CIB. En la figura se indican los dominios presentes en las diferentes proteínas: dominios *reiceicer* (REC), dominio de unión a ADN (HTH), dominio sensor (PAS), dominio histidín-quinasa (HK), dominio diguanilato ciclasa (DGC) y dominio fosfodiesterasa (EAL). Se detalla la posición de los aminoácidos que delimitan cada dominio dentro de las proteínas. Se representan los residuos de histidina (H) y ácido aspártico (D) susceptibles de ser fosforilados durante la cascada de fosforilación asociada a la transmisión de la señal correspondiente. Los residuos que caracterizan a los dominios DGC y EAL también se indican. El diferente rayado de los PAS (PAS-1 y PAS-2) indica que no poseen una identidad de secuencia significativa.

1.5 El gen *tol*.

El gen *tol* codifica una proteína de función desconocida que presenta una arquitectura modular compleja y peculiar, que no se ha descrito en ninguna de las proteínas caracterizadas hasta la fecha (Fig. 12). El extremo N-terminal de la proteína Tol posee un dominio PAS (Fig. 12) que posee una elevada identidad de secuencia (62%) con el correspondiente dominio PAS de la proteína TutC de *T. aromatica* T1, presunta histidín-quinasa implicada en la regulación del catabolismo aeróbico del tolueno, y con los dominios PAS que presentan las histidín-quinasas sensoras TodS y StyS de las rutas de degradación aeróbica de tolueno y estireno, respectivamente, en distintas cepas del género *Pseudomonas* (Leuthner y Heider, 1998). La región central de

la proteína Tol posee similitud significativa con el dominio HK y el dominio REC de las distintas histidín-quinasas híbridas TutC, TodS y StyS (Leuthner y Heider, 1998). Finalmente, la región C-terminal de la proteína Tol se estructura en dos dominios que guardan identidad con dominios diguanilato ciclasa (DGC) y fosfodiesterasa (EAL) implicados en la síntesis y degradación, respectivamente, del dinucleótido cíclico di-GMPc (Fig. 12) (Galperin ,2004). Mientras que el dominio EAL de la proteína Tol está perfectamente conservado, el dominio DGC posee una secuencia de reconocimiento (GDREF) que difiere en varias posiciones de la secuencia consenso GGDEF. Proteínas con motivos GGDEF alterados generalmente carecen de actividad guanilato ciclasa (Christen y col., 2005).

Con el fin de comprobar la participación de la proteína Tol en la degradación anaeróbica de tolueno, se procedió a realizar una mutación que hacía perder la funcionalidad del gen *tol* en *Azoarcus* sp. CIB (Fig. 13).

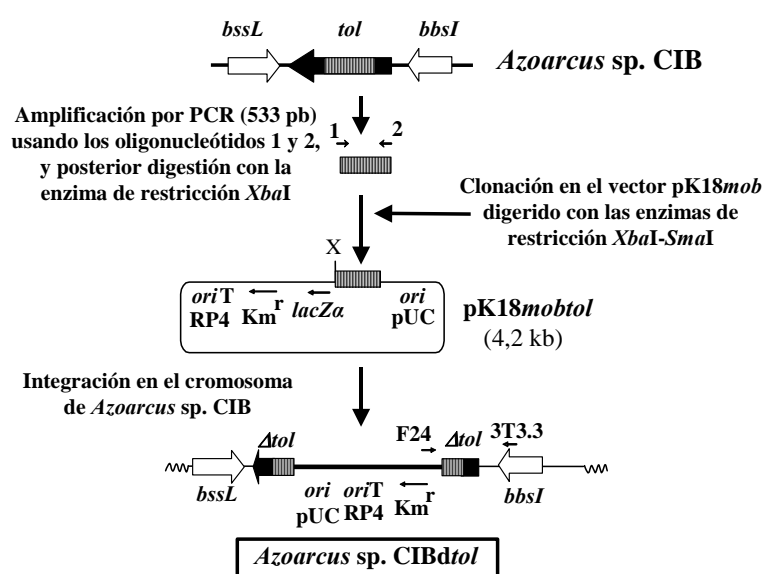


Figura 13. Estrategia seguida para la construcción del mutante *Azoarcus* sp. CIBdtol. 1 y 2 representan los oligonucleótidos TutC3' y 3T3.4, respectivamente (Tabla 4), usados para la amplificación por PCR del fragmento interno del gen *tol*. Los oligonucleótidos F24 y 3T3.3 (Tabla 4) sirvieron para comprobar por PCR la correcta disrupción del gen *tol* en el cromosoma de *Azoarcus* sp. CIB. Se representan los orígenes de replicación (*ori*pUC) y de transferencia (*ori*TRP4) del plásmido suicida pK18mob. *Km^r* representa el gen que codifica la resistencia a kanamicina. X, diana de restricción *XbaI*. *lacZα*, secuencia que codifica el fragmento α de la β -galactosidasa de *E. coli*.

Dado que la cepa mutante *Azoarcus* sp. CIBdtol fue capaz de crecer anaerómicamente en medio mínimo MC con tolueno como única fuente de carbono, el gen *tol* no parece ser imprescindible para el crecimiento anaeróbico en este hidrocarburo aromático.

Puesto que la proteína Tol posee dominios presuntamente implicados en el metabolismo del di-GMPc, y este segundo mensajero bacteriano está involucrado en el control de diferentes procesos fisiológicos tales como movilidad, formación de

biopelículas, virulencia, diferenciación, etc. (Tamayo y col., 2007), se decidió investigar el posible papel del gen *tol* en el control de alguno de estos procesos en *Azoarcus* sp. CIB. Así, se comparó la quimiotaxis de la cepa parental CIB y la del mutante CIB*dtol* frente a tolueno y otra fuente de carbono no aromática como condición control. Cuando se realizó un ensayo de quimiotaxis a tiempos cortos utilizando la técnica de la gota de agarosa (*Agarose Plug Assay*; ver apartado 4.1 de Materiales y Métodos), se observó que las células de *Azoarcus* sp. CIB se aproximaban rápidamente a la gota de agarosa que contiene tolueno generando un halo refringente alrededor de dicha gota visible al microscopio de contraste de fases. Sin embargo, la cepa mutante *Azoarcus* sp. CIB*dtol* no mostró la formación de ningún halo refringente alrededor de la gota de agarosa con tolueno ni siquiera después de 25 min de contacto con el hidrocarburo (Fig. 14). Por el contrario, la respuesta quimiotáctica de las dos cepas fue similar cuando se utilizó piruvato como atrayente. Estos resultados sugieren que la mutación del gen *tol* afecta a la respuesta quimiotáctica de la cepa CIB al tolueno.

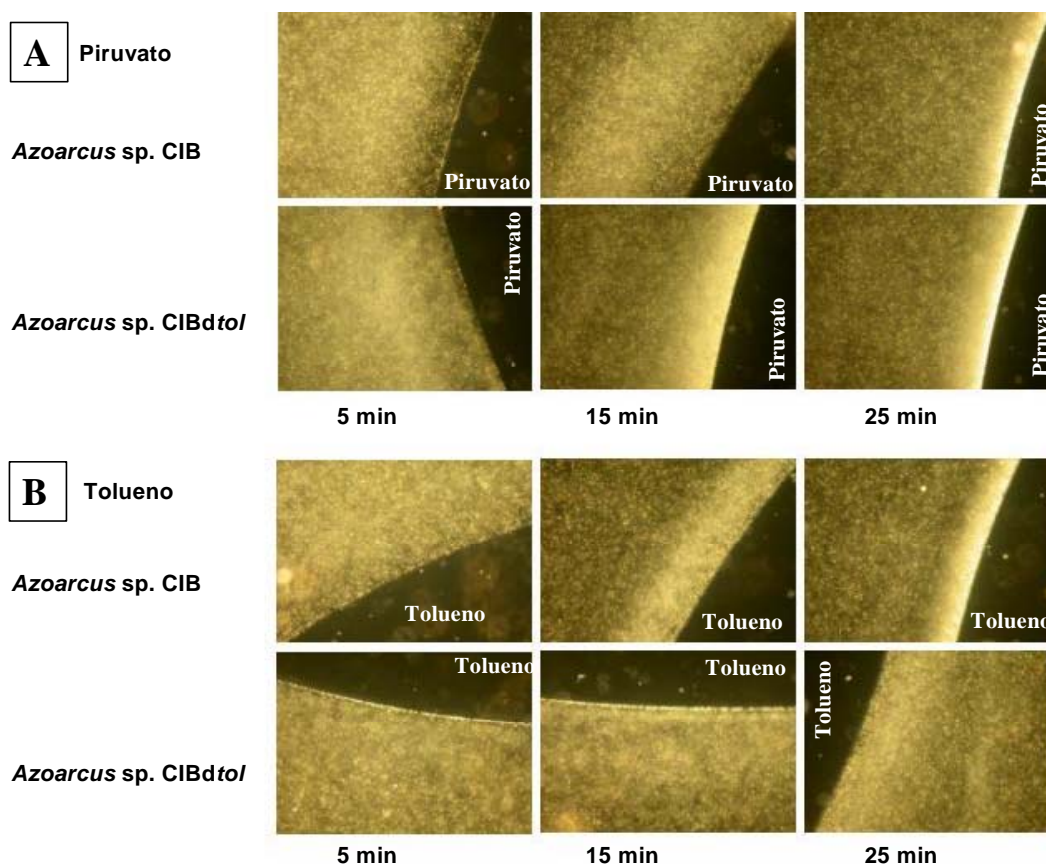


Figura 14. Quimiotaxis de las cepas *Azoarcus* sp. CIB y *Azoarcus* sp. CIB*dtol*. Las células se cultivaron anaeróbicamente en medio MC con piruvato 0,2% como fuente de carbono y se sometieron al ensayo de *Agarose Plug Assay* con piruvato 0,2% (panel A) o tolueno 1mM (panel B) como atrayentes en la gota como se indica en el apartado 4.1 de Materiales y Métodos. Las fotografías se tomaron al microscopio de contraste de fases (x10 aumentos) al borde de la gota de agarosa a los 5, 15 y 25 min de ensayo.

En un ensayo de quimiotaxis en placas de agar al 0,3% (ver apartado 4.2 de Materiales y Métodos), se observó al cabo de 17 días de crecimiento anaeróbico un halo de crecimiento significativamente mayor para la cepa parental *Azoarcus* sp. CIB que para la cepa mutante *Azoarcus* sp. CIB $dtol$ (Fig. 16), lo cual confirmaba los resultados obtenidos con el ensayo de la gota de agarosa (Fig. 14). Para confirmar que el defecto en quimiotaxis al tolueno del mutante *Azoarcus* sp. CIB $dtol$ podía ser revertido al introducir de nuevo el gen *tol* nativo, se construyó el plásmido pIZTolR que expresa el gen *tol* bajo el control del promotor *Ptac* en el plásmido promiscuo pIZ1016 (Fig. 15).

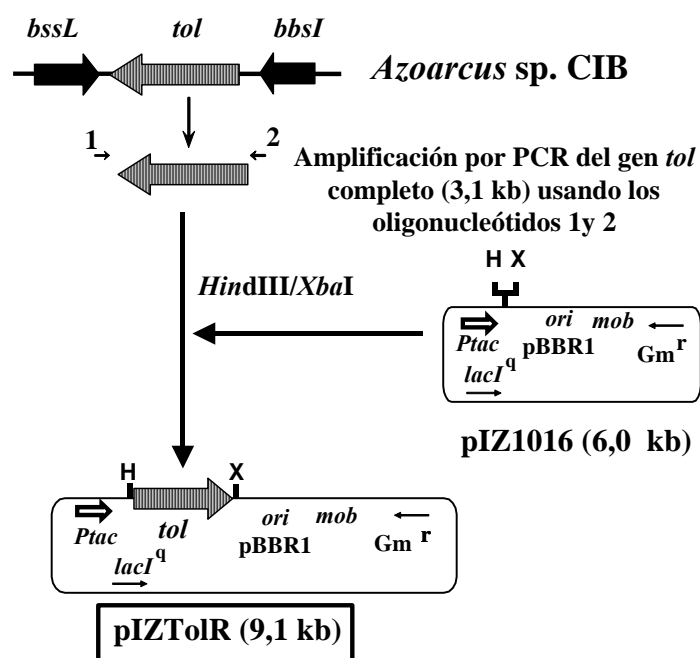


Figura 15. Esquema de la estrategia seguida para la construcción del plásmido pIZTolR. En la parte superior de la figura se esquematizan los genes *bssL*, *tol* y *bbsI*. Las flechas con número 1 y 2 representan los oligonucleótidos *gtolSR3'* (1) y *gtolSR5'* (2) (tabla 4), usados para la amplificación por PCR del gen *tol*. Se representa el origen de replicación (*ori*pBBR1), la región de movilización (*mob*) y el gen que codifica la resistencia a gentamicina (Gm^r) del plásmido pIZ1016 y sus derivados. Se representa con una flecha blanca el promotor *Ptac*. *lacI^q*, gen que codifica el represor LacI. Las dianas de restricción utilizadas para la clonación se indican como: *HindIII* (H) y *XbaI* (X).

El plásmido pIZTolR se transfirió por conjugación biparental desde la cepa donadora *E. coli* S17-1 λ *pir* (pIZTolR) a *Azoarcus* sp. CIB $dtol$ siguiendo el protocolo descrito en el apartado 9.2 de Materiales y Métodos. La selección se llevó a cabo en placas de medio mínimo MC suplementado con succinato 0,2% como fuente de carbono y con gentamicina y kanamicina, obteniéndose la cepa *Azoarcus* sp. CIB $dtol$ (pIZTolR). En el ensayo de quimiotaxis en placas de agar al 0,3%, la cepa complementada con el gen *tol* clonado en pIZTolR mostró un halo de crecimiento mayor que el de la cepa mutante CIB $dtol$, aunque todavía ligeramente inferior al mostrado por la cepa parental

CIB (Fig. 16), lo que volvía a confirmar la participación del gen *tol* en la respuesta quimiotáctica de *Azoarcus* sp. CIB al tolueno.



Figura 16. Quimiotaxis de las cepas *Azoarcus* sp. CIB, *Azoarcus* sp. CIBdtol (pIZTolR) y *Azoarcus* sp. CIBdtol cultivadas anaeróbicamente en medio sólido con agar al 0,3%. Se midió el halo de crecimiento anaeróbico tras 17 días de incubación en medio mínimo MC conteniendo piruvato 0,02% y tolueno 1mM como fuente de carbono, y suplementado con nitrato 10mM e IPTG 5mM.

Dado que la proteína Tol posee un motivo EAL conservado y presuntamente implicado en la hidrólisis del segundo mensajero di-GMPc, se procedió a investigar la posible actividad fosfodiesterasa de la proteína Tol. Para ello se utilizó la cepa *E. coli* TRMG (Tabla 2), que posee una disrupción del regulador global *csrA* implicado, entre otras funciones, en la regulación negativa de guanilato ciclasas responsables de la síntesis de di-GMPc. La cepa mutante *E. coli* TRMG posee niveles elevados de di-GMPc y, por ello presenta un significativo incremento en su capacidad de formar biopelículas respecto a la cepa parental (Jonas y col., 2008).

Se transfirieron por electroporación los plásmidos pIZTolR y pIZ1016 (control) a la cepa mutante *E. coli* TRMG, y se analizó la capacidad de formación de biopelículas de las cepas resultantes *E. coli* TRMG (pIZTolR) y *E. coli* TRMG (pIZ1016) (ver apartados 9.1 y 5 de Materiales y Métodos). Se observó que la cepa complementada con el gen *tol* clonado en pIZTolR mostraba una significativa menor capacidad de formación de biopelículas cuando en el medio había tolueno que la cepa control conteniendo el plásmido pIZ1016 (Fig. 17). Este resultado sugiere que el gen *tol* de la cepa CIB es capaz de disminuir los niveles de di-GMPc en *E. coli* TRMG y, por tanto, reducir su capacidad de formación de biopelículas. En este sentido, un resultado similar se obtuvo cuando la cepa *E. coli* TRMG contenía el plásmido pJN2133 que expresa el gen *PA2133* de *P. aeruginosa* el cual se ha demostrado que codifica una fosfodiesterasa de di-GMPc (Kulesekara y col., 2006) (Fig. 17).

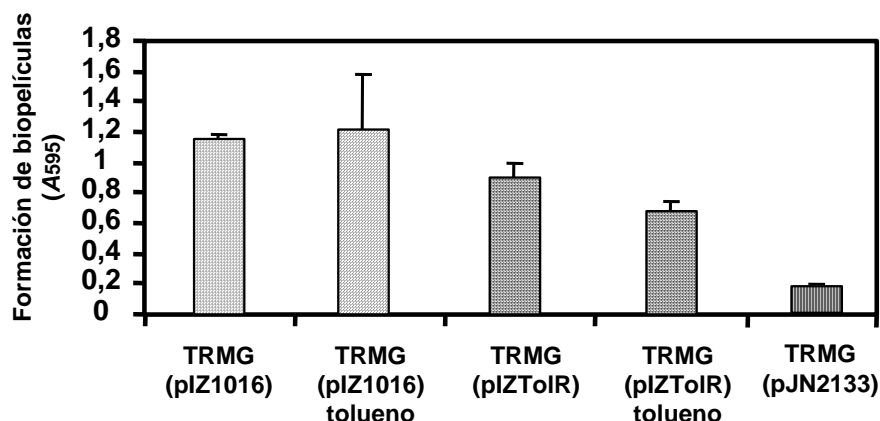


Figura 17. Formación de biopelículas por las cepas *E. coli* TRMG (pIZ1016) y *E. coli* TRMG (pIZToIR). La formación de biopelículas se midió mediante tinción con cristal violeta y posterior cuantificación de la absorbancia a 595nm. Las células fueron cultivadas anaeróbicamente en medio CFA con IPTG 1 mM (plásmidos pIZ1016 y pIZToIR) o arabinosa 50 mM (plásmido pJN2133) y en ausencia o presencia de tolueno (0,5 mM) a 30°C durante 24 h en placas multipocillo. Las barras muestran la desviación estándar de los resultados de tres experimentos independientes.

En conclusión, los resultados presentados en esta Tesis sugieren que el gen *tol* de *Azoarcus* sp. CIB codifica una fosfodiesterasa de di-GMPc cuya actividad responde a la presencia de tolueno en el medio, y está implicado en procesos de adaptación al consumo de este hidrocarburo tales como la respuesta quimiotáctica a esta fuente de carbono.

1.6. El cluster *bss/bbs* está implicado en el catabolismo anaeróbico del *m*-xileno en *Azoarcus* sp. CIB.

Como se ha indicado en el apartado 4.2 de la Introducción, la ruta de degradación anaeróbica del *m*-xileno ha sido mucho menos estudiada que la del tolueno, en parte debido a que la mayoría de los microorganismos degradadores de tolueno no utilizan *m*-xileno anaeróbicamente como fuente de carbono. Estudios realizados en *Azoarcus* sp. T han revelado que la etapa inicial de la degradación anaeróbica de *m*-xileno implica su activación a 3-metilbencilsuccinato por la bencilsuccinato sintasa (BssABC) (Krieger y col., 1999; Achong y col., 2001; Verfürth y col., 2004).

Para determinar la implicación de los genes *bss* y *bbs* en el catabolismo anaeróbico del *m*-xileno, se procedió a analizar el crecimiento anaeróbico de *Azoarcus* sp. CIB y las cepas mutantes *Azoarcus* sp. CIB*tdiR*, *Azoarcus* sp. CIB*dbssD*, y *Azoarcus* sp. CIB*dbbsB* en *m*-xileno como única fuente de carbono. Al igual que ocurrió con el tolueno, ninguna de las cepas mutantes fue capaz de degradar *m*-xileno,

indicando que tanto los genes *tdiSR*, como los *bss* y los *bbs* son necesarios para el catabolismo anaeróbico del *m*-xileno. Estos resultados confirman la participación de la bencilsuccinato sintasa (genes *bss*) en la etapa inicial de la activación del *m*-xileno a 3-metilbencilsuccinato, y constituyen la primera demostración experimental de que los genes *bbs* (o al menos alguno de ellos) son imprescindibles para la oxidación del 3-metilbencilsuccinato. Por otro lado, los productos génicos *tdiSR* también parecen responder al *m*-xileno (o a alguno de sus intermediarios catabólicos) y son necesarios para la degradación anaeróbica de este hidrocarburo. En conclusión, el cluster *bss/bbs* de *Azoarcus* sp. CIB está implicado tanto en la ruta periférica anaeróbica del tolueno como en la del *m*-xileno, y muy probablemente ambas rutas comparten muchas (o todas) de las actividades enzimáticas necesarias para la conversión de estos hidrocarburos en intermediarios centrales.

Por lo que respecta a la implicación del gen *tol* en la degradación anaeróbica del *m*-xileno, la cepa mutante *Azoarcus* sp. CIB*dtol* fue capaz de utilizar este hidrocarburo como fuente de carbono, por lo que el gen *tol* no parece ser indispensable para el crecimiento anaeróbico en *m*-xileno, al igual que ya se había observado con el tolueno.

1.7. Análisis transcripcional del cluster de degradación de tolueno y *m*-xileno en *Azoarcus* sp. CIB.

El análisis de la organización génica del cluster *bss/bbs* de *Azoarcus* sp. CIB revela que los genes se disponen en tres orientaciones distintas y están separados por distancias que oscilan entre 0 y 839 pb, siendo las mayores regiones intergénicas *bssG-bssI* (839 pb) seguida de *tol-bbsI* (231 pb). Esta organización génica permite especular con la presencia de, al menos, tres promotores, *PtdiS*, *PbssD* y *PbbsA*, que controlan la expresión de tres posibles operones *tdi*, *bss* y *bbs*, respectivamente (Fig. 18).

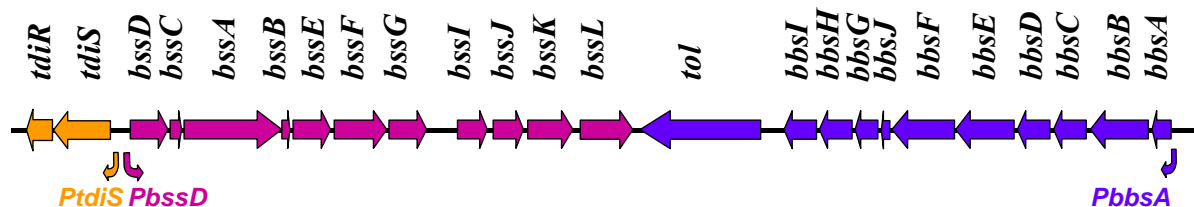


Figura 18. Esquema de la posible organización transcripcional de los genes implicados en la ruta periférica de degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB. Se representan los genes responsables del catabolismo anaeróbico del tolueno y *m*-xileno indicando en color naranja los posibles genes reguladores *tdiR/tdiS*, en color morado los genes del posible operón *bss*, y en color azul se muestran los genes del presunto operón *bbs*. Las flechas curvadas representan los posibles promotores *PtdiS*, *PbssD* y *PbbsA*.

cluster responsable del catabolismo anaeróbico de tolueno/*m*-xileno presente en *Azoarcus* sp. CIB.

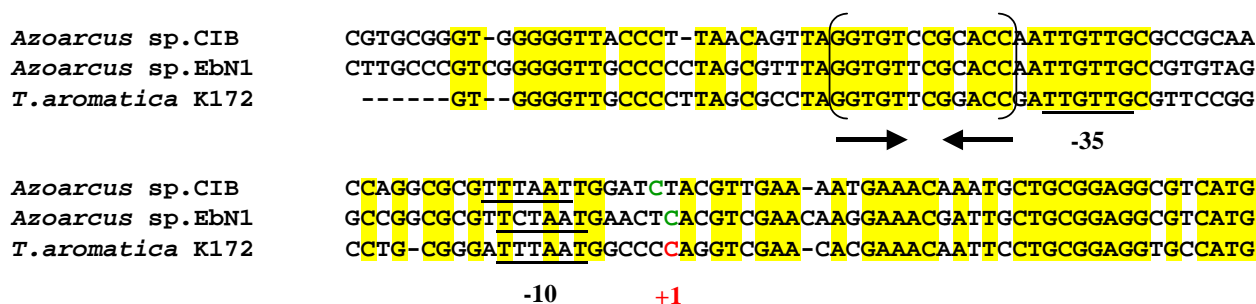


Figura 20. Comparación de las secuencias correspondientes a las regiones promotoras del gen *bbsA* de *Azoarcus* sp. EbN1, *T. aromatica* K172 y *Azoarcus* sp. CIB. En los alineamientos se indican las regiones -10/-35 de interacción con la subunidad σ^{70} de la ARNP (subrayados), así como el inicio de transcripción del gen *bbsA* mapeado en *T. aromatica* K172 (en rojo) y los posibles inicios de transcripción en *Azoarcus* sp. EbN1 y CIB (en verde). Las regiones sombreadas en amarillo representan los nucleótidos conservados en las tres secuencias. Entre corchetes se muestra una región palindrómica que se ha propuesto como la caja operadora de TdiR (Kube y col., 2004).

1.7.2 Estudio de la expresión del cluster génico *bss/bbs* en *Azoarcus* sp. CIB.

Para estudiar la expresión del cluster génico *bss/bbs* en la cepa *Azoarcus* sp. CIB, se analizó si estos genes se inducían cuando la cepa era cultivada en presencia de tolueno y *m*-xileno. Se realizaron experimentos de RT-PCR de los genes *tdiR-tdiS*, *bssA* y *bbsA* con ARN total aislado de *Azoarcus* sp. CIB creciendo en condiciones desnitrificantes y utilizando como única fuente de carbono succinato (condición control), tolueno o *m*-xileno (Fig. 21).

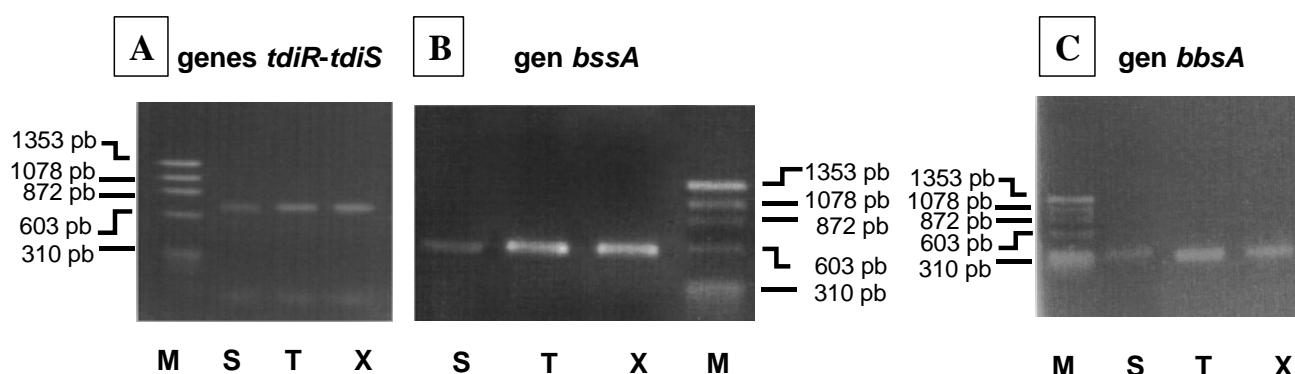


Figura 21. Electroforesis en geles de agarosa de los productos de RT-PCR de los genes *tdiR-tdiS* (A), *bssA* (B) y *bbsA* (C), en *Azoarcus* sp. CIB. Los oligonucleótidos utilizados para las reacciones de amplificación fueron: *tdiRint5'* y *tdiSF.3'* para los genes *tdiR-tdiS*, *bssA5'new* y *bssA3'new* para el gen *bssA*, y *bbsA331.3'* y *bbsA5'new* para el gen *bbsA* (Tabla 4). Las células de *Azoarcus* sp. CIB utilizadas para la extracción del ARN total se cultivaron en condiciones desnitrificantes hasta alcanzar la mitad de la fase exponencial de crecimiento usando succinato 0,2% (calles S), tolueno 250 mM (calles T) o *m*-xileno 250 mM (calles X) como únicas fuentes de carbono. El tolueno y *m*-xileno se añadieron en una fase de HMN. La extracción del ARN y las reacciones de RT-PCR se realizaron como se indica en el apartado 7 de Materiales y Métodos. Se utilizó ADN del fago ϕ X-174 digerido con la enzima de restricción *HaeIII* como marcador de peso molecular (calles M). En el margen izquierdo del panel A se indican los tamaños (en pb) de los marcadores.

Al analizar los productos de las reacciones de RT-PCR mediante electroforesis en geles de agarosa, se observa que el nivel basal de expresión de los genes *tdiR-tdiS*, *bssA* y *bbsA* en células cultivadas en succinato se incrementa significativamente cuando éstas se cultivan en tolueno o *m*-xileno (Fig. 21). Estos resultados confirman que la transcripción del *cluster bss/bbs* es inducible y dependiente del catabolismo de tolueno/*m*-xileno en *Azoarcus* sp. CIB.

Para estudiar la organización transcripcional del *cluster* que controla el catabolismo anaeróbico de tolueno/*m*-xileno, se analizó la cotranscripción de distintas parejas de genes adyacentes incluyendo todos aquellos que poseen las distancias intergénicas mayores. Para ello, se realizaron experimentos de RT-PCR de las correspondientes regiones intergénicas con ARN total aislado de *Azoarcus* sp. CIB cultivado en condiciones desnitrificantes, utilizando como única fuente de carbono tolueno o *m*-xileno (Fig. 22).

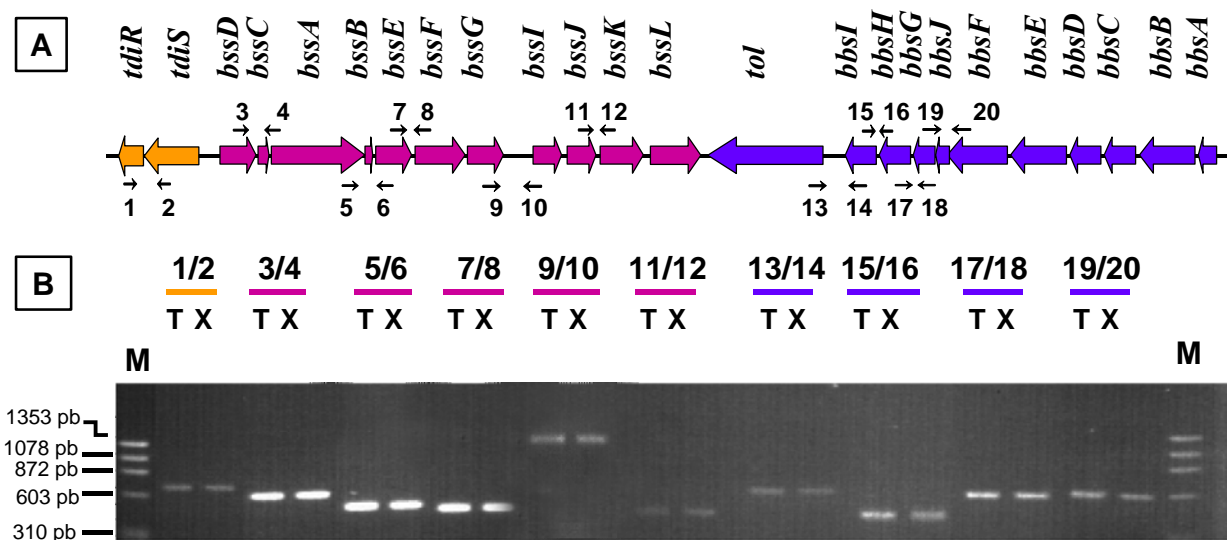


Figura 22. Electroforesis en geles de agarosa de los productos de RT-PCR obtenidos a partir de regiones intergénicas del *cluster bss/bbs* de *Azoarcus* sp. CIB. (A) Esquema de la posible organización transcripcional del *cluster bss/bbs* según se detalla en la Fig. 18. Debajo de la representación del *cluster* se indican con flechas negras los oligonucleótidos utilizados para las reacciones de amplificación de las regiones intergénicas mediante RT-PCR: 1 (*tdiR*int5'), 2 (*tdiS*F.3'), 3 (3T7.3), 4 (*bssA*int3'), 5 (3T7.7), 6 (*bssE*int3'), 7 (3T79), 8 (*bssF*int3'), 9 (*bssG*int5'), 10 (3T3.14), 11 (*c2A200*int5'), 12 (3T3.12), 13 (*TolSR*int3'), 14 (3T3.3), 15 (*bbsI*int3'), 16 (3T3.2), 17 (*bbsH*333.3'), 18 (*bbsG*.F5'), 19 (*bbsG*int3'), 20 (*bbsF*int5') (Tabla 4). (B) Las células de *Azoarcus* sp. CIB utilizadas para la extracción del ARN total se cultivaron en condiciones desnitrificantes hasta alcanzar la mitad de la fase exponencial de crecimiento usando tolueno 250 mM (calles T) o *m*-xileno 250 mM (calles X) añadidos en fase de HMN como únicas fuentes de carbono. La extracción del ARN y las reacciones de RT-PCR se realizaron como se indica en el apartado 7 de Materiales y Métodos. Se utilizó ADN del fago ϕ X-174 digerido con la enzima de restricción *Hae*III como marcador de peso molecular (calles M). En el margen izquierdo de la figura se indican los tamaños (en pb) de los marcadores.

Tras analizar los productos de las reacciones de RT-PCR mediante electroforesis en geles de agarosa, se observaron fragmentos de ADN del tamaño esperado en todas las regiones analizadas (Fig. 22), lo que sugiere la cotranscripción de los genes estudiados y la existencia de, al menos, tres operones dentro del *cluster bss/bbs*: un operón regulador, *tdiSR*, y dos operones catabólicos, *bssDCABEFGIJKL* y *bbsABCDEFJGHItol* (Fig. 18).

1.7.3. Identificación de una cuarta unidad transcripcional en el *cluster bss/bbs*.

La localización del gen *tol* en el extremo 3' del operón *bbs* y su expresión en presencia de tolueno y *m*-xileno (Fig. 22), sugieren su implicación en la ruta periférica de degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB. Así, aunque los resultados obtenidos con el mutante *Azoarcus* sp. CIB*dtol* revelan que la inactivación del gen *tol* no afecta de forma significativa al crecimiento anaeróbico en tolueno o *m*-xileno, sí es posible observar un defecto en quimiotaxis del mutante (apartado 1.5 de Resultados), lo que hace sospechar que el producto del gen *tol* estaría interviniendo en otros procesos más generales y en la posibilidad de una regulación adicional a la del resto de los genes *bss/bbs*.

Para confirmar los resultados obtenidos mediante RT-PCR y demostrar la existencia de un único operón catabólico *bbs* que incluye al gen *tol* en *Azoarcus* sp. CIB, se analizó la expresión de dichos genes en la cepa *Azoarcus* sp. CIB*dbbsB* portadora de una mutación polar en el gen *bbsB* (Fig. 11). La cepa mutante fue cultivada anaerómicamente en medio mínimo MC suplementado con piruvato como fuente de carbono y tolueno como inductor. La expresión de los genes *bbs* localizados en posición 3' del presunto operón catabólico *bbs* se analizó mediante RT-PCR tal y como se indica en la figura 23. Se detectó un fragmento del tamaño esperado para la región amplificada del gen *tol*, mientras que no se observó amplificación de los genes *bbsI* y *bbsH* (Fig. 23). Este resultado sugiere que los genes *bbsA-bbsI* forman un único operón (operón *bbs*) controlado por el promotor *PbbsA*, y que si bien el gen *tol* se cotranscribe con los genes *bbs* (Fig. 22), también es capaz de transcribirse de manera independiente a dichos genes, sugiriendo la presencia de un promotor adicional que se ha denominado *Ptol* (Fig. 23).

Cuando se analizó por RT-PCR la expresión del gen *bssA* en células de *Azoarcus* sp. CIB*dbbsB*, se observó que ésta era inducible en presencia de tolueno, como ocurría en la cepa parental (Fig. 21), lo que indica que la mutación del gen *bbsB* no afecta a la

expresión de los genes *bss* que se transcriben desde una unidad transcripcional diferente (Fig. 23).

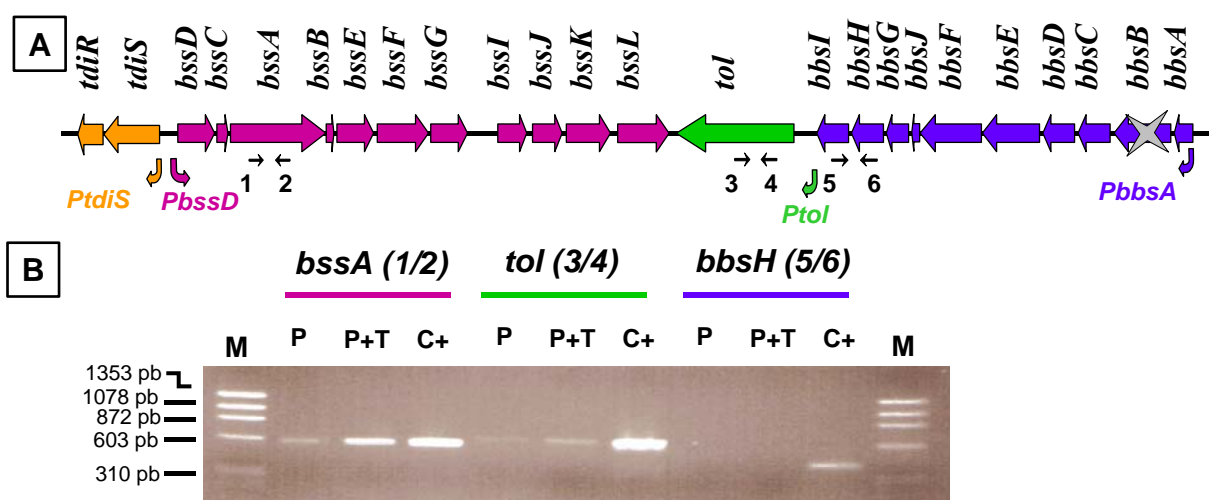


Figura 23. Electroforesis en gel de agarosa de los productos de RT-PCR obtenidos a partir del ARN total de *Azoarcus* sp. CIBdbbsB. (A) Esquema de la organización transcripcional del cluster *bss/bbs* detallándose con flechas curvadas los cuatro promotores presentes en el cluster. La cruz gris representa las mutación polar realizada en el gen *bbsB*. Se indican con flechas convergentes los oligonucleótidos utilizados para las reacciones de RT-PCR: 1 (*bssA*5'new), 2 (*bssA*3'new), 3 (*tolSR2244.3'*), 4 (*3T3.6*), 5 (*bbsIint3'*) y 6 (*3T3.2*) (Tabla 4). (B) Las células de *Azoarcus* sp. CIBdbbsB utilizadas para la extracción del ARN total se cultivaron en condiciones desnitrificantes hasta alcanzar la fase estacionaria de crecimiento usando como fuente de carbono piruvato 0,2% (calles P), o piruvato 0,2% más tolueno 250 mM en fase HMN (calles P+T). Se detallan los productos de amplificación de los genes *bssA*, *tol* y *bbsH-I*. Los controles positivos (calles C+) consistieron en amplificaciones por PCR con los mismos oligonucleótidos usando ADN genómico de *Azoarcus* sp. CIB como molde. Se utilizó ADN del fago ϕ X-174 digerido con la enzima de restricción *Hae*III como marcador de peso molecular (calles M). En el margen izquierdo de la figura se indican los tamaños (en pb) de los marcadores.

Por lo tanto, los resultados obtenidos permiten sugerir que el cluster catabólico de degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB posee al menos cuatro operones: el operón *tdiSR* que incluye los posibles genes reguladores, los operones catabólicos *bss* y *bbs* (incluyendo al gen *tol*) y un cuarto operón que incluye sólo al gen *tol* (Fig. 23).

1.8. Estudio de la regulación transcripcional de los operones catabólicos *bss* y *bbs*.

Los análisis de comparación de secuencias de los genes *tdiSR* (ver apartado 1.4) y su localización adyacente a los genes catabólicos *bss* y *bbs*, sugiere que codifican un sistema regulador de dos componentes (TdiS, proteína sensora; TdiR, regulador

transcripcional) encargado de la regulación transcripcional de la ruta periférica del tolueno y *m*-xileno en *Azoarcus* sp. CIB. Además, el hecho de que mutaciones que inactivan los genes *tdiS* y *tdiR* impidan el crecimiento anaeróbico de las cepas mutantes en tolueno y *m*-xileno (ver apartados 1.3 y 1.6), sugiere que TdiR es un activador transcripcional de los promotores catabólicos *PbbsD* y/o *PbbsA*. Para tratar de confirmar estas hipótesis, se procedió al estudio de la regulación de los promotores *PbbsD* y *PbbsA*.

1.8.1 Clonación y actividad del promotor *PbbsA* en *E. coli*.

Inicialmente se procedió a la clonación de la región 5' del gen *bbsA* (*PbbsA*) en el vector de búsqueda de promotores pSJ3, generando el plásmido pSJ**PbbsA** (Fig. 24). El plásmido recombinante contiene el fragmento de ADN clonado de tal manera que el codón de inicio de la traducción del gen *bbsA* se dispone en fase con el gen reportero *lacZ* del vector, generándose así una fusión traduccional *PbbsA::lacZ*.

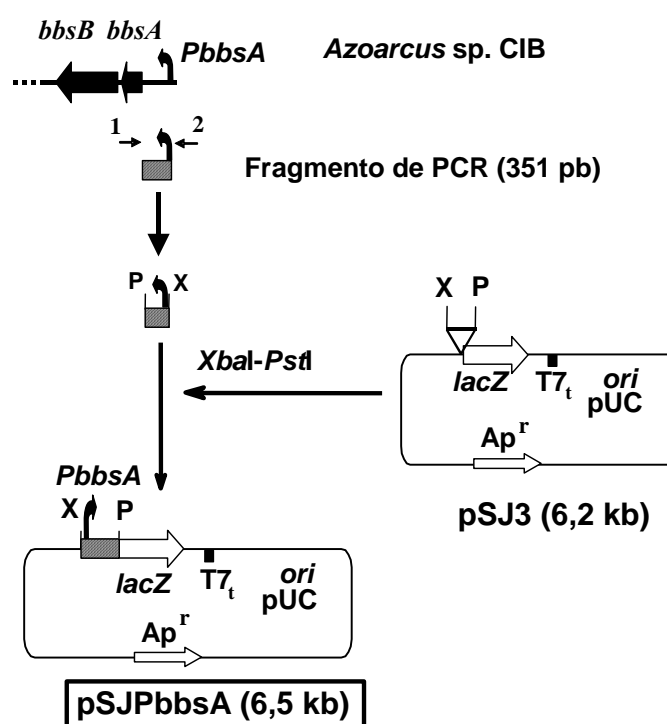


Figura 24. Esquema de la construcción de la fusión traduccional *PbbsA::lacZ* en el plásmido pSJPbbsA**.** Las flechas negras indican los genes *bbs*. La flecha curvada representa el promotor *PbbsA*. Las flechas con número 1 y 2 indican los oligonucleótidos *pbbsA*3' (1) y *pbbsA*5' (2) respectivamente (Tabla 4, Anexo I) que se utilizaron para amplificar por PCR la región promotora *bbsA* y los dos primeros codones del gen *bbsA*, y posteriormente clonarla en fase con el gen *lacZ* del vector de búsqueda de promotores pSJ3. Con flechas blancas se representan el gen *lacZ* y el gen de resistencia a ampicilina (Ap^R). En posición 3' del gen *lacZ* se encuentra un terminador fuerte de la transcripción del fago T7 ($T7_t$). Se indica también el origen de replicación del plásmido (*ori* pUC). Las dianas de restricción utilizadas para la clonación se indican como: *Xba*I (X) y *Pst*I (P).

Dado que la cepa *E. coli* MC4100 (pSJPbbsA) cultivada anaeróbicamente no mostró actividad β -galactosidasa, parecía razonable asumir que el promotor *PbbsA* podría necesitar una proteína reguladora específica para su actividad. Con el fin de comprobar si el presunto sistema regulador de dos componentes *tdiS* y *tdiR* localizado en la posición 5' del *cluster* jugaba algún papel en la expresión del promotor *PbbsA*, se clonaron ambos genes en el plásmido pVTRtdiRS compatible con pSJPbbsA (Fig. 25).

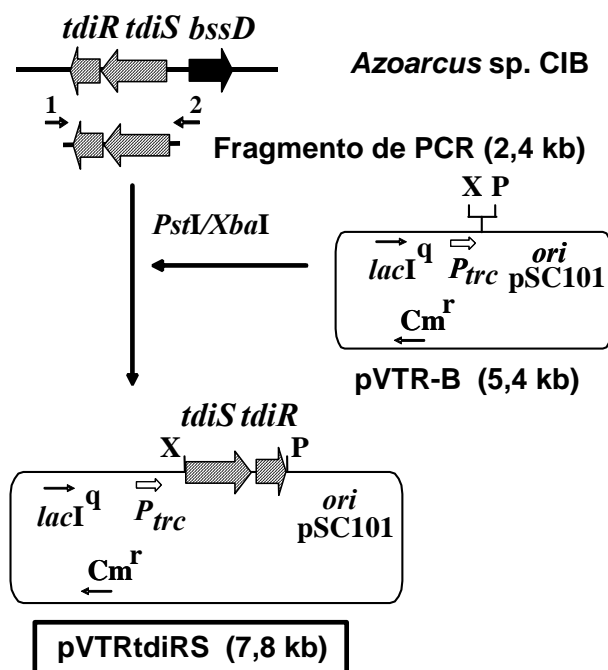


Figura 25. Esquema de la estrategia seguida para la construcción del plásmido pVTRtdiRS. Las flechas rayadas indican los genes *tdiS* y *tdiR*. La flecha negra corresponde al gen catabólico *bssD*. El promotor *P_{trc}* se muestra con flechas blancas. Las flechas con número 1 y 2 indican los oligonucleótidos *tdiRS3'* y *tdiRS5'*, respectivamente (Tabla 4) usados en la amplificación por PCR. Se indica el origen de replicación del plásmido (*ori*pSC101) y el gen *lacI^q*. *Cm^r* representa el gen que codifica la resistencia a cloranfenicol. Las dianas de restricción utilizadas para la clonación se indican como: *XbaI* (X) y *PstI* (P).

La cepa *E. coli* MC4100 portadora de los plásmidos pSJPbbsA y pVTRtdiRS fue cultivada anaeróbicamente en medio M63 suplementado con glicerol 30 mM, como fuente de carbono, IPTG 1 mM y tolueno 1 mM o *m*-xileno 1 mM, como compuestos inductores. Como control se cultivó en las mismas condiciones la cepa *E. coli* MC4100 portadora del plásmido pSJPbbsA y el plásmido parental pVTR-B. No se observó, ni en la cepa portadora de los genes *tdiS*-*tdiR* ni en la cepa control, actividad β -galactosidasa a lo largo de la curva de crecimiento. Este resultado sugería que o bien las proteínas TdiSR no participaban en la activación del promotor *PbbsA*, o que ni el tolueno ni el *m*-xileno eran las moléculas inductoras del sistema de dos componentes. Sin embargo,

tampoco podía descartarse que la ausencia de actividad fuese debida a que ambas proteínas reguladoras no fuesen funcionales en el sistema heterólogo de *E. coli*, y/o que fuese necesario algún otro factor regulador ausente en esta bacteria huésped. Por ello, se decidió estudiar la actividad del promotor *PbbsA*, así como la del promotor *PbssD*, directamente en *Azoarcus* sp. CIB.

1.8.2 Actividad de los promotores *PbbsA* y *PbssD* en *Azoarcus* sp. CIB.

Para investigar la actividad del promotor *PbbsA* en el sistema homólogo *Azoarcus* sp. CIB, se subclonó la fusión *PbbsA::lacZ* en el plásmido de amplio rango de huésped pBBR5T (Tabla 3), generándose el plásmido pBBRPbbsA (Fig. 26).

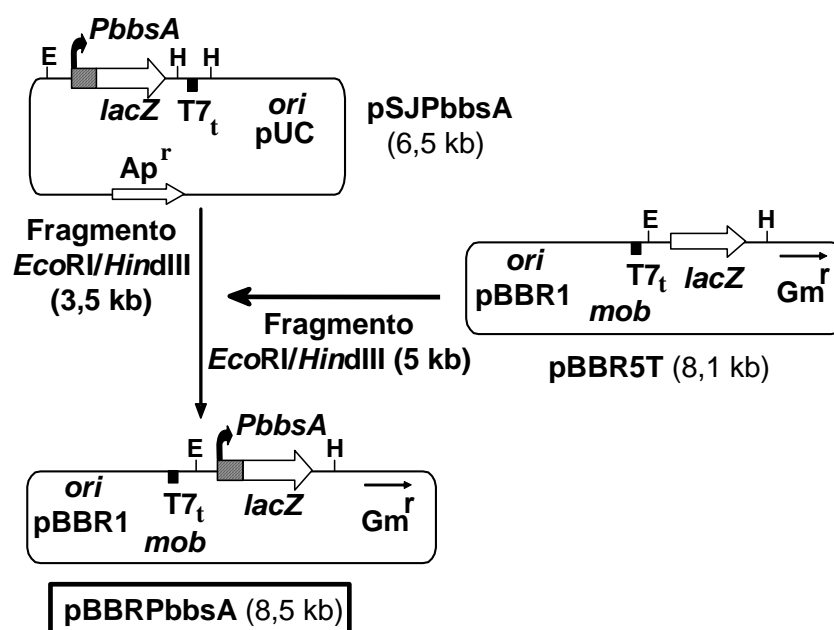


Figura 26. Esquema de la estrategia seguida para la construcción del plásmido pBBRPbbsA. La flecha curvada representa el promotor *PbbsA* y la flecha blanca representa el gen *lacZ*. Se representan los genes que codifican resistencia a ampicilina y gentamicina como *Ap^r* y *Gm^r*, respectivamente. Se indican las funciones de replicación (*ori pBBR1* y *ori pUC*) y de movilización (*mob*). Se muestra también un terminador de la transcripción del fago T7 (*T7_t*). Las dianas de restricción utilizadas para la subclonación se indican como: *HindIII* (H) y *EcoRI* (E).

Para estudiar también la actividad del promotor *PbssD* en *Azoarcus* sp. CIB, se construyó una fusión traduccional *PbssD::lacZ* en el plásmido de amplio espectro de huéspedes pBBRPbssD (Fig. 27).

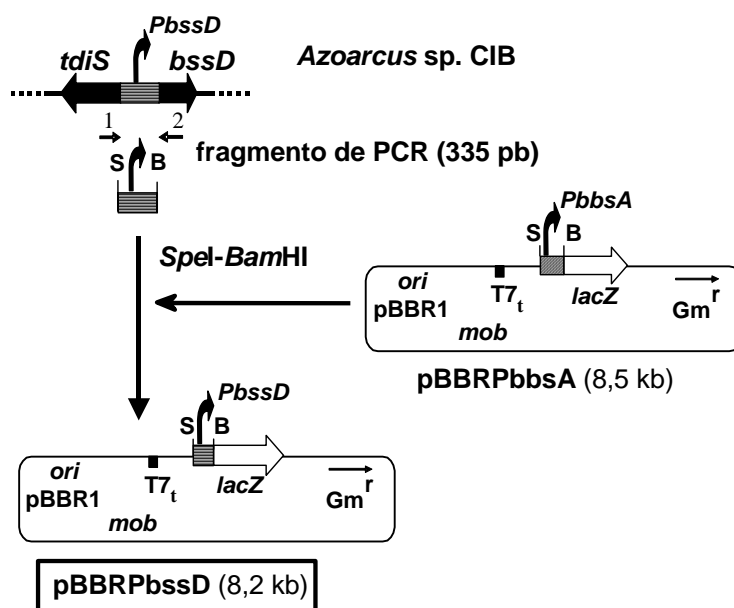


Figura 27. Esquema de la construcción de la fusión traduccional *PbsD::lacZ*. En color negro se indican los genes *tdiS* y *bssD*. En rayado se representa la región intergénica que contiene el promotor *PbsD* (indicado con una flecha curvada). Las flechas 1 y 2 indican los oligonucleótidos ptdiS5' y pbsD3' (Tabla 4, Anexo I) usados en la amplificación por PCR del promotor *PbsD* y los dos primeros codones del gen *bssD*. Gm^r representa el gen que codifica la resistencia a la gentamicina. Se representan las funciones de replicación (*ori* pBBR1) y de movilización (*mob*). En posición 5' del gen *lacZ* se localiza un terminador de la transcripción del fago T7 ($T7_t$). Las dianas de restricción utilizadas para la subclonación se indican como: *Bam*HI (B) y *Spe*I (S).

Para estudiar la expresión de las fusiones traduccionales generadas se cultivaron anaeróbicamente células de *Azoarcus* sp. CIB (pBBRPbbsA) y *Azoarcus* sp. CIB (pBBRPbssD) en piruvato como única fuente de carbono ó piruvato como fuente de carbono suplementado con tolueno como inductor. Las actividades de los promotores *PbsD* y *PbbsA* se ensayaron valorando las actividades β -galactosidasa asociadas a las fusiones traduccionales. Los resultados obtenidos revelan una activación de 6 veces tanto del promotor *PbbsA* como del promotor *PbsD* en presencia de tolueno (Fig. 28). Además, la actividad del promotor *PbsD* fue considerablemente mayor que la del promotor *PbbsA* (Fig. 28), lo que está de acuerdo con la significativa mayor expresión de los genes *bss* respecto a los genes *bbs* detectada en los experimentos de RT-PCR (Fig. 21). Un dato remarcable es el hecho de que los valores de actividad del promotor *PbsD* sin inducir (condiciones de piruvato) son similares a los niveles de actividad del promotor *PbbsA* en presencia de tolueno. Los resultados presentados revelan, por lo tanto, que las fusiones *PbsD::lacZ* y *PbbsA::lacZ* son funcionales en *Azoarcus* sp. CIB y confirman la inducción de los genes *bss* y *bbs*, respectivamente, cuando las células metabolizan tolueno.

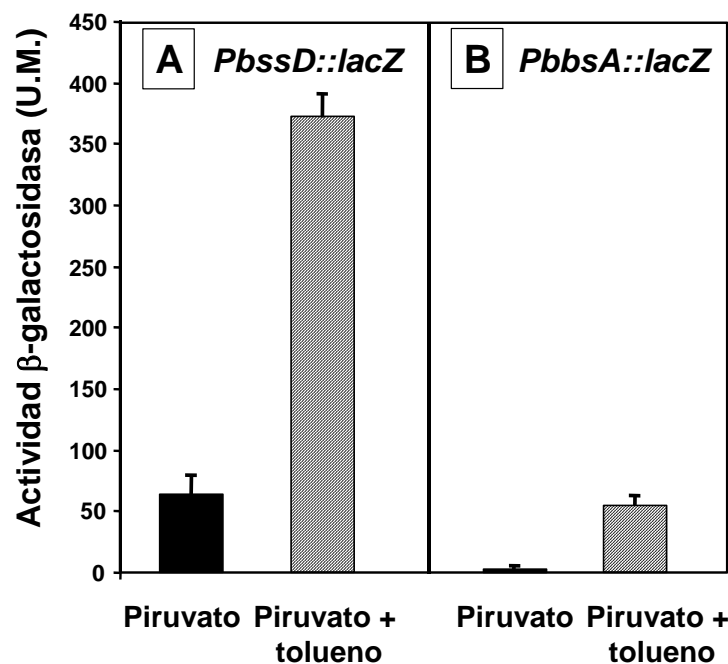


Figura 28. Expresión de las fusiones traduccionales *PbsD::lacZ* y *PbsA::lacZ* en *Azoarcus* sp. CIB. Medida de la actividad β-galactosidasa (en unidades Miller, U.M.) en células de *Azoarcus* sp. CIB que contienen las fusiones traduccionales *PbsD::lacZ* (plásmido pBBRPbsD) (A) o *PbsA::lacZ* (plásmido pBBRPbsA) (B). Las células fueron cultivadas anaeróbicamente hasta alcanzar la fase estacionaria en medio mínimo MC con piruvato 0,2% como fuente de carbono o piruvato 0,2% como fuente de carbono más tolueno 250 mM (en fase de HMN) como compuesto inductor. Las barras indican la desviación estándar de tres experimentos independientes.

1.8.3. TdiR es un activador transcripcional de los genes *bss* y *bbs*.

Con el fin de confirmar si el producto del gen *tdiR* estaba implicado en la activación de los promotores *PbsD* y *PbsA*, se valoró la expresión de las fusiones *PbsD::lacZ* y *PbsA::lacZ* en un fondo génico carente del gen *tdiR*, i. e., en las cepas *Azoarcus* sp. CIB*tdiR* (pBBRPbsD) y *Azoarcus* sp. CIB*tdiR* (pBBRPbsA). Se analizó la actividad β-galactosidasa de estas células cultivadas anaeróbicamente en medio MC con piruvato como fuente de carbono (condición control) y con piruvato como fuente de carbono más tolueno como inductor (Fig. 29). La cepa *Azoarcus* sp. CIB*tdiR* (pBBRPbsA) no presenta actividad β-galactosidasa en ninguna condición y, por lo tanto, no hay actividad del promotor *PbsA* aunque en el medio haya tolueno (Fig. 29). Por el contrario, la cepa *Azoarcus* sp. CIB*tdiR* (pBBRPbsD) presenta los niveles basales de actividad β-galactosidasa cuando se cultiva en piruvato y estos niveles se incrementan en presencia de tolueno (Fig. 29). Estos resultados sugieren que el producto del gen *tdiR* es imprescindible para activar al promotor *PbsA* pero no para

la activación del promotor *PbssD*, si bien la actividad máxima de este último se ve reducida a la mitad en el mutante *tdiR* respecto a la cepa parental CIB (Fig. 29).

Por lo tanto, los dos operones catabólicos *bss* y *bbs* de degradación anaeróbica de tolueno en *Azoarcus* sp. CIB muestran una regulación diferencial. Mientras que el operón *bbs* está estrictamente regulado (activado) por la proteína TdiR, el operón *bss* tiene un mayor nivel de expresión basal y su activación en presencia de tolueno está sólo parcialmente mediada por el regulador TdiR.

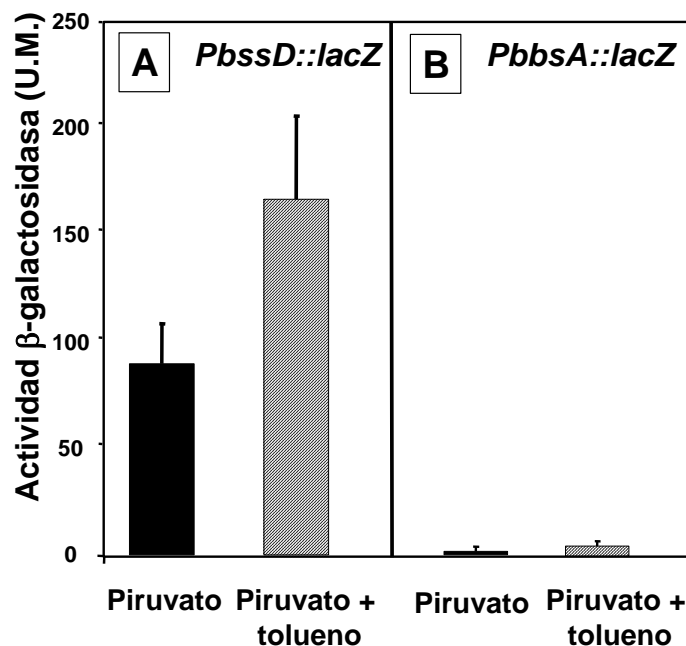


Figura 29. Expresión de las fusiones traduccionales *PbsD::lacZ* (A) y *PbsA::lacZ* (B) en la cepa mutante *Azoarcus* sp. CIB*tdiR*. Medida de la actividad β-galactosidasa (en unidades Miller, U.M.) en células de *Azoarcus* sp. CIB*tdiR* que contienen las fusiones traduccionales *PbsD::lacZ* (plásmido pBBRPbsD) o *PbsA::lacZ* (plásmido pBBRPbsA). Las células utilizadas en este experimento fueron cultivadas anaeróbicamente hasta alcanzar la fase estacionaria en medio mínimo MC con piruvato 0,2% como fuente de carbono o piruvato 0,2% como fuente de carbono más tolueno 250 mM como inductor (suministrado en fase de HMN). Las barras indican la desviación estándar de tres experimentos independientes.

1.8.4 Caracterización de los inductores del operón catabólico *bbs*.

Como se ha indicado en el apartado 1.7.2, la ruta periférica de la degradación anaeróbica de tolueno es inducible en *Azoarcus* sp. CIB y la expresión de los genes *bss/bbs* se produce cuando las células se cultivan en presencia de tolueno o *m*-xileno (Fig. 21).

Para determinar si estos dos hidrocarburos son los verdaderos inductores del promotor *PbsA* dependiente de TdiR, se valoró la expresión de la fusión traduccional *PbsA::lacZ* en las cepas mutantes *Azoarcus* sp. CIB*dbssD* y *Azoarcus* sp. CIB*dbbsB* que tienen bloqueada la activación del tolueno a bencilsuccinato (incapaces de

metabolizar tolueno) y la β -oxidación de este último a benzoil-CoA (acumulan bencilsuccinato), respectivamente.

Se realizaron experimentos de actividad β -galactosidasa en las cepas *Azoarcus* sp. CIBdbssD (pBBRPbbsA) y *Azoarcus* sp. CIBdbbsB (pBBRPbbsA) cultivadas anaeróbicamente en piruvato como única fuente de carbono o en piruvato como fuente de carbono más tolueno o *m*-xileno como inductores. Se observó que la cepa *Azoarcus* sp. CIBdbssD (pBBRPbbsA), incapaz de iniciar el catabolismo anaeróbico del tolueno y *m*-xileno, no presentaba actividad β -galactosidasa en ninguna de las condiciones ensayadas, sugiriendo que ninguno de estos dos hidrocarburos aromáticos son los verdaderos inductores del operón *bbs* (Fig. 30).

La cepa *Azoarcus* sp. CIBdbbsA (pBBRPbbsA), que conserva el operón *bbs* activo y, por tanto, es capaz de activar el tolueno y *m*-xileno a bencilsuccinato y 3-metilbencilsuccinato, respectivamente, sí presentaba actividad β -galactosidasa cuando las células habían crecido en presencia de tolueno o *m*-xileno. Por el contrario, otros hidrocarburos aromáticos tales como *p*-xileno, *o*-xileno o estireno fueron incapaces de activar el promotor *PbbsA* (Fig. 30). Todos estos resultados sugieren que los verdaderos inductores del operón *bbs* son los primeros intermediarios de la ruta de degradación anaeróbica de tolueno y *m*-xileno, es decir, el bencilsuccinato y el 3-metilbencilsuccinato, respectivamente.

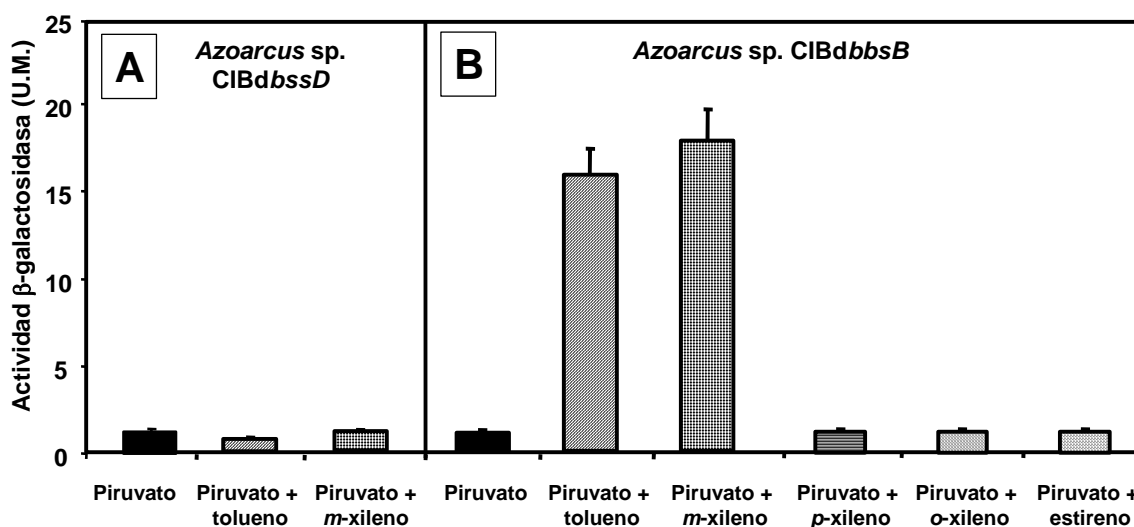


Figura 30. Expresión de la fusión traduccional *PbbsA::lacZ* en las cepas mutantes *Azoarcus* sp. CIBdbssD (A) y *Azoarcus* sp. CIBdbbsB (B). Se muestra la actividad β -galactosidasa (en unidades Miller, U.M.) de células de *Azoarcus* sp. CIBdbssD y *Azoarcus* sp. CIBdbbsB que contienen la fusión traduccional *PbbsA::lacZ* (plásmido pBBRPbbsA). Las células fueron cultivadas anaeróbicamente hasta alcanzar la fase estacionaria en medio MC con piruvato 0,2% como fuente de carbono o en piruvato 0,2% como fuente de carbono más tolueno, *m*-xileno, *p*-xileno, *o*-xileno o estireno 250mM (en fase de HMN) como inductores. Las barras indican la desviación estándar de tres experimentos independientes.

1.8.5 Caracterización de la inducción independiente de TdiR del operón catabólico *bss*.

Para la caracterización de la inducción independiente de TdiR del operón catabólico *bss*, se valoró la expresión de la fusión traduccional *PbssD::lacZ* en la cepa mutante *Azoarcus* sp. CIB*tdiR*. Se realizaron experimentos de actividad β -galactosidasa de la cepa *Azoarcus* sp. CIB*tdiR* (pBBRP*bssD*) cultivada anaeróticamente en piruvato como única fuente de carbono o en piruvato como fuente de carbono más tolueno, *m*-xileno, *p*-xileno o estireno como inductores (Fig. 31). Se observó que la cepa *Azoarcus* sp. CIB*tdiR* (pBBRP*bssD*), incapaz de iniciar el catabolismo anaeróbico del tolueno y *m*-xileno, presentaba mayor actividad β -galactosidasa en presencia de cualquiera de los hidrocarburos ensayados. Este dato sugiere que existe una regulación del promotor *PbssD* independiente de TdiR pero que responde a la presencia de aromáticos. Dado que los mejores inductores de la actividad independiente de TdiR del promotor *PbssD* son hidrocarburos aromáticos que no pueden ser metabolizados por *Azoarcus* sp. CIB, por ejemplo *p*-xileno y estireno, la inducción observada podría ser el resultado de una respuesta al estrés general ocasionado por los hidrocarburos aromáticos en la cepa CIB.

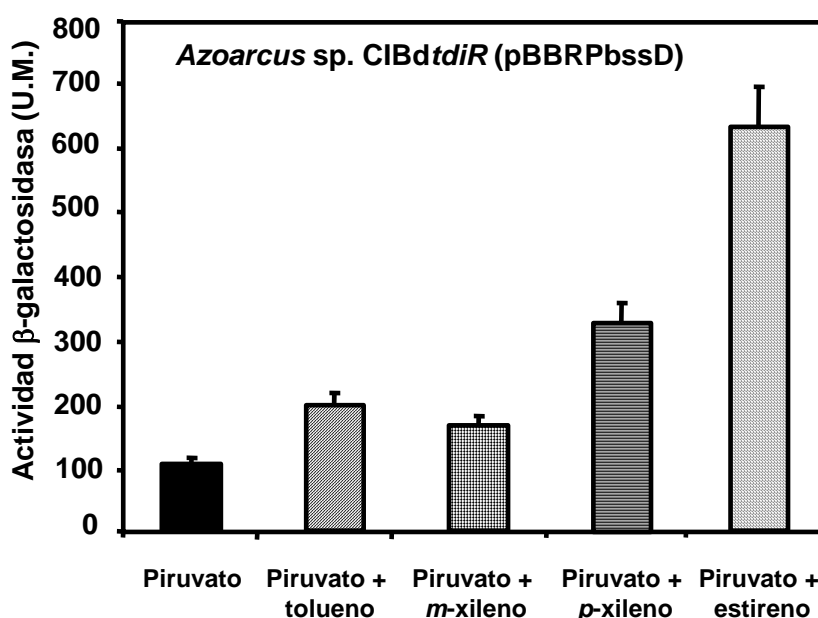


Figura 31. Expresión de la fusión traduccional *PbssD::lacZ* en la cepa mutante *Azoarcus* sp. CIB*tdiR*. Se muestra la actividad β -galactosidasa (en unidades Miller, U.M.) de células de *Azoarcus* sp. CIB*tdiR* que contienen la fusión traduccional *PbssD::lacZ* (plásmido pBBRP*bssD*). Las células fueron cultivadas anaeróticamente hasta alcanzar la fase estacionaria en medio MC con piruvato 0,2% como fuente de carbono, o piruvato 0,2% como fuente de carbono más tolueno, *m*-xileno, *p*-xileno o estireno 250mM (en fase de HMN) como inductores. Las barras indican la desviación estándar de tres experimentos independientes.

1.9. Regulación sobreimpuesta de los genes *bbs* en *Azoarcus* sp. CIB.

Como se ha indicado en el apartado 3.2 de Introducción, a la regulación específica que controla la expresión de rutas catabólicas se superpone una regulación más general que depende del estado fisiológico global de la célula y que condiciona la expresión de los genes catabólicos a la existencia de otras fuentes de carbono preferentes (represión catabólica) o a otros factores ambientales tales como la disponibilidad de oxígeno.

Para determinar si los genes *bbs* de la ruta periférica de degradación anaeróbica de tolueno/*m*-xileno en *Azoarcus* sp. CIB estaban sometidos a una regulación sobreimpuesta, se procedió a analizar la expresión de dichos genes en presencia de fuentes de carbono adicionales y bajo condiciones aeróbicas.

1.9.1 Represión catabólica de los genes *bbs*.

Para determinar si la expresión de los genes *bbs* en *Azoarcus* sp. CIB estaba sometida a represión catabólica por fuentes de carbono preferenciales, se monitorizó la actividad del promotor *PbbsA::lacZ* a lo largo del tiempo en células de *Azoarcus* sp. CIB (pBBRPbbsA) cultivadas anaeróticamente en medio mínimo MC conteniendo una mezcla de tolueno y distintas fuentes de carbono, tales como piruvato, alanina, glutamato, glutarato y benzoato (Fig. 32).

Los resultados obtenidos muestran que existe inhibición de la actividad del promotor *PbbsA* por todas las fuentes de carbono probadas (en la figura 32 se muestran los datos obtenidos con piruvato). La inhibición se observa desde el inicio de la curva de crecimiento y hasta que se alcanza un consumo significativo de la fuente de carbono preferencial, i.e, 72 h de incubación, lo que se corresponde con una fase exponencial tardía de crecimiento. A partir de este momento, la actividad del promotor *PbbsA* se hace patente (Fig. 32). Resultados similares se obtuvieron cuando en lugar de tolueno se utilizó *m*-xileno como segunda fuente de carbono.

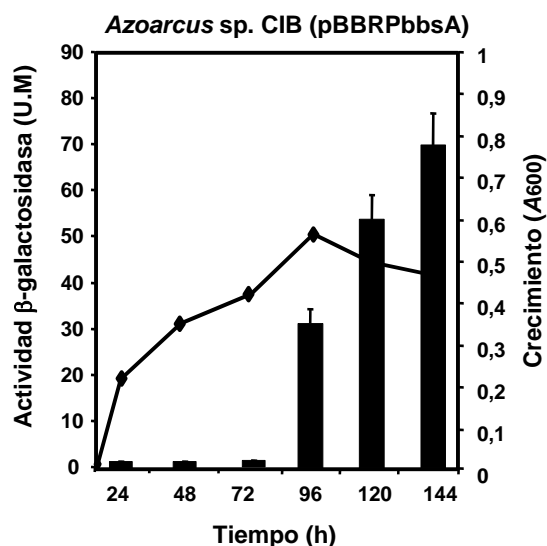


Figura 32. Actividad del promotor *PpbsA* a lo largo de la curva de crecimiento. Se monitorizó la actividad β-galactosidasa a lo largo de la curva de crecimiento (rombos) de *Azoarcus* sp. CIB (pBBRPpbsA) cultivado anaeróbicamente en presencia de piruvato (0,2%) más tolueno (250 mM en fase de HMN). A las 72 h de incubación, se adicionó más KNO_3 a los cultivos.

Estos resultados indican, por lo tanto, la existencia de represión catabólica de los genes *bbs* por todas las fuentes de carbono probadas, lo que confirma que el tolueno/*m*-xileno no son sustratos preferentes para *Azoarcus* sp. CIB.

1.9.2 Expresión de los genes *bbs* en presencia de oxígeno.

Se procedió también a estudiar si la expresión de los genes *bbs* está bajo regulación por oxígeno. Para ello, se analizó mediante experimentos de RT-PCR la expresión del gen *bbsB* en células de *Azoarcus* sp. CIB cultivadas en tolueno como fuente de carbono tanto en condiciones aeróbicas como anaeróbicas (Fig. 33). Los resultados obtenidos muestran que la expresión del operón *bbs* está reprimida en presencia de oxígeno y sólo en condiciones de anaerobiosis se produce su transcripción. Un resultado similar se obtuvo para los genes *bss* (datos no mostrados).

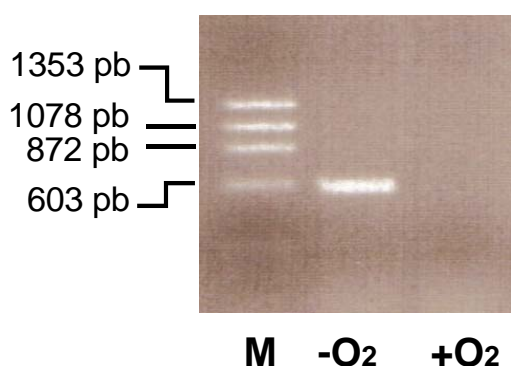


Figura 33. Electroforesis en gel de agarosa del producto de RT-PCR del gen *bbsA* de *Azoarcus* sp. CIB cultivado en tolueno. Los oligonucleótidos utilizados para las reacciones de amplificación mediante RT-PCR fueron: *bbsA*331.3' y *bbsA*5'new (Tabla 4). Las células de *Azoarcus* sp. CIB utilizadas para la extracción del ARN total se cultivaron en medio MC hasta alcanzar la mitad de la fase exponencial de crecimiento usando tolueno 2 mM en condiciones aeróbicas o tolueno 250 mM (en fase HMN) en condiciones anaeróbicas. La extracción del ARN y las reacciones de RT-PCR células cultivadas en ausencia o presencia de oxígeno se realizaron como se indica en el apartado 7 de Materiales y Métodos. Se utilizó ADN del fago ϕ X-174 digerido con la enzima de restricción *Hae*III como marcador de peso molecular (calle M). En el margen derecho de la figura se indican los tamaños (en pb) de los marcadores.

2. CARACTERIZACIÓN Y REGULACIÓN DEL GEN *gcdH* IMPLICADO EN LA RUTA BAJA DE DEGRADACIÓN ANAERÓBICA DE TOLUENO EN *Azoarcus* sp. CIB.

Como se ha indicado en la Introducción, el catabolismo anaeróbico periférico del tolueno genera finalmente el intermediario central benzoil-CoA. La ruta alta del benzoil-CoA transforma éste en un compuesto alifático, i.e., pimelil-CoA (en la bacteria fotosintética *R. palustris*) o 3-hidroxpimelil-CoA (en el resto de los organismos anaerobios degradadores de benzoil-CoA descritos hasta la fecha) (Harwood y Gibson, 1997; Gibson y Harwood, 2002; López-Barragán y col., 2004a; Harrison y Harwood, 2005). A la serie de reacciones que transforman el pimelil-CoA/3-hidroxpimelil-CoA en CO₂ y acetil-CoA se le ha denominado ruta baja del benzoil-CoA (Fig. 4). En dicha ruta, la oxidación del glutaril-CoA a crotonil-CoA, catalizada por la enzima glutaril-CoA deshidrogenasa (GcdH), es una de las etapas limitantes y la que conecta la serie de reacciones de β -oxidación de ácidos dicarboxílicos de cadena media-larga con la ruta general de β -oxidación de ácidos grasos de cadena corta (Fig. 4) (Gallus y Schink, 1994; Harrison y Harwood, 2005). Hasta la fecha no se había caracterizado ningún gen bacteriano responsable de la actividad GcdH y no había ningún dato experimental concluyente de que esta actividad fuese esencial para el crecimiento anaeróbico en compuestos aromáticos. Por ello, en esta tesis se decidió caracterizar el gen que codifica la actividad GcdH de *Azoarcus* sp. CIB.

2.1 Estudio de la inducción de la actividad glutaril-CoA deshidrogenasa (GcdH) de *Azoarcus* sp. CIB.

Dado que *Azoarcus* sp. CIB es capaz de crecer anaeróbicamente en compuestos aromáticos, tales como tolueno y benzoato, así como en otras fuentes de carbono alifáticas, e. g., pimelato y glutarato, y alicíclicas, e. g., ciclohexano carboxilato, que generan glutaril-CoA en su catabolismo, se decidió investigar la actividad GcdH de *Azoarcus* sp. CIB cultivado en dichas fuentes de carbono. Con este fin, se cultivaron anaeróbicamente células de *Azoarcus* sp. CIB en medio mínimo MC suplementado con diversas fuentes de carbono tales como pimelato, benzoato, piruvato, succinato, acetato ó una mezcla de piruvato y pimelato, y se midió su actividad glutaril-CoA deshidrogenasa. Extractos celulares de *Azoarcus* sp. CIB cultivado anaeróbicamente en

pimelato o benzoato presentaron una actividad GcdH significativa, en comparación con la obtenida con extractos procedentes de células cultivadas anaeróbicamente en piruvato, acetato o succinato (Fig. 34). También se observó una significativa actividad GcdH en células cultivadas en piruvato más pimelato (Fig. 34). Estos resultados muestran la existencia de una actividad GcdH en *Azoarcus* sp. CIB que se induce específicamente cuando las células metabolizan anaeróbicamente pimelato o benzoato, compuestos que originan glutaril-CoA en su degradación. Además, el hecho de que haya inducción de la actividad GcdH cuando las células crecen en piruvato y pimelato, sugiere que el piruvato no ejerce represión catabólica sobre la producción/actividad de esta enzima.

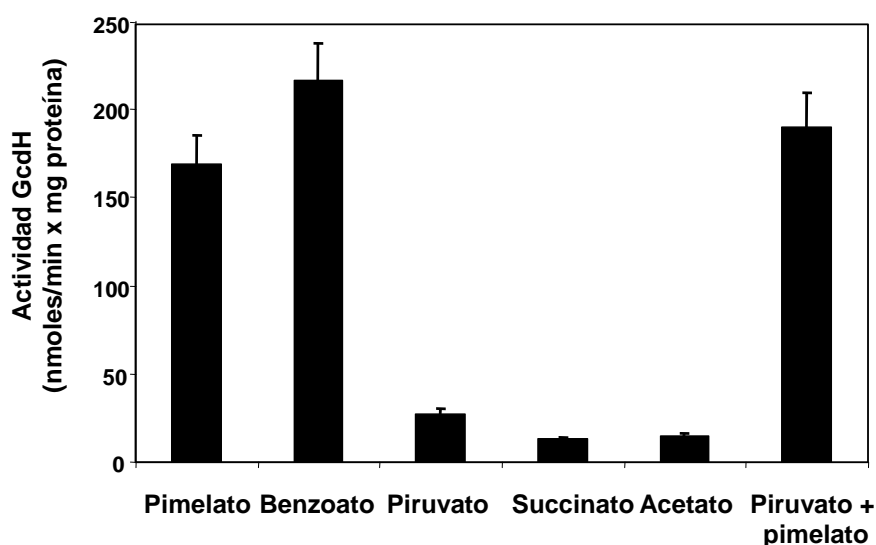


Figura 34. Actividad glutaril-CoA deshidrogenasa de *Azoarcus* sp. CIB. Células de *Azoarcus* sp. CIB fueron cultivadas anaeróbicamente hasta la mitad de la fase exponencial de crecimiento en medio mínimo MC conteniendo diferentes fuentes de carbono (pimelato 5 mM, benzoato 3 mM, piruvato 0,2%, succinato 0,2%, acetato 0,2%, ó piruvato 0,2% más pimelato 10 mM). La actividad GcdH se midió en extractos celulares tal y como se describe en el apartado 13.2 de Materiales y Métodos. Las barras indican la desviación estándar de tres experimentos independientes.

2.2 Identificación del gen *gcdH* de *Azoarcus* sp. CIB.

Tras comprobar que existía una actividad GcdH inducible en *Azoarcus* sp. CIB, se planteó identificar al gen *gcdH* responsable de la misma. Dado que en el genoma de *Azoarcus* sp. EbN1 se había identificado un ORF anotado como un presunto gen *gcdH* (*ebA2993*), se empleó esta secuencia como referencia para el diseño de dos oligonucleótidos (3'GcdEbN1 y 5'GcdEbN1) que permitieron la amplificación

mediante PCR de un fragmento interno del presunto gen *gcdH* de la cepa CIB. Dicho fragmento interno se clonó en el vector suicida pK18*mob*, obteniéndose de esta forma el plásmido pK18*mobgcdH* (Fig. 35) Este plásmido se transfirió mediante conjugación biparental a la cepa *Azoarcus* sp. CIB, seleccionándose la cepa mutante *Azoarcus* sp. CIB*dgcdH* (Fig. 35).

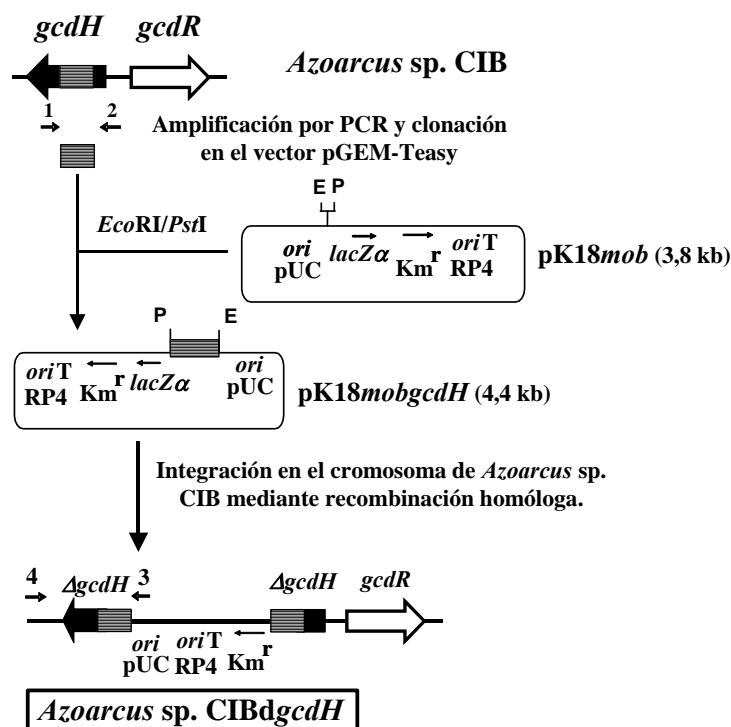


Figura 35. Esquema de la construcción de la cepa mutante *Azoarcus* sp. CIB*dgcdH*. En la parte superior de la figura se esquematizan los genes *gcdH* (negro) y *gcdR* (blanco) (cuyas secuencias se indican en la figura 53 del Anexo II). Las flechas con número 1 y 2 representan los oligonucleótidos 5'GcdEbN1 (1) y 3'GcdEbN1 (2) (Tabla 4), respectivamente, utilizados para la amplificación por PCR del fragmento interno (rayado) del gen *gcdH*. Las flechas con número 3 y 4 representan los oligonucleótidos R24 (3) y 3GcdH (4) (Tabla 4) que se emplearon para comprobar por PCR la correcta inserción del gen *gcdH* en el cromosoma de la cepa mutante *Azoarcus* sp. CIB*dgcdH*. Se representan los orígenes de replicación (*ori*pUC) y de transferencia (*ori*TRP4) del vector suicida pK18*mob*. *Km^r* representa el gen que codifica la resistencia a kanamicina. *lacZα*, secuencia que codifica el fragmento α de la β -galactosidasa de *E. coli*. Las dianas de restricción utilizadas para la clonación se indican como: *Eco*RI (E) y *Pst*I (P). Δ , indica gen truncado.

La cepa mutante *Azoarcus* sp. CIB*dgcdH* fue incapaz de crecer anaeróbicamente en medio mínimo MC cuando se utilizaron como fuentes de carbono compuestos aromáticos, tales como benzoato (3mM), tolueno (250 mM en fase HMN) ó 3-hidroxibenzoato (3 mM), compuestos alicíclicos, tales como ciclohexano carboxilato (3 mM), o ácidos dicarboxílicos, tales como el pimelato (5 mM) o glutarato (5 mM) (Fig. 36). Además, no se observó actividad GcdH en extractos crudos de células de *Azoarcus* sp. CIB*dgcdH* cultivadas en piruvato más pimelato o benzoato. Estos resultados

sugieren que el gen *gcdH* codifica la actividad GcdH en *Azoarcus* sp. CIB y, por primera vez, se aporta un dato experimental que indica que esta actividad es esencial para el catabolismo anaeróbico de compuestos aromáticos que se degradan vía benzoil-CoA, de compuestos alicíclicos como el ciclohexanocarboxilato y de ácidos dicarboxílicos de número impar (e igual o superior a 5) de átomos de carbono.

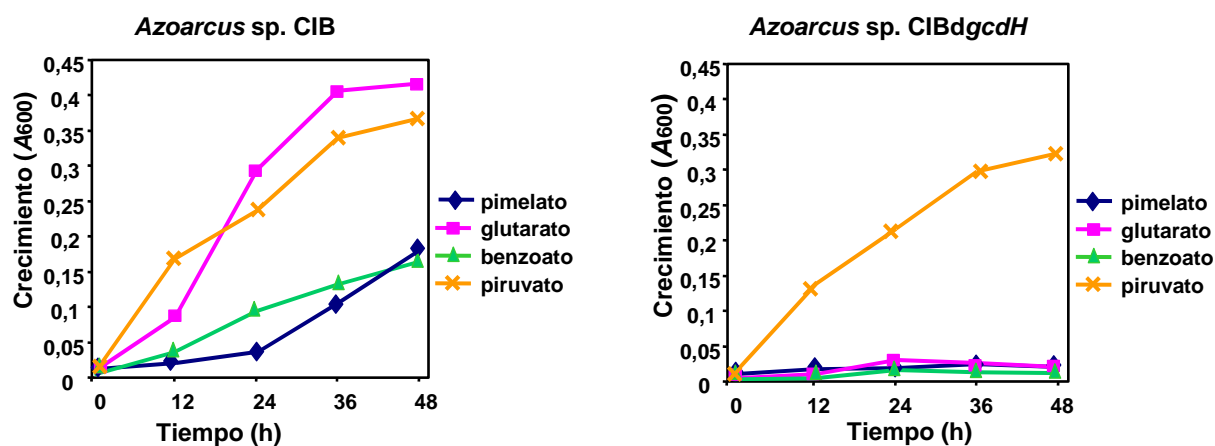


Figura 36. Crecimiento anaeróbico de *Azoarcus* sp. CIB y *Azoarcus* sp. CIBdgcDH utilizando diferentes fuentes de carbono. Se valoró el incremento de absorbancia a A₆₀₀ de *Azoarcus* sp. CIB y *Azoarcus* sp. CIBdgcDH cultivado anaeróbicamente en medio mínimo MC con KNO₃ 10 mM como aceptor final de electrones, y con pimelato 5 mM, glutarato 5 mM, benzoato 3 mM o piruvato 0,2% como únicas fuentes de carbono.

Llama la atención el hecho de que mientras que *Azoarcus* sp. CIBdgcDH no fue capaz de crecer anaeróbicamente en medio mínimo MC suplementado con tolueno (250 mM en fase HMN) si fue capaz de hacerlo utilizando *m*-xileno (250 mM en fase HMN) como única fuente de carbono. Este resultado sugiere que mientras que el catabolismo del tolueno a través del benzoil-CoA produce glutaril-CoA, la degradación anaeróbica del *m*-xileno a través del 3-metilbenzoil-CoA (Krieger y col., 1999) genera un intermediario alifático cuyo posterior metabolismo no requiere la enzima GcdH, sino otra enzima distinta y específica para el catabolismo de este hidrocarburo.

2.3. Clonación y caracterización del gen *gcdH* de *Azoarcus* sp. CIB

Para la identificación del gen *gcdH* en la genoteca fágica de ADN de *Azoarcus* sp. CIB (López-Barragán y col., 2004a), se utilizó como sonda el fragmento interno del gen *gcdH* tal y como se describe en el apartado 6.3 de Materiales y Métodos. Se seleccionó un clon positivo, λgcdH4, que contenía el gen *gcdH* tal y como se confirmó

por PCR empleando los oligonucleótidos 3'GcdEbN1 y 5'GcdEbN1. Se secuenció un fragmento clonado en el fago λ gcdH4 y el análisis de la secuencia reveló la existencia de dos ORFs, *gcdH* y *gcdR*, con orientación divergente. Esta secuencia fue depositada en la base de datos del GenkBank con el número de acceso EF152775 (ver Fig. 52 del anexo II).

El producto del gen *gcdH* muestra una significativa identidad (mayor del 85%) con otros genes bacterianos de *Azoarcus* spp. y *Pseudomonas* spp. que codifican posibles glutaril-CoA deshidrogenasas. El producto del gen *gcdR* presenta elevada identidad con posibles reguladores transcripcionales de la familia LysR (Tabla 6).

Tabla 6. Análisis del fragmento de ADN que contiene los genes *gcdH* y *gcdR* de *Azoarcus* sp. CIB. En la tabla se detallan los genes *gcdH* y *gcdR* de *Azoarcus* sp. CIB, su contenido G+C, la longitud de aminoácidos (aa) y la masa molecular (kDa) de los productos génicos, así como la función que se ha asignado a cada uno de ellos. En los productos génicos relacionados se detalla el nombre, la longitud en aminoácidos (aa), la posible función, el grado de identidad (Id(%)) que presenta con las proteínas de *Azoarcus* sp. CIB y su N° de acceso en el GenBank. En el anexo II se detalla la secuencia completa del fragmento de ADN de 5,3 kb, así como los oligonucleótidos utilizados para su secuenciación y diversas dianas de enzimas de restricción.

Gen	G+C %	Producto génico aa/kDa	Función	Productos génicos relacionados				
				Nombre/aa	Función	Organismo	Id %	N° Acceso
<i>gcdH</i>	66,5	395/43,3	Glutaril-CoA deshidrogenasa	GcdH/395	Hipotética glutaril-CoA deshidrogenasa	<i>Azoarcus</i> sp. EbN1	93	YP_158711
<i>gcdR</i>	69,4	303/33,9	Regulador transcripcional	EbA2995/303	Posible regulador transcripcional de la familia LysR	<i>Azoarcus</i> sp. EbN1	83	YP_158712

Para comprobar experimentalmente que el gen *gcdH* codifica una proteína con actividad glutaril-CoA deshidrogenasa, se procedió a clonar este gen, expresar su producto génico y valorar su actividad enzimática. Para ello, se amplificó mediante PCR el gen *gcdH* del cromosoma de *Azoarcus* sp. CIB y el fragmento resultante (1,3 kb) fue clonado en el vector de amplio rango de huésped pIZ1016. El plásmido resultante, pGCDH, contiene el gen *gcdH* bajo el control del promotor *Ptac* inducible por IPTG (Fig. 37).

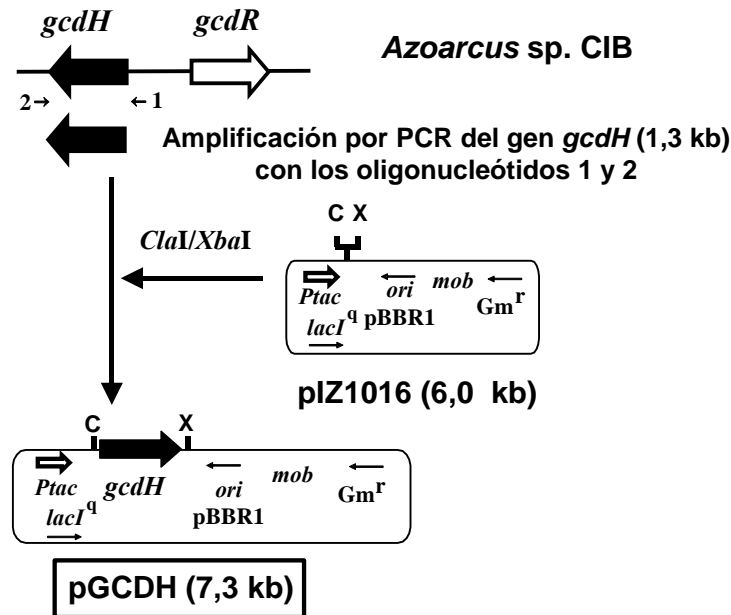


Figura 37. Esquema de la estrategia seguida para la construcción del plásmido pGCDH.

En la parte superior de la figura se esquematizan los genes *gcdR* (blanco) y *gcdH* (negro). Las flechas con número 1 y 2 representan los oligonucleótidos 5GcdH (1) y 3GcdH (2) (Tabla 4), utilizados para la amplificación por PCR del gen *gcdH*. Se representa el origen de replicación (*ori*pBBR1), la región de movilización (*mob*), el gen que codifica la resistencia a gentamicina (*Gm^r*) y el gen *lacI^q*. Se representa con una flecha blanca el promotor *Ptac*. Las dianas de restricción utilizadas para la clonación se indican como: *ClaI* (C) y *XbaI* (X).

La cepa *E. coli* DH5 α (pGCDH) fue cultivada hasta una A_{600} de 0,4 en medio mínimo M63 con glicerol 30 mM como única fuente de carbono, e inducida con IPTG 1 mM durante 2 h, hiperexpresando una proteína de aproximadamente 45 kDa, que se corresponde con la masa molecular predicha para el producto del gen *gcdH* (43,3 kDa) (Fig. 38).

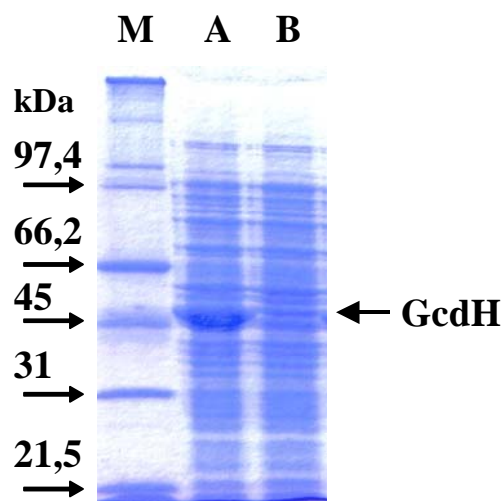


Figura 38. Hiperproducción de la proteína GcdH en extractos de *E. coli* DH5 α (pGCDH). Análisis mediante SDS-PAGE al 12,5% de extractos celulares de la cepa *E. coli* DH5 α (pGCDH) (calle A) y la cepa control *E. coli* DH5 α (pIZ1016) (calle B). La calle M corresponde a los marcadores de peso molecular "Broad Range" (BioRad). En el margen izquierdo de la figura se detallan los pesos moleculares (en kDa) de las proteínas marcadoras. Con una flecha se señala la posición de la banda de hiperproducción correspondiente a la proteína GcdH.

Los extractos celulares de la cepa *E. coli* DH5 α (pGCDH) presentaron actividad glutaril-CoA deshidrogenasa (6,5 μ moles/min x mg proteína), sin embargo esta actividad no se detectó en la cepa control *E. coli* DH5 α (pIZ1016). Estos resultados sugieren que el gen *gcdH* es capaz de expresarse en *E. coli* y codifica una proteína con actividad glutaril-CoA deshidrogenasa.

Con el fin de confirmar que el gen *gcdH* clonado en *E. coli* es el responsable de la actividad glutaril-CoA deshidrogenasa de *Azoarcus* sp. CIB, se transfirió el plásmido pGCDH por conjugación biparental desde la cepa donadora *E. coli* S17-1 λ pir (pGCDH) a *Azoarcus* sp. CIB*gcdH*, la cepa mutante que carece de actividad GcdH y, por tanto, es incapaz de catabolizar tolueno, benzoato, ciclohexano carboxilato, pimelato o glutarato (ver más arriba). La cepa complementada, *Azoarcus* sp. CIB*gcdH* (pGCDH), recuperó la capacidad de utilizar las fuentes de carbono anteriormente mencionadas (Fig. 39) y presentaba unos niveles de actividad glutaril-CoA deshidrogenasa (6,8 μ moles/min x mg proteína) similares a los observados en *E. coli* DH5 α (pGCDH) y aproximadamente 30 veces superiores a los detectados en células de *Azoarcus* sp. CIB (Fig. 34), lo que está de acuerdo con el hecho de que el gen *gcdH* se encuentra clonado en multicopia y bajo un promotor fuerte en el plásmido pGCDH.

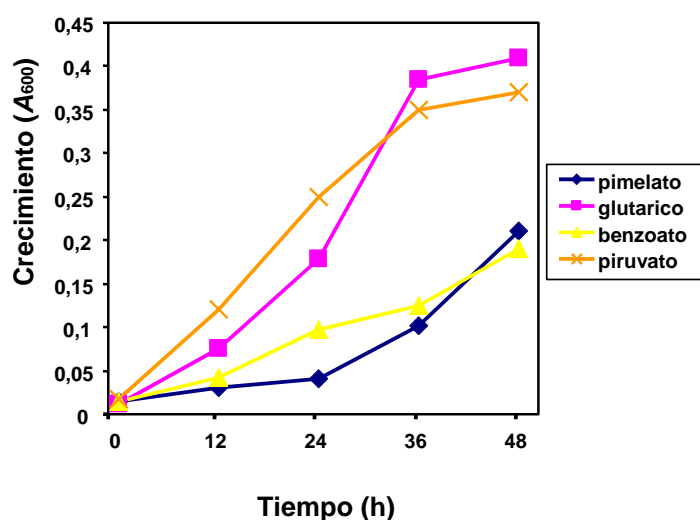


Figura 39. Crecimiento anaeróbico de *Azoarcus* sp. CIB*gcdH* (pGCDH) utilizando diferentes fuentes de carbono. Se valoró el incremento de absorbancia a 600 nm de *Azoarcus* sp. CIB*gcdH* (pGCDH) cultivado anaeróbicamente en medio mínimo MC con KNO₃ 10 mM como aceptor final de electrones, y con pimelato 5 mM o benzoato 3 mM como únicas fuentes de carbono.

Estos resultados confirman que el gen *gcdH* es el responsable de la actividad glutaril-CoA deshidrogenasa de *Azoarcus* sp. CIB, y es esencial para la degradación de compuestos aromáticos que son metabolizados por la ruta central del benzoil-CoA así como de ácidos dicarboxílicos y otros compuestos tales como el ciclohexano carboxilato que generan glutaril-CoA en su degradación.

2.4 Estudio de la regulación del gen *gcdH* de *Azoarcus* sp. CIB.

Dado que el gen *gcdH* de *Azoarcus* sp. CIB es el primero que se describe en bacterias y dado que no existe ningún dato descrito en la literatura acerca de la regulación del gen *gcdH* bacteriano, se decidió estudiar la expresión del gen *gcdH* de la cepa CIB.

2.4.1 Estudio de la expresión del gen *gcdH* en *Azoarcus* sp. CIB

Para estudiar la regulación del gen *gcdH*, se realizaron experimentos de RT-PCR a partir del ARN extraído de células de *Azoarcus* sp. CIB cultivadas anaeróbicamente en diferentes fuentes de carbono, incluyendo aquellas en las que se había observado actividad GcdH en extractos celulares. La figura 40 muestra que el gen *gcdH* se induce en células cultivadas en presencia de benzoato o tolueno, mientras que no existe inducción del gen *gcdH* en células cultivadas en succinato. Este resultado confirma que la inducción de la actividad glutaril-CoA deshidrogenasa observada en células que crecen anaeróbicamente en compuestos aromáticos es debida a una inducción de la transcripción del gen *gcdH*.

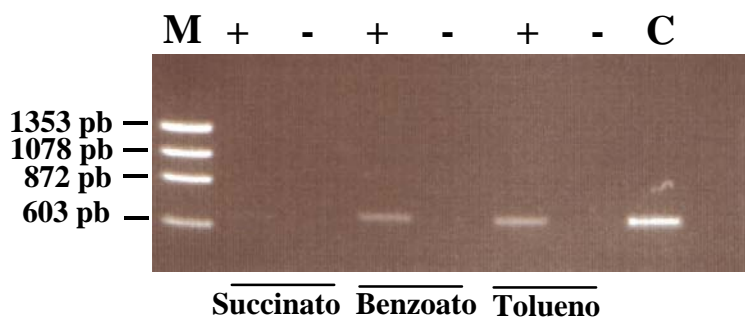


Figura 40. Expresión del gen *gcdH* de *Azoarcus* sp. CIB. Electroforesis en gel de agarosa de los productos de RT-PCR del gen *gcdH* obtenidos a partir del ARN total de *Azoarcus* sp. CIB cultivado en distintas fuentes de carbono. Los ARNs fueron extraídos de células de *Azoarcus* sp. CIB cultivadas en condiciones desnitrificantes hasta la mitad de la fase exponencial utilizando succinato 0,2%, benzoato 3 mM ó tolueno 250 mM (en fase de HMN) como fuentes de carbono. Las reacciones de RT-PCR se llevaron a cabo con los oligonucleótidos Gcd3' y Gcd5' (Tabla 4), tal y como se describe en el apartado 7 de Materiales y Métodos. Las reacciones de retrotranscripción se llevaron a cabo en presencia de transcriptasa reversa (+), o, en su ausencia como control negativo (-). Se indica la PCR llevada a cabo con ADN genómico como control positivo (C). La calle M corresponde al marcador de peso molecular (ADN del fago Φ X174 digerido con *Hae*III). Los números de la izquierda indican los tamaños de los marcadores.

Tal y como se ha indicado anteriormente (apartado 2.3) adyacente al gen *gcdH*, y en posición divergente (Fig. 37), se localiza el gen *gcdR* que codifica un posible

regulador transcripcional que podría participar en el control de la expresión del gen *gcdH*. Por lo tanto, para continuar con el estudio de la regulación del gen *gcdH*, se procedió a investigar el papel del gen *gcdR* en *Azoarcus* sp. CIB.

2.4.2 El gen *gcdR* codifica un activador transcripcional del gen *gcdH*.

El gen *gcdR* codifica una proteína de 303 aminoácidos cuya estructura primaria muestra similitud con la de reguladores transcripcionales de la familia LysR (Tabla 6). Así, GcdR presenta una identidad del 83% con la proteína ortóloga de la cepa *Azoarcus* sp. EbN1, y un 69-73% con proteínas ortólogas de distintas cepas de *P. aeruginosa*, *P. putida* y *P. fluorescens*.

Para determinar si realmente el producto del gen *gcdR* es el regulador transcripcional que controla al gen *gcdH*, se construyó mediante inserción disruptiva un mutante de *Azoarcus* sp. CIB en el gen *gcdR* (Fig. 41).

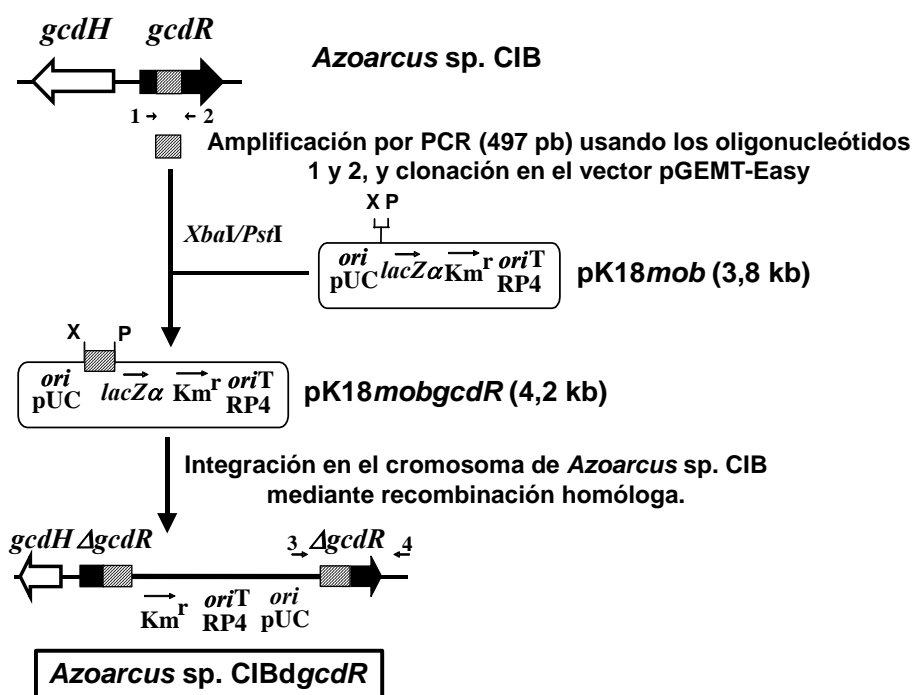


Figura 41. Esquema de la construcción de la cepa mutante *Azoarcus* sp. CIBΔ*gcdR*. En la parte superior de la figura se esquematizan los genes *gcdR* (negro) y *gcdH* (blanco). Las flechas con número 1 y 2 representan los oligonucleótidos GcdR5' (1) y LysRint3' (2) (Tabla 4) utilizados para la amplificación por PCR de un fragmento interno de 497 pb del gen *gcdR*, y los oligonucleótidos R24 (3) y gcdRHis3' (4) (Tabla 4) utilizados para comprobar por PCR la correcta inserción del gen *gcdR* en el cromosoma de la cepa mutante de *Azoarcus* sp. CIB. Se representan los orígenes de replicación (*ori* pUC) y de transferencia (*oriT* RP4) del vector suicida pK18*mob*. *Km^r* representa el gen que codifica la resistencia a kanamicina. *lacZα*, secuencia que codifica el fragmento α de la β-galactosidasa de *E. coli*. Las dianas de restricción utilizadas para la clonación se indican como: *XbaI* (X) y *PstI* (P). Δ, indica gen truncado.

La cepa mutante *Azoarcus* sp. CIBd*gcdR* fue incapaz de crecer anaeróbicamente en medio mínimo MC, empleando pimelato, glutarato o benzoato como únicas fuentes de carbono, si bien mostró un crecimiento similar al de la cepa silvestre en fuentes de carbono, tales como piruvato, que no requieren la actividad GcdH para ser metabolizadas (Fig. 42).

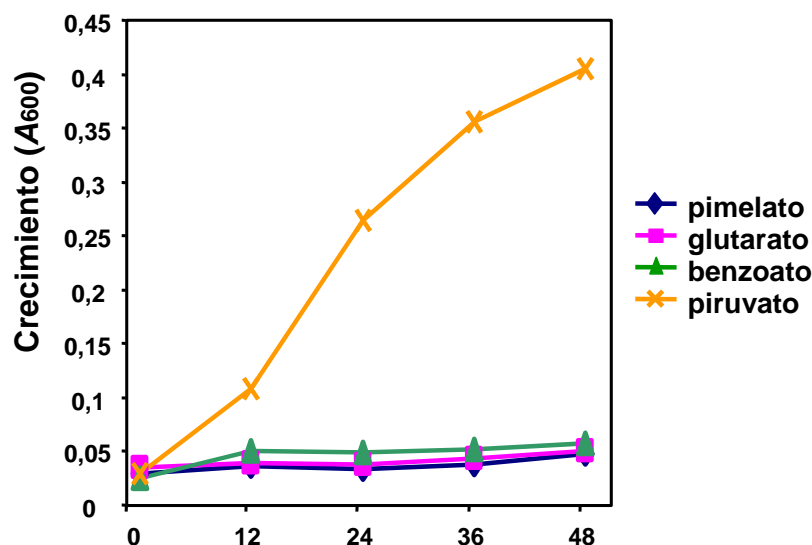


Figura 42. Crecimiento anaeróbico de *Azoarcus* sp. CIBd*gcdR* utilizando distintas fuentes de carbono. Se valoró el incremento de absorbancia a 600 nm de *Azoarcus* sp. CIBd*gcdR* cultivado anaeróbicamente en medio mínimo MC con KNO_3 10 mM como aceptor final de electrones, y pimelato (5 mM), glutarato (5 mM), benzoato (3 mM) o piruvato 0,2% como únicas fuentes de carbono.

El hecho de que la cepa mutante *Azoarcus* sp. CIBd*gcdR* fuese incapaz de crecer en fuentes de carbono que precisan de una actividad glutaril-CoA deshidrogenasa para su mineralización, tales como el glutarato, pimelato o benzoato, sugiere que el gen *gcdR* está involucrado en la expresión del gen *gcdH* y su función podría ser la de un activador transcripcional.

2.4.3 Caracterización del sistema regulador GcdR/Pg en *E. coli*.

Para confirmar que el gen *gcdR* es un activador transcripcional del gen *gcdH* y caracterizar el inductor del sistema, se decidió estudiar la interacción entre la proteína GcdR y el promotor *Pg*, que controla al gen *gcdH* en *E. coli*. Para realizar este estudio, se construyó una fusión traduccional del promotor *Pg* al gen reportero *lacZ* dando lugar al plásmido pBBRPg (Fig. 43).

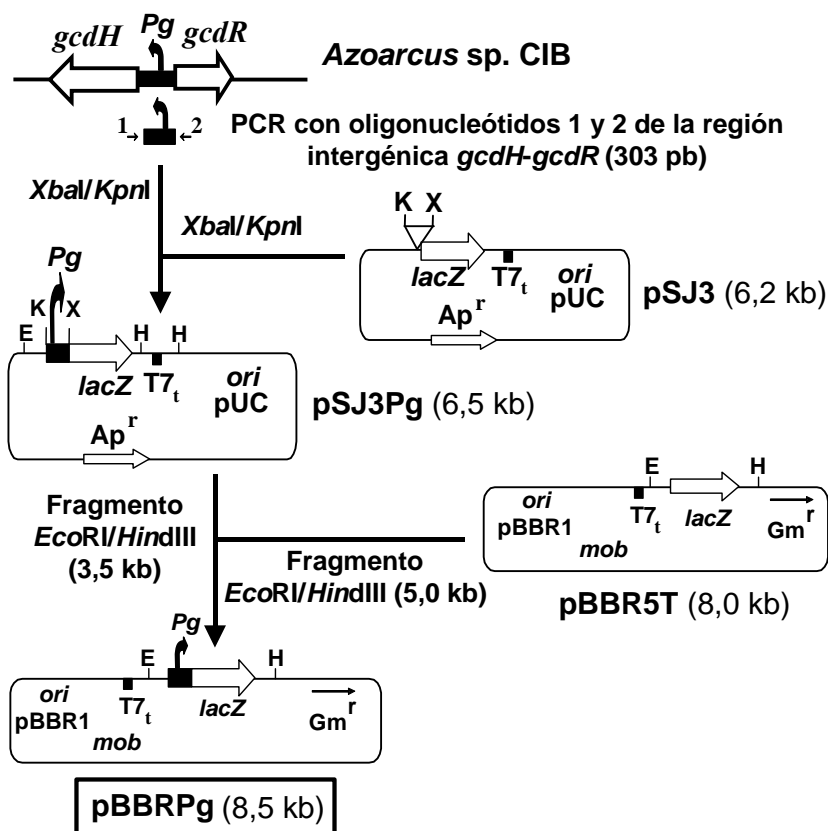


Figura 43. Esquema de la estrategia seguida para la construcción del plásmido pBBRPg. Las flechas con número 1 y 2 indican los oligonucleótidos Pgc3' (1) y Pgl5' (2) (Tabla 4) que se utilizaron para amplificar por PCR la región promotora *Pg* y el codón de inicio de la traducción del gen *gcdH* (303 pb). Este fragmento se digirió con las enzimas de restricción *XbaI/KpnI* y se clonó en el vector de búsqueda de promotores pSJ3 para generar el plásmido pSJ3Pg que contiene la fusión traduccional *Pg::lacZ*. El plásmido que contiene la fusión traduccional *Pg::lacZ* fue digerido con las enzimas de restricción *EcoRI/HindIII* y la fusión subclonada en el vector pBBR5T digerido con las mismas enzimas. Se obtuvo así el plásmido de amplio rango de huésped pBBRPg. Se representan las funciones de replicación (*ori pBBR1* y *ori pUC*) y de movilización (*mob*) de los plásmidos empleados. El gen *lacZ* se representa con una flecha blanca y se detalla también el terminador de la transcripción del fago T7 (*T7_t*). *Ap^r* y *Gm^r* representan los genes que codifican las resistencias a ampicilina y gentamicina, respectivamente. Las dianas de restricción utilizadas para la clonación se indican como: *EcoRI* (E), *HindIII* (H), *KpnI* (K) y *XbaI* (X).

La cepa *E. coli* MC4100 conteniendo el plásmido pBBRPg (*Pg::lacZ*) cultivada aeróbicamente en medio mínimo M63 con glicerol 30 mM como fuente de carbono, no presentaba actividad β -galactosidasa. Este resultado sugería que el promotor *Pg* era inactivo y que era necesaria la presencia de un factor regulador no presente en *E. coli*, probablemente la proteína GcdR, para su activación. Para confirmar esta hipótesis, se amplificó por PCR el gen *gcdR* y se clonó en el vector de bajo número de copias pCK01 originando el plásmido pGCDR (Fig. 44).

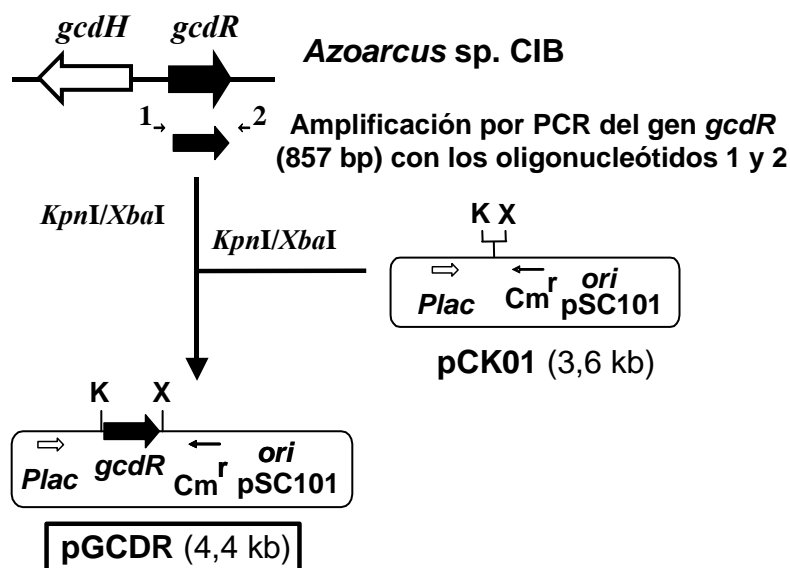


Figura 44. Esquema de la estrategia seguida para la construcción del plásmido pGCDR. Las flechas con número 1 y 2 representan los oligonucleótidos gLysR5' (1) y gLysR3' (2) (Tabla 4) utilizados para la amplificación por PCR del gen *gcdR*. Se representan el origen de replicación (*ori*pSC101) y el promotor *Plac* (flecha blanca) del vector pCK01 y sus derivados. *Cm^r* representa el gen que codifica la resistencia a cloranfenicol. Las dianas de restricción utilizadas para la clonación se indican como: *KpnI* (K) y *XbaI* (X).

La cepa *E. coli* MC4100 conteniendo los plásmidos compatibles pBBRPg y pGCDR tampoco presentó actividad β -galactosidasa, lo que sugería que para la activación del promotor *Pg* además de la proteína reguladora GcdR era necesaria la presencia de una molécula inductora o de algún otro factor no presente en *E. coli*. Para tratar de caracterizar el inductor del sistema, se realizaron experimentos de actividad β -galactosidasa en la cepa *E. coli* MC4100 (pBBRPg, pGCDR) cultivada en medio mínimo M63 con glicerol como fuente de carbono y con glicerol más distintos ácidos orgánicos (α -cetoglutarato, valerato, crotonato, fumarato, succinato, *cis,cis*-muconato, pimelato o adipato) o aminoácidos (L-glutamato o L-glutamina), como moléculas inductoras. Sólo se observó una significativa inducción de la fusión *Pg::lacZ* cuando las células se habían cultivado en presencia de glutarato o glutaconato, y esta inducción era estrictamente dependiente de la presencia del gen *gcdR* ya que no se producía en la cepa control *E. coli* (pBBRPg, pCK01) (Fig. 45). Este resultado indica que el gen *gcdR* codifica el activador transcripcional del promotor *Pg*. Además, el activador transcripcional GcdR muestra un perfil de inductores muy restringido, reconociendo solamente ácidos dicarboxílicos de 5 carbonos no sustituidos saturados (glutarato) o insaturados (glutaconato) como inductores.

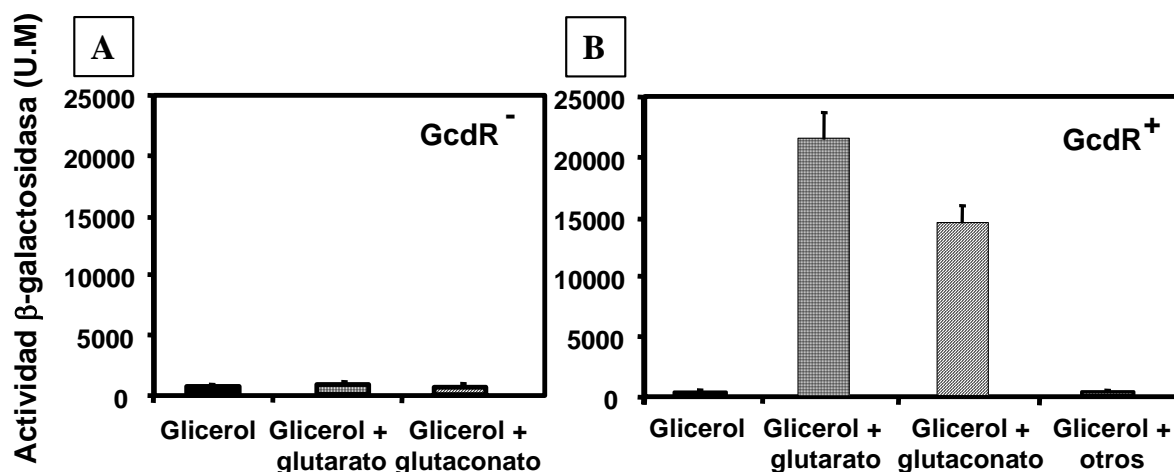


Figura 45. Expresión de la fusión traduccional *Pg::lacZ* en *E. coli* MC4100. Células de *E. coli* MC4100 (pBBRPg, pCK01) (panel A) o *E. coli* MC4100 (pBBRPg, pGCDR) (panel B), se cultivaron en medio mínimo M63 con glicerol 30mM como fuente de carbono y glutarato 1 mM, glutaconato 1mM, y otros compuestos 1mM (valerato, pimelato, L-glutamato, L-glutamina, succinato, α -cetoglutarato, crotonato, fumarato, adipato y *cis,cis*-muconato), como posibles moléculas inductoras. Los ensayos de actividad β -galactosidasa en Unidades Miller (U.M.) se realizaron con células en fase exponencial de crecimiento según se describe en el apartado 8.1 de Materiales y Métodos. Las barras muestran la desviación estándar de tres experimentos independientes.

2.4.4 Determinación del sitio de inicio de la transcripción en el promotor *Pg*.

Para caracterizar mejor el promotor *Pg*, se determinó el sitio de inicio de la transcripción del gen *gcdH* mediante experimentos de extensión por cebador (*primer extension*) a partir del ARN total extraído de células de *E. coli* MC4100 (pSJ3Pg, pGCDR) cultivadas en medio mínimo con glicerol como fuente de carbono y glutarato como inductor (Fig. 46). El sitio de inicio de la transcripción (+1) es una adenina localizada a 59 nucleótidos del codón de iniciación de la traducción (ATG) del gen *gcdH* (Fig. 47).

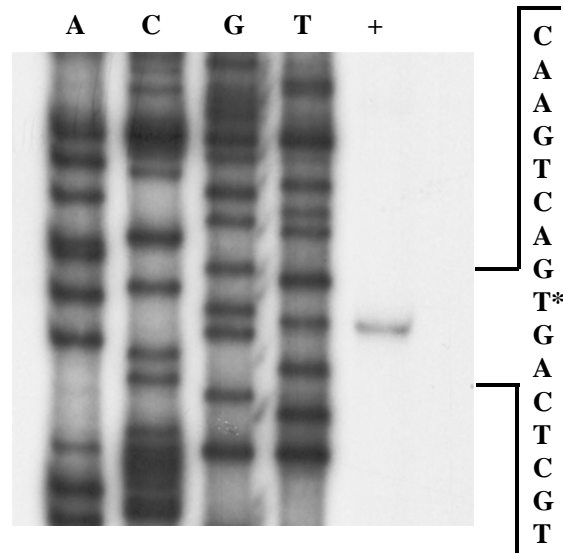


Figura 46. Determinación del sitio de inicio de la transcripción en el promotor *Pg*. A partir del ARN total aislado de células de *E. coli* MC4100 (pSJ3Pg, pGCDR) cultivadas en medio mínimo M63 con glicerol 30 mM como fuente de carbono y glutarato 1mM como inductor (calle +), se realizó la técnica de *primer extension* según se detalla en el apartado 8 de Materiales y Métodos utilizando el oligonucleótido Lac57 (Tabla 4) como cebador. El tamaño del producto de extensión se determinó por comparación con el patrón de secuenciación del promotor *Pg*, utilizando el oligonucleótido Lac57 (calles A, C, G y T). En el lateral derecho de la figura se muestra la secuencia de nucleótidos que flanquea el sitio +1, indicado por un asterisco en la hebra no codificante.

La identificación del sitio +1 del promotor *Pg* permitió definir las posibles cajas -10 y -35 de interacción con la subunidad σ^{70} de la ARNP. También se identificaron las posibles cajas de reconocimiento (RBS) y activación (ABS) que definen la zona de interacción del regulador GcdR con el promotor *Pg*. La caja RBS presenta la secuencia consenso T-N₁₁-A que caracteriza a los promotores regulados por activadores de la familia LysR (Fig. 47).

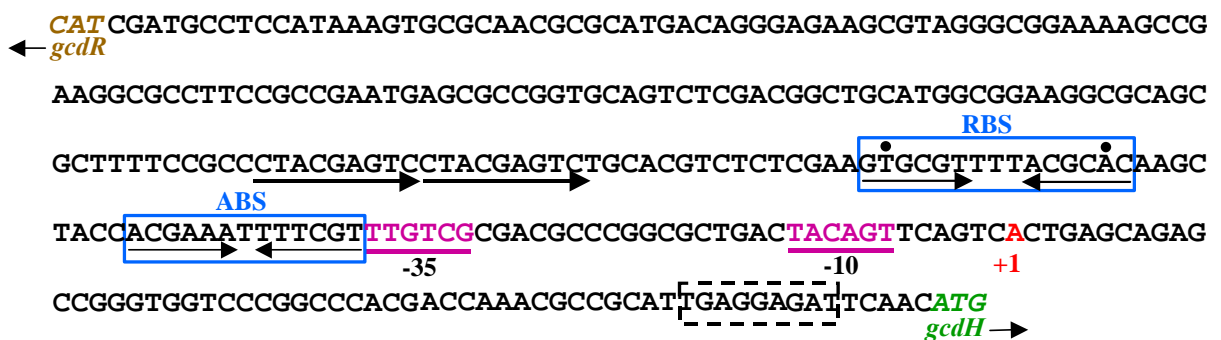


Figura 47. Secuencia de la región promotora *Pg*. Se muestra la región intergénica *gcdR-gcdH*, señalando en *itálicas* los codones de inicio (ATG) de los genes *gcdR* (marrón) y *gcdH* (verde). La hipotética caja de unión del ribosoma en el gen *gcdH* se recuadra con líneas discontinuas. Se indica el sitio de inicio de la transcripción (+1) en color rojo, y las hipotéticas cajas -10 y -35 de interacción con la subunidad σ^{70} de la ARNP subrayadas en morado. Las posibles regiones RBS y ABS de interacción de GcdR con el promotor *Pg* se recuadran con línea continua y color azul. Los nucleótidos conservados T-N₁₁-A del RBS se señalan con un punto. Las repeticiones directas e invertidas detectadas se señalan con flechas debajo de la secuencia correspondiente.

2.5. Clonación y regulación de otros genes *gcdH* bacterianos: el gen *gcdH* de *P. putida*.

Numerosos genes *gcdH* bacterianos han sido identificados y anotados como tal en las bases de datos, si bien ninguno de ellos ha sido confirmado experimentalmente. Para demostrar que estos genes realmente codifican glutaril-CoA deshidrogenasas, se seleccionó para su estudio el gen *gcdH* de *P. putida* KT2440, cuyo producto presenta una identidad del 85% con la proteína GcdH de *Azoarcus* sp. CIB. Para ello, se amplificó dicho gen (*gcdH_{KT}*) y su correspondiente promotor (*P_{gKT}*) por PCR utilizando el genoma de *P. putida* KT2440 como ADN molde, y se clonó en el vector de amplio rango de huésped pBBR1MCS-5, dando como resultado el plásmido pGCDHKT (Fig. 48).

La cepa *Azoarcus* sp. CIBd*gcdH* conteniendo el plásmido pGCDHKT fue capaz de recuperar la capacidad de utilizar anaeróbicamente pimelato y benzoato como únicas fuentes de carbono (Fig. 49).

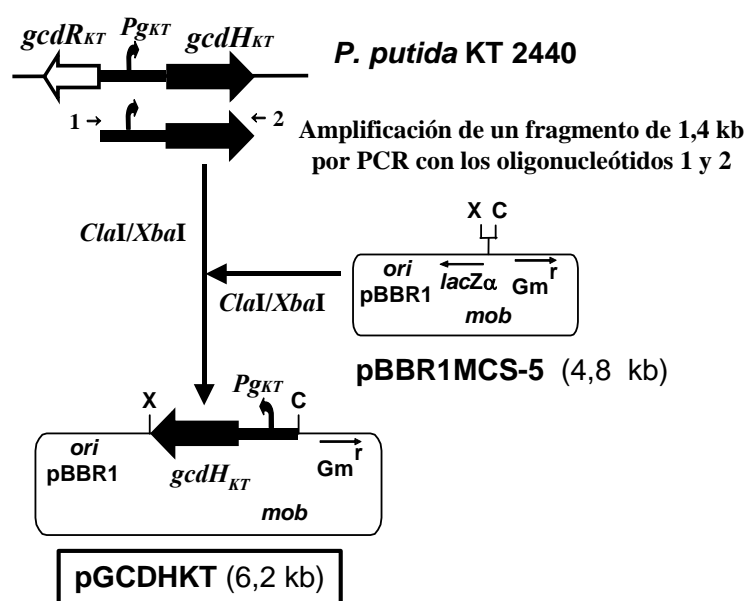


Figura 48. Esquema de la estrategia seguida para la construcción del plásmido pGCDHKT. En la parte superior de la figura se esquematizan los genes *gcdR_{KT}* (blanco) y *gcdH_{KT}* (negro). Las flechas con número 1 y 2 representan los oligonucleótidos Ppgcd5' (1) y Ppgcd3' (2) (Tabla 4) utilizados para la amplificación por PCR del gen *gcdH_{KT}* con su correspondiente promotor *P_{gKT}*. Se representa el origen de replicación (*ori_{pBBR1}*) y la región de movilización (*mob*) de los plásmidos, así como el fragmento génico que codifica la subunidad α de la β -galactosidasa (*lacZ α*). *Gm^r* representa el gen que codifica la resistencia a gentamicina. Las dianas de restricción utilizadas para la clonación se indican como: *ClaI* (C) y *XbaI* (X).

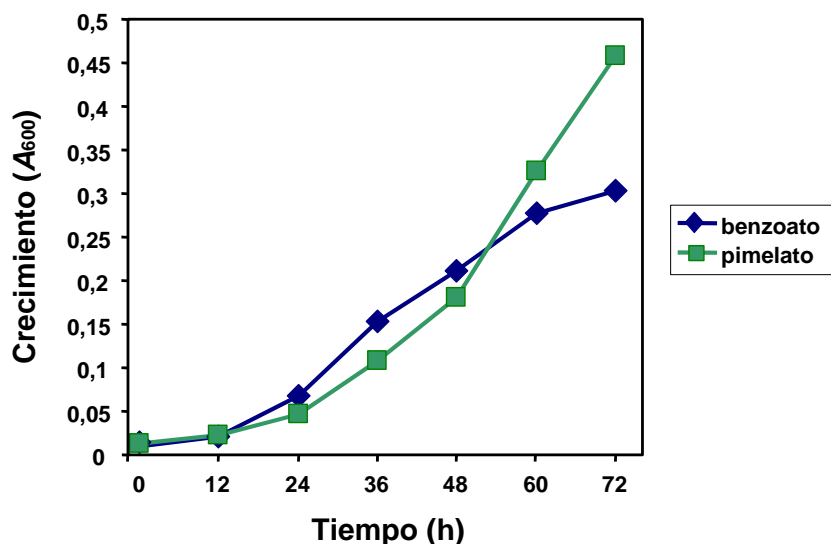


Figura 49. Crecimiento anaeróbico de *Azoarcus* sp. CIBdgdH (pGCDHKT) utilizando benzoato o pimelato como únicas fuentes de carbono. Se valoró el incremento de absorbancia a 600 nm de *Azoarcus* sp. CIBdgdH (pGCDHKT) cultivado anaeróbicamente en medio mínimo MC con KNO₃ 10 mM como aceptor final de electrones, y con benzoato (3 mM) o pimelato (5 mM) como únicas fuentes de carbono.

Además, los valores de actividad glutaril-CoA deshidrogenasa de la cepa recombinante eran similares a los de la cepa silvestre *Azoarcus* sp. CIB, y esta actividad era inducible en células cultivadas en presencia de benzoato (Fig. 50). Estos resultados demuestran que el gen *gcdH* de *P. putida* puede reemplazar eficazmente al gen *gcdH* de *Azoarcus* sp. CIB y, por tanto, codifica la actividad glutaril-CoA deshidrogenasa predicha. Además, la inducción del gen *gcdH_{KT}* en presencia de benzoato sugiere que el gen *gcdR* de la cepa CIB controla de forma eficaz la actividad del promotor *Pg_{KT}*.

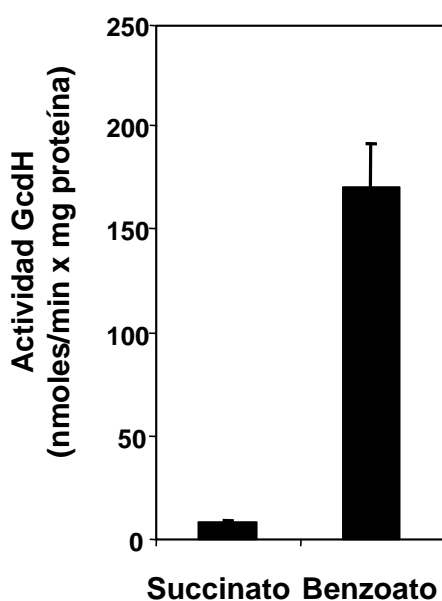


Figura 50. Actividad glutaril-CoA deshidrogenasa en *Azoarcus* sp. CIBdgdH (pGCDHKT). Las células de *Azoarcus* sp. CIBdgdH (pGCDHKT) fueron cultivadas anaeróbicamente hasta la mitad de la fase exponencial de crecimiento en medio mínimo MC conteniendo succinato 0,2% o benzoato 3 mM, como únicas fuentes de carbono. La actividad GcdH se midió en extractos celulares tal y como se describe en el apartado 13.2 de Materiales y Métodos. Las barras indican la desviación estándar de tres experimentos independientes.

Al igual que ocurre en *Azoarcus* sp. CIB, el gen *gcdH* de *P. putida* KT2440 (*gcdH_{KT}*) se localiza adyacente y divergente a un gen cuyo producto muestra similitud significativa con GcdR de la cepa CIB (69% identidad) y con otras proteínas reguladoras de la familia LysR y que, por ello, se denominó *gcdR_{KT}* (Fig. 48). Para demostrar si el gen *gcdH_{KT}* está bajo el control de la proteína reguladora GcdR de *Azoarcus* sp. CIB, se transfirió el plásmido pGCDHKT a la cepa *Azoarcus* sp. CIB*gcdR* mediante conjugación biparental. La cepa *Azoarcus* sp. CIB*gcdR* (pGCDHKT) presenta un fenotipo similar al de la cepa *Azoarcus* sp. CIB*gcdH*, siendo incapaz de crecer en compuestos que generan glutaril-CoA. Por lo tanto, este experimento demuestra que existe una regulación cruzada entre el regulador GcdR de *Azoarcus* sp. CIB y el promotor del gen *gcdH* de *P. putida* KT2440, y sugiere que la regulación del gen *gcdH* se encuentra conservada en varias proteobacterias.

V. ANEXO I

Nuno →

GATCAACTGTTTCGACGTGCCATACATCCGACGGTAAATGGCCCTTAGCCCTTTTCTCAAT 60
 CTAGTTGACAAGCTGCACGGGTATGTAGGCTGCCATTTACCGGAATCGGGAAGAGGATTA

ATCATCTCCGGTACCGAAAAATCGATGGACACAAGAATAAGCAATAGGACATGTGCGG 120
 TAGTAGAGCCATGGCCCTTTTACGTACCTGTGTTCTTATTCGTTATCTGTACGACGCC

CTTAGTCAACCCCTAGGAGTGTTCATGTTGATGAATTTACCAGCTTCGCGTCGGCGAG 180
 GAATCAGTGGGGATCCTCAGAACCTACAACACTTAAAGTGTGCGAAGCGCAGCCGCTC

CGGACCCGCGTTCGCCATGTACTACTCGATCGGATCAACAATGTACTTTTCTGCTTCCTC 240
 GCCTGGCGCAAGCGGTACATGAGTAGTGCCTAGTGTGTACATGAAAAGACGGAGG

AAGGAGCATCAGCATGATTGCGTCTTTGGTCTTGAATAAGAAATAGATCGCACCCCTTGGT 300
 TTCTCTGTAGTCTACTAACCGCAGAAACGAACTTTATCTTTATCTAGCGTGGGAACCA

CAATCTGCTCTTTCGCCGATCATGTGATCGTGTTCGCGACGTAACCATTTTTCGACGAA 360
 GTTAGGACGAGAACCGCGCTAGTACAGCTAGCAACAGCGTGCATTTGTAATAAATCTGCTT

AAGATCGAGCGCACAGCCAGTATATCTCGTACTGGCCTTCTCTGCTCTCTCTTCTT 420
 TTCTAGCTTCGCGTTCGGTTCATATAAGACCATGACCGGAAGGAGACGAGAAGGAAGAA

← ctoI3'

GTTACGCTTTTCTCATGATGTCTGGAGATGTTCCAATGGATTGCTCTCGGCACCTTCAGA 480
 CAAGTCGAAAGAGTACTACAGACCTCTACAAGTGTACTCAACGAGAGCCGCTGAAGTCT

GAAGTTGGCAATTTAAGTGAAGACGCGCATTCTGAACGTTTCGCCAGCCAGTGCACCT 540
 CTCCAACCGTTAAATCTACTCTTGGCGGTAAGACTTGAAGCGGTCCGTGACGTGGA

CGGGTTGTCGAAAGCGGTACGCGGCCCGCTCCGCGGATTGGCAACAACCATTTCC 600
 GCCCAACAGGCTCTTTCCGCAATGCGCGGTCAGAGGCGCTCAACCGGTTGTTGTAAGG

ATCAGCACCTTCGCCCGCTTGTCCAGTTTCTGCTACCCAGCGCCATCCCCGAAAATTC 660
 TAGTCGTGGAAGCGGCGAACAGGTCAAAAGAGCAGTGGGTCCGCGTAGGGGCTTTAAGG

TCCGCCACCCAGGCTTCGTCGCCAATCGTACCCAACTGTAAGAGTAGACGCAATTT 720
 AGCGCGTGGTCCGGAAGCAGCGGTTAGCCAGTGGGTTGACATTTCTCATCTCGCTAAAA

ACGCTGAGTTTACAATCGCTTGTGTAAGTGAAGGTAATGGATATTTTCGCTCGCG 780
 TCCGCTCAAAATGTTAGCAGAACCAATCTCACTAACTTCCAGTACATAAAGCAGCGC

CGGTTGGTGGTGAAGAGTTGTGTATAATAATAATAATAAAGAGATTTTCGCTGATGTAC 840
 CGCAACCACCACTTCTCAACACATATTTATATATTAATGTTCTCATAAAGCACTACATG

PstI → tdiRS3' → tdiRT3' →

GGAATGATGGATCCTGACTTCGAGCATCACCGCGCGGGCGGGATGCCACTTACACT 900
 CCTTAACCTATAGGACTGAAGCTCGTAGTGGCGCGCCCGCCCTACGGGTGAATGTGA
 * R A P R S A W K C

CGAGCAGTGGCTTTTCGTGAGCATCGACAGTTGAATCAGCTCGCCAAACGAGCGGACGC 960
 GCTCCGTCACCGAAAAGCACTCGTAGCTGTCAACTAGTTCGAGCGGGTGTCTCGCTGCG
 E L C H S K T L M S L Q I L E G L S R V

tdiRint5' →

CCAGCTTGTCTATCAGCGCCCCCTATGCGCTCCACGGTTTGGGGCTGATCCCGAGAA 1020
 GGTCGAACGAGTAGTCGGCGGGGATACGCGGAGGTGCCAAAACCCGACTAGGGCTCTT
 G L K S M L R G R H A E V T K P S I G L

GTCGGCGGATTTCTTGTGCTATTTCCATCGACGACGAGATCGAAGATTTCTTTTCCC 1080
 CAGCCCGCTAAAGGAACGAGATAAAGGTAGCTGTGCTAGCTTTAAAGGAAAAGGG

← tdiRext5'

L R A I E K S S N G D V V L D F I E K E

TCAGTGAAGGCTTTCCAGCAATCACGATATTTTCGCGCTCACTTTGCTTGGCGATCG 1140
 AGTCACTCTCCGAAAGTCCGTTAGTGTCTAATAACAGCGCGAGTGAACGAACCGCTGAC
 R L S L S E L C D R Y K D R E S Q K G I

CCTCCTTACTCTGAGTACGCGCAGCATTGACCGCATCAAGCAGATCTGATTGCGATAGG 1200
 GGAGGAATGAGACTCAGTCGCGTCTGTAAGTGGCGTAGTTCGCTAGGACTAACCGTATCC
 A E K S Q T L A A N V A D L L D Q N C R Y

GCTTCTGCAAGAACTATTGCGCCCTTCTGATCGCTGCGGCCCATCTGCGCATCCC 1260
 CGAAGACTTCTTTAGATAACCGGGAAGAAGTACGCGGACGCGGGTAGACGCTAGGG
 P K Q L F D I A G K K I A Q A G M Q A D

CGTATGCCGAGAAAAATGATCGGCAAAATTGAAGCGCGCTGGCACAAGATCTCTTGAA 1320
 GCATACGGCTCTCTTTTACTAGCCGTTAACTTCGGCGGACCGTGTCTAGAGAAGCTT
 G Y A S L F I I P L N F G R Q C L I E Q

GCTGACGCCGCTAAGCGCTGATCTTCACTGTAAGGATCAGGCAACCGTAGGAACCTG 1380
 CGACGTCGGCGATTCGCGGACGTAAGAGTGCAGTTCCTAGTCCGTTGGCATCTTTGGAC

← tdiR131.5'

L Q L G S L A Q M R V D L I L C G Y S G

AAATATCCGCGTCGAGAAACGCTTCGCGGAGTGTACTACTTTCAGCTTACGCGATCG 1440
 TTTATAGCGCAGCTCTTTGCGAACCCGCTCAGCATGTAGAAGTGCAGATCGCGCTAGC

tdiR131.5' XbaI

S I D A D L F A Q P S D Y M K V D L A I

AAACAAGCAACAGCTCAACGAATGGCGGACCGACTTTCATCATCGACGACGAACGCG 1500
 TTGGTTGCTGGTTCGAGTGTCTTACCCTGCTGAGTGTGAGTGTGCTGCTGCTGCTGCGC
 S G L L W S L S H R V S E E D D V V F V

TCGGCGAAAGGTTTCGACCTGCATTGTCTCTCTCTTTGTTTCATGGCCTTCGCCGCCCGA 1560
 AGCCCGTTCCAAAAGCTGGACGTAACACAGAGGAGGAAACAAGTACCGGAAGCGCGGGCT
 T P S L N E V Q M ← tdiR

GACTCTTCGGTCTCATGTGCGCGGGACCTCTCTACATCGCGCAAGGTGAATGAAACC 1620
 CTGAGAAGCCAGAAGTACAGCCGCGCTGGAGAGGATGTAGCCGTTCCACTTAACTTTGG
 * M P L T F Q F

GTGATCCCCCCTGCGGATCCTCGAAAACGAAAGCTGCCCGCGTGCAGCTCGACGATCG 1680
 CACTAGGGGGGGACGGCTAGGAGCTTTTGTCTTCGACGGCGGCACGCTGAGCTGCTAGC

← tdiS.F3'

R S G G R G I R S F S L Q G G H S E V I

AGCGGCAGATCGTCAAGCCAATGCCATCCCATCGGGCTTCGTGGAAAAGTATGGTTCAA 1740
 TCGCCGTCTAGCAGTTTCGGTTACGGGTAGGGTAGCCGGAAGCACCTTTTCATACCAAGTT
 S R C I T L G I G M G D P K T S F Y P E

ACACAGCTTCGCGGACCGCTTCTCCAGGCCGCTGCCGTGGTCTGAGCAAAAACCTTCA 1800
 TGTGTGCGAAGCGCTTCGCGAAGAGTCCGGCAGCGGACCGACAGGACTCTTTTGAAGT
 F V R K A S R K E L G T G H D Q V F V K

TCGCCCCACCTCGCGACTATCTCGTTCGATGACGAGCTTGCAGCGCTCGGATGGCA 1860
 AGCGGGTGGGAGCGGCTGATAGAGCGAGCTACTGCTCGAACGCGCGGAGCTACCGT
 M A G G E G V I E S G I V L K R A E S F

GGCCGTCATGCGCTCCATGCCGTTCTCATCAGGTTGAACAAGACTTGTGGATCTCTA 1920
 CCGGCGATGAGCGGAGTACGGCAAGAAGTAGTCCAACCTTGTCTGTAACGACCTAGAGAT
 L G T M A E M G N K M L N F L V Q Q I E

CCTTGACACGCGCACCGCGGATCCTCCAGTAATCGGAGTTCGACACTGACATCAT 1980
 GGAACGCTGCGCGTGGTTCGCCACCTAGGAGTCAATAGCCCTCAGCTGTGACTGTAGTA
 V K C L P V L P P D E L L R L D V S V D

GACGCGGACTCGAAATGAGGAACGAGAGCGCATCTGATCACTCGTGTGATGTTCT 2040
 CTGCTGCTGGAGCTTTAAGTCTTGTCTCTCGADAGAACTAGTGGAGCACTACAAGA
 H R R V E F N L F S L C D K I V E N I N

CGGGCTCTTGTCCGGTTTATGCTTGCAGACGAACTCTTTCAGTGGCGCAGTATCTCGC 2100
 GCCCGAGGAACGCGCAAAATACGACGCTGTGTTGAGGAAGTGCACCGCTCATAGAGCG
 E P E K D P K H K C V F E K V H R L I E

CCGCTTGTGAGATACTTTCAGCGGAACCCCATGGACTCCGAAATCTCCCCACGCGCTT 2160
 GCGAAGCAGCTCTATGAAGTCCGCTTGGGGTACTGAGGCTTTAGAGGGGTGCGGAA
 G A Q E L Y K L A F G M S E S I E G V G

GCACCCGCTCGAGACGAGCATGACGCGCTTCAAGTAATGACCGCGACACCAAGCGGCT 2220
 GCTGGCGGAGCTTGCCTCGTACGTCGCGCAAGTCCATTAAGTGGCGGCTGTGTCGCGA
 E V R E L R L M C G N L Y N V A S V L P

GCCCATCTGTTGCGCGAGGGCGCGCCATTTCCGCGCATCGATGATCCGCGGAGCC 2280
 CGGGGTAGACACGCGCTCCCGCGCGGTAAGAGCGGTCAACTAAGCCGCGCTCGG
 Q G M G Q H A L A A A M E A M S N I R G L

GCGTGAAGTCCGCGTGGCGAGCGGCTTCGAGCTCTTCGATGCGCTTGGCGTAGGTAATG 2340
 CGCACTCGAGCGCACCGCGTGGCAAGCTCGAGAGTACGCGAACGCCATCCATTACA
 R T L E A H R L R E L E E I R K R Y T I

CGTGAAGGCGCGACGACGAACTGCGCGCTTGCCTCGATGACGACGAGCTGACGC 2400
 GCCACTTCCCGCTGCTGCTTTCGCGCGGAGGAAAGTCAAGTCAAGTCTGCTGCTGCTGCG
 D T F A A V V F Q R D K R E I L S S S V

TCAGCCAGCGAGCTTTCGCGGCTTCTGTCGCGCATCCGATTTTCATGTTGTTGATCG 2460
 AGTCCGTCGCGTGAACCGCGGAGAGCAGCGTAGGCTAAGTCAACAACACTAGC
 S L W R L K R A K E D R M G I E I N N I

tdiS3'

AATCCTTCTGAAAATCGCGGATCTGCCACGCGAGGCGGTGCCCGGGATCTCCGTTCCGT 2520
 TTAGGAAGACTTTTAGCGCCTAGACGCTGCGCTCCGCCAAGCGCGCCTAGAGGCAAGGCA
 S D K Q F D R I Q W P L R H R P I E T G

← tdiS3'

CGTGCAGGAAGCGCAGCGGAGTTTCGCGCAGCTTCGCGGAGCGCGCGCAAGC 2580
 GCACGCTCTTCTTTCGCGTCCGCTCAAAGCGGTCGAGAGCCCGCTCCGCGCGGCTTCG
 D H L F F R L P L K R W S E P S A R R L

CGCCATCTCCGCGAATTGGCAATTGGCCCTCGAGGATGTCGCCCTGCGGATCGGTGATCG 2640
 GCGCGTAGAGCGCTTAAACCGTTAACCGGAGCTCCTACAGCGGGACGCTAGGCCACTAGC
 G A M E A F Q C N A E L I D G Q P D T I

AAATCCGATCGCGAGCATGTGGAAGGACTTCTGGTATTTCTCGGATTTTCTGTTTCGCG 2700
 TTTAGGCTAGCGCTGACACCTTCTCGAAGACCATAAAGAGCGCTAAAAGCAAGCGCC

← tdiS566.5'

S I G I P L M H F L K Q Y K E R N E N A

CCTCTCGCGCGCTGGCGGCTGCGACTTCTCTCCAGGCTCGCAACGTCGCGGCTAA 2760
 GGAGGAGCCGCGACCGCGCAGCGCTGAAGAGGAGTCCGAGCGTTGACGCGCGGCTATT
 A E E A R Q R T A V E E E L S A V H R S

GCTCCACCGTCCGCTGTCACGACGATTTTCGAGCTTCTCGAGCTCGGCGCTCAATTCAT 2820
 CGAGTGGCAGGCGACGACGCTGCTGCTAAAGCTCGAAGAGCTGACGCGGAAAGTTAAGTA
 L E V T R Q Q V V I E L K E V D A K L E

2T7.2
 CCTCGATCCGCTTGTGCTCGGAGATGTCCTGCCACCCAGATACCTCCGAGTTTCAGGCT 2880
 GGAGCTAGGCGAACACGAGCCTCTACAGGACGCTGGTGGTCTATGAAGGCTCAACGCTCA
 D E I R K H E S I D Q V V W I S G L Q L
 2T7.2
 TGCGCGGTCGAGCAGCCTGCCAGTGCAGATGCACAGATCAAGGTGCCATCCTCCGAA 2940
 ACGCCACAGCTCGTCGACGGTCACTGCTACGTGGTCTAGTTCCAGGTTAGGAAGGCTT
 N R P D L L T G T V I C W I L T G D K R
 CGAACGGACGCTCGTCTTCAAAGTGTCTGCTCGCGCTCGAAAAATGGCCGACTTCCGGC 3000
 GCTTGCCTGCGAGCAGGAACCTTTCACAGGACGCGAGCTTTTACCGGCATGAAGGCCG
 V F P R E D K F T H E R E F F H G Y K R
 CGATTGTGCAAAACGATTTCGTCGCTCGGGTAGTGCACCCGACCGATTGGTTGTCGAGTT 3060
 GCTAACAGCTTTTGTAAAGCAGGAGCCATCACGTGGGCTTACCAACAGCAGCTCAA
 G I T S F S E D S P Y H V R V S Q N D L
 CTCGGGCTCGTATCCGAACAATTCTCGTATCGTCGGTTACAGCGCTTGTATCGTCGGT 3120
 GAGGCCGAGCATAGGCTTTGTAAGGAGCATAGCAGCCAAATGTCGCGAATAGCAGCCCA
 ← tdiS5'
 E G P E Y G F L E E Y R R N C R K I T R
 TTCGAGATGGCAGATGGCCACGGGCGTGTTCGCGAAGATCGCCTCGTAGCCGACCGAGA 3180
 AAGCCTCTACCCTTACCCTGACCGGTCGCCGACACAGCCTTCTAGCGGAGCATCGGCTGGCTCT
 tdiS5'
 N R L H C I A V P T N R F I A E Y G V S
 SpeI ptdiS.5'
 CGTCCTCGCCCTGCTCATCGTCAATCCACGCGGACCACTGCTTCTTATCGAGCACT 3240
 CGAGGACGGGACGAGTAGCAGCTTAAGGTGGCGCTGTCAGTAAGAAGTACCTCGTGA
 ← tdiRS5' XbaI
 V D E G Q E D D F E V A V V D N K M ← tdiS
 SfiI 2T7.3
 ACCTCCTCGAACCCGCAACCCGGGAACCGGTCTATCTAACCGCGCGCTGGCCCCATTGG 3300
 TGAGGAGCGTTGGCCCTTGGGCCCTTGGCCAGATAGATTGCGCGCGCACCGGGTAACCC
 2T7.3
 CCGCCGATCCGACCCCGTCAACCTGAGCAGTCTCCGATTGTGCGCCGTGACGACGCGGC 3360
 GCGCGTAGGCTGGGCGAGTGGGACTCGTGAAGGCTAACACGCGGCGACTGCGTCCGCC
 Afl III
 CGCACTAGGACAACCCCTCTTTGCTTAAGTGTTCGCAACCAATTGGTCAAGTTCGCGGAA 3420
 GCGTATCCCTGTTGGGGAGAAACGAATTCACAAGCGTGGTTAACAGTGAAGCGCGCTT
 3T7.2
 TGGTCTTAAATCAATTCAACCGGGCCAGACTCGCGCTGCCAACCTGACTCCTCCGGTG 3480
 ACCACGAATTTAAGTTAAGTGCCTCCGCTGAGCGCGAGCGCTTGGACTGAGGAGGCCAC
 3T7.2
 CATGATACCTCGTCTGCTTCCAAAATGTTTGGTTTGCATGAACCGCAGAGGCTGCTGCTG 3540
 GTACTATGACGAGCAAGGTTTAAACAACCAAGCTTACTTGGCGTCTCCGACAGCAC
 ← pbssD3'
 bssD → M
 AGAATTCCTCTTGTACCAGAAATACAGAGGTTTCAGCCTGCAGGACGGGCCCGGATCAGG 3600
 TCTTAAGGAGAACAAATGGCTTTATGCTCCAAGTCGACGCTCTGCCGGCCCTAGTCC
 pbssD3' BamHI
 R I P L V T E I Q R F S L Q D G P G I R
 ACGACCATCTTCTGAAGGGTGCCTGCTAGTGTGCGGTTGCCACAACCCCGAAACC 3660
 TGCTGGTAGAAGGACTTCCGACGGCGGATGCAACGGGACACCGGTGTTGGGGCTTTGG
 T T I F L K G C P L R C P W C H N P E T
 CAGGACGTGAGACAGGATTTACTATTACCGGGCGGCTGCGTCCGCTGCGGCCGCTGCG 3720
 GTCTGCACTCTGCTCAAGATGATAATGGCGCGGACGCGAGCGGACCGCGGACG
 Q D V R Q E F Y Y P A R C V G C G R C
 ATGGCCGCTGCTCCCGCGGGAAACATCCCGCTGGTGCACAACCTCGGACGGCCGACGATC 3780
 TACCGGACAGCGGGGCGCCCTTGTAGGCGGACACGCTGTGAGCCTGCCGGCTGCTAG
 M A V C P A G T S R L V H N S D G R T I
 PstI bssD5'
 GTCGAACCTGACCGCAGGACTGTCAGCGTTGCATGCGTTGCGTCCGCGCTGCCTGACG 3840
 CAGCTTGAAGCTGGCGTGCCTGACAGTGCACAGCTGCAACGTCGCAACGCGAGCGGACGACTGC
 V E L D R T D C Q R C M R C V A A C L T
 bssD5'
 GACCGCGCAACCCAGTCCGGCAACGCATGAGCGTCGAGGAAATCCTGCGGAAAGCGCTG 3900
 CTCGCGCGTGGTGGCCAGCCGCTTGGTACTGCGAGCTCCTTTAGGACGCGCTTCCGAGC
 D A R T T V G Q R M S V E E I L R E A L
 TCCGACTCCCCTTTTACCAGAACGCGCGGCGAGTCCACATCAGCGTGGCGACCCCT 3960
 AGGCTGAGGGGAAAATGGCGTGTGCGCCGCGCTCAGTGTAGTGCACCACCGCTGGGA
 S D S P F Y R N S G C C G C V T I S G G T D P
 CTCTTCCATCCTTCGTTACATTTGAGCTGGCGCGAGAATCAAAGAGCGCAACGCTCCAC 4020
 GAGAAGGTAGGAAGCAAGTGAACCTGACCGCGCGCTTCTAGTCTCGCTGTCAGGTTG
 L F H P S F T L E L A R R I K E R N V H

GTCGCCATCGAGACTTCTCTGCTTCCCAAGCGTGGGAGGTTATTACGCCGCTGCTGGAA 4080
 CAGCGGTAGCTCTGAAGGACGAAAGGGTTCGCAACCCCTCCAATAAGTCCGCGACGACCTT
 V A I E T S C F P K R W E V I Q P L L E
 2T7.4
 TTTGTGCTGCTGTTTCATCGTCGACCTAAAATCGCTGAATCCAAAGAACAATGAAGAAGTC 4140
 AAACAGCTAGACAAGTAGCAGCTGGATTTTAGCGACTTAGGTTTCTTGTACTTCTTCAG
 F V D L F I V D L K S L N P K K H E E V
 2T7.4
 GTTGGTGGCCACTTCAGCCGATACTCGACAACATCGAGCATCTCATGCGATCCAAGGCC 4200
 CAACCAACCGGTGAAGTCCGCTATGAGCTGTTGAGCTCGTAGAGTACGCTAGGTTCCGG
 V G W P L Q P I L D N I E H L M R S K A
 AACATCCGCATCCATATTCCCGTATCCCGTGAATCAATGATTCCCGCCAGGATTTCGAC 4260
 TTTGAGGCGTAGTATAAGGGCACTAGGGCACTAAGTTTCAAGGGGGTCTTAAAGCTG
 N I R I H I P V I P G F N D S P Q D F D
 3T7.3
 GACTATATTAGTTATTGAGTCCGCATACCATCGAGCTGGAGGGTTCGATATTCTAAAT 4320
 CTGATATAATCAATAACTCAGCGGTATGGTACGTCGACCTCCACAGCTATAAGATTTA
 D Y I S Y L S R H T M Q L E G V D I L N
 SspI
 TACCAGCTCTATGGAGAAGAAAATATCGCTCCCTGGCGAGGAAAAGAATATCAATAT 4380
 ATGGTGCAGATACCTTCTCTTTATAGCGAGGACCCGCTCCTTTTCTTATAGTTATA
 Y H V Y G E G K Y R S L G R E K E Y Q Y
 TTAGGTGTTGAAGAGAACCAGCCCGGAGAGGTTAGTACCTTGTCAAAGGTTTGAAGCTT 4440
 AATCCACACTTCTCTTGGCGGGCTCTTCCATCATGAGAAACAGTTTCCAAACTTCGAA
 L G V E E N P P E K V V P L V K G L K L
 GCAGGCATCAAGAACGTAACGATTGGCGGTTGGTTGGCATCACGGCGGACAAAGACGAG 4500
 CGTCCGTAGTCTTGCATGCTAACCGGCAACCAACCGCTAGTGGCCGCTGTTCTGCTC
 ← bssD3' BamHI
 A G I K N V T I G G L V G I T A D K D E
 ACTGATCGCGATGCCCCACAAGGTGTATTACATAAACAATAGGAGACTCATCTATGGAA 4560
 TGACTAGCGCTACGGCGGTGTTCCACATAATGTTTGTATCTCTGAGTAGATACCTT
 T D R D A A T R C I T *
 bssC → M E
 ACCACCACATGTAAGCAGTGGCAGAACTTCTTCTCTTCCCAAGGATGCCGATGATTAC 4620
 TGGTGGTGTACATTGCTCAGCGTTTGAAGAAAGGAGAGGGTTCCTACGGCTACTAATG
 T T T C K Q C A N F F P L P K D A D D Y
 GAAGCCGCAAGGACAGATTGCGTGCAGAAAGGAAAGTACAAAGGCAAACTACTGGCTC 4680
 CTTCCGCGCTTCCGCTAACGACGCGCTTTTCCCTTACTGTTTCCGTTTATGACCGGAG
 E A G K A D C V R E K E D D K G K Y W L
 TCCAAACTATATTCCGAAACAGCAGCGGATGTAAGCCTTTCACGCGAAACGATAAAAC 4740
 AGGTTTGGATATAAGCTCTTGTGCTGGCTACACTTCGGAAGGTCGCTTGTGCTATTTG
 S K P I F E N S T R C E A F H A K R *
 CAAAGTTCGAGGAGACTGCCATGAATGACATCGCAAGCGCAAGGACTTGAATAACAAG 4800
 GTTTCAGCTCCTGACGGTACTTACTGATGATGATGATGATGATGATGATGATGATGATG
 bssA → M N D I A S A K V L E Y K G
 bssA5' new
 GAAAAACGCTCAATTTACACCTGAAGATCCGCGGAGCAAGATTCCATCCGACGAGT 4860
 CTTTTCGAGTAAAGTGGACTTCTAGGCGGCTTCCGTTTCTAAGGTAGGCTGCTCA
 ← bssAint3'
 K T L N F T P E D P A E A K I P S D E L
 TGCACGAGCATCTGCAAAAGCCGTCGACCGCAAGAACCAACCGCTGAAGGAGCGTTGCC 4920
 ACGTGCTCGTAGACGTTTTCCGAGCTGGCGTCTTGGTTTGGCGACTTCTCGCAACCG
 bssAint3'
 H E H L Q K P S T A R T K R L K E R C R
 3T7.4
 GCTGAAACACGATCCGCGAGGCTTATCGAAAGAGCGTCAAGCCCGCATCGAGC 4980
 CGACTTGTGCGTAGGCGTCCGCTCAAGTAGCTTTCTCGCAGTCCGCGCTGAGCTCG
 W K H A S A G E F I E K S V T A G I E R
 3T7.4
 GCATGGCTATCTGACCGAAGCACAAAGCTAGCGAGGTTAGCGGAAAGTCAATCCGCG 5040
 CGTAGCCGATAGACTGGCTGCTGTTGTTGCTGCTCCAGTCCGCTCCAGTCCGCTCAGTAGGCGG
 M R Y L T E A H K A S E G Q P E V I R R
 GCGCGCTCGGTTGGCAATGTTCTGAACAAGTCAACCTGGTCTTCAAGAGGACGAAT 5100
 CGCGCGGACCGCTTACAAGACTTGTTCAGTTGGGACCCAGGATTCCTCTGCTTA
 A L G L A N V L N K S T L V L Q E D E F
 TCATCGTGGCTACACCGCAAGATCCGAACATGTTCCGCTGTATCCGAAATTTGCC 5160
 AGTAGACCGGATGGTGGCGCTTCTAGCTTGTACAAGGGCGCATAGGGCTTAACAGG
 I V G Y H A E D P N M F P L Y P E L S H
 ACATGGCGGTTGAGGACTACTGAGGAGCGACTCTCGCTCAACCGCCGACGAGGCTG 5220
 TGTACCGCAAGTCTGATGGACTCTCGCTGATGAGCGGAGTTGGCGGCTGCTCCGAC
 M A V Q D Y L R S D Y S P Q P A D E A A

CCCGCATCAATGATTACTGGAAGCCGACAGCTGCAAAAGCAAGTCCAGCCCTACTTCCG 5280
GGCGTAGTTACTAATGACCTTCGCGCGTGCAGAGTTTGGTTCACGGTCCGGATGAAGC
A I N D Y W K P H S L Q S K C Q P Y F D

ACCCGGCCGACCTCCGGCGTATGTACCAGTTCAGCAGCATGGAGGCCCCCTCCGCGT 5340
TGGCCCGCTGGAGCCGCATACATGTCAGCTGCTACCTCCGGGGGGAAGCGCA
P A D L G R M Y Q V S S M E A P P F A S

CCGGTTACAACAGTATCGTACCGCCCTACGAAACCGTCTTGGAGACGGCTGCTGGCGC 5400
GGCCAATGTTGTCATAGCATGGCGGATGCTTTGGCAGGAACCTCCGCGGACGACCGCG

← bssA3' new

G Y N S I V P P Y E T V L E D G L L A R

GCATCAAGCTCCGCGAGAAACATATCGCCGAGGCCAGGCCGACATGTCCACCTTCCCGT 5460
CGTAGTTCGAGCGGCTCTTTGTATAGCGCTTCGGGTCGGCTGTACAGGTGGAAGGCA
I K L A E K H I A E A Q A D M S T F P W

GGAAACGGACCAAGGCGCTAGACAACATCGCAAGATCGACCACTGGAAGGCGATGTTCA 5520
CCTTCCCGTGGTTCGCGGATCTGTTGAGCGGTTCTAGCTGGTGACCTTCCGCTACCA
N G T K G L D N I A K I D H W K A M V I

TCGCCTGCAAGCGGTCATCAGCTGGCGCGCCGCTCAAGGCCGCTCTGTGCCGATCGTCC 5580
AGCGGACGTTCCGCGAGTAGTCGACCCGCGCGGAGTTCGCGCAGACACGCGCTAGCA
A C K A V I S W A F Q G R L C R I V A

CCGAGAATTTTGAACCGATCCGAAAGCGCTCAGGCCAGCTGCTCGAAGTTGCTGACATCT 5640
GGCTCTTAAACTTTGGCTAGGCTTCGCGAGTCCGCGCTCGACGAGCTTCAACGACTGTAGA
E N F E T D P K R Q A E L L E V A D I C

GCCATCGCGTTCGCCGCGAGCCCTGCAAGGGCCCTCAAAGCGCGATGCAAGCGAAGTTCT 5700
CGTAGCGCAAGGGCGGCTCGGAGCGTTCGGGTTCTCGCGCTACGTCGCGCTTCAAAGA
H R V P A E P C K G L K D A M Q A K F F

TCACCTTCTGATCTGCCACGCCATCGAGCGCTACGCGAGCGGCTACGCGCAGAAGGAAG 5760
AGTGAAGGACTAGACCGTCCGATAGCTCCGATCGCTCGCGATGCGCGTCTTCCCTTC
T F L I C H A I E R Y A S G Y A Q K E D

→ 3T7.5

ACACCTGCTGTGGCCGTAATAAGGCGCTCGGTATCGACAAAAGTTCCAGCCGATGG 5820
TGTGGACGACACCCGGCATGATGTTCCGGAGGAGTAGCTGTTTTCAGGTCGCTACCA
T L L W P Y Y K A S V I D K K F Q P M D

ACCACATGGGCGCGTGAAGTCTCGAGATGGAACGTTTGAAGATTTCCGAGCACGGTG 5880
TGGTGTACCCCGGCCACCTTGACGAGCTTACCTTGCRAAACTTCTAAAAGCTCGTGCCAC
H M G A V E L V E M R L K I S E H G A

CCGGCAAGTCGCGCCCTATCGGAAATCTTCCCGGGCTCGAACGACCTGTTTCATCTCA 5940
GGCGTTTCCGCGCGGATAGCGCTTTTAAAGGGCCGAGCTTGTGACAAAGTAGGAGT
G K S R A Y R E I F P G S N D L F I L T

CCGTCCGCGGACGAAACGCAAGGCGGAGGACGATGCAACGACATGACCGACGCCATCC 6000
GGCAGCCCGGCTGTTGCGGTTCCCGCTTCCGATGCTTACTGCTGCTGCGGTAGG
V G G T N A K G E D A C N D M T D A I L

TGGAAGCGGCAAGCGGATCCGCACTGCTGAGCCCTCCATCGCTTCCGCTACAGCAAAA 6060
ACCTTCGCGGTTCCGCTAGCGTAGCAGTTCGAGGAGTAGCAGAAAGCGATGCTGTTTT
E A A K R I R T A E P S I V F R Y S K K

AGAGTCGCGAAAAGCGCTGCGATGGGTTTTCGAGTGCATCCGCGACGGTCTAGGCTATC 6120
TCTCAGCGCTTTTTCGCGAGCTACCCAAAGTCACTACGCTAGGCGCTGCGAGATCGATAG
S R E K T L R W V A E C I R D G L G Y P

CGTCGATCAAGCAGCAGATCGCACGCGCAGATGAAGAAATATGCCAAATTCAGCC 6180
GCAGCTAGTTCGCTGCTTACCGCTGCGCGCTTACTTCTTATACGTTTAAAGTCCG
S I K H D E I G T A Q M K E Y A K F S L

TCAACGGCAACGCGCCACCGACGAGGAAGCCCAACTGGTCAACGCTGCTGTGCATGT 6240
AGTTGCGGTTGCGCGGCTGCTGCTTCCGCTTCCGCTGTTGACCCAGTTGCACGACAGTACA
N G N G A T D E E A H N W V N V L C M S

CGCCCGGCTTACGCTGCGCGCAAGACACAAAAGCCGATCGGAAGGCGGAGTTCGG 6300
GCGGCGGAAGTGCAGCGGCTTCTGTTTCTTGGGCTAGCCTTCCGCGCTCAAGCC
P G L H G R R K T Y T T R S E G S S V

TCTTCCGCGCAAGTGTGGAATACAGCTCAACGACGCTACGATTGTTGCTGCTACGCGG 6360
AGAAGGCGCGTTCACGACCTTTAGTGCAGTTCCTCCGATGCTAACACGATGCGGCC
F P A K V L E I T L N D G Y D W S Y A D

ACATGCACTTAGGCCGAAGACGCGTGAACCTCTCTCGTTGAAGACCTTCGAGGACATCT 6420
TGTACGTCATCCGGCTTCTGCCCACTTGAAGGAGCACTTCTGGAAGCTCCTGTAGA
M Q L G P K T G E L S S L K T F E D I W

GCGGAGCCTTCCGCAAGCAGTATCAATACGCGATCAACTTGGCATCAGCACAAGGACG 6480
CCCTCCGGAAGCGTTCGTCATAGTTATGCGCTAGTTGGAACCGTAGTCTGTTGCTCC
E A F R K Q Y Q Y A I N L G I S T K D V

→ 3T7.6

TGTCGCGCTACTTCCGAGCAGCGTACCTGCACTTACCTTTCGCTGCGCAATCGACGACG 6540
ACAGCGGATGAAGCTCGTCGCGATGACGCTCAAGTGGAAAGCACAGCCGTTAGCTGCTGC
S R Y F E Q R Y L Q L P F V S A I D D G

GCTGCATGGAATTCGGCATGGACGCTTCCGCGCTGTCGCAACCAACCGCTGGCACA 6600
CGACGTACCTTAAAGCGTACCTCGAAGCGGACAGGCTTGTGGTTCGCGGACCGGCT
C M E F G M D A C A L S E Q P N A W H N

ACGAGGTCACGACGGTCTCGCGCGAAGTCCCTCGTGGCGATCAAGAAGCTGGTGTTCG 6660
TGCTCCAGTGTCCAGCAGCGCGCTTGGAGGAGCACCGCTAGTTCTTCGACCAAGC
E V T T V V A A N S L V A I K K L V F E

AGGAGAAGAAGTACACCTTGGAAACACTCAGCCAGGACCTGAAGGCGAAGTGGGAAGGCT 6720
TCCTCTTCTTCAATGTGGGACTTGTGAGTCCGCTGACTTCCGCTTGAACCTCCGA
E K K Y T L E Q L S Q L K A N W E G F

TCGAGAAAATGCGCGTCAAGTCAAGCGCGCGCAAGTGGGCAACGACGACGAATACG 6780
AGCTCCTTACGCGAGTGAAGTTCGCGCGCGTTCACCCGCTTGTGCTGCTTATGC
E E M R V D F K R A P K W G N D E Y A

CCGACAGATCTCTCCGATTCTACGAGAAGTCACTCGCGCGAATTCGCGCAAGATCA 6840
GGCTGCTGACGAGGGGCTAAGATGCTCCTTCTAGTAGCCGCGCTTAAACGCTTCTAGT
D S I V S R F Y E E V I G G E L R K I T

CCAACTACTCTGTTGGCCCGTCTCGCGACCGGTCAGGCTGTCGCGCTGTACATGGAAG 6900
GGTGTAGAGACACCGGCGCAGGACGGTCCGAGTCCGACAGCCGACATGTACCTTC
N Y S G G P V L P T G Q A V G L Y M E V

TGGGTTCCGCGATGGGCCGACCGCGCGGCTTCCGCGCGAAGCGGCTGACGACG 6960
ACCCAAGCGGCTACCCGGCTGCGGCTGCGCGGCGGCGCGCTTCCGCGACTGCTGC
G S R M G P T P D G R F G G E A A D D G

GTGGCATTCTCCCTACATGGGACCGCAAGAAGGCGCGACGCGGTTGTCGCTCGG 7020
CACCCTAAAGAGGATGTACCGTGGCTGTTCTTCCGCGCTGTCCGCAACCGGACGCC
G I S P Y M G T D K K G P T A V L R S V

TGTCGAAGTGCAGAAGAACCAGAAGGCCAATCTGCTGAACCGCGCTTGTGCTGACCCA 7080
ACAGGTTCCAGCTTCTTGTGCTTCCGCTTACGAGTGTGGTTCGCGAAGCAGGCTGGGT
S K V Q K N Q K A N L L N Q R L S V P I

TCATCGCTCCAAGCAGGTTTCGAGATCTGGAACCGCTACATGAAGACCTGGCAGCAAG 7140
AGTAGCGGAGTTCGCGCAAGCTTAGACCTTCCGCAAGTACTTTCGCGCGGCTG
M R S K H G F E I W N A Y M K T W H E L

TGAACATCGATCATGTTCAAGTCTGACGACCGGCAAGTTCGCGCGGCGGCGGCTG 7200
ACTTGTAGCTAGTACAAGTCAAGTTCGAGGAGCTGTGCTGCTTACCGCGCGCGCTGC
N I D H V Q F N V V S T D E M R A A D R

GCGAACCGAGAAGCACAATGATCTCATCGTTCGCGTTCGCGCTACAGCGCCCGGTTCCG 7260
CGCTTGTCTCTTCTGTTACTAGAGTAGCAAGCGCACGCGGATGTCCGCGGCGCAAGC
E P E K H N D L I V R V S G Y S A R F V

→ 3T7.7

TCGACCTTCCAACCTATGGGAGAACCATATTGCCCGTCAGGAACAGGACTTCAGCG 7320
AGCTTGAAGGTTGGATACCGCTCTGTTGAGTAAACCGGACGCTTGTCTGAGTCCG
D L P T Y G Q N T I I A R Q E Q D F S A

→ 3T7.7

CGTCCGATCTCGAATTCCTAAACGTCGAAATCTGAAAAGCAGCCAAACGAGGGGCG 7380
CGAGGCTAGAGCTTAAAGATTGACAGCTTTAGACTTTTTCTGCTGGTTTGCCTCCCGCTC
S D L E F L N V E I *

CGTTCTGCGCCCTCCACCCATAGTCAAACCAATATATTGGAGCCATAACATGGAAGCG 7440
GCAAGACCGCGGGAGTGGGTATCAGTTTGGTTTATATAACCTCGGATTTGACCTTCGC
bssB → M E A

GAAAAGAGCTTGAACACAGCCGATACCGAAGTCGCGACAGCGAAGCCCTGCCGATC 7500
CTTTTCTCGAACGTTTGGTCCGATATGGCTTACGCGTGTGCTTCCGGAGCGCTAG
E K S L Q N Q P Y T E V G T A K P C R I

TGCAAGTGGCAAAACCCCTGACCCACCGATCCGACCGTGGGAGTGCACCGCAACCGC 7560
AGCTTACCGTTTGGGACTGGGTTGGCTAGGCGTAAACCGCCCTCACGTCGCGGTTGGCG
C K W Q T P D P T D P H R G Q C T A N R

CACGCCATGGGCGGAGTCTGGAACCGTGGCTTAGGGAGCTCGAGAACAGCAGCTGCTCC 7620
GTGCGGTACCCGCTCAGACCTTTCGACGCAATCCCTGAGCTCTTGTGCTGGACGAG
H A M G G V W K R W L R D V E N T T C S S

AGGCATGAGGAAGGCAAACTGAGCTTCCGCGACCGCTGGAACACCGGACAGAAGTGGC 7680
TCCGACTCCTTCCGTTTACTCGAAGGCGCTGGTGCAGACTGTGGCCTGCTTCCAGC
R H E E G K L S F R D H V *

TTACCTCCCGACCGGTTGACGATAGATCATGAAAACCTCTCCTAGCACAAATGGCTTGT 7740
AATGGAGGGCTGCGCAACTGCTACTAGTACTTTTGGAGAGGATCGTGTTTTACCGAACA
← bssEint3'

→ bssE → M K T S P S T N G L F

TCATCCCGGAAGTGGACCCCTACTACTATGTAAGTACGGAAAACAGAGCTTCTCGATA 7800
AGTAGGGCTTACCTGGGATGATGATACATCATGCTTTTGGTCTCGAAAGAGCTAT

I P E V D P Y Y Y V S T E N Q S F L D K

AATTGCCAAGATATCCAAAAGCACCCCGTCAACGCTTCTGGTGGTTCGGCAACAGGCT 7860
TTAAGCGGTTCTATAGTGTCTTCTGGGCGACTGTCAGAACCCAGCCGTTTGTCCCGA
G A K I S K K H P V N V L V G K Q G C

→ 3T7.8

GCGGCAAGTCTCTGTTGCGGCAATACCGCGCGTCAACAGGCTGCGCTGGGCACT 7920
CCGCGTTCCAGGAGACACCGCTTATCGCGCGGAGTTCGCGACGGGACCGGCTGGA
G K S S L V R Q Y A A V N R L P L A T F

3T7.8 →
 TTCAGATCGGCATCCTGTCCGAGCCGGGACAAATTGTTTGGTGAATACCGCTGGAGAACC 7980
 AAGTCTAGCCGTAGGACAGGCTCGCCCTGTTAAACAAACCACTTATCGCGGACCTCTTGC
 Q I G I L S E P G Q L F G E Y A L E N G

GGGAAACCGCTACAAGCAGTTCTCTTCCCTCAGGCCATCCAGACGCCCACTCGCTCA 8040
 CCCTTTGCGCGATGTTCTGCAAGGAGAAGGAGTCCGCTAGTCTGCGGTTGACGCACT
 E T R Y K Q F L F Q T A I Q T P N C V I

TCCATCTTGAGGAGATCAATCGCCCCGAACCCGAAAGCGGTTGAACATGCTGTTCTCCA 8100
 AGGTAGAACTCCTTAGTTCAGCGGGCTTGTGGCTCCGCAACTGTACGACAAAGAGGT
 H L E E I N R P E H P K A L N M L F S I

TCCTCTCCGATGATCGACAGGTATGGATGACGAACTGGCGCTGCTGCAAGTGGCGCCG 8160
 AGGAGAGGCTACTAGCTGTCCATACCTACCTGCTTACCCCGGACGACGTTACCCGCGGC
 L S D D R Q V W M D E L G L L Q V A P G

GGTCTGTTTCTTCGCGACGCTGAACGAGGGCGCGAAATTCGTCGGCAGGAGCTACTCG 8220
 CCACGAAAAGAAAGCGCTGCGACTTGTCCCGCGCTTAAGCAGCCGTCGCTCGATGAGC
 V V F F A T L N E G A E F V G T E L L D

ACCCGCGCTCGCGACGCTCTATGTCACGACCATGGATTTCTGCCGAACGAAGTCC 8280
 TGGCGGGGACGCGCTGGCGAAGTACAGTCTGTGACTTAAAGACGGCTTGCCTCAGC
 P A L R D R F Y V T T M D F L P N E V E

AAATGGAGGTGCTGAAAAGAAAGACCGCGTAACAAACGAGCAGGCCAGGAAATCATCA 8340
 TTACTCTCCACGACCTTTTCTTCCGCGCATGTTGCTGCTGCGGCTCCCTTTAGTAGT
 M E V L E K K T Y T N E Q V L A R E I I T

CGTGGCGAACACGACTTCGCGCAATGCCGACTCGCGCTCGACGTTTCCACCGCAAGA 8400
 GCCACCGCTTGTGTAAGCGCGTACGGCTGGAGCGCAGCTGCAAAAGTGTGGCTTCT
 V A N S I R A N A D L G V D V S T R K I

TTCTGATGCTCGCGGAGATGATCGCCGCGGAGTACGCTGCGCGAGGCCATCGTGACGA 8460
 AAGACTACGACCGCTCTACTAGCCGCGGCTCTGCGACGCGCTCCGCTGACTGACTGCT
 L M L G E M I A A G G T L R E A I V T S

GTCTCAAACCGACAAGAAGCGCTCGAGTCCGTTTGTGCTGCTGCGACGCTCAATCTGG 8520
 CAGAGGTTTGGCTGTTCTTCTGCGAGCTAGCGCAAAACGACGCGACGTCAGTGTAGAC
 L Q T D K K T L E S V L L S L H V N L G

3T7.9 →
 GGAAGTGGAAAAGAGCAAGACCGAATACGTTTCACTGATGATGCTTCCGCTTAAAGT 8580
 CCTTCCACCTTTTCTCGTCTGCTTATGCAAGTCAATGTAACCGCGGATTTCCAAAAGTA
 K V E K S K T E Y V Q Y I A A *

GGCAAGAACAAGGACACCACTTCGCGTATGAACAGCGCGGAGACGTCGACGCGCTT 8640
 CGGTTCTTGTCTGCTGTTGTAAGCGCACTTGTCCGCGCCCTCTGCAAGTTCGCGAA
 bssF → M N S A G D V K R F

CGTCATCCAGGTGAGGAAGGCTATTTCGGAATCTTGGCGTCCGCAAAATCCCCGATCGA 8700
 GCAGTAGGCTCCACTCCTTCCGATAAGCCTGAAGACCGCAGCGCTGTTTAGGGCTAGCT
 V I P G E E G Y S D F W R R D K S P I E

GCTCCGCGAGTTGGTGAAGCTATTGGTCCGCGATTTCGCAACTGTCGACCTTCATCGGACG 8760
 CAGGCGAGCTCAACCACTTCGATAAACCAGCGCTAAGCGTTTGACAGCTGGAAGTAGCCTCG
 S V E L V K L L V A I R K L S T F I G R

CAACGTCGCGAAATCGTCTGGTCCGGAATGGAATCGCAATCGCCTCGCCCTCGATCC 8820
 GTTGCGCCGCTTTAGCAGACCGCTTACCTTGTAGCTGTTACCGGAGCGGGAGCTAGG
 N V G E I V W S G C M E L D N A L A L D P

AAACCCGATCATGGCACGATCCGGTTCGCGCGGAAAAACCGATCTGATGTTGCGCAT 8880
 TTGGGCTAGTACCCGTCATAGGCAAGGCGCCCTTTTGGCTAGACTACACGCGTA
 T P I M G T Y P V P A G K T D L M V G I

CATGTTCAAGAGCGTACAAGCGTGTGAGTGGTTCAGAGCGTCTGCGCGAGATGCTCAG 8940
 GTACCACGTTCTCCGATGTTCCGACAGCTACACAGTCTCGCAGACGCGCTTACGAGTC
 bssFint3' ←

M V Q E A Y K R V E W S E R L R E M L R

GCTGCGCGTCCAGCCCCGACGCGAGTACGAATAACAAGTTCGACATGTTCTTACCCCTTG 9000
 CGACGCGCAGCTCGGGGCTCGCTCATGCTTATGTTCAAGCTGTACAAGAAGTGGAAAC
 bssFint3' ←

L R V Q P P T Q Y E Y K F D M F F T L C

CGAGTCCGCTACGTCGACGCGCTGGCCAAAGAGCGTCTCGTTACTACCGCGAGGT 9060
 GCTCAGGCAAGTGCAGTCCGCGACCGGTTGTTCTCGCAGGCAATGATCGGCTCCA
 E S V Y V D G L A N K S V L G Y Y A E V

GGCGCGTACTGGCGTATCGTCAAGAGCGTGAAGAGCGCTGATCAAGCCGCCACCCCTC 9120
 CGCCGCTGACCGCATAGCAGTTCGCGACTTCTCGGACTAGTTCGCGGTTGGGAGAG
 A R D W R I V K T L K S L I K P P T L S

3T7.10 →
 CGAGATGTTGACCTGTGGTGGCGCTTGGCGGCCATCGGACCCGGAATCTACAAGCA 9180
 GCTCTACAACGTGACACCCGCAACCCGCGGAGTCCGCTGGCCCTTGAGATGTTCTG
 E M L H L W W R L A A H R D P E L Y K Q

CGGCTCACAGCATCTCACCCCTGGTGGCGCTGATGCGGGGTAGCCTGGACAGTATTA 9240
 CCCAGTGTGCTAGAGTGGGACCCACCGCACTACCCGCCATCGGACCTGGTGCATAAT
 G H S D L T L G G L V M R G S L D Q Y Y

CAGCAAGCCGATGCAGACCATGAACAGCATCGTGCCCGCCCTGCGCCACGACTGCCCGCC 9300
 GTCGTTCCGGCTACGCTGTTGTTCTGCTAGCAGCGGGCGGACCGGGTGTGACGGGCGC
 S K P M Q T M N S I V P A L R H D C P A

GCTCTCGAGCGTCAAGCAGTTCGGCGGATTTCCGCTTGTATCTTATGAAAACTCTGGTG 9360
 CGAGAGCTCGCAGTTCGCTAGCCGCTTAAAGCGGAACTAGAGATACTTTTTGAGACC
 L S S V S D R R D F F R L D L Y A K T L W C

CGAAGTGTCAAACCACTCCGCTTTCGCGCGGGGATCGCAGCGATCGTTTCATGATCCC 9420
 GCTTACAGAGTTTTTGTAGCGGAAGACCGGCGCCCTAGCGTTCGCTAGCAAAAGTACTAGG
 E V L K N I R F W P G D R S D R F M I P

CGACATGGCGGATGATGAACTGGCCAGGAAGGCGGAGCAGCAGTCAAGGCCAC 9480
 GCTGTACCCGCTACTACTTACCGGGTCTTCTCCGCTCGTTCGCTGCTGCTGCTGCTG
 D M G D D E L A Q E E A E Q A A V K A T

CATCGTCAATACGCCAACCTGATCGAGCGCGCGTCCGCGAGAAGAACCGCGATTCAC 9540
 GTAGCAGTAAATGCGGTTGGACTAGCTAGCTCCGCGGACGGGCTCTTCTGGCGCTAAAGT
 I V N Y A N L I E A A L P Q K N R D F T

CGATCAGATCAAAGGAAAGCGTGGCAATTTTCGAGAAGCTGCGCCGAGTTCGAGGCAACA 9600
 GCTAGTCTAGTTCCTTGCACCGTTTAAAGCTTTGACGCGGCTCAACTCCCGTCT
 D Q I K G N V A N F E N V A R V E G N D

CATCGTCAATGAGCGCGCAACCGCGTTCGATCGCCACCTTTTGCACAAGTTCGACAACT 9660
 GTAGCAGTACTACCGCGGTTGGCGGAGTTCGAGGTAAGGAACTGTTGCAAGTTCGTTCA
 I V M M A R N R V D R H L L H K L E Q V

GGTGGCGAACCGCAAGTTCGCGCGAGCTTCAACCGCGGCTCAGTTCGCGGAGAGAT 9720
 CCACCGCTTGGCGTGGCTAGCGCGCTGCAAGAAAGTTCGCGGAGTCAAGCCCTTCTA
 V R N A T D R R S V F N R G L S S G K I

3T7.11

TCATAGTCCGCGGCTTTACCGCGCCACACCACCGCGAGTGTTCACAAAAGAAATCA 9780
 AGTATCAGCCGCGAAATGGCGGGTGTGGTGGCGCGTCAAGGTTGTTTCTTAGT
 H S R R L Y R A H T T G A V F Q Q K N H

3T7.12

TGAGTTCGACATCGCAAGATGTCGACTGCTGGTGGACGCGAGGGGTCGATGGCCGA 9840
 ACTCAAGCTGACCGGTTCTTACAGCATGACGACCACTCGCTGCGCCAGTACCGGCT
 E F D M R K N V V L G L V D A T G S M A D

CCGACGCAATGGGACCGGAAATGATCACCAGACGCTGTTCCCGGATTTCTGGA 9900
 GGGCTGCGTTACCCGCTGCGCTTTACTAGATGTTGCGCAAGTGGCGCTAAGACCT
 P T Q W D Q A E M I Y Q T L F T A I L E

ATACACTAGCAACCGCGACTATTGCGCTACAACGATGTCAAGAAGCGGTCGCCGATCA 9960
 TATGATGATCGTTCGCGCTGATAAGCGGATGTTGCTACAGTCTTTCGCGACCGGCTAGT
 Y T S N A R L F A Y N D V K N A C R I T

CGAGATCTATCGTGGTGGCGCATGCTGACGCTGCTGCGCAGCGAAGACTGCTTCCG 10020
 GCTCTAGATAGCACCACCGCGTACGACTGCGCAGCAGGCTGCTTCTGACGCAAGGC
 E I Y R G G R M L T V L P H G R T A S G

CGAAGCCATCATCGCCAGCGATTGAATACCCGACGCGCGGAAAAAACCGTCTGAT 10080
 GCTTCGGTAGTACCGGTCGCTAATGTTGGCGTGGCGCCCTTTTTCGAGGACTA
 E A I I A T A L N T R T P G K K T L L I

CCACATCCGACCGCGCTCGAATGGGGTGTGGCGTGGCGGATGCCATCAAGTATG 10140
 GTGTTAGTGGCTCCGCGGAGCTTAAACCCACCGCGCTACGGCTACGGTTCATAAC
 H I T D G A S N W G C G V A D A I K Y C

CAAGGCAACCGGTATCAGCCTGCTCAGCTCGGCATCAGTTCGAGTCTGTCGCGAAAGCA 10200
 TTTCCCGTTGCCATAGTCGACGAGTGGAGCGTAGTCAACGTCAGACGCGGTTCTGT
 K G N G I S L L T L G I S C S L S A K Q

ATCGTACGGGACGAATACCGCACTCTCGTGAAGTTCGTCGACAAGCCGAACAATTGCC 10260
 TAGCATGCCCTGCTTATCGCGTGGAGCACTCAAGGCTGTTCTGGCTGTTTAAAGC
 S L R D E Y G T L V L V D K T E Q L P

GAAGTGTGTTGGCGAGTTGATCATCAGTGAATGCGTGAATCGAGGACCCACAGAAGTG 10320
 CTTCAACAAACCGCTCACTAGTAGTCACTTTACGCACTTAGCTCTCGGCTGTCTTAC
 K L F G E L I I S E M R E S R T P Q K *
 bssG → M

AACTGCTCTTGTGCGACCGTGTGGAAGCCGAATGGCAGATGTTCTGCGCGTCAAG 10380
 TTGAGCAGGAACGAGCTGGTGCAGCACTCGGCTTACCGTCTACAAGCAGCGCGAGTCC
 N S S L L D H V L E A E W Q M F V R V R

AGCGCGGCGACCGCCCTGTCAGAGCGCTCAGACAACCTTCAAGACGATCGCTCCAGT 10440
 TCGCGCGCGTGGCGGACAGTCTCGCGAGTCTGTTGAAGTTCGCTAGGCGAGGTC
 S A R H A P C Q S A P D N F K T I R S S

CTGTTGAGACGTTGACGCAACCGATGCTGCGCTCTATCTTCCGATCTGGAACAGGCG 10500
 GACAAGCTTGCACCTGCGTTGGCTACGAGCGGAGGATAGAACCGCTAGACCTTTGTCGG
 L F E T W T Q P M L A S Y L A D L E Q A

3T7.13

GAGGCGGTCGGCGCAACCTGCTTGGGAAAAATACGCGCGATGGACAACCTGATTCCA 10560
 CTCGCGCGCGGCTGGACGACGCTTTTATGCGCGGCTACCTGTTGGACTTAAGGT
 E A V G R N L L A E K Y A R M D N L I P

CCACTGTCAAACAGCCCGTTGATGACATCATGTCAGATGGAAGCAATGGCAGGAA 10620
 GGTGACAGTTGTGGCGCAACTAGCTAGTAGCAEWGCTACTCTTCTGTTAACCCGCTT
 P L S N S P L I D I I V T M E S N W Q E

GAGTTGGGACGTGCTATCCCGCCCTTTATCGCGATGTTGCCGACGATGGATGAAACC 10680
CTCAACCCCTGCAGCATAGGGCGGAAATAGCCGTACAACCGCGTGTACTACTTTGG
E L G R R Y P A L Y R R C C R S M D E T

GGCGACGGCGGAACTTCGAAATCTATCTGGGCTGCGAACTGAAACCTACGGCGATCAT 10740
CCGCTGCCCGCCTTGAAGCCTTAGATAGACCCGACGCTTGACCTTTGGATGCCGCTAGTA
G D G R N F G I Y L G C E L E T Y G D H

bssGint5' ->

ACATTGCAGCTGATTTTCGAGAATATCGAGGACGCGCTGCCGCCAATCGCAATCTCGCG 10800
TGTAACGTCGACATAAAGCTCTTATAGCTCCGTCGCGAGCGCGGTTAGCGTTAGAGCGC
T L Q L Y F E N I E A A V A A N R N L A

CTCGAGGCCCTGCGCTCGACTGGTATGCAAGAAGCGGATCTCGGACCTGGATCACGCCGAG 10860
GAGCTCCGGGACGACGCTGACCATACGTTCTTGCCGATAGCGCTGGACCTAGTGGCGGCT
L E A L R R R L V C K N G Y R D L D H A E

AGCGTATTGAAATACGGAGAGACGGCATGAAACATCGACGTCAGGCTGAATCTGCTCGCG 10920
TCGCATAACTTTATGCTCTCTGCGTACTTGTAGCTGCAGTCCGACTTAGACGAGCCGC
S V L K Y G E T A *

TATTGGGCATCAAAACAAGCGAGCTTGGCCATCGATACACCGGGCAATGCCACGGTCGCCA 10980
ATAACCCGTAGTTTGTTCGCTCGAACGGTAGCTATGTGGCCGCTTACGGTGCACGGGT

CGTTGATCAGCTTGATCGACAGCAACAATCCGGCTTCGCGCGGCACTGACGAACGACA 11040
GCAACTAGTCGAAGTGTGCTGTTAGGCGCCGAAGCGCGCGTGTGCTGCTGCTGT

GCGAGGCACTCGCCCCAGTTCTGATTTTCATCAATGCGGAAACCGCTACATCTGCG 11100
CGCCTCCGTAGAGCGGGTCAAGCATAAGAAGTAGTTACCGGCTTTGCGGCATGTAGACC

ATGGCATGAACACAGACTCGACGACGGCATGCTGCTCAACGTTGATTCGCGCCATTGCGG 11160
TACCGTACTTGTGGTCTGAGCTGCTGCGCTACAGCAAGTTCGCACTAAGCGCGGTAACGCC

GAGGCTGAAAGGGGGCGTACATGATCGTCGACCGCGGCAATCGCAITTTCACTCTCGTC 11220
CTCCGACTTTCCCGCAGTGTACTAGCAGCTGGCCGCGCTTAGCGTAAAGTAGAAGCAG

GTTCTCTGATGATCGCTGAGTACCGTTACAACATACATCAACCTGCGCGGCAACCGCTG 11280
CAAGGAGACCTAGTGCAGCTCATGCAATGTTGATGATGTTGGACGCGCGCTTGGCGAC

ACCAACAAGCAATATCTGGAACATTTGGGAAAATGGCTTGTGTTGCGGCTGCGCGAGGAA 11340
TGTTGTTCTGTTATAGACCTTTGTAACCCCTTTTACCGAACAACAAGCGGACGCTCCTT

GTCAGGAAATGCGAAGGAGCTGACCCCTTATGTCGAGGAAAGAAGATTCCCGGCTC 11400
CAGCTCCTTGACCGTTCCTCGACCTGGGAATACAGCTCCTTTCTTAAAGGGCGCCAG

AAATACGACCGCAAGCTGATCACTGAATCCAGCTCAACCGCTGCGTCACTGTGTCTAT 11460
TTTATGCTGCGGTTGACTAGTACTTAAAGTTCGAGTTCGCGACGAGTACACACAGATA

3T3.15

TGCCACGATGAAACCAAGGACGAGTTTGGGAAAATCTGCGCGGCTTGGCGTCTAGGAC 11520
ACGGTGTACTTTGGTTCCTGCTCCAAACCCCTTTAGGACCGCGCAACCCGAGTACCTG

3T3.15
AAGGCATGGATGTACGAACGGGAGACGCTGGAGAAATGGCTGCCCGCGCGCTCAATCTG 11580
TCCGTACTTACATGCTTGCCTTGCAGCTCTTTACCGACGGCGCGCGCAGATTAGAC

GAAAAATGGATCCAGGGACGCGGGTGGATCATGAACAAGCTGAAACGCTACGCGCGGAT 11640
CTTTTACCTAGTCCCTGCGCCGACCTAGTACTTGTTCGACTTGCAGTACGCGCGCTA

GCACGTGCGAAGTTCAAGAAAATGTTGCGCCGACGAGAATGAAATTTTACCGCGCTCTAT 11700
CGTGACGCTTCAAGTCTTTTACAAGCGCTGCTCTTACTTTTAAAGTGGCGCGAGATA

CAGTAGTCGAAACCGTGTTCGCGCAAGCATGCCCCTTCCCGCCCCGGCGCGGAAACCG 11760
GTCATACGCTTGGCAACAAGCGGTTCCGTTACGGGAAGGGCGGGCGCGCTTTGGC
bssI -> M P L P A P G R E T E

AGTGGATCGCCGACGCGCGAAGCCTCGCGCTACGCTGCTTCCGAACACGATGATCCGCT 11820
TCACCTAGCGGCTGCGCGGCTTCCGAGCGGATGACGACAGGCTGTGACTAGGCGA
W I A A R A E A S R Y V L S E H V I R S

CGTGTGCGCGGATCCGTCACCGGGCGAGATCGAAGCGGCTACGACCGCGCGCA 11880
GCGACTACCGGCTAGGCGGTTGCCCGGCTAGCTTCCGCGGATGCGTGGCGCGCT
L M A G S V N G A Q I E A A L R T G R I

TCATCGAGGACACAGGACGCTTGGAGGAGTCCCGGCTACCTGCTGTGCGCGCTCCACG 11940
AGTAGCTCCTCGTGCCTGCAACTCGCTACGGCGCGGATGGAACGACCGCGCAGGTC
I E E H R H V E R V P A Y L L C A V H D

ACGGCAAAGCGGTGACGCTGATCGTGCACACAGCGGCTGCGGGCTTGTGCTCACG 12000
TGCCGTTTCCGACAGTGCAGTACGAGCTGGTGTCCGGCTACCGCCGCAACGACGAGTGG
3T3.14

G K A V H V I A A P Q A D G G L V V T H
AGCCTACGTCGCCGCGCGCCTTGTGGCGACCGCCTTGCATCGCTCCGAGGGGATAG 12060
TGCGGATCAGGGCGCGCGGGAACCCGCTGGCGGAACGAGGCTCCCTATC
A Y V P A P P L W R T A L H R S E G I A

CTGCAATGCTCGATCCGATCAGCAGCTGCTACTTCTGCGGGAGGCCATCAAGCAGGTTGA 12120
GAGCTTACAGGCTAGGCTAGTGTGCTGGAGCATGAAGACCGCGCTCGGTAGTTCGTCACCT
A M S D P I T T C Y F C G G A I K Q V T

CTGTCGGCAATTTGACTACCGCCTGGAAGGACGCTGTACGCTCATCAAGAAGTGGCCG 12180
GACGACCGTTAAAGCTGATGGCGGACCTTCTGCGGACATGCAAGTGTCTTCCACGGCG
V G N F D Y R L E G R L Y V I K K V P A

CCGGACTGTGCCAGCAGTGGCGTAAAAAATACGTCGACGCCAAGGTCGGCCGACGGCTCG 12240
GGCCTGACACGGTCTGACCGCACTTTTATGACAGTCCGGTTCCAGCCGGCTGCGGAGC
G L C Q Q C G E K Y V D A K V G R R L D

ACGCACTGATCGCGCAGCAGGCTTCCACCGCAGCGAGACCGTCCGGCTGATCGACTTCG 12300
TGCGTGTACTGCGGCTGCTCCGGAAGTGGCGTCCGCTGCGCAGCCGACCTAGTGTGAAG
A L I A Q Q A F T G S E T V G V I D F A

CGGCTGCGCTGTGAACCGCCTGACCTCGCCGAGCGGATTTTCGAGCGTTACGGCGGCGA 12360
GCCGACGCGACTTGGCGGACTGGGAGCGGCTCGCCATAAAGCTCGCAATGCCCGCGCT
A A L *
bssJ -> M N G L T L A E R Y F E R Y G G D

CCTGCTCGATGGTGAATTCGCCGCTGGCGGAGCGCATCGCGGTCGGACTCGTGGCGA 12420
GGACGAGCTACCTTAAGCGGGGACCGCGCTCGCTAGCGCAGCTGAGGACCGCT
L L D G E F A A W R E R I A G V G L A G D

TGGCTCCGACTGCTCGGGTTCGACGATGAGCAGTCCGCGGACCGACTGGGGTCCGGG 12480
ACCGAGGCTGACGGAGCCAGCTGCTACTGCTGACGCGCTGGTGTGACCCAGGCGCC
G S D C L G F D D E Q S R D H D W G F G

CTTCTGCTGTGGCTCACCGACGAGACCGCGCGGATCGCGCGCGCTGCGAGCGGGC 12540
GAAGACGGACCGAGTGGCTGCTGCTGTTGGTGGCGCTAGCCGCGCGCAGCTGCGCGC
F C L W L T D A D H A A I G A A L Q R A

CTACGAGCGGCTGCCAAGTCTGTCGCGGCTTGGGCGCGCGGTCACCAATGGGGGCG 12600
GATGCTCGCGACCGGTTTCAGCAAGCGCGAAGCCCGCGCAGCTGGCTTACCCGCTC
Y E R L P K S F A G F G R A V T E W G S

CGGACGGTCCGGGTCGAGACCGCGCTTGGTACAAGCAATACGTCGCCACCCCGA 12660
GCCTGCCACGCCACGAGCTGCGCCGCAACCATGTTGTTATGACAGCGGTTGGGGCT
3T3.13

G R V G V L E T G A W Y K Q Y V R H P D
CGGCCCGCAGACGCTTTTCGACTGGTCCGCAATCCCGAAAAGGATCTGGCAGCTCCAC 12720
GCGCGGCTGTCGAAAAGCTGACCGACGCGTAAGGCTTTCTTAGACCGTCCGACGCTG

3T3.13
G P Q T L F D W L R I P E K D L A A C T
CGCGCGCGGCTTCCACGACCGCTCGACGACTTCTCGCGCGCGCGCGCGGACTACT 12780
GCGCGCGCGCAGAAGTGTGGCGAGCTGCTGAAGAGCCCGCGCGCGCGGCTGATGA
A G L V F H D P L D D F S R R R R L L

CTCGTCTATCCGACGATGTCAGGCTCGCAAGATCGGTGCCCGCTGCATGTCGATCG 12840
GAGCAAGATAGGCTGCTACAGTCCGAGCGGTTCTAGCCACGGGCGACGTCACAGTACCC
S F Y P D D V R L A K I G A R C M S I G

CCAGTCCGGGCACTACAACCTTCCCGCGGCTCGCACGTCGGGAGGTTTCGCGCGCGG 12900
GGTCAGCCCGCTATGTTGAAGCGGGCGCAGGCGTGCACCGCTCCACAAGCGCGCGC
Q S G Q Y N F A R V L A R G E V F A A R

c2A200int5'
ATACGCGGAGACGAAGTCTGCAACGATCTGATGTCCTGCTACCTGCTGAACCGCGC 12960
TATGCGCCTCTGCTCAAGACGTTGCTAGACTACAGGAGCAGATGAGCAGCTTGGCGGC
Y A E T K F C N D L M S L V Y L L N R R

c2A200int5'
CTATGACCGTATTACAAGTGGTGCATCGCGGCTGCTCGCTGCTGCGCGGCTGGGCGG 13020
GATACGTTGCAATAATGTTTACCGACGATGACCGCGGACGAGCAGCGCGCGCCAGCCG
Y A P Y Y K W L H R G L L A L P R L G R

CGTCTGATCGCGAGGTCAGGCTCAAGCCCTGGTGTCTGCGGGGATGCCGACGAGAAGCGGGC 13080
GCAGCAGTACGCTCCAGTTCCGGGACACAGACCGCGCTACGGCTGCTTTCGCGCG
V V H A Q V K A L V S A G D A D E K R A

GCACATCGACGCGCTGTCGCGGACGAGCAGCGCAACTGACGGCGGAAGGGCTGACCC 13140
CGTGTAGCTCGCGACACGCGCGCTGCTGACGGCTTGAAGTCCCGCTTCCGACTGCTC
H I D A L C A A A I A E L Q A E G L T S

CTCAGACAGCGCTTCTGCTGATCAGGTCGCTGATCCACGAGCGCATCCACGAGCC 13200
GAGTCTGTCGGCGAAAGCAGCTAGTGCAGCGGACTAGTGTCTGCTGCTGCTGCTG
S D S R F L V D H G P L I H E R I H D P

CGCTTTCGCGCGCTGAGCGTGTGGTCCGCTGAGCGAGCGCAACCGGAGACCCCTTCA 13260
GCGAAACCGCGGACTCGCACGACCGGACTCGCTCGCGGTTGCTTGGGAAGTAC
A L R R L S V L V G *
bssK -> M

ACGAAACCGGCTAGTGGTATCGCGCGGCTATTTTTCCGGCGGGTATTCGTCGAGGAA 13320
TGCTTTGCGCATGACCACTAGCGCGCTCAATAAAAAGCGCGCCATAAAGCTTCTT
T K R V L V I G G S Y F S R V F V E C T

GCGCTGAAGATGCCCGCGGAGCTGACGCTTCAACCGCGGCAATTTCCGCTGCGC 13380
CGCGACTTCTACGGCGCGGCTGACGCTGACGAGTGGCGCGGTAAGGGCGACGCG
3T3.12

A L K M P G A E L H V F N R G H F P L R
ATGGAAGCGCTACCGGACGATGTCGGCAACCGCGAGCACCCCGAGGAGTCCGGGATCG 13440
TACCTTCCGAGTGGCTGACAGCGCTTGGCGCTGTTGGGGCTCTCCAGGCGCTACGC
M E G V T E H V G N R E H P E E V R D A

ATCCCGGTGGCGAATGGGACGCGGTTGTGCGATTTCGCGCTACACGCCCGCTCACGTT 13500
 TAGGGCCACCGCTTACCTTGCGCCAACAGCTAAAGACGCGGATGTGCGGGCAGTGCAA
 I P G G E W D A V V D F C A Y T P A H V

GAAACGCTGTCCGGAATTCGCGCGGTGGGTTCCGCAAGTTTCTGCTGATCAGCAGCAGC 13560
 CTTTTCGACGAGCCCTTAGACGCGCCAGCCAAAGCGTCAAAAGACGACTAGTCTGCTGCT
 E T L L G N L R G R Q F L R L A I S T T C

ACCGTCTACCGGAACCCAAAGCGCGGGCCGCTCGACGAGAACGCCCGCTGCTCGACGGT 13620
 TGGCAGATGGCCTTGGGTTGCGCGCCGGGAGCTCTTGGCGGGGACGAGCTGCCTA
 T V Y R N P S G G P L D E N A P L L D G

CCGACGCGAACTGGGCGAATACGCGGTTACGCGTATGAGAAATGCTCCGCGAAGAC 13680
 GCGTCCGCTTACCGCGCTTATGCGGCGAATGCCGATACTCTTTACGGAGCGGCTTCTG
 P Q P E L G E Y A G Y G Y E K C L A E D

GCGCGCGCGCGAATGCGAGGACTGGGGATCGCCCTTACTGTGCTGAGCGCCGCGATC 13740
 CGCCGGCCCGCTTACGCTCGCTGACCCCTAGCCGAAATGACACGACTCCGCGCGCTAG
 A A R R E C E R L G I G L T V L R P A I

ATCTATGGCTATTACAACACGCGCGGAGAGCTATTTTTTTGACCGCTCGCTAAC 13800
 TAGATACGATAATGTTGATGCTGGCGGCTCTGCATAAAAAAAGTGGCGAGCCATTG
 I Y G Y Y N Y A P R E T Y F F D R L R L

CGCGAACCATCGTGATACCGACCCACGCGTCTGCTGTTCAATTTCACTGCGGTCGTG 13860
 CGCTTGGGTAGCACTATGGCTGGGGTGGCGCAGCAGCAAGTAAAGTAGACCCAGCAG
 R E P I V I P D P T R S S F N F I W V

GATATGGCGCGCTCTGTGGCGCTGCATCGGAGAGCCGGCGTGGTGGGGGAGACCTTC 13920
 CTATACCGCGGAGGACACCGCGCTTCCGCTTCTCGGGCCGACCCGCTTCTGGAAG
 D M A R L L W R C I G E P G V V G E T F

AACTCGCTCGGGGAGGCGGTGACCTACTCGCGCATCGTGGAGCGCTCGCGGGGATC 13980
 TTAGACGGAGCCCGCTCCCGCACTGGATGAGCGGTAGCAGCTCCGCGAGCCGCTTAG
 ← 3T3.11

N L A S G E A V T Y S R I V E A L G G I

GTCGCAAGCGCGTGGAAACCTTGCCTGCGGTCGCGGTAACACGCGCGCAACATT 14040
 CAGCGCTTGGCGCAGCTTGGAAACGCGCAGCGCCCTTATGTCGCGGGCTTGTAA
 V G K P V E T L P L P V G E I T R R N I

CCGCTGCCCTTCCCGCTCGACGCGCGCTTCTTTACGCGCGCGAAGATCGACCGTCTT 14100
 GCGCAGCGGAAGGGCAGCTGCCCGCGAAGAAATGTCGCGCGCTTCTAGCTGGCAGAA
 P L P F P L D G P L L Y S G A K I D R L

Scal

TTCCGGTTCGAGCACACGCCCTTCCACGTCGCGCTGCTGAGCGCTGAAGTACTACCTG 14160
 AAGCCAAAGCTCGTGTGCGGGAAGGTGACCGCGGACGACTCCGCGACTTCATGATGGAC
 F G F E H T P F H V G L R E A L K Y Y L

ATGACGAAACGACGCGAGGCGCGAGTTGGCGTGGCGCCATGAATGCCGACTGATCCTTT 14220
 TACTGCTTTTCCGCTCCGCGCTCAACCCGACCCCGGTACTTACGGCTGGACTAGGAAA
 M T K R R E A A V G V A P *
 bssL → M N A D L I L C

GCAATCTCGTTCGCTTACCGCCGAGCGCCGACTACGCTGCGGACCTCGTTCGCTATCC 14280
 CGTTAGACCAACGAAAGTGGCGCTCCGCGTGTAGTCCACCGCTTGAGCAGCGGTAGG
 N L V A F T P G A P T H G A D L V A I R

GCAATGGTCTGATTCGACGTCGGAAACCGCTCGCGCCGCGCGTCCCTCACAGGCCCG 14340
 CGTTACAGCATAGGACGCTGCACCTTTTGGCGAGCCGCGCGCAGGAGTGTTCGCGGGC
 N G R I L H V G N R S A R A S L T G P H

ACACGAGGTCGTGACTGCCAGGGGCAACCTGATTCCGGGTTTCAACGACGCGCACT 14400
 TGTGCGTCCAGCACCTGACGGTCCCGCTTGGGACTAAGGCCCAAAGTTGCTGCGCGTGA
 T Q V V D C Q G A T L I P G F N D A H C

GCCACCGCGTCCGCTTTCGATGACAAACCGCCATGTGATTCGCGGCGCGCGCGCTCG 14460
 CGGTGGCCAGCGAAACGCTACTGTTGGCGGTACAGCTAACCGCCGCGCGCGCGCAGC
 H P V A F A M T T R H V D C G P P R V A

CGAGTATCGCGATCTCGTCGACGCCCTGCGCGCCGTCGCGTGGAGACGCGCCCGGCA 14520
 GCTCATAGCGGTAGAGCAGCTGCGGGACGCGCGGACCGCAGCTCTGCGGGGGCGCT
 S I A D L V D A L R A R A V E T P P G N

ACTGGCTCGCGCGCGAATTCGACCCCGGACTCTTGGCGAGCGAGCGCTGCCGAAC 14580
 TGACCGACGCGCGCTTACGCTGGGGCTGAGAACGGCTCGCTGCCGACGGCTTG
 ← 3T3.10

W L R A A N C D P A T L A E R R L P N R

GCTGGAGCTTGACCTCGCGTCCCGCCCATCCGGTATGCTGTCGAGCGCAGCGGGC 14640
 CGACCTCGAATGAGCGCAGCGGGCGTATGGCCACTAGCAGCAGCTCGCTCGCCGQ
 W E L D L A S P A H P V I V E R S G Q

AACATTGCGTCTCAACAGCGCGCTCGCGTATGCGGATCGATGAGTCCGCGCGCT 14700
 TTGTAACGACAGTGTGTCGCGCGGAGCGGATACCGCTAGCTACTCAGCGCGCGA
 H C V L N S R A L A L C G I D E S A R S

CGTCCGACGCGCGCAATTCACCGGACTCCGCGAGCGCCACGCTCAACCGCATCGTAT 14760
 GCAGCGTCCGCGCGTAAAGTGGCGTGGAGCGCTCGCGCGTGCAGTTGCGGTAGCATA
 S D A G A I H R D S A S G H V N G I V S

CCGCAACCACCGCCAGTCCGCGCGCGCTGCCCGCTCGCGCCGACGAGGTCGAGG 14820
 GGCGTTGGTGGCGGTTCCAGCGCGCGCGGCGCGCGCGCGCGGCTGCTCCAGCTCC
 G N H A Q V A A A L P P L A P D E V E A

CTGGATTGCGCGTCCGCGCGCTTCTCGAATCGGGGATCACCTCCCTGCAGGACA 14880
 GACCTAACCGCGCGCGCTGGCGCAGAAGGAGCTTAGCCCTAGTTSGLQDGT
 G L R A A D R V F L E S G S L Q D T

CGAGCTGTCGAACACGCTGCACATTGGCACGCGATGGCGACTACAAGCGCGCGGCC 14940
 GCTCGACCACTTGTGCGGACGTAAACCGTGCCTACCGGCTGATGTTGCGCGCGCGC
 S W S N T P A H W H A M A D Y K R R G L

TTGTACACCGCGCTATCCTTGTGCGTGGCGCGCAGCAGATCGAAGCGTTCGCGAGC 15000
 AACAGTGTGGCGCATAGGAACACGCGCGCGCTGCTGATGCTTCCGACGCGGCTCG
 V T P R V S L C V G A D S I E A F A E L

TGGGTTGCGCACCGGAAGCCCATCGCGACTGCGCGCGGAAGCGTTCGCGCATCGCG 15060
 ACCCCAAAGCTGGCTTCCGGGGTAGCGCTGACCGCGCGCTTTCGCGAGCGTAGCGCG
 G L R T G S P H R D C A P E A L R I G A

CGCGAAGATCGCGCTCGACGAGAGCAGCGGCGATCCGGATCCCGCGCAGGCGTCTCG 15120
 GCCGCTCTAGCGCGAGCTGCTCTGTCGCGCTAGGCGTAGGCGGCTCCGCGACGAGC
 A K I A L D E S T G D P D P P Q A L L D

ACGCCACCGCGCTCGCGCGCACTGTGCGGGTTCAGCTCGCATCCACGTCGCGGACC 15180
 TGGGTGGCGGAGCGCGCTGACACGCGCCAAAGTGCAGCGTAAAGTGCACGCGCTGG
 ← 3T3.9

A T A L A A H C A G F Q L A F H V P D L

TCGTGCTGCTCGACCGCTCGCTCAACACGCTCGCATGTTGGACGAGCGCGCTCACAC 15240
 AGCACGACGAGCTGGCGAGCGAGTGTGCGGAGCTACCAACCTGCTCCGCGCGAGTGTG
 V L L D R S L N T L A W L D E A G L T P

CGTGCCTGCGCGCGCTTCCGAGCACTGCCCGCTGCCCGCGAATGGCTGGACCGCG 15300
 GCAGCGACCGCGCGAAGCTGTCGACCGCCGCGGCGCGCGCTTACCAGCTGGCG
 S L R P R F E H C P A C P R E W L D R V

TGCGCGGAGCGCGCATCGTCTGCGCCACCGCGCTGCTGCGCTGGCGCGCGCGG 15360
 AGCGCGCTCGCGCGCTAGCAGCAGCGGGTGGCGCGCAGCAGCGCGGACCGCGCGCGC
 A S G A I V V A Q P A L L P L A A P A

CGTGCATGCTCTCCGCGGACCGCGCAGCCCGCTTCCCGTTCGCTCCTTCC 15420
 GCAGCTACGAGAGCGCGCTGGCGCGCTGGGGCGGAGAAGGCAAGCGAGGAAGG
 L H A S P P D P A Q P P L F P F R S F L

TGCGCCACCGCATCCCGCTCGCGTTCGGCTCCGATGCGCGCTCGCTCCTGTGCGCGC 15480
 ACGCGTCCGCTAGGGGAGCGAAGCCGAGGCTACCGCGCGAGCGCAGGACACCGCGCG
 R H G I P L A F G S D A P L A S C A P L

TCGAAGCGCTGAAAGCGCGCTCACACGCGCGCAGATGACGCGCGCGGTTGCCCGCG 15540
 AGCTTCGCGACTTTCGGCGCGAGTGTGCGCGCTGCTACGCTCCGCGCGCAACCGCGCG
 E G L K A A V T R R D D A G R A V A A E

AAGAAGCGATTGCGCTCGCGGAGCGCTGCTCTACACCGCGCGCGCGCTGCGCGA 15600
 TTCTTCGCTAACCGGAGCGCTCCGCGACCGAGATGGGCGCGCGCGCGCGCGCGCT
 E A I A L A E A L R L Y T R G G A C A S

GCGCGATGAGGACGCGACCGGCGATCGCAGCGCGCGCGCGCGCGCTGCTGCTGA 15660
 CGCGCGCTTCTGCGCTGCGCTGCTAGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCT
 G D E D A T G T I A A G R R A D L V L I

TCGAAGCGCGAGCGATTGGGCTGGAGCGCGCGCACTGGCGCGGCTCGCTGCGGATGA 15720
 AGCTTCCGCGCTCCGCTAACCGGACTCCGCGCGCTGAGACCGCGCGCGCGCGCGCTACT
 ← 3T3.8

E G A G D W A G A A D L A R A R V A M T

CCCTGATCGCGGAGCGGTGGTGTGGCGCGCTAGCGCGCAAGCGCATCCGCTGCGCAC 15780
 GGGACTAGCGCTGCCACCAACCGCGGATCGCGCGCTTCCGCTAGGCGACGCGGTG
 L I G G R V V W A R *
 gto1SR3' → NotI HindIII pETol3'
 XbaI GGGGCCATACGGAAACCGCGCGCGGCAAGCGCGCTTACCGCGCTACTGCCCGTTCGCCC 15840
 CCGCGTATGCCCTTGGCGCGCGCGTCCGCGCGCAATGCCCGGATGACGGGCAAGCGGG
 * Q G N A

pETol3' →
 TGAGCGCGCGCGCGCGCTTACCCATCCGTCAAAGTCCGCCACCGGCATCGCGCGCGC 15900
 ACTCGCGGTGCGCGCGGAACTGGTAGGCACTTTCAGCGGGTGGCGCGCGCGCGCGCG
 R L P W P A A K V W G D F D G V P M P P

CGACGTAATAGCTTGCCTCGTGCATCCGAGGCGATCAGCGCTTCGAAGATTCC 15960
 GCTGCATTTCGGAACGCGGACGCGTAGGCTCCGCTAGTCCGCAAGCTTCAAAGCT
 G V Y Y G Q A E D C T A G L R E F I E

GGTTCGACGCTTCCGCAATCACCGGATACCGAGCTGTGGCGAGCTCTACTGCTGG 16020
 CCAAGAGTCCGGAAGCGGTAAGTGGCGCTAGGCTCCGACACCGCGCTCGAGATGACCC
 R N E V G E A I V R I G L S H G L E I T

AATGACGATCACTCCGCGATCGCGCTTTCGAGCATCGGAAGCAGCAAGACTTGTGCA 16080
 TTACGCTAGTGGCGGTAGCGCGCAAGCGTGTAGCTGAGCTTCTGCTGCTTGCAGACCT
 S H V I V G A D R D Q L M P L V F S K D

TCTTGATCGCATCGATGGGCACTTTCTGCACATACGCGAGCGACACCAGCCGGTGC0GA 16140
 AGAACTAGCCGTAGCTACCCCTCAAAGACGCTGTATCGGCTCGCTGTGGTGGCCACCGGCT
 I K I A D I P L K Q V Y A L S S W G T G

AATCGTCGATGACGATGCGAAAGCCCGGTCGCTCAGGCGACCGAGCGCTGTTTGGGCGAG 16200
 TTAGCAGCTACTGCTACGCTTTCCGGCCAGCGAGTCCGCTGGCTCGCAGCAAAACCCGCT
 F D D I V I R F G R D S G R G L T T Q A

CATGGCGATCCACCATGATCGAGCTCTCCGTGATCTCGAGTTCGATCCAGCGCGCTCCG 16260
 GTACGCCCTAGGTGGTACTAGCTCGAGAGGCATAGAGCTCAAGCTAGGTGCGCGGAGGC
 A H P D V M I S S E T I E L E I W R A E

CGCCCCAGGTGATCAGCGGCTTCTCAGGATATTCTGTTTCCAGCGCGGATCGCCAGATTGC 16320
 GCGGGTCCACTAGTCCGCAAGAGCTCCATAAGCAAGTCCGCGCTAGCCGCTTAACG
 A G W T I L A N Q L Y E N L R P D A L N

tolrev3' →

GCGCCAGCAGGTTGATCGCGATCGGACCGGCAACGCTTCTGCTGCCAGCGTGGCAGG 16380
 CGCGCTGCTCAACTAGCGTACCGTCCGCTTTCGGAAGACGACGCTCGCCACCGTCC
 R A S L N I A I P V R L R K Q Q W R H C

CCGCCAGAGCGCGCTCTGCGACCCAGCGCTGAGCGCGCGATGAGTCCCGTCTCTCGA 16440
 GCGGCTCTCGCCGACAGCGCTGGTTCGCGCACTCCGCGCTACTCAGGCGACGAGGCT
 A A L A A D A V W R T L P A I L G T S E

TGAGCGGACGAAACCATCGGGAGAAATGAGGCGGTAGAAGCGATGCTGCCAGCGCTCA 16500
 ACATGCCCTGCTTTGGTAGCCCTTTTACTCCGCTTTCGCTACGACGCTCGCGCAGT
 I L P V F G D P S I L G Y F P H Q W R T

GGGCTTCGAGCGGACGATTTCCGCGGTTATTGAGATCGAGCTTGGGCTGGCAGTAGAGCA 16560
 CCGAAGCTCGCGCTGCTAAAGCCCAATACTTAGCTCGAACCAGCCGCTCATCTCGT
 L A E L A V I E R T N L D L K P Q C Y L

3T3.7 ←

CCAGTTCGCGCTTGCAGTCCCGCGCGCAATCCCCCATCAGCTGCAAGCGCGCGGAT 16620
 GGTCAGCGGAAACGCTAGGGCGCGGTTAAGGGGTAGTTCAGCTTCGCGCGGCTA
 V L E R R A I G A R L E G M L Q L R R P

TGTACGGATCTGCTCGCGCTGTAGAGCGAGTGGACTTTCGCTCTGCTGCGCGGCT 16680
 ACATGCTAGGACGAGCGCGACATCTGCTCACCTGAAAGCGAGGACGACGCGCGGCA
 N Y P D Q E A S Y L S H S K G D Q Q A R

tolSR2244.3' →

AGAGCGGACTTCCCGTGGCGCATCAGGACTCCCGCTTCGTCGCCATGACCGGAAAGA 16740
 TCTCGCGTGAAGGCGACCGGCTAGTCTGAGGGCGAAGCAGCGGTACTGGCCCTTTCT
 Y L A V E A H R M L V G A E D G H G P F

tolSR2244.3' →

TCGGGAGCGGACGTTGGCGCGATCTCAATCGGAAATCGGCAAGCTCGAACGCTTCTCT 16800
 AGCGCTCCCGTCAACCGCGCTAGAGTTACGGCTTTAGCGCTTCGAGCTTGGCAAGGA
 I A L G V N A G I E I G F D A L E F A E

CGAAGCGGCGGACGAGCGCTCGGCCAGTTCGAGGCGCGCTGGGCGCGAGACCGGGAC 16860
 GTTCCCGCTGCTCCGCGAGCCGTTAAGGCTCCGCGGACCCGCGGCTCTGCGCCCTG
 E F R A V L R E A L E S A A Q A G L G P

AGAGCACGCGAATTCGCGTCCCGCAATAGCGAGCGGCTGCTCCGCGTCAAGTACGG 16920
 CTCGTGCGCCCTTAAGCGCAGCGGCTTATCCGCTCGCGCAGCAGGCCACCGTATGCC
 C L V A F E R D G C Y A D L P H L D L N L A

ACCGGATCCGCGGCGATATCTGTAGCAGCGGTCGCGTCTGCGCCACGCTGT 16980
 TGGCTAGGGCCCGGCTATAGACATCTCGCGCAGCGGCGCAGACCGGCTCGCACA
 S R I T P G I D Q L R D G H T Q G L T

AGTGTATCTCTTGAACGCTCCAGCGAAACGATCAGCAAGCGAATGGCCGCGCGCG 17040
 TCCACTAGAGGAACTTTCGAGGTCGCTTTGCTAGTCTGCTTCCGCTTACCAGCGCGCGC
 Y T I E K F R E L S V I L L A F P R G A

CACGCGCGGCGGACCGCTGCTGACGATCTCCGTGAGCTCGCGCGATGGGCGAGT 17100
 GTGCGCGCCCGCTGGCGGACGAGCTGTAGAGGCACTCGACGCGGCTAACCCTGCTA
 A R A A A V A Q Q L M E T L Q A R N P L

CGTGGTGGCTCGAAGAACGCGAGGTGCCGACCCCTTCTCGCGACCTTGCAGGCT 17160
 GCCACCACCGAGCTTCTGCGCTCCAGGCTGGGAAAGGAGCGCTGGAACGCTCGGA
 D T T A D F F A L H R V E A V K R A

TGAGGCTGCGTGCAGCGACCGCTTGTTCGAAATGAGTCCGCGGCAAGCGCAGAGGT 17220
 ACTCCGACGCGCTGCTGGCGAACAAGCTTTACTCGACGCGCCGCTCCGCTGCTCCA
 K L S R Q L S R K N S I L Q R A L A V L

pETeal5' ←

pIZeal5' ←

HindIII ←

CCACGTTCTGGTTCGCGAGCTTCTGCTGCAAGGCTCGCGCGCTGCTTCACTCGTACCA 17280
 GGTGCAAGCCAGCGCTCGAGAACGCTTCCGAGCGCGGACGAAAGTACGAGTGGT
 D V N Q T A L E Q Q L A D R A Q K M T V

GGTTCGCGCGGCGCGCTGCTGCGCGGACGAGCGCTTGGTGCATAGTCTCGTCA 17340
 CCAAGCGTGGCGCGCTCGACGACCGGCTTTCGGAACCACTGTATCAGGAGCT
 L N R V R A R L E H P S F P K T V Y D Q

CCGATTTCGGCCAGCAGTTTCAAGCGCAGCTCGGCATCGGCGCTCGCCGAAAGCATCAGCA 17400
 GGCTAAGCCCGCTGCAAAAGTTCGCGTTCGAGCCGTAGCCGCGAGCGGCTTTCGTAGTCTGT
 3T3.6

V S E A L L K L R L E A D A S A S L M L

CCGGCAGCTGGGCGAGCGCCGGGGGGCGCAGTTTCGCGCACAGTTGATCGCCGTCGA 17460
 GGGCTCGACCCGCTCGCGGCGCCCGCGCGCTCAAGCGGTTGTAACCTAGCGCGAGCT
 V P L Q A L A P R A R L E R V L Q D G D

CGGTGCGCATCACCAGATCCGTGACGACGAGTCCGGCGCTCGCGAGCGCCAGCGCGA 17520
 CGCAGCGTGTGGTCTAGGCACTGCTGCTCAGCGCCGCGAGCGCTCGCGGTGCGGCT
 L T P M V L D T V V L D P P E A L A L A

GGGCGCGCCCGCATCGCGAGCGCGCGCGATGATGCGGCGCGAGCACCTCGACGA 17580
 CCGCGCGCGGGTAGCCGTCGCGCGCTCGCTATCAGCCCGGCTCGTGGAGCTGCT
 L A A A G D A A P A V R Y D A G L V E V

TGTATTGCGCATGTCGGCATTGTCTCGCGACGATGACGAGCGCAGGTCGTCGCGCG 17640
 ACATAACGGCTACAGCCGTAACAGGAGCCGCTGCTACTGCTCGCGTCCAGCAGCGCGC
 pETrec5' ←

I Y Q R M D A N D E A V I V L P L D D A

CGGGCCGATCGCGCTCAACCGCTTCCAGCGCTGAAAGGCTTCCAGCGCCCGATGACCG 17700
 GCGCGCTAGCGGAGTGGCGGAAGTGGCGAGCTTCCGAAAGTCCGCGGCTACTGGC
 pETrec5' Ndel

A P R D R E V A K V A D F P K V A G I V

CATCGATGTCGCCAACTCGTCCGCGCGAGTCCGGCTTCGCGCCCGGGCGCAGCCCGG 17760
 GTAGCTACAGGGCGTTGAGCAGCGCGCTCCAGCGCAAGCGCGCGCCCGCGCTGGCGG
 A D I D R L E D A A P G A E R G P A V A

ATCGCGCGCAGTTGCACCTCGAACAGCGCACCGCTTCGGGCGCTCCGATACCGCCA 17820
 TAGCGCGCGCTCAACGTTGAGCTTGTGCGTGGCGAAGCCCGCGCAGGCTATGGCGG
 S R P L Q V E F L A G E R V A D S V A

CGGTCCCGCGTCAAGTTCGCGCAATCTCTTCCAGTTCGCGAGGCGCAGTCCGGTTCGCC 17880
 GGCAGGGGCGCAGTCAAGCGCTTAAGGAAGTCTAGCGCTCCGGCTCGAGCGCAAGGGG
 V T G G H L E A F E K V I A L G L G T G

CGAATCCGCGATTTCGCGCCCGCGCCCTGTCGCAATCGATCGAAGTACCGGGCGCA 17940
 GCTTAGCGCGTAAAGCGCGGGCGGGGACACGCTTAGCTAGCTTCTACTGCGCCGCGT
 G F G R M E G G R G Q A F R D F I V P R

TGTCCGGCGCACGCGCGCGCTATCGCTCACCGGAGCAGCAGCGGCTGTCGCCCA 18000
 ACAGCGCGCGTCCGCGCGGCGATAGCGAGTGCCTGCTGTCGCGCAGCAGCGGGGT
 M D P A V G P G S D S V R L L V R D A G

GGGCTTCGAGCGCACATCGGACTCGCCCGCGATCGCGCTGAATTTGAAGCGCTTGGACA 18060
 CCGAAGCTCGCGTGTAGCTGAGCGGGCGCTAGCCGACTTAAACTTCGCGCAACCTGT
 3T3.5

L A E L A C R V R G G I P T F K F A N S

XbaI →

TutC3' →

ACAGTTTCAAGCAGTTCGCGTCACTTTTCGGAATCCGCTCGCGGCAAGCGTACGCG 18120
 TGTCAAAGTCTGCTACGCCAGTAGAAAGCCTTAGCGGACACCGCTGTTTCGCATCGCC
 L L L V I R T M K E S D A A C V L T L A

CGTTCGCGCACTCGTAAGCAGCGAAGCTGCGGCGGAGTGTTCGAAATGGGCGCAA 18180
 CGCAGCGGTAAGCATTGCTGCTTTCGCGCAGCGCGCTCAGCAAGCTTTACCGCGGCT
 P T A V E Y A V S R Q P A L A E F H A A

CGTTCGCGACAGATGTGCAAGTTCGACCTGCGAGCAGTTCGCGCGCGCTGCGCGCGT 18240
 GGCAGCGCTGCTACAGCTTTCGCTGAGCAGTGTCAAGCGGGCTCGACCGCGCGCA
 V T R V L H A L D V Q L C N A R L Q G A

CGAGCTTTCGAGGTCGAGCAACTCATTGACCTGCTGAGCAGCATCGCCGCTGTCGAT 18300
 GCTCGAAACGCTCCAGCTCGTTGAGTAACTGAGCAAGCTGCTGTCGAGCGCGCAACGCTA
 D L K A L D L L E N V Q K L L M A A N R

GGATGACTCCGAGCTCGCGCGCTGCAAGTTCGCTGAGATTTCCCGTTCGCGCAACATCG 18360
 CCTACTGAGGCTGAGCGCGCGAGCTTACGCACTTAACAGGGGACCGGCTTGTAGC
 H I V G V D R R R Q L D T L N A G H A L M

ACTCGCGAGCGCCATGATCAGCGCGCGCGCTGCGCAACTCGTGGCTCAGGTTGGCGA 18420
 TGAGCGCTCGGGTACTAGTTCGCGTCCGCGCAGCGCTTGGCAGCGAGTCCACCGCT
 S E A P G M I L A L P T R L E H S L N A

AGAAACCGTTCGAGCGCTGAGCTGACGCACTTGTCAATGAGCTGCTGCAATTCG 18480
 TCTTTGGCAAGAAGTTCGCGCAGCTCGACTCGTGAACAGTACTCGACGAGCTTAAGGC
 F F G N K L A D L Q R V K D I L Q Q L E

CGTTCGTCGCGGAGCTGATCTGCGTCTTCTTCGCTGATGTTGCGGCTTCCG 18540
 CCAAGAACACCGCTCGAGCTAGAGACGACGAAAGAGCCACTACAACCGCGAAGGC
 A N K H A L E I E A H K K E T I N R G E

GCAGCAGAAAGACAGCCCGCCCTGCTGTCGCGCAGCGCTGACGAGAAAGTTCGATGA 18600
 CGTCTGCTTCTGCTGGCGGGGACGAGCAGGCGCTGCGCCAGCTGCTTTCAGTACT
 3T3.4 ←

P L L F V V R G Q E D R V P Q L S F D I

CGACCCGCTCCTTGCCTCCGCGGCGCCCGCTACACTTCGAGATCGCATCGGACGAATTCCG 18660
 GCTGGCGAGGAACGGCGCCCGGGGCATGTGAAGCTCTAGCGTAGCCTGCTTAAGCC

3T3.4
 V V R E K G A A A G Y V E L D C R V F E

CTGTCGCGCCCGTCGCACGAGATCATGCGCGACCTCGCGCAACTCTTCCGACACGGGCC 18720
 CAGACGGCGGGCAGCGTCTTAGTAGCTAGCGCGGAGCGCGTTGAGAAGCTGTGCGCGG
 G T R A R R V L D H A V E R L E E S V A

ACCAGCGCGTCCAGAAAGGCTTGCCTCGATGTCATCGAGCTGCACGCGGCGCCGCT 18780
 TGTGCGCGAAGGCTCTTCCGAAACGGGCATACAGTAGCTCGACGTGCGGCGGGGCA
 W W R A E W F P K G R I D D L Q V G A G

CGAGCGCCCGCGGGAGATATCGACGATCCGGCCTTCCGGTTCGAGCAGCCCGCAACT 18840
 GCTCGCGGCGCCCTCTAGTAGCTAGCGCGAAGCCAGCTCGTCCGGGCTCTTGA
 D L A R R S I D V I R G E P D L L G V F

toSRint3' →

CGAATCGCGTCCGAGGATGATCGGTGCGAGTTTCTGCCAAATTCGCGATCGCTCGTCGG 18900
 GCTTGTACCGCAGCTCTACTACGACGGTCAAAGACGGCTTACAGCGTAGCAGCAGCC
 E F M G D L I I R A L K Q R F T A D D D

ACTCGTGTGAGCAGCAATCGCAACGCAACCGTTGACGCGCTCGGAGAATAACTGACGAT 18960
 TGAGCAACTCTGGCTTAGCGCTTGCCTTGGCAACTCGGAGCCTCTTATTGACTGCTA

← gTolSR5'

S E N L V S D R V R V T S A E S F L Q R

CGAACGTGCAGATGCGTTACCTTCTACGCGCAACCTTGCCTCGGCTTCCCGGGTTCAC 19020
 GCTTGCAGCTGTACCAATGGAAGATGCGCGCTTGGGACGGCGCGGCGGCGCCAGTG

gTolSR5' HindIII
 D F T S M ← tol

GTTACGAGGGATCCGATCGCACTGTCACCTTCCCGGCCACAATGTGCTGCGAGGCGCC 19080
 CAATGCTCCCTAGCTATCGCTCAACGTGAAGGCGCGGTGTACACGACGCTCCGCGG

GTTCCGAGGTAATGAACGAGGCGTGGTGGTCGAGCAGCCATATCGTTTGTTCGACAATT 19140
 CAAGCTCCATTACTTGTCTCCGACCCAGCCTGCTGGTATAGCAAACAAGCTGTAA

GCTCGCGCGATTGATTCGCGGTTCCGCATATGCGCCTCGCGCGCCAAAGGCACTCCC 19200
 CGAGCCGGTAACAAAGCGCCAAAGCGTATGACGCGGAGCGCGCGGTTCCCGTGAGGG

TGGCTCAGCGCGCTGCGCGCGGATTAAGACTGTAGCGTGCCTGAGGACGCGCTGGC 19260
 ACCGAGTCCGCGCGCACCGCGCCCTAATTCGACATCGCAACCGCACTCCCTCGCGGACCG
 * G A S R R P N L S Y R A T L S A Q

CGAGGACATGCCCTTCGATCGTCCGAGCAGCTCCGCGCGGTAACCGCCCTCCAGCG 19320
 GCTCTGTACGGGAAGCTAGCAGGCTGTGACGCGCGCCACTTGGCGGGGAGGTGCG
 G L V H G E I T R L V D P G T F R G E L

GACAACGATCGAAATCGACCGCTAGAGCGTGAAGACGTAGCGGTGCACGCGCTCGTCT 19380
 CTGTTGCTAGCTTGTAGCTGCGCATCTCGACTTCTGCATCGCCACGTGCGCGGAGCAGCA
 P C R D F D V A Y L T F V Y R H V R E D

TCCACGGCGCGCACGGACCGTCTAGCGCAAGTACAGCCCGCATGTGCTCGTCCGCGG 19440
 AGGTGCCCGCTGCTGCGACGATCGGCTCATGTGCGGGCGGTACACGAGCAGCGGCC

← 3T3.3

N W P P C P G D Y G F Y L G A M H E D G

CAAAGCCGCGTGAAGTCTGTAGCCCGTGCCTGCGCGCTGCCGCGCGCGGGCCCCG 19500
 GTTCCGCGCGCACTCAGCAACTCGGCGACGGCACGCGCACGCGCGCGCGCGCGGGC
 A F G R T F D N L G H R A G H R A A P G

GCTTCCACGCGCATGACGCGCTCGGAGAATTTCCCTCGCGCACGCGCGGTCACGCGG 19560
 CGAAAGTTCGCGCTACTGCGGCGCTCTTAAAGGAGCGCGCTGCGGCGCATGCGCC
 P K G R P M V G D S F K G E A V S G R Y A

CAGGCAGATCGACGACCCAGTGAAGAAGTGCACGCGCGCATCGACTCGGGAATCA 19620
 GTCCGTAGCTGCTGCTGGGTACCTTCTAGCTGCGCGCTAGCTGAGCCCTTAGT
 A P L D V L V W H F F D V R P M S E P I

CGATATCGTCCGCGTGTATGCTCGAAACGGCTCGGCACATCCGGATCGTGGCACAGCA 19680
 GCTATAGCAGCGCAACTAGACGAGCTTTCGCGAGCCGCTGAGGCTAGCACCGTGTCTG
 V I D D R N I Q E F R S P V D P D H C L

GCACGAAGAACGCGTGCCTGCGCGCGGCTCGTTCCAGTGAAGCGCGGCTCGGTTGG 19740
 CTGCTTCCCTTTCGCAACGCGCGCGCTGCGGCTTCCAGTGAAGCGCGGCTCGGTTGG
 L V F S R T G A P P E N W T F A P S R N

← bbslint3'

CGCCGAACGGAACGCGCTTCCGCGTCCACACCGCAAGGCGTTCGCGTCCGCGAGCG 19800
 GCGCTTCCGCTTTCGCGCAAGCGCGCGCTGCGGCTTCCGCAACGCGAGCGCTCCG
 A G F R F R R E A D V V A F A N R D P L

GCCGTTCTCGCGAACTGCTGTCAGACGATGGTGTCTCTGCAACGCGAAGCGCG 19860
 CGCCCAACGCGGCTTTCGACACGATCTCGTACACACAGAGGACGTGCGCTTCCGCG
 R R N D G F S S T L R M ← bbsI

← bbsH3'

CCTCGCCGCGCGCTCCGCCATTAGGTGATGCTGTGAGCGCGCGGAATTCGGGCGC 19920
 GGAGCGCGCGCGGAGGCGTAAGTCCACTACGACAGTCCGCGCGCTTAAAGCCGCG
 * A G R F E P A

CGCCCGTTCGCCAAGGCGCTTGAACCTTCTTTCGCGTCTCGGTATCGCGCAGCACGCC 19980
 CGCGCAAGCGGCTTCCGGAACCTTGGAAAGAACCGCAGGACCATAGCGCGTCTGTCGG
 R R E G F A K F G E K R D E T D R L V G

CCACAGCAGGTGCGAGTACAGGATCCCATGCTCGAGCGGCATATCCAGGCCAGCACGCC 20040
 GGTGTCTGCCAGCTCATGTCTTAGGCTCAGAGCTCGCCGATAGGTCGGGTCGTGGV
 W L L H S Y L I G H E L P M D L G L V A

GGTGTCTGGTGCCTGCACGCTCAGCGCGCGCGCTGGCGATCTTCTCCGCGTACGT 20100
 CCGAACCGCAGCGGACGTGCGAGTTCGCGCGCGCGGACCGCTAGAAGGCGCATGCA
 A Q K T A Q V S L P A A S A I K E A Y T

TCCGCGCAGGTCCGCGAGCGCGCGGCTCGACCACTCGCTGACAAGCCCGTAGCGCAG 20160
 AGCGCGTCCAGGCGCTCGCGGCTGCCGAGCTGGTGCAGCGACTGTTCCGGCATCGCGTC

← 3T3.2

R A L D A L A S P E V V D S V L G Y R L

CGCTCTCCGAGAGATCATCTCGCCGTCGACGACATCTTATCGCCACCGCTGCGG 20220
 GCGGAGGAGCGCTCTTAGTAGAGCGGAGTCTGCTAGAACTAGCAGGCGCGGACGCC
 A E E A S I M E G T L L M K M A V A Q P

GATCGCGCGGCGCAGCGCTGCGTGCCTTGGAGCGCGGAGGCTCGCGACCTTCACTTC 20280
 CTAGCGCGCGCTCGCGGACGACGCGCACTCCGCGCGCTCCGAGCGCTGGAAGTGAAG
 I A R P L A Q T G N L G A L S A V K V E

bbsH333.3'

GGTACGCCCAACTTGGCAITTCGACGAGCGGATCCGCATGTCGACGCAAGCGCATCTC 20340
 CCAGTCGGGCTTGAACCGTAAAGTCTCGCTAGGCGTACAGCGTGCCTTCCCGGTAGAG
 T L G F K A N S S A I R M D C A L A M E

bbsH333.3' →

GAGCCCGCACCCACCGTAGCCGTTGATCGCGCGATGATCGGCTTCCACATCTTCA 20400
 CTGCGCGTGGTGGCGCATCGCAACTAGCGCGCTACTAGCAGGAGGTGTAGAAATA
 L G G G V A Y G N I A A I I P K W M K M

bbsHint3' →

GTGCGCAGGATCGGCTGCCCTCGCGCAGATAGGTCGCTGCCATGCACCTCGTGGGCGG 20460
 CACGCCCTCTAGCCGACCGGAGCGCTCTATCCAGCGACGCTACGTGAGCCACCGCC

← bbsH5'

H P L I P Q G E R L Y T A A M C E T P P

CGTGCTCTTCTATGTCGCTGCTGTGCGAAGCGCTTTTCCGCGCAGCCGACACAC 20520
 GCGACGGAAGTACAGGACCGGACGCTTTCGCGAAGAGCGGTGCTGGCCTGTGTG
 P A K K M D T G T G F A K E G A G S L V

GGCCACGCGGATGTCGGGTCGTCACGACGCGCGCCAGATTTCGCGGAGATCGCAGGT 20580
 CCGGTGCGCTCAGCGCCCGCAGTGTGCGCGCGGCTTAAAGCCGCTTACGCTCCA
 A V R I D P D T R V R A W I E A L D C T

CGATTCCGGGTCGAGCGAGTTCATCGCTTCCGCGCGGTCAGGTCGAGTGGCGAGCTG 20640
 GCTAAGGCCAGCTCGCTCAAGTAGCGAAGCGCTCCAAAGTCCCACTGCATCCGCTGCAC
 S E P D L S N M A E P R N L T V Y A V H

ATTCTTACTTCCAGCGTTACCGGCTTGTCTTCTCTCTTCTTACTTCCGCGCGG 20700
 TAAGAAATGAAGTTCGCAATGGCGGTAACGACGACGAGGAGAGAAATGAGCCCGGCG
 N K V E L T V P M ← bbsH

bbsG3' →

CCTCAGCGCGCCATGAACGAGTGTAGTGCAGCGGCTCCGCGAGCAACTTCTTAGCGATC 20760
 GGAGTCCGCGGCTTGCCTGACATCCACGCGCGGAGCGGCTCGTTGAAGAATCGCTAG
 * A A M F P S Y T R G D A L L K K A I

GACATGCGATGAATTTACAGATGCGCGCTCGACGATCCGCGAGATCCGACCATCCGCGG 20820
 CTGTACGCTACTTAAAGTCTACGCGGAGCTGCTAGGCGGCTTAGGCTGGTAGGCGCTC

← bbsGint5'

S M R H I E S A G E V I R W I R V M R S

ACGTATTGATGCCAGCTCGCGGACAAATCCGAGGCCACCGTGCAACTGGATCGCGCGG 20880
 TGCAATAGCTACGGGTCGAGCGGCTGTTAGGCTCCGCTGGCACGTTGACCTAGCGCGCC

← bbsG.F5'

V Y E I G L E R S L G L G G H L Q I A R

TCCGCGACACGCGTGCAGTTCGCTCGCGCGGATCTTCCAGGACGACGCGCTCGATGCGC 20940
 AGCCGCTGCGCGCATCTGAAAGCCGCGCGCTAGAAGTCCCTGCTCGGAGCTACCGG

bbsG.F5' →

D A V R T L M E T A A I K L S S A E I R

AGATCTTGTGGCCCTGATCCGACTTCCAGCGGCTGATGATAGAGTGCAGCGGACCTGC 21000
 TCTAGGAACACCGGACTAGGCTGAAGTCCGCGCACATCATCTCGACGCTCGGCTGGAG
 L D K H G Q D T S K W A T Y Y L Q L R V Q

TCGAGTTCGACGCTGCAATCGCGGATCCAGTTTTCACCGTCTGACGCTCGCGGAGCGGC 21060
 AGCTCAAGCTGCGAGCTTAGCCGCTAGTCAAACGTCGACGAGCTCGAGCCGCTCCGCG
 E L E V T S D A I W N Q V T Q R E A L P

TCGCGGAAGGTTCTCCGAGGTTCCGCTGCTGATCATCATCTGGATCAGCGCTCGCGC 21120
 AGCGGCTTCCAGAGGCGTCCAAGCGGACGAGTGTAGTAGACCTAGTCCGCGAGCCGG
 E G F T K R L N A Q E I M M Q I L R E A

ATGCCGGTGCAGTGCGCCGCGAFTTCGATCCGCGCCACGCGGAAGCGGTTCTGCAGCGGA 21180
TACGCCACCTACCGCGCCCTCAAGCTAGCGCGGGTGCAGGCTTCGCCAAGACGTCGCCT
M G T C H A A L E I R R V G F R N Q L P

ATGAAGCATCGCCGACCTCGCCAACTCGCTCGTCCGACCCGACGTTATCGAGA 21240
TACTTCCGTAGCGGTCGAGCGGTTGTAGCAGCAGCGGCTGGCGTGAATAGCTC
I F A D G V E G L V A D G V R V N D L

TAGACGGTCCAGTTCGCGCATCGCGCCGATCACCGAACTCGCGCCGATGCGCAGTCCC 21300
ATCTGCCAGTTCAGCGGTAGCGGCTTAGTGGCTTAGAGCCGCGGCTACGCGTACGGG
Y V T W T P M A G I V P I E A G I R L G

GGCGTATCCTTGTGAGGATGAAGCACTCAAGCGCGCTTGTGCGCCCTCCCTGTCA 21360
CCGCATAGGAACAGCTCTACTTCCGTGAGTTCGCGCGAACCAGCGCGGAGGACAGT
P T D K D L I F A T L R R K T A A E R D

GTCACCGCGATCGAGATGAAGAACAGATCGTCTTGTGCGCACTGCTGATGAAGATCTT 21420
CAGTGGCGTACGTCTACTTCTGTAGCAGGAACAGCGTGAACGACTACTTCTAGAAG
T V A I C I F F L D D K D C K S I F I K

GACCCGTTGATGACCCACTGCCCCTCAACCGCCACCGCGTGGTTGTGAGGCGCGTATG 21480
CTGGCAACTACTGGGTGACGGCGAGTTCGCGTGGCGGACCAACACTCCGCGACTAC
S G N I V W Q G D V R V A T T T L G S I

TCGGAGCCCGCGCCGTTGCGTCCCATTCGCGGATTCGATCTCCCGCGCGAGTAG 21540
AGCTCGGGCGCGGCCAAGCCAGCGGTAGTAACGCCCTAGCTAGAGGGCGCGCTCATC
D S G A K P E T A M M M A S E I E G R C Y

CGTCCAGGTATTTCTCACGCTGCGCGCGCGTCCGCACTCGTCCGAGTAAATAGAGTTA 21600
GCAGCGCTATAAAGAGTGCAGCGCGCGGAGCGCGTGAACGACTCCATTTCTCAAT
R R L Y K E R Q A P T G C E D L Y Y L N

GGGGCATCGGGCGAGCGTGAAGCCGTGATCGGAAACCGATCAACGACTTCGACATT 21660
CCCCGTAGCCCGCTGCGCACTTCGCGCACTAGCGTTCGCGTGGTGTGTAAGCGTAA
P A D P P L T F G H H S F G I L S K S M

TCCTCGAACAAGGGTCTTGGCGAATCGCGAGCCCTGTCCACCAAGTTTCTCGTCC 21720
AGGAGCTGTTGCTCCAGAACCGCTTGTACGGCTCGGGGACAGTGGTTCAAAGAGCAGC
E V V L T K A F M G L G Q G G L K E D

ACCTCGATGCCCAAGAACCACTCCCTGGTTCATGTCACCGAGCATGATCGCGG 21780
TGGAGCTACGGGCTTTCGGGTAGAGGACAGTACAAGTAGTGGCTCGTACTAGCCGC
V E I G W F G M E R T M N M L R A H D A

GCGCGCAGCAGCGTGTACCCGTCGGTCCACATCCGCATCTCCCGCTCCAGCAAGCCTT 21840
CGCCGCTCGTCGCACATGGGACCCAGGTGTAGGCGTAGAGGGCGAGTCTGTGGAAA

A P L L T Y G D T W M R M E R E L L V K
bbsG5'
bbsGint3'

TCCAGCGCATCAGCTCGTTGGCGGTGAACCGCGGGCCACTCTCTTGAGCATCATCTGT 21900
AGTTCGCGTAGTTCGAGCAACCGCCACTTGGCGCGCGGTGGAGGAACCTGTAGTAGACA
E L P M L E N A T F R R A V E K L M M Q

TCTTCGATAAGGAAAATCAATTTTTCGCTCTTGTCTGATTTATCGGGCCGATC 21960
AGAAGCTTATCTCTTTAGTAAAGAAAGCAGGACAGGACATCAATAGCCCCGGCTAG
E E S L S F D M ← bbsG

CGTCGAATGCTTCTTCACTCGCGCTCGGAGCGCATCAAGACTCGAATGTCGCGGAT 22020
GCAGCTTGACAGAAGAAGTACCGCGAGCTCGCGTGTCTGAGCTTGTACAGGCGGTA
* E A D S R M L S E F M D A I

CGACGTGTCGCGTGTGTCGCGCATTTCTCGGCGCGGCTTGACGCCCTCCGAAAAC 22080
GCTGCACGACGGCAACAACCGCTTAAAGACCGCTCGCCGAACTGCGCGAGGAGCTTTT
S T S G N N R C K E A A A K V G D E F V

AGCCATCAGCGCTCGCGGATCATCGCGAGTCTGACAGCGCGCACCATGAGGAGGC 22140
TCGGTAGTTCGCGGAGCGCTAGTAGCGGCTCAGCACTGTCGCGGCTGGTACTCCGTC
A M L A E R I M A S D H C G G V M L C A

CGCATAGGGTGGTATTCCGCCATGGCCACCGTATGCGCGCACTCGGGCACTCGTGTAT 22200
GCATATCCCAACCAATAAGCGGTTACCGGTGGCAGTACGCGGTGAGCCCCGTGAGCACTAC
A Y P H Y E G M A V T M G C E P C E H H

CGAGATACCGGCTGCGCGGAATTCGCCAGCTGCGTGTGTCGAGTTGTTGCTTGTACAT 22260
GCTCTATGGCCACCGCGCTTAAAGCGCTGCGACCGAAGCGTCAACAACGACATGTA
S I G A P A S N G V S A Q R L Q Q K Y M ← bbsJ

GATTCGCTCTCTCGGCTCGATCCGCGTCCGTCACGCACTGCTCGAGCTCGCCG 22320
CTAAGCAGGAGGAGACCGAGCTAGGCGCAGGAGTGCCTGACGGAGCTGCCACGGGC
CATCACTGGTCTCAACAGTCAACCAACCTCTGACTGACCAAGTCTCGATTTCTG 22380
GTAGTGACGAGCAGTTGCTAGTGGTTGTGGAGGACTGACTGGTCTAAGAGCTAAAGCA
* V L V E Q S V L N E I E

CCTCGGAATAGCCGAGCATCTCGCCGAGCACATCGCGGGTGTGCTGCCCGATCGCCGCG 22440
GGACGCTTATCGGCTCGTAGAGCGGCTCGTGTAGCGCCACACAGCGGGTACCGGCGCG
← bbsFint5'

D Q S Y G L M E G L V D R T H Q G I A P
CAGCGGTCTTCATTTCAGCGGGGTACCGGAGAACACGATCGGTGCGCGGTTGTAGAGCG 22500
GTCGCGAAGTAAGGTTCGCCCCATCGCTCTTGTGTAGCCACGCGCAACATCTCCG

bbsFint5' bbsF5'deg
A A T K M E L P T R S F V I P A R N Y L
NdeI

TGACCGCGACTCGGGGTGATCGAGATAAGCCCAAGACCGCGCGCTCGCATATGCTCAT 22560
ACTGCGGTGGAGCCCCACTAGCTCTATTCCGGTCTTCGGCGCCGACGGTATACAGGTA
T V Y G E P H D L Y A W F G R R R M H E

CCTCGATCGCTCGCGGCATCCCGCACCTCGCGTCCCGGGATCCCACTTCAGCAGCG 22620
GGAGCTAGCGGAGCCCGTAGGGGTGGAGCGGACCGCTAGGGTGGAAAGTGTGCGC
D E I A E R A D R V E G A P I G V K L L

CCTCCATCAACCAATCGCCGACTGCTGCGCAGTCCAGCATTTCATGCGCTCGTTCAGCT 22680
GGAGGTAGTTGGTGTAGCGGCATGACGACCGCTCAGTCTGAAGTACCGGAGCAAGTCA
A E M L W D G Y Q Q A T W C E I R E N L

CGACCTCATACCGCGTTCGATTCGCGGAGTTCGCGAAACCGCTCTCTCGGCCACCGCG 22740
GCTGGAGTATGGCCGACGTAACCGCTCAGCGCTTTCGAGGAGGAGCGCGGTGCCGC
E V E Y R R R M A A T A F R E E A W P

GATTGCCATCAGCGCGCGAGCGCGCCACTGAAGATCGTGAAGACCGCGATCGCGA 22800
CTAAGCGGTAGTCCGCGCGCTCGCGCGGCTGACTTCTAGCAGCTTCTGGCGCTAGCGCT
P N G M V R R L A A W G T D F V A I A

TCCATTTCCGTTAGCCAGGTCGTATAGTACCGTTCGCGGACTGCTTCCGGATCGCCGT 22860
AGTAAACCGCTCCGGTCCAGCATATCTATGCAACCGCTGACGAAGCGCTAGCGCA
I W K R Y G L T T Y I G H P V A E P D G

AGCCGAGCGCGCGAGCGCTCGCCGTTGGCGCGAAGGCCATTGGCGCGCTCGCGTCA 22920
TCGGCTCGCCGGCTCGCGGAGCGCAACCGCGCTTCGGTAAACCGCGGACCGCGAGT
Y G L P G L A E G N A A F A M P A S P T

TGCAGATCGCGATTCGAGCTGCGACAGGTCAGCGTTCGCGCGCGCCGCTGACCTCC 22980
ACGCTAGCGGCTAAGCTCGACGCTGTCGCGTGCAGACCGCGCGGCGACTGGAGGG
M C I A S E L Q S L S V T Q G R G S V E

bbsF3'
GCTCAAGCAGCGCCCATGATCCAAACAAGGATGTGTGCGTACCATACGTCATCGG 23040
CGAGTTCGTCGCGGGTACTAGGTTTGTTCATACACAGCCATGTTAGTGCATAGCC
R E L L A G M I G F L T H T P V M V H D

TATAGTTCGTCGCGGTCCGAAGGCAACTCGCCGGGAAAGCGCGCACCGCGCTCAGCC 23100
ATATCAAGCAGCGCCACCGGCTTCCGTTGAGCGGCCCTTCGCGGTGCGCGGACCGCG
T Y N T G T G F P L E G P F G A R A T L

CGCAGCGCGTGGAGTTACCGCGTACGCCATGTCGCGCTGTCGCGGCGATCGGTCG 23160
GCGTTCGCGCAACTCCAAAGTGGCGCATCGCGTACATCGCGACAGCCGCGGTCAGCG
G C L A N L N V G Y G M Y R S H P G D T

CCTGCATGCTCATCGTACAGTACGCGCGGATTGATCTTCTGCACGCTCTCCACG 23220
GGACCTGACAGTACAGTGCATCTAGCGCGGCTTAACTAGAAGACGTCGAGAGGGTCC
G Q M S M T V Y I A R P N I K Q V D E W

CGAGCTTCCATTTCTCCATCTGCGCAGCGGAAAGTTATTGATGACGATGTCGCTTTCG 23280
GCTCAAGGTAAGAGGTAGACCGGCTCGCGCTTCAATAACTACTGCTACAGCGAGAAGC
G L K W K E M Q G V R F N N I V I D S K

CGATCAGCCGACCGCCACCTCGCGCGAAGCGGTTGTTATGTCGAGCGCATGTCG 23340
GCTAGTTCGCGTGGCGTGGAGCGCGCTTCCGCCAACAAGTACAGCTCGCGCTACAGGG
A I L R V A V E R A L P N N M D L A I D

GCTTGTTCGCTGCGCGCGCGAAGTACCGCTCGCGTTCGAGGCGGTCGCGATGCGCT 23400
GCAACAGCGCAACCGCGGCTTCACTCGCGCAGCGCAGCTCGCGCACCGCTACGGC
R K N A N R A A F Y G S R D L G T G I G

CCTTGAACGCTCGCCACCGCGCATGTCGCGGACGCTTTCGAGCTTTCGATTTTGATA 23460
GGAACCTGCGGAGCGGTCGCGCTCATACAGCCCTGCGAAGTTCGATATGCGGGGG
D K F P E G R R L I D P R K V S E I K I

CGTCCGCGCGCATTCGCGCAGCAGCTCGTGGCGATCGCGCCAGCGCGCACCCACGAAA 23520
GCAGGCGCGGTAACCGCGTCTGCGAGCACCGCTAGCCGGTTCGCGGCTGGGTTGCTTT
V D A G C Q A L V S T A I P G A G V W S

AATCGCACACGCAATGCCCTCCAGGCAACCTCAACGCTTTCACATATACCGCCCC 23580
TTAGCGTGTGCTTACGGGAGGTCCCGTGGGAGTTCGACAAAGTTGTATATGGCGGGG
* R Q E V Y V A G
F D C C V I G E L A G E V S N L M ← bbsF

GTTCGACGAGTTCATCGATCTCGAGTCCCGAATAGCCGATTCGGTTAGTATTTCCCGC 23640
CAAGCGCTCAAGTAGCTAGAGTCAAGGCTATCGCGTAAGCAATCATAAAGGGCG
T R V L E D I E L G S Y G C E T L I E R

GTGTGCTCACCAGTCCGCGGGGCGATTGCGTCCGAGTTGCCATTGCAACTCGCCGAACTCG 23700
CACACAGTGGCTACGCGCCCGGTAAACGAGGCTCAACCGTAACGTTGAGCGGCTTGAGC
T H E G I R P A N R G L Q W Q L E G F E

TACGGCGCGCCGGTAGTAAATGTCGCGACCCGAGCGTTTCGTTGAACTGAGTCTTCCAG 23760
ATGCCGCGGGCCATCCATTACAGCGTGGCTCGCAAGCAACTTGACTCAGAAGTGC
Y P A G P Y T I D A G L T E N F Q T K W

AAITTCGCGTGCAAGAGCTGGGGTTCGCGAGCAGATCCGCGCGTTACTGACTGGCGTC 23820
TTAAAGGCCAGTTCGACCCCAAGCGCTCGTCTAGGGCGGCAATGACTGACCGCGAG
F K R H L L Q P N A L L D R G N S V P T

ACCAGAGCGCTGAAGCGCTGCCCATTTTCATACAGGTAGCCCTTGTCTGGGGTTCTGTA 23880
TGGCTCTGCGACTTCGCGAGCGGGTAAAGTATGTCATCCGGAACGACGCCCAAGACCAT
V S V S F R Q G M E Y L Y A K S R T R T

bbsE3' deg →

TAGCGCTCAAGACCGCGCAGAAGTTCGCGTAGCCCTCGTTCGAGCACCGGTACGCGAAA 23940
ATCGGAGCTTCTGCGCGCTTCCAGCGCATCGGAGCCAGCTCGTGGCCATGCGCTTT
Y R E F V R C F T A Y G E T S C R Y A F

TGCATCCACTTGTGCTGCTGAGAGAAGCTCCATTCTCGACCCCTTCGGCTCGAGCCAG 24000
AGCTAGGTGAACAGCAGCAGCTCTTCGAAGTAAGGAGCTGGGAAGCCGGAGCTCGGT
D I W K D D D L L K W E E V G E A E L W

CGCACGAAGGCGCTCCACATGCGCTTGTTCGCGCCATGATGGCGACCACTGATGTGAT 24060
R V F P T W M G K N R G M I A V L M I Y

CCGTCGACGACGGATGCAAGTTCGCGCAGCCGCTCGCGCCCGCTCGCGCCCGGATT 24120
GGCAGCTCGTCTACGTCACAGGGCGCTCGGCGGAGCGCGGGCGCAGGCGCGCCCTAA
G D V C P H L T G C G A E R G R G R I

TTGCCCTCGAGTCCAGAACTGCGCGCATTTCTGAGCGCCATCGCCTCGGCTCGATG 24180
AACGGAGCTCCAGGCTTTCGAGCGCGTAAAGCTCGCGTAGCGGAGCCGAGCTAC
K G E L D W F Q A A N E L A M A E A E I

CAGGCCACCTGCGACCACTGGCCCCCGCGGTACGCTCGCGCTGAACAGCGCGATCGCA 24240
GTCCGGTCAAGCTGGTGGACCGGGCGGCGATGCGACGCGCAGCTTGTGCGCTAGCGT
C A V D V V Q G R G T R Q A S F L A I A

CTACCGACGGCCGATACCGCTCGGCCATCCGGTACGCGCTGGTGTGCGCGCGCGCACC 24300
GATGCTCGCGCGTATCGGAGCCGGTACGGCCATGCGGACCAACAGCCGCGCGCGTGG
S G V A A Y A E A M R Y A Q N D P A R V

GGCTTGTGCGCATCTACGCCCGGAGATAGAGAAAGCGCTCAGCGCAACAACGTCOAGA 24360
CGGAACAGCGGTAGATGCGGGCGCTCTATCTCTTCGCGGAGTCGCTTGTGCGAGTCT
P K D G D V G A L Y L F G S L A S C T L

TCCGAACCGGGATAAGCCGCGAGCGCCCGGTGCGCCCGAAGCGGAGTACCGAGGTTGC 24420
AGGCTTGGCCCTATTTCGCGCGTTCGCGGGCCAGCGGGGCTTGCCTCACTGGCTCCAGACG
D S G P Y A A L P G T R G F P T V S T Q

ACCAGACTGCGGTTGTGCGGACTCAGCACATCGTAGGACAGCGCGAGCCCGTCCAGATA 24480
TGCTGTACCGCAACAGCGCTGAGTCTGCTGAGTCACTCTGTCGCGCTCGGCGAGGCTATT
V L S A N D R S L V D Y S L G L G D L Y

CCGAGCGGCGAGCTCTCAACAGCAGGTCGCGGCTGCGCGATAGCCGCGGAACACCGCA 24540
GGCTCGCGCTCGAGAGTTCCTCCAGCGCGCAGCGGATTCGCGCCCTTGTGCGCT
G S R C S E V L L D A S G C L R R F V A

CGCCCTCTCAITTCGAGGTCCACCGCGAGCCGCGCTTGCCTGTTGTAATAGAGG 24600
GCGGGAGGAGTAAAGCTCCAGGTGCGCTCCGCGCGAACCAGCGCAACATTATCTCC
R G E E N E L D V A L G R K G A N Y Y L

TACGGTAGACTCGCCTGACACCGGGTTCATTTCGCGAAATAGGGGCGCAGCGCGGCATC 24660
ATCCCATCTGAGCGGACGCTGCGCCCAAGTAACCGCTTTATCCCGCGCTGCGCCGCTAG
Y P L S A Q V G P E N R F Y P G V R R M

GGATCGCCCGCGGCGACTCCAGTGGATCAGCTCAGCACCCAGCCGCGGAACATCTTG 24720
CCTAGCGCGCCGCTCAAGCGGCTGAGTCTGCTGAGTCTGCTGAGTCTGAGTCTGAGT
P D G A P S E V H I V D A G L G A F M K

GCCGCTACGCGCCGAGTTCGCGGTCAGTCCAGCACCCGTAAGTCGCGAAAAGTCTGCT 24780
CGCGCATGCGCGGCTCAAGCGGCTGAGTCTGCTGAGTCTGAGTCTGAGTCTGAGTCT
A A Y P G L E G T L D L V R F R S F D Q

CCCATCTCACATTTCTCCTGCGCTGAGGGTTCGCGCGCAACCGCGCTGAGCCTGCGAAC 24840
GGTGTAGTGTAAAGAGGACGGCGAGCTCCAGCGCGCGTGGCGGCGAGTCCGAGCCTTG
G M ← bbsE * G A F

← bbsD3' nuevo → bbsD3'

AGGCTTCGTCGCGCCGACAGTACGACCTTGGCGGTGACGAAACCCGCTCTGCTGCTG 24900
TCCGAAGCAGCGGGCTGCTATCTGTCGAGCGGCACTGCTTGGCGAGCAGCAGC
L S R G G C V Y L V Q G T V F G T E D D

CGCGACAACAGGAAGTCCCGATCCTTCTCCGGCAGTTCCTCCACATCGGCGTGTGGATC 25020
GCGCTGTTGCTTTCGCGGCTAGGAAGAGGCGCTCAAGGAGGTTGATGCCGACACCTAG
R S L L F Q R D K E P L E E W M P T H I

AGCCCGGAGCAACGAATTGACGGTGTATCCGCGCGCGCGAGCTCGATCGAAGCGTA 25080
TCGGGCGCTGTTGCGTTAAGTCCCACTAGGCGCGCGCGCTCGAGTTCGCGCT
← bbsD524.5'
L G P A V C N V T I G A R G L E I A L T

GCGCTCATGCCGACGACCGCGGCTTCGCGCAGCAATACGAGTCTGCCCCGCGCGCG 25140
GCGAGTACGGCTGCTGCGCGCGAAGCGGCTGCTTATGCTCAGACGCGGGCGCGCGC
R T M G V V G A K A S S Y P T Q G A G G

AGCCATGCGCGGGAGCGATGTTGACGATGCGCGCGTCTGTTCTCGACCATGTGGCCG 25200
TCGGTACGCGCCCTCCGCTACAAGTCTACGCGGKCAAGCAAGAGCTGTTACACCGC
L W A R S A I N V I R G H K N E V M H G

TGCACCGCTCGGAGCAGGAAGGTGCGCTTTCAGATTGACATTGAGCGTGACATCCAG 25260
ACGTGGCGAGCTCGTCTTCCACGCGAAGTCTAAGTAACTCGACTCGACTGTAGGTC
H V A Q S C L F T G K L N V N L T V D W

TCCGCTCGTAAGCTTCCGACGCGCGCGCGCGCTCCATGCCCCGCTTATTACAGAGG 25320
AGGCGGAGCGATTGCAACCGCTCGCGCGGGCGGAGGTACGCGCGCAATAAGTGTCC
D A E S L K R L A G A R E M G A N N V L

ATGTCGATGCGCGCAAGGCATTACCGCTGCTGACCATCGCTCGACCGCGCGCTTG 25380
TACAGTACCGCGCTTCCGTAAGTGCAGACGAGTGGTAGCGAGCTGGCGCGGAA
← bbsD5'EbN1
I D I R G F A N V T Q Q V M A D V A A K

TCGCTAATGTCGCGGACGGCGCGGCGCGATGTCCTTTCGCGTGAATCATCGACC 25440
AGCGCTTACCGCTGCGCGCGCTCGCGCGCTACAGGGAAGCCGACTTGAAGTAGCTGG
← bbsD5'EbN1
D S I D A V A G L A R H G K R S F E D V

GTCGCTGACACTTTCGCGGTTGATGCTGTTGATAACGACGCGCGCACCCTGCTCGCG 25500
CAGCGCAGCTGGAAGAGCGCAACTACAGCAACTATTGTCGCGCGTGGGACGAGCGC
T A H V K E A N I D N I V A S D S A L E A

AAGCGCAATGCGCTCTGCTACCCATGCGCTCGCGCACCCTGTATCAGCGCAGCCCTG 25560
TTCGCGTTACCGCAGACGGATGGTACGCGGAGCGGCTGGGACACTAGTTCGCGCTGGGAC
F R L A T Q R G M G S A S G T I L A V R

TTGCAATTCCTTTCGCGCCCGGCTGATCAAATCGACATCGAAAAGTACGCGCTTCTTC 25620
ACGCTTAAAGGTAACCGCGCGGCACTAGGTTAGCTGAGCTTTTCATGCGCGCAGAAA
N R I G M ← bbsD * I S M S F Y A S K

CGCCACAGACGAAAAGCGTCTGCGCGTCACTTTCGAATCTCGCTCGCAAGAAAGC 25680
GCGGTCTGCTTTTCGCGAGCGGGCAGTGCATGAAGCTTGTATAGCGAGGCTTCTTCG
G G V F L T T Q G T V Y K S S D S A L E A

CGACTGCTTGGCAACGTCAGAGGCGCTGCCAAGCGCTTTCAGGGAATGCGCGCGCGA 25740
GCTGACGCAACCGTTCAGCTTCCGCGAGCGGTTGCGAAGTGCCTTACGCGCGCGCT
G V A N A V D S P R G L R K V P I G A A

GCTTTTCCCGCTCCGCGCAGCGCTCGCTGCTGCTGAGATGCGCCATGACGACCGTGC 25800
CGAAAAGGGCGAGCGGTGTCGCGAGCGCAGCCACAGCTTACGCGGTAAGTGTGGCAGC
← bbsC.F5'
L K E R E A V P T A D T D L D G M V V T

TCACCGTAACCGCTCTCGCGCGATTTCGAGCGGAGCGCCGCTGTGAGCCGAAACGAC 25860
AGTGGCAATTGCGCGAGCGCGCTAAGCTCGCGCTCGCGGCGCACTCGGCTTGTCTG
bbsC.F5'
S V T V G D R A S E L A L A R T L G F L

CGCGCAGCGCAGCGCGAGCTTGGCGCTCGCGCGAGCGCGAGATAACGAATATCGCTGA 25920
GGCGTGGCGCTCGCGCTGCAACCGCGAGCGCGCTCCGCTGATGCTTATAGCGACT
G A R A A A V N A S A A L G L Y R I D S

TATTGACCAAGCGCCGAGCGCACTTGGCGATCAGCGGAACGACTTTCGCGCAGTGT 25980
ATAACTTGGCGCGGCTTTCGCTGTAAGCGGTAAGTCTGCTTTCGTAAGCGCGCATCA
I N V V R G F G C K R M L P V V E R L Y

GGAACGGGCGCGATGCGCGCGCGAGCGCGCGCGAGCTCCGATCGGAAATGCTCT 26040
CCTTTCGCGCGCTACGCGCGGCTCGCGCGCGCTCGCGCGGCTGAGCGCTTACAGGA
H F P G A I G A G L A A G V D A S I D

GCAGCGCTTGGCGGCTCAGTACGCTTTCGTTCTGACCAAGATGTCGATCCGCGCTAGC 26100
CGTCCGCAACCGCAGTATGCAAGCAAGAGCTGCTTACAGCTAGGCGCGCGCTCG
Q L A K A T L V S E N Q V L I D I R G Y

GCGCCATCACTGCGGACCGCATCGCGGATCGAACCGGCTCGACTGCGCCGCTACCA 26160
CGCGTGTGAGCGCGCTGCGCTAGCGCTAGCTTGGCCGAGCTGACGCGCGCAGTGGT
R A M V Q A V A D R I S G P E V A G T V

CCGCGTACCGCGCACCCCTTCTCGCGCAGCGGATGCGGAGCGCTGAGCGCGCGCG 26220
GGCGCTAGTGGCGTGGGGAAAGCGGCTCGCGCTACCTCTCGCAGCTCGGCGCGC
V A I V A V G K E A L R S A L A D L R G

CGTCCGCGCCACGAGCGCGAGTTGCACACCCGGCGGGGAGCCGCGAGCGCAACCCGCT 26280
 GCAGCGCGGGTCTCGCGCTCAAGGTGGGCGCGCCCTCGCGTTCGGCGTGGCGGA
 A D A G V L A L Q V G P A A L R L A V A

CGCCACCCGATCGTCCGATGACGATCAGAGCCACCCCTTCTTTCGATTCGATTCCA 26340
 CGCCGTGGCGTAGCAGCCGTAAGTCTAGTCTCGGTGGGAAGAAACGGAAGCTAAGGT
 E G V A D D A N V I G L A L V R E K G K S E

TGCGCGTCTCCAGTTACCAGCCGCGCACCCAGCATGTGCATCGTCAAGCCGCGTGTTC 26400
 AGCGCAGAGGGTCAATGGTGGCGCGTGGTGTACACGTAGCAGCTTCGCGCGCAAGG
 M ← *bbsC* * W G R V L M H M T C A A H E
 ← *BamHI* *bbsB3'*

GTGCCAGACAGCCCGCCACCCGTAACGTGCGTCTATCGCCACCGCGCCAGCGGATCGACC 26460
 CACGGTCTGTCGGCGGTGGGCATTGCACCCAGTAGCGGTGTGCGCGTGGCCTAGTCTG
 ← *bbsB.F5'*

T G S L G G G T V H T M A V R P V P D V

TGATAGCCGGGACTGGTTGCGCATCTGCTTACGCTCGCGCGATCTGGCGGCCCC 26520
 ACTATCGCCCTGTGACCAACCGTAGACGAAGTCCGAGCGCCGCTAGACCCCGGGGG
 Q Y G P C Q N R M Q K V S A A I Q A A G

GAAGCGCGATCGGGTGACCATACGAGAGCATGCCGCCACGCGGACTGACGACACTTG 26580
 TACAGCAGCGGCTAGCCACTGGTATGCTCTCGTACCGGCTTCGCGAGCAGTCAAGCTGTGTAAC
 S A G I P H G Y S L M G G R P S V V C K

CCCCGTGCTGCTGTTGCGGTGCGCGAGAACTTACGCCCTTCGCCCTTCGGACACAGG 26640
 GGGGCAACCGCAACCGCAGCGCTCTTGAAGTCGGGAAGCGGAAGCCTGTGTC
 G G H T T N G D R L G L E R L G K P C L C

CCGAGGCATTCGTAGTACAGCAGCTCGGCAATCGTAAAGCGTGTGACGTTTCGACGATA 26700
 GGCTCCGTAAGCATCATGCTCGTACGCGTTAGCACTTCGCGAGCAGTCAAGCTGCTAT
 G L C E Y Y L L E A I T F A D H L E V I

TTCACATCTCCAGGACCGATGCCGACTCTTCATAGAGCTGCGCTGCGGTTTCCCTCGTG 26760
 AAGTGTAGAGTCTTGGCTACGGGCTGAGAAGTATCTCGACCGCAGCCAAAGGAGCCAC
 N V D E P G I G S E E Y L Q A A T E E T

ATGTCGTCGCGGTTATGTCGCGCGCGGTTGGTAAAGCCCGCAGCCGACGACCGG 26820
 TACAGCAGCGGCAATACAGCGCGCCGCAACAGCATTCCGGGCTGGCGTCTGGCTC
 I D D G T I D R P R N H Y P G S R V V S

CCCGCCACAGAATCGGCTTTCGATGCCCACTTCTTCCACCATCTCTCCGAGCACAG 26880
 GGGCGTGGTCTTAGCCGAACAGCTACGGTGAAGAAGTGGTAGAAGGCTCGTGTGTC
 G A V L I P K D I G L K V M E E S C V
 ← *PstI*

ACGACCGCCCGCGCATCGCGATGCCACGATTCCTGACGCTGACGCGGAGACCG 26940
 TGCTGGCGGGCGCGGTAGCCGCTACGGTGTGTAACGACGTCGAGTCCGCTTCGGC
 V V A G A G D A I G C C Q Q L T L P S A

ACGACCGCGCAATTGACGACCTCTCGAGCAAACTTCCACCTTGAACACGCGATACGG 27000
 TGCTGGCGGCTTAACTGCTGGAGAGCTGCTTTCAGTGGGAACCTGGTGGGATGCTT
 V V R S N V V E E V S I E G K F W A Y P

TTGTTGGTTCGCGGCGGCTTCTTCCAGGTGATCATGCCAGATCTCGGCCGTGGCC 27060
 AACACACCGCCACCGCGCAAGAAAGTGCACATAGTACGGTCTAGGAGCGGCAACCGG
 ← *bbsB5'*

N N T A H R R N K V T I M A L D E A T A

CGGTCTCGTACATGTAACGCTGCGCGCATCGGTAAGTTCGCGGTCATGACCCGCGCG 27120
 GGCCAGGATGATGATGCGACGCGCGGCTAGCCGATGAACGGCCAGTACTGGCGGCG
 G T E Y M Y R Q A R M A Y K G T M V A G

TGCACCGTTTCGATATCCGACATGGCCGTTGCAAAAGCAGTGCCTTTCGCGCATGTGCG 27180
 ACCTGGCAAGCTATAGGCTGATACCGGCAACGTTTCCGTCACGGAACGCGGTACACGTC
 H V T E I D S M A T A F A T G Q A M H L

GTGTATCTCGCAGCGGATGCCGATGCGCAGCTCGGCAAGCCCATCGCCAGTCCCGC 27240
 GCACTAAGAGCGTCCGCTACGGTACGCGTGCAGCCGTTTCGGGTAGCGGTGACGGCG
 T H N E C G I G I A V D A L G M A V D R

ACCGCCAGATGCACTGCCATACCCCTGCGGAGCATGCGCTTTCGACTGTGATGATTGGA 27300
 TGGCGCTACGTGACGGTATGGGGGACCGCTGACGCAAGAGCTGACAGTACTAACCCT
 V A L H V A M G G A C S E V T M I P

AGATGTCGACATGCCAGATCCTTGAAGCAGGTTTACCGATCACCATTCGCTTGGTG 27360
 TCTACAGCGGTGACGGTCTAGGAACCTGCTGCCAACTGGCTAGTGGTAAGGCAACCC
 L H G C M G L D K L V T Q G I V M G N T

CCATTGCCGACGTAGGCGCTGTCACCATCGTCCGACGGTGCATATTCGACGCGCTTCAG 27420
 GGTAAACGCTGATCCGCGAGACGTTGAGCAGCTTACGCTAAGCTGCGGAAGTCG
 G N G V Y A S Q V M T P R D I N S A K L

CGCTGAAATGCCCGGACCGCCAGCACATCGAAATCGACCTCGTCCCGCCGAAATTC 27480
 CGGAGCTTACGGCGCGTCCGGTCTGTTAGCTTACGCTGAGACGCGGGCTTAAAG
 A E F A A A R G L V D F D V E H R G F K

GTTCGCGACCCCGCGGATATAGACCTTGGCGTGCAGTTTTCATTTTTCACCCACCG 27540
 CAAAGCGGCTGGGGCGCTATATCTGGAACGCGAGCTCAAAGTAAAGAAAGTGGGTGGC
 T E G V G A I Y V K R Q L K M ← *bbsB*

bbsA3'new → *bbsA331.3'*

AGATTTATGCTTGGCGAACTTGTAGCTGATCAGCGGAACCGCGTTCGCGGAA 27600
 TCTAAATACGGAACCGCTTGAACATCGACTAGTGCCTTTCGCGCAGCGCCAAACCGCTT
 * A K R F K Y S I L P V G D R N D R

bbsA331.3' →
 CGGGCCCCACGGTCAAGTTCGACCTCTCTCGTCCGAGCGGAACTCGCCGGCTCGCCCTTCCA 27660
 GCGCGGGTGCAGTCAAGCTGGAGGACGAGCGTGCCTTCGACGCGCGAGCGGAAGGT
 ← *bbsAint5'*

V P G V T L E V E E D C R F S G P E G E

GCTCGCGAAAATCCGCGAGTTCGCGGAGATCGATGAGCGAAGCAATATGGCGTTC 27720
 CGACCGCTTTTAGCGTCCAAGAGCCGCTAGTACATCGGCTTGTCTTATACCCGAAA
 ← *bbsAint5'*

L Q A F I R L N E P L D I Y G F S Y P T

TCATTCGCGCTGACCAATGAAATGTCGCTGAAGCTGTAGAGCTTCCCGCTCGGCTTA 27780
 AGTAAGGGCGGACTGGTACTTTAACAGCAGCTTCGACATCTCGAAGGGCGCAGCCGAAT
 K M G A Q G I F N D S F S Y L K G R R S

GCGGCACGACATCGAACTCTCGCCCCAGCACTCGGGCAGCGACTGAGTGGGAAAT 27840
 CGCCGTGCTGAGTTCGAGGAGCGGGTCTGAGGCCGCTGAGTGCAGAACCCCTTTA
 L P V V D F E E G W C S P C P S L K P F

bbsAint3'
 CCAGCTGGCGCACCCCTTGCAGGATACCCCTTGAATACGCGCGCGCCGCTCGCGG 27900
 GGTGACCGCGCTGGGGAACGCTGATGGGGAACCTATGCCGCGCGCGCAGCGCGC
 ← *bbsA5'new*

D L Q G C G K C R Y G K L Y P A G G D A

bbsAint3' →
 GCACCTCGAGCAGGTCGGATGAAAGAATGTGATATCCGCTTCTTCCGCTTCTTGT 27960
 CGTGGAGCTGTCAGGCTACTTCTTACACTATAGCCGGAAGAGAGCCAAAGAACA
 P V E L L D P H F F T I D P K K E P K K

PstI → *pbbsA3'*
 TCGGCTTTTGCATGACATGACGCTCCGCGAGCATTTGTTTCATTTTCAACGTAGATCCA 28020
 AGCCGAAAATAGACTGTAAGTTCGCGGAGCGCTGTAACAAAGTAAAGTTGCATCTAGGT
 N P K Q D S M ← *bbsA*

ATTAAACCGCGCTGGTTCGCGCGCAACAATGTCGCGACACCTAAGCTTAAAGGTAAC 28080
 TAATTTGCGCGGACCAACCGCGCTTGTAAACACGCTGTGATTGACAAATCCCATTTG

CCCCACCGCACCGCCAAAGCGGAAATCACGTCGAGCGCGGTGTCGCAATCTGGCC 28140
 GGGTGGGCGTGGCGGTTCCGCTTAAAGTGAAGCTCGCCGCAAGCTGTAGACCGG

GCACAGGCGGATACCCGCAATTTTCGCTGTAATAAAGCGCGGAAAGCGCGCA 28200
 CGTGTCCCGCTATGGGACGTTAAAGCCGACATTTTTTTTCGCGCGCTTCCGCGCGT
 ← *c1A68.5'*

AAGTTCACCGGAGGAGGAACCCCGAGAAAATCTTCTACTCAGTTTCCCGCACGCA 28260
 TTCAAGGTTGCCCTCTCTCTTGGGCTCTTTTGAAGATAGTCAAAGCGGCTGCGT
 * N G G R

TGTTTCGAGCTTAAACAGGCTGACCGGACCATCTGACGCTCCAGTTGCGCCTTGAGCT 28320
 ACAAGACTGCAATTTGTCGACTGGCCCTGTAAGTACGAGCTCAACCGGCAACTCGA
 ← *cbbs5'* ← *pbbsA5'*

M N Q V N F L S V P V M Q A D L Q A K L

GTCCGCTGTCGAGTCTTCGCTGGATATCGATCATGGTGAACGATCGAAGCGGAGTTA 28380
 CAGGCCACGCTACGAGCGGACCTATAGCTAGTACCCTTGCCTAGCTTGCCTCAAT
 ← *pbbsA5'* ← *XbaI*

Q G T T C H K A Q I D I M T F A D F P T

TGCCGCTGCCCTTGTCTTGGCCAGGTGGAAGTCGGGATGCGTCCATCCGACAACATA 28440
 ACGGCGCGGAAACAAGAACCGGTCACCTTCAGCCCTACGCGAGTGGCTGTTGTATGC
 I G S G K N K A L H F D P H T W G V Y V

TCTTCCAAAATCACCTTGGCGTATGCAACAGCAGATGCTAATACCGAGGTTTGG 28500
 AGAAGGTTTGTAGTGGGAACCGCAGTACGTTGCTGTACGATTAATGGCTCCAACCG
 T K W L D G K R D H V L S M S I V S T Q

CGTCCATGTAATGACCCGCTGCTCAACCGGATGATTGGTTCGATCGGAGACGATTGCA 28560
 GCAGGTACATACGTTGGGACAGAGTGGCTACTAACCAACAGTACGCTCTGTAAGCT
 A D M Y H V R R S V P H N T D I P S S E

GGATGTCACCTTGGCAACTGCCAGTGCAGCTGGGGAAGCACCCGCTTTCGCGCGA 28620
 CCTACGCTGGAACCGGTTGACGCTGACGACCCCTTCTGTTGGGGAAGGCGGCT
 L I H V K R L Q W T V S P F C G G K G G

AGGCGAGTACTTGTAGCGTCCGCTCTTGTATTCATCGGACGCTTACATGCTGTTAT 28680
 TCCGCTGATGAACATCGGCGCGGAGAACATAAGTACGCTGTCGAAGTACAGCAATA
 F A V Y K Y G D P E K Y E D S L K M D N

GGTGTAAAAGGCAATCAGTATGTCGCGCTGCGCTGCTACTTCCAGTTCAATGTCGAAA 28740
 CCAACATTTTCCGTAGTCATAAAGCGGACGGGACCATGAAGTCAAGTACAGGCTTT
 H N Y F P M L I N A T G Q Y K W N M D S

TACGGCGCTTATAGCCCTCGAAGTCTCGATCATGATGTCGAGCGGAAAGGCGTCCG 28800
 ATGCCGCAATATCGGAGCTTACGAGCTTAGTACTACAGACTCGCTCTTCCGAGCG
 I R G N Y G E F D E I M I D S G L F A D

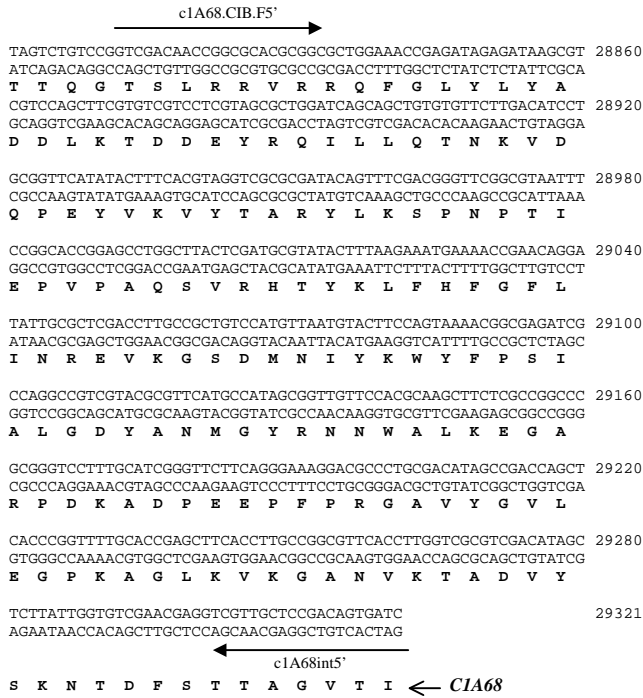


Figura 51. Secuencia de nucleótidos y secuencia deducida de aminoácidos del cluster de genes *bss/bbs* de *Azoarcus* sp. CIB y regiones adyacentes. Las flechas cortas al comienzo de las secuencias deducidas de aminoácidos indican la dirección de la transcripción de los genes que codifican dichas proteínas. Las flechas negras más largas indican los oligonucleótidos utilizados para la secuenciación del cluster, indicándose en rojo la diana de restricción. En azul se indican algunas de las dianas de restricción presentes en el cluster.

VI. ANEXO II

AAGGTCGCGCCCTGGCCCGGTTGCTGACCGCCAGACCCGTTCCACCCATGCGCTGCACT 60
 TTCAGCGCGGGACCGGGCAAGCGACTGCGGGGCTGGCAAGGTGGGTACGCGCACTGCA
 GCCTTGGCTACCAGCGGAGCCCTATGCGCGGTGCCGGGATAGTCTTCGGCGCGATGCA 120
 CGAAACGATGTTGCGGCTCGGGATACGGCCACGGCCCTATCAGAAGCCGCGCTACGCTC
 CGCTGGAATATCTCGAAGATGCGATCATGATACTTTCATGTGCAAGCCACGCGGTTGTCC 180
 GGACCTTATAGAGCTTCTACGCTAGTACTATGAAGTACAGCTTCGGGTGCGGCAACAGG
 GCCACCGTCAGGACGATGTACGGTACGCGCGGTACGCATCGATCGTGTGCGCGCTGCG 240
 CGGTGCGAGTCCCTGCTACACTGCGAGTTCGGCCATGCGGTAGCACTACGCGCGAGCC
 GCGTGGCGCGGGTGAATTTGAGGGCGTGTGCTGATGAGGTTGCGCAGCACCAGCGTAGC 300
 CCGCACCGCCCACTTAACTCCCGCAACGACTACTCCAACGCGTGTGGTGCAGCTCG
 CGCTCGGGTTCGCGGATCTCGCGCGGACGCTCATTGTTGACGACCTCGATTCCGCGG 360
 GCGAGCCACGACCGCTAGACGCGCGCTGAGTAACCACTGCTGGAGCTAAGGCGCC
 GTCTGGATCTCGTGTGCGAAGCCGTCGAGCAGCCCTGCACGACGGCGGGCAGATCGACG 420
 CAGACTAGAGCACACGCTTCGGCAGCTCGTGGCGGAGCTGCTGCCGCCGCTTAGCTGTC
 GCGCGGATGTCAGGCTGACGCTCGATACGCGAGTAGGCGAGCAAGTCTCGATCAGC 480
 CGCGCTACAGGTTCCAGACTGCGAGCTATGCGCTCATCCGCTCGTTCAGGAGCTAGTCG
 TCGTGCATGTTGACGCGCGCGCGGATGCTCATGAGGAAGTGGCGCGCTCCTCGTTCG 540
 AGCAGTACCAGTTCGCGCGCGCGCCCTACGAGTACTCCTTACCGCGCGCAGGAGCAGC
 AGCGGTCGCGTACTGCTGAGCAGCAGCTGCTGTAGCCCTCGATGCCGCGAGCGCG 600
 TCCGCGACGCCATGACCGAGCTCGTCTGACGCGACATCGGGAGCTACGGCGCTCGCCG
 GCCTTAGGTCGCGGAGCGGTGTAGGAAACGACCGCAGTTCCTGTTGGCGCGCTGCG 660
 CGGAACCTCAGCAGCTCTGCCACATCCTTGTGGGTCAAGGACCAACCGCGCGAGC
 AGTTCGCGGTCGCTTCGGGACCGGCTTCCACGCTGGCGTGTGAGCGCGCGCACCTCG 720
 TCGACGCGCCACGCAAGCCGTCGCCAGAAAGTGCACCGCACTCCGCGCGTGGAGC
 GCCTCGCGCGCTTGGTTCGCGGTGATGTCGAGCAGCAGCCGCTGATCTCGTCCAGCGG 780
 CGGACCGCGCGAACCAGCGCCACTACAGCTCGTGTGCGCGACCTAGAGCCAGTGGCGC
 CCCTGACGTCGCGCGTGGACAGGTGCGGTCGTCAGCCAGATCCACGTCGCCGAGGCGG 840
 GGGACATGACGCGCGCACTCGTCCAGCGCAGGTCGCTAGGTGACGCGCTCCCGC
 ebA2987_5'
 GTGCGGATGCGGTTATTCCTGGCGGTAGTCTCGGACCGTGGCGAGGTGTTGCGCGGACG 900
 CAGCGCTACGCCATAAGGACCGCCATCAGCAGGCGTGGCAACCGCTCCACAAGCGCTGC
 TCGCGCTCGATGCGCGGGCGGTGCTCGCGCTGGATCAGGTCGGCAGTGGAGCTTGGCCG 960
 AGCGGAGCTACCGCGCCGCGCAGCAGCGCACTAGTCCAGCGTATCACTGCAACGCG
 GAGATGAAGTATCGCGGCTGTAGCCCAATTGAGCGATGCGCTGCTGACGAAGCCGACC 1020
 CTCTACTTACAGTACGCGCAGATCGGGTAACTCGCTACGGCAGCGACTGCTTCGGCTGG
 Sau3AI
 GGCCAACCGGGCTGGTTACGCCATTTCGATCGCCACCCCGGAGCGGTTGATCAGCGCG 1080
 CGGTTGGCCCGACCAATGCGGTAAGTCCGCGTGGTGGCGCTCGCCAACTAGTTCGCGC
 GTGATGCGGGCGCGCTCCGTCGCGTGTGACGCTGCGGTGCGATCCGTCGTTTTTCCATC 1140
 CACTACGCGCCCGCAGCAGCGCAACTGCGACGCCAGCTAGGCGAGCAAAAAGGTAG
 GTGCTCTCCGCGCTGGCGGTGCGGGACCGCCAGTTCGGGGCGCCCGGGTCTGTTGTC 1200
 CACGAGGAGCGCGACCGCGCAGCGCCCTGGCGTCAAGCCCGCGGGCCGACGAAACAG
 GGAGAATTCACCGTCTGCGCGGCGATGCTGTGGCGTGTGCGCACCAATGACCGCGG 1260
 CCTCTTAAGATGGCAGGACGGCCGCTACGACACCGCAACAACGCGTGTACTGCGCGCC
 TGTGCGGAATCGGGTAATGATGAGACTATGTTGCGCGATGGATGAGCCGCGACCGCG 1320
 ACACGCGCTTAGGCCATTACTACTCGATACACAACGGCCACTACTCGCGCTGGCCG
 XbaI 3GcdH
 GGGTGTGCGCGGTGCGCTCCGAGGAAAGAAAGAACCCCGTGGCGGGGTCGTCGCGC 1380
 CCCACGACGGCCAGCGCAGGCGCTCTTCTTCTTGGGGCCACCGGCCACGACGCGGG
 TTGCCCGGCGCCGACCCGATGAGGTCGCGTGGCGGGCGGCGAGTTGCGGAACTCAG 1440
 AACGGCCCGCGGGTGGGCTACTCCAGCCAGCGCCCGCGCTCAAGCGCTTGGATC
 FG5'
 *
 TTCGAGAACCGCGGAGGCGGTTGATGGCGCGCCGAGGATCAGCGCGTGGATGTCGTCG 1500
 CACGGGAGCATCCACAAGTGGTGAAGTCCCAAGTGTCTCCACGCGCGCTACGCTTTGAGC
 N S F A A L G T I A R G L I L A H I D H
 3'GcdEbN1
 GTCCCTCGTAGTGTTCACCACTTCGAGGTTACGAGGTTGCGCGCGATGACAGAATCG 1560
 CACGGGAGCATCCACAAGTGGTGAAGTCCCAAGTGTCTCCACGCGCGCTACGCTTTGAGC
 T G E Y T N V V E L N V L H R A I C F E
 3'GcdEbN1
 TCCGAGATGCGGTTGCCGCCGAGCATGTCGGGGCGACGCGGGCGATGTCGAGCGACTTG 1620
 AGGCTCTACGCGAACCGGGCTCGTACAGCGCCCGCTGCGCCGCTACAGCTCGCTGAAC
 Gedinv5'
 D S I G N G G L M D R A V R A I D L S K

Gcd3'
 CCGCAGGAGTTGCGCTTCATGATCGAGGTCAGTTCGACCGGCGGTTGCCCTCGTCTTG 1680
 GCGTCTTCAACCGGAAGTACTAGCTCCAGTCAAGCTGGCCGCGCAACGGGAGCAGGAAAC
 G C S N R K M I S T L E V P A N G E D K
 PstI
 AGGGCGCCAGCGCGCAGGACGCTTCGAGGCGGCGGCGGCTGATCTCGGTGATGTCGGCG 1740
 TCCGCGGGTTCGCGTCCGTCGCGGACGCTCGGTCGACGACGAGTTCGATGACGCGCGC
 L R G L R L C G Q L G L T I E T L M D A
 AGCTTCTTCTGGATCAGTGGTTCGCGGCGAGCGGGCGGCGAAGTGTTCGCGTCCATC 1800
 TCGAAGAGACCTAGTTCGACCAAGCGCGCTCGCCCGCGGCTTGAACGAAACGCGAGTAG
 L K K Q I L Q N A A L P R G F Q K R D M
 GTGTACTGGGTCGCGTCTGATGACGAGGCTTCGCGGGCGCGGAGCGCGCCACGCGATG 1860
 CACATGACCGCACGCGAGCATCTCCGAAAGCGCGCGGCTCGCGGGGTCCGCTAC
 T Y Q R A T E Y C A E A A G L A G W A I
 CCGTAGCGGGCGAGTTCAGCAGGTTGAAGGGCCCTTTGAGGCGCGCACGCTCGGGAAAG 1920
 GGCATCGCCGCGCTCAAGTCCGCTCACTCCCGGAAACTCGGCGCGTGCAGCGCTTC
 G Y R A S N L C T F P G K L G R V T P F
 GCGTTTTCTTCGGGGCAGAACCTCGTCCATGACGATCTCGCGGTGATCGAGGCGCGC 1980
 CGCAAAGAACCGCGCTTGTGGAGCAGGACTGCTAGAGCGGCCACTAGCTCCGCGCG
 A N E E P C F V E D M V I E G T I S A R
 AGGCGACCTTCGCGTGGATGCGCGGGCGGACAGGCGCTTCACGCGCTTTTCCAGCAGC 2040
 TCCGCGTGAACGCGCACCTAGCGCGCCCGCTGTCGCGGAAGTTCGCGAAAGTTCGTC
 L G V K G H I A P A S L G K W G K E L V
 AAGCCGCGGATGTCGCGCGGCTGCTCCTTCGCGCACGCAAGACGTCGCGGATCGGG 2100
 TCGCGCGCTACAGCGCGCAGCAGGAAAGCGGTTGCTGCTTTCGCGCGCTAGCGC
 F G R I D G A D D K A W V V F V D A I P
 CTGTTGTTGATCCACATCTTGGCGCGGACAGGCTGATGCGCGCTGACCTTCTGCGG 2160
 GACAACCACTAGTGTAGAACCAGCGGCTGTCGACACTCGGCGAGCTGGAAGGACCGC
 S N T I W M K A G S L Y G G D V L K R A
 CCGGTGACCATGCTGCGCGGGTGGAGCGGTTGGGCTCGGTGAGCGCAAGCAACCC 2220
 GCGCACTGGTACGACGCGCGCCGCTCGGCGCACCAACCGAGCGCTCGGCTCGGTTGGG
 R T V M S G P D S G H N P E T L G F C G
 Gedinv3'
 ACCCATTTCGCGGTTGGCGAGTTCGGCGAGTATTTCTGCTTGGTTTCTCGGAGCCGAAT 2280
 TGGGTAAGCGGCCACCGCTCGAACCCTCCATAAAGCAAGCAACAAAGGAGCCTCGGCTTA
 Ged5'
 V W E G T A L L K P L Y K Q K T E E S G F
 TCGTTGATCGGCACCATCACGAGCGAGCTCTGACGCTCATCATCGAGCGGTAGCCGGAA 2340
 AGCAACTAGCGGTTGCTGCTCGTACGAGCGTGGAGTAGTACTGCTCGCCATCGGCCTT
 5'GcdEbN1
 E N I P V M V L S S Q V S M M S R Y G S
 TCGACGCGCTGACTTTCAGCGGCGATCAGGCGGTAGCAGCAGTTCATCGCGGAAACCG 2400
 AGCTGCGGAGCTGAAGTCCCGCTAGTTCGCGCATCGTGTGATCAAGTACGCGCTTGGC
 D V R E V E R A I L G Y C V Y N M G S G
 CCGTACTGCTCGGGATCGTTGCGCGGAGCAGGCGGCTTCGCGCAATTCGTTGAAGATC 2460
 GGCATGACGAGGCGCTAGCAACCGCGCTCGTCCGGTCAAGCGGGTAAAGCAACTTCTAG
 G Y Q E P I T A G L L G L E G M E N F I
 Ged109.3'
 GCGCGTGGTCTTCTGCGGGAAGGCTTCGTCACGCGCGGCGAGCAGGCGCTCTGCG 2520
 CGCGCCAGCAGGAGCAGCGCGCTTCGCAAGGAGCTGCGCGCTGCTCGCGAGGAGC
 A R D T K E H R F A E Q V R P L L R E Q
 Ged109.3'
 GAGTAGGCGCGGCGCTGTCGCGCACCATGCGTTCGTTGCGTGGTGGTGGCTGTCGAGA 2580
 CTCATCGCGCGCGCGCAGCAGCGGTCGTCGCGCAAGCCACGCGCACTCAACCGCAGCTCT
 S Y A R A S D R V M R E T D T L Q S D L
 GedF3' XbaI Pgc3'
 TACAGGGGTCGCGCCAGTCAAGGATCAGCGGTTGTTGCCATGTTGAATCTCTCCTCAAT 2640
 ATGTCCCGACCGCGGTCAGCTTTCAGTGGCGCAAGCAAGCTCAACTAGAGGAGTTA
 5GcdH Clal
 Y L P D A W D F T V R N T A M ← gdh
 Pgc3'
 GCGGCGTTTGGTTCGTTGGGCGGACCGCGCTCTGCTCAGTGAAGTGTAGTCTGAG 2700
 CGCCGCAAAACGACCGCGCGCTTGGGCGGAGCAGGACTGACTGACCTGACACTGACACT
 CGCGGGCGTGCAGCAAAACGAAATTTTCGTTGAGTGTGTCGCTAAAACGCACTTCGAG 2760
 GCGGCGCGAGCGCTGTTGCTTTAAAGACCACTCGAAGCAGCAATTTTGGCTGAAGCTC
 AGACGTGACAGCTCTGAGTCTGATGAGGCGGAAAGCGCTGCGCTTCCGCCATGACGAC 2820
 TCTGACGCTGTAGCATCTGAGCATCCCGCTTTTCGCGACGCGGAAGGCGGTACGCTG
 Pgc5' KpnI
 CGTCGAGACTGCACCGGCGCTCATTTCGCGGAAGCGCGCTTCGCGCTTTTCCGCCCTACGC 2880
 GCAGCTCTGACGTGGCGCGAGTAAGCGCGCTTCGCGGAAGCGGAAAGGCGGGATGCG

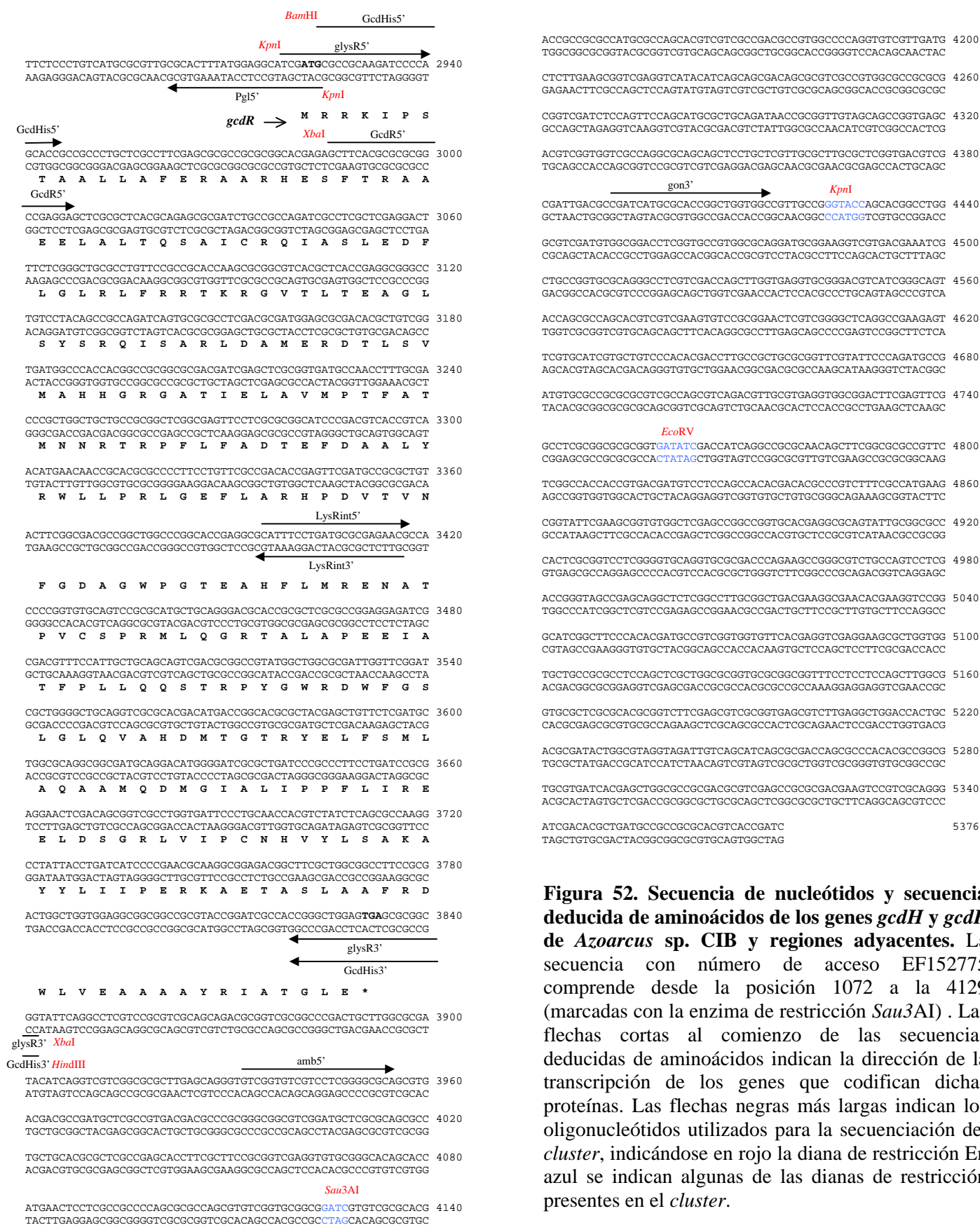


Figura 52. Secuencia de nucleótidos y secuencia deducida de aminoácidos de los genes *gcdH* y *gcdR* de *Azoarcus* sp. CIB y regiones adyacentes. La secuencia con número de acceso EF152775 comprende desde la posición 1072 a la 4129 (marcadas con la enzima de restricción *Sau3AI*). Las flechas cortas al comienzo de las secuencias deducidas de aminoácidos indican la dirección de la transcripción de los genes que codifican dichas proteínas. Las flechas negras más largas indican los oligonucleótidos utilizados para la secuenciación del *cluster*, indicándose en rojo la diana de restricción. En azul se indican algunas de las dianas de restricción presentes en el *cluster*.

VII. DISCUSIÓN

Si bien el tolueno y otros compuestos BTEX, como los xilenos, son importantes contaminantes y su degradación aeróbica ha sido ampliamente estudiada en numerosos microorganismos, el catabolismo anaeróbico de estos hidrocarburos, que juega un papel capital en el ciclo del carbono en ambientes anóxicos contaminados, ha sido mucho menos estudiado (Foght, 2008).

Aunque se han descrito algunos de los *clusters* génicos (genes *bss* y *bbs*) involucrados en la ruta periférica de degradación anaeróbica de tolueno a benzoil-CoA en distintos organismos, tales como distintas cepas de los géneros *Azoarcus*, *Thauera* y *Magnetospirillum* (anaerobios facultativos) y *Geobacter* (anaerobios estrictos) (Leuthner y Heider, 2000; Achong y col., 2001; Kane y col., 2002; Kube y col., 2004, Shinoda y col., 2005), la mayor parte de los estudios realizados están relacionados con la caracterización bioquímica de las proteínas implicadas en la activación del tolueno a bencilsuccinato (BssDCABE) y de algunas de las enzimas implicadas en la β -oxidación del bencilsuccinato a benzoil-CoA (BbsABCDEFGH) (Leuthner y Heider, 2000; Achong y col., 2001; Kube y col., 2004). Sin embargo, el análisis transcripcional y la regulación de los genes *bss* y *bbs* sólo se ha abordado de forma preliminar y parcial en unos pocos microorganismos (Coschigano y Bishop, 2004; Rabus, 2005). Por otro lado, si bien se habían implicado los genes *bss* en la activación inicial del *m*-xileno a 3-metilbencilsuccinato en la ruta periférica del *m*-xileno (Krieger y col., 1999), otros estudios sugerían la implicación de isoenzimas distintas de las proteínas Bbs como las responsables de la β -oxidación del 3-metilbencilsuccinato hasta 3-metilbenzoil-CoA (Leutwein y Heider, 2001; Leutwein y Heider, 2002). Para profundizar en el análisis transcripcional de los genes responsables de la ruta anaeróbica del tolueno, y para identificar si dichos genes están también implicados en la degradación anaeróbica del *m*-xileno en el mismo microorganismo, en esta Tesis se procedió a la caracterización de los genes responsables de la ruta periférica de degradación anaeróbica de tolueno y *m*-xileno, así como al estudio del gen responsable de una de las etapas clave de la ruta baja del benzoil-CoA, la mediada por la enzima glutaril-CoA deshidrogenasa, en una bacteria desnitrificante, *Azoarcus* sp. CIB, capaz de degradar anaeróbicamente distintos compuestos aromáticos y que ha sido objeto de numerosos estudios en nuestro laboratorio (López-Barragán y col., 2004a; Durante-Rodríguez y col., 2006; Blázquez y col., 2008).

Aunque ya había sido descrita la capacidad de *Azoarcus* sp. CIB de metabolizar anaeróbicamente tolueno y *m*-xileno (López-Barragán y col., 2004a), la cinética de este

crecimiento era desconocida. Al analizar el crecimiento anaeróbico de la cepa CIB en tolueno o *m*-xileno añadidos a una concentración de 250 mM a una fase orgánica de HMN (0,4 mM tolueno y 0,15 mM *m*-xileno en fase acuosa), se observó un comportamiento claramente diferente (Fig. 8). Así, el tiempo de duplicación de *Azoarcus* sp. CIB es de 12 h en tolueno y de 24 h en *m*-xileno. Estas tasas de duplicación son similares a las que presentan otras cepas desnitrificantes degradadoras de hidrocarburos aromáticos tales como *Azoarcus* sp. T, para la que el tiempo de duplicación es de 15 h en tolueno y 35 h en *m*-xileno (Verfürth y col., 2004), *Azoarcus* sp. EbN1, con un tiempo de duplicación de 10 h en tolueno (Wöhlbrand y col., 2008) y *T. aromatica* K172 y *Thauera* sp. DNT-1 con un tiempo de duplicación de 10-12 h en tolueno (Shinoda y col., 2004; Duldhardt y col., 2007).

Cultivos de *Azoarcus* sp. CIB crecidos en *m*-xileno alcanzan menores valores de A_{600} que los que crecen en tolueno, lo que está de acuerdo con el hecho de que el *m*-xileno es más tóxico para la célula que el tolueno, tal y como se demuestra en los experimentos que revelan una concentración mínima inhibitoria de 500-600 mM (0,8-0,9 mM en fase acuosa) para el tolueno y de 300 mM (0,19 mM en fase acuosa) para el *m*-xileno (Fig. 9). Valores similares se han descrito para otros degradadores anaeróbicos de tolueno (Wöhlbrand y col., 2008) y revelan que la sensibilidad a los hidrocarburos aromáticos es mayor para los organismos anaerobios que para los aerobios (Duldhardt y col., 2007).

Es sabido que la adaptación bacteriana a los hidrocarburos aromáticos implica distintos programas celulares, por ejemplo, cambios en la composición de los lípidos de membrana, bombas de extrusión del hidrocarburo, degradación de dicho compuesto, etc (Carmona y col., 2008). En esta Tesis se discuten algunas de las bases moleculares que han sido identificadas y que son responsables de algunas de las respuestas de *Azoarcus* sp. CIB cuando se cultiva anaeróbicamente en tolueno o *m*-xileno como únicas fuentes de carbono y energía. Inicialmente se procedió a la identificación de los responsables genéticos implicados en la ruta periférica de degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB.

1. Identificación y caracterización genética del *cluster* responsable de la ruta periférica de degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB.

Dado que *Azoarcus* sp. CIB*bzdN*, una cepa mutante que no expresa el *cluster bzd* y, por tanto, incapaz de degradar anaeróticamente benzoil-CoA (López-Barragán y col., 2004a), no mostró crecimiento anaeróbico en tolueno, era previsible la existencia en la cepa CIB de una ruta periférica que convirtiese el tolueno en benzoil-CoA similar a la ruta *bss-bbs* descrita en otros organismos desnitrificantes degradadores de tolueno (Leuthner y col., 1998; Leuthner y Heider, 2000; Kube y col., 2004). Utilizando las secuencias conocidas de los genes *bssA* de *Azoarcus* sp. T y *bbsA* de *T. aromatica* K172, se identificaron los genes ortólogos en una genoteca fágica de *Azoarcus* sp. CIB. La secuenciación de un fragmento de ADN de 29.321 pb (Nº de acceso DQ988527; Anexo I) reveló la existencia de un *cluster* génico compuesto por 24 genes organizados en tres presuntos operones: *tdiSR*, *bssDCABEFGIJKL* y *bbsABCDEFGHIJHItol*. La construcción de cepas mutantes de *Azoarcus* sp. CIB portadoras de una disrupción insercional en distintos genes del *cluster* (*tdiR*, *tdiS*, *bssD* y *bbsB*), y la incapacidad de estos mutantes para utilizar tolueno como única fuente de carbono, indicaba que el *cluster* identificado estaba efectivamente implicado en la degradación anaeróbica de tolueno en *Azoarcus* sp. CIB. La asignación de funciones a cada uno de los genes del *cluster* (Tabla 5) y la ruta bioquímica propuesta para el catabolismo anaeróbico del tolueno en *Azoarcus* sp. CIB (Fig. 6), se realizó en base a los estudios bioquímicos y genéticos realizados previamente sobre la ruta periférica equivalente de las cepas *T. aromatica* T1 y K172 y *Azoarcus* sp. T y EbN1 (Leuthner y col., 1998; Leuthner y Heider, 2000; Achong y col., 2001; Kube y col., 2004; Coschigano y Bishop, 2004).

Los genes *bssC*, *bssA* y *bssB* codifican las subunidades γ (BssC), α (BssA) y β (BssB) de la enzima heterohexamérica bencilsuccinato sintasa ($\alpha_2\beta_2\gamma_2$) (Leuthner y col., 1998), las cuales presentan una significativa identidad de secuencia con sus ortólogos en otros organismos degradadores de tolueno (Fig. 53). La enzima bencilsuccinato sintasa cataliza el primer paso de la degradación anaeróbica del tolueno y *m*-xileno con la incorporación al fumarato del grupo metilo del tolueno para formar bencilsuccinato (Leuthner y col., 1998; Coschigano y col., 1998; Beller y Spormann, 1999; Krieger y col., 2001), o de uno de los grupos metilo del *m*-xileno para formar 3-metilbencilsuccinato (Krieger y col., 1999; Achong y col., 2001). La bencilsuccinato sintasa requiere anaerobiosis para su actividad (Biegert y col., 1996; Beller y Spormann,

1999; Leuthner y col., 1998; Selmer y col., 2005). En *T. aromatica* T1, la mutagénesis dirigida de los genes *tutF* (*bssC*), *tutD* (*bssA*) y *tutG* (*bssB*) ha permitido determinar los residuos críticos para la actividad de la enzima bencilsuccinato sintasa (Bhandare y col., 2006). Los residuos del centro activo de la bencilsuccinato sintasa están perfectamente conservados en las subunidades BssA de los distintos organismos, y se corresponden con Cys-9, Ala-10, Cys-496 y Gly-828 de la proteína BssA de *Azoarcus* sp. CIB.

El gen *bssD* codifica un activador enzimático de la bencilsuccinato sintasa en una reacción dependiente de S-adenosilmetionina (Selmer y col., 2005). La proteína BssD de la cepa CIB presenta una elevada identidad con sus ortólogos en otros organismos degradadores de tolueno (Fig. 53). Todas las proteínas BssD presentan tres motivos en su región N-terminal ricos en residuos Cys y característicos de proteínas dependientes de S-adenosilmetionina (Sofía y col., 2001). La mutagénesis dirigida de estos residuos Cys impide la activación de la proteína bencilsuccinato sintasa por el activador BssD (Bhandare y col., 2006). El primero de los motivos (Cys-X₃-Cys-X₂-Cys) se localiza en las posiciones 29-36, los otros dos motivos (Cys-X₂-Cys-X₂-Cys-X₃-Cys) se localizan en las posiciones 55-65 y 89-99 de la proteína BssD de la cepa CIB.

El producto del gen *bssE* presenta una alta identidad con proteínas ortólogas de otras bacterias (Fig. 53) y contiene un dominio tipo "Walker" de interacción con nucleótidos (ATP/GTP), similar al que presentan chaperonas que intervienen en el ensamblaje de complejos proteicos (Neuwald y col., 1999), lo cual sugiere que puede ser requerido para el ensamblaje de la bencilsuccinato sintasa heterohexamérica (Hermuth y col., 2002; Boll y col., 2002). Al realizar mutaciones en los residuos de Gly, Lys y Ser (posiciones 52-54) del posible dominio ATP/GTP en el producto del gen *tutH* de *T. aromatica* T1, se originan cepas mutantes incapaces de utilizar tolueno, demostrando que esta proteína es esencial para el metabolismo anaeróbico de este hidrocarburo. (Bhandare y col., 2006).

Los productos de los genes *bssF*, *bssG*, *bssI*, *bssJ*, *bssK* y *bssL* codifican proteínas de función desconocida. Algunos de estos genes también están presentes en otros *clusters bss/bbs* (Fig. 53), si bien hasta la fecha no se sabía si formaban parte del mismo operón *bss*. Los resultados presentados en esta Tesis revelan que dichos genes forman parte, junto con *bssDCABE*, del mismo operón catabólico (Fig. 22) lo que sugiere fuertemente su implicación en el catabolismo anaeróbico del tolueno.

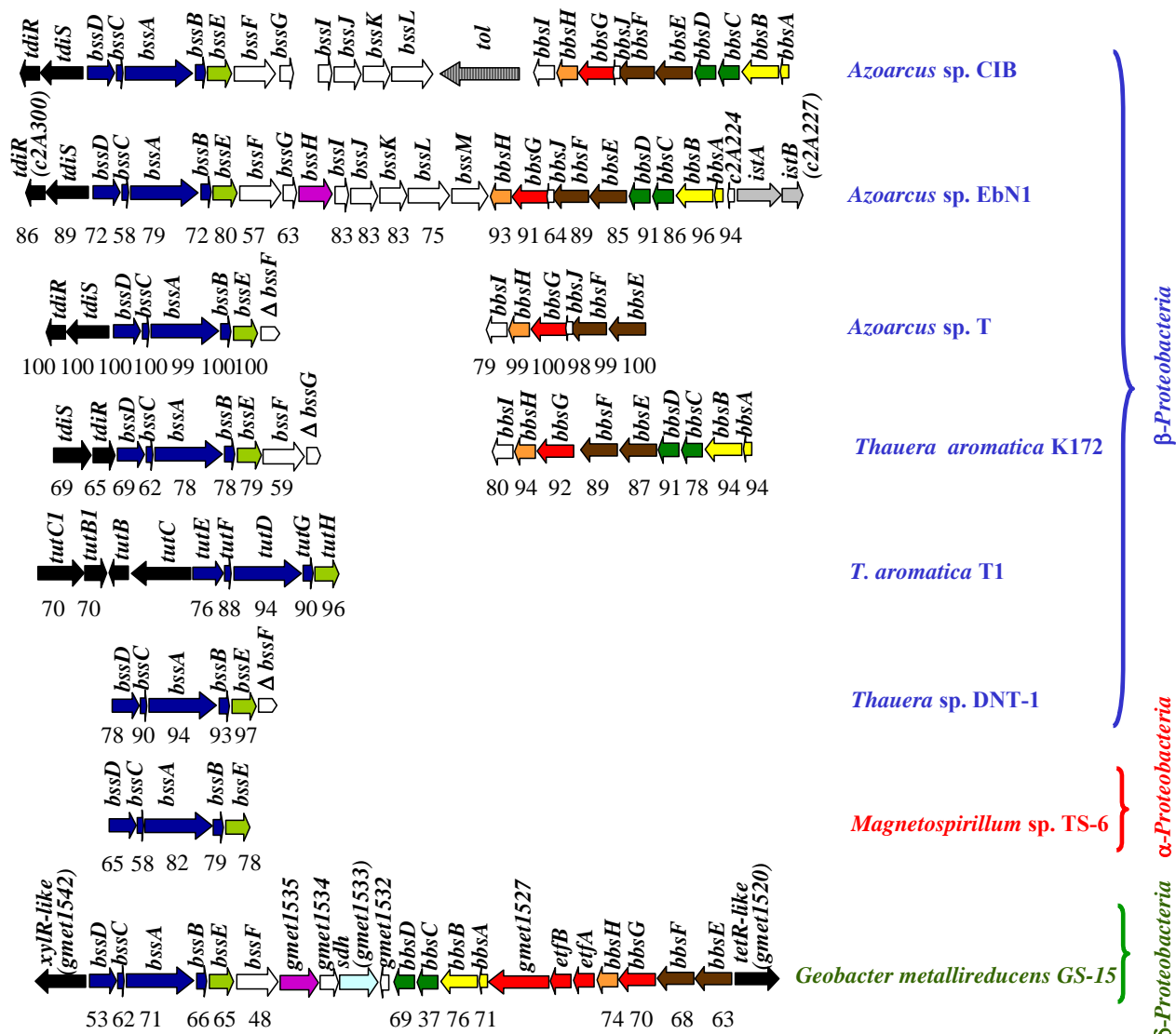


Figura 53. Organización de los clusters génicos responsables del catabolismo anaeróbico de tolueno en las cepas *Azoarcus* sp. CIB, *Azoarcus* sp. EbN1 (NC_006513), *Azoarcus* sp. T (AY032676, AY732214), *T. aromatica* K172 (AJ001848, AF173961), *T. aromatica* T1 (U57900, AF113168), *Thauera* sp. DNT-1 (AB066263), *Magnetospirillum* sp. TS-6 (AB167725) y *G. metallireducens* GS-15 (NC_007517). Los genes están representados por flechas: azul oscuro, genes que codifican la enzima bencilsuccinato sintasa; marrón, genes que codifican la succinil-CoA:(R)-bencilsuccinato CoA transferasa, rojo, genes que codifican la (R)-bencilsuccinato-deshidrogenasa (*bbsG*) y un posible sistema acoplado receptor de electrones; naranja, genes que codifican la posible fenil-itaconil-CoA hidratasa; verde oscuro, genes que codifican la posible 2-[hidroxi(fenil)metil]-succinil-CoA deshidrogenasa; amarillo, genes que codifican la posible benzoilsuccinil-CoA tiolasa; verde claro, genes que codifican una posible chaperona de la bencilsuccinato sintasa; azul claro, genes que codifican una posible succinato-deshidrogenasa, violeta, genes que codifican un hipotético sistema de transporte de tolueno; negro, genes reguladores; gris, genes que codifican transposasas; líneas verticales, actividad fosfodiesterasa; blanco, genes con función desconocida. Las regiones entre los clusters *bss* y *bbs* de *Azoarcus* sp. T y *T. aromatica* K172 no han sido secuenciadas aún. El gen *bbsJ* de *Azoarcus* sp. T se ha identificado en este trabajo y está localizado dentro de la región anotada como *orfA* (Nº. Acceso AY732214). Debajo de los genes se indica el % de identidad de secuencia del correspondiente producto génico con su ortólogo en la cepa CIB. Δ , indica que el gen no se ha secuenciado completamente.

Los genes *bbsF* y *bbsE* se prevé que codifiquen las dos subunidades de la enzima succinil-CoA:bencilsuccinato CoA transferasa en *Azoarcus* sp. CIB. Esta enzima interviene en el primer paso de la β -oxidación del bencilsuccinato activando éste a benzoilsuccinil-CoA mediante la adición de un grupo CoA (fig. 6) (Leutwein y Heider, 2001). La enzima muestra especificidad hacia el enantiómero (*R*)-bencilsuccinato. La actividad succinil-CoA:bencilsuccinato CoA transferasa ha sido descrita en diferentes cepas bacterianas cuando se cultivan en tolueno o *m*-xileno, siendo esta actividad mayor cuando la cepa utiliza el *m*-xileno (Leutwein y Heider, 2001).

El producto del gen *bbsG* posee una elevada identidad con la (*R*)-bencilsuccinil-CoA deshidrogenasa de otros microorganismos degradadores de tolueno (Fig. 53). Esta enzima sería la encargada de transformar el bencilsuccinil-CoA en fenilitaconil-CoA (Leutwein y Heider., 2002).

El producto del gen *bbsH* muestra una identidad de secuencia significativa con miembros de la familia enoil-CoA hidratasa/isomerasa (Fig. 53). La proteína BbsH de *T. aromatica* se ha predicho como la fenil-itaconil-CoA hidratasa encargada de la adición de una molécula de H₂O al fenil-itaconil-CoA para producir 2-[Hidroxi(fenil)metil]-succinil-CoA (Leuthner y Heider., 2000).

Los genes *bbsC* y *bbsD* codifican presuntamente las dos subunidades de la enzima 2-[Hidroxi(fenil)metil]succinil-CoA deshidrogenasa. Se ha propuesto la participación de ésta en el penúltimo paso de la β -oxidación oxidando el 2-[Hidroxi(fenil)metil]succinil-CoA y transformándolo en benzoilsuccinil-CoA (Fig. 53) (Leuthner y Heider, 2000).

Los productos de los genes *bbsA* y *bbsB* presentan una elevada identidad (94%-96%) con proteínas ortólogas de las cepas *Azoarcus* sp. EbN1 y *T. aromatica* K172 (Fig. 53). Se ha propuesto que BbsA y BbsB son las dos subunidades de la benzoilsuccinil-CoA tiolasa responsable del último paso de la β -oxidación, generando succinil-CoA y benzoil-CoA (Fig. 6) (Leuthner y Heider, 2000; Kube y col.,2004).La proteína BbsB posee un residuo Cys (Cys-84 en la proteína de la cepa CIB) conservado en todas las tiolasas y que se corresponde con el sitio de formación del intermediario covalente de la reacción.

El producto del gen *bbsI* codifica una proteína con función desconocida, aunque posee un dominio conservado de unión a fosfatidiletanolamina (PEBP) (Serre y col., 1998) Si bien la proteína BbsI presenta una identidad del 79% y 80% con sus ortólogos

en *Azoarcus* sp. T y *T. aromatica* K172 (Fig. 53), en *Azoarcus* sp. EbN1 la proteína que muestra una mayor identidad con BbsI está codificada por un gen (ebA5090) que no se encuentra dentro de ningún *cluster* de degradación de compuestos aromáticos.

Particularmente interesante fue observar que ninguno de los mutantes de *Azoarcus* sp. CIB en los genes *tdi*, *bss* y *bbs* fue capaz de utilizar el *m*-xileno, lo que implicaba que estos genes son también imprescindibles para la degradación anaeróbica de este hidrocarburo. A pesar de que inicialmente se consideró la posibilidad de que los genes *bbs* eran los encargados de transformar el 3-metilbencilsuccinato hasta 3-metilbenzoil-CoA, estudios bioquímicos posteriores sugerían la existencia de diferentes isoenzimas Bbs para la β -oxidación del bencilsuccinato y del 3-metilbencilsuccinato, lo que implicaría la participación de genes adicionales, y distintos a los *bbs*, para el catabolismo anaeróbico del *m*-xileno (Leutwein y Heider, 2002). Los resultados de esta Tesis confirman la participación de la bencilsuccinato sintasa (genes *bss*) en la etapa inicial de la activación del *m*-xileno a 3-metilbencilsuccinato (Achong y col., 2001), y representan la primera demostración experimental de que al menos alguno de los genes *bbs* son necesarios para la oxidación del 3-metilbencilsuccinato. Como se había sugerido previamente (Krieger y col., 1999), el producto final de la β -oxidación del 3-metilbencilsuccinato es, muy probablemente, 3-metilbenzoil-CoA (Fig. 54). La formación de benzoil-CoA y 3-metilbenzoil-CoA como productos finales de la ruta periférica de tolueno y *m*-xileno, respectivamente, está de acuerdo con el hecho de que la cepa mutante *Azoarcus* sp. CIB*dbzdN* no es capaz de crecer anaeróbicamente en tolueno pero sí en *m*-xileno, lo que indica la existencia de una ruta central adicional, y distinta a la ruta *bzd*, para la degradación anaeróbica del 3-metilbenzoil-CoA. En este sentido, la reciente secuenciación del genoma de *Azoarcus* sp. CIB ha permitido la identificación de un nuevo *cluster* génico (genes *mbd*), localizado en las proximidades del *cluster* *bss/bbs* que codifica la primera ruta central que se describe para la degradación anaeróbica del 3-metilbenzoil-CoA y que, por lo tanto, está implicado en el catabolismo anaeróbico del *m*-xileno en la cepa CIB. Como cabría esperar, un mutante de la cepa CIB portador de una disrupción del *cluster* *mbd* es incapaz de utilizar 3-metilbenzoato y *m*-xileno como únicas fuentes de carbono (J.F. Juárez, Tesis doctoral), lo que refuerza la hipótesis de que la degradación anaeróbica de tolueno y *m*-xileno involucra distintas rutas centrales para el metabolismo del benzoil-CoA y 3-metilbenzoil-CoA, respectivamente (Fig. 54). La necesidad de una ruta central específica del 3-metilbenzoil-CoA y distinta a la ampliamente extendida ruta central del

benzoil-CoA permite explicar que se hayan aislado numerosos organismos anaerobios degradadores de tolueno pero que tan sólo unos pocos sean capaces de mineralizar también *m*-xileno (Foght, 2008).

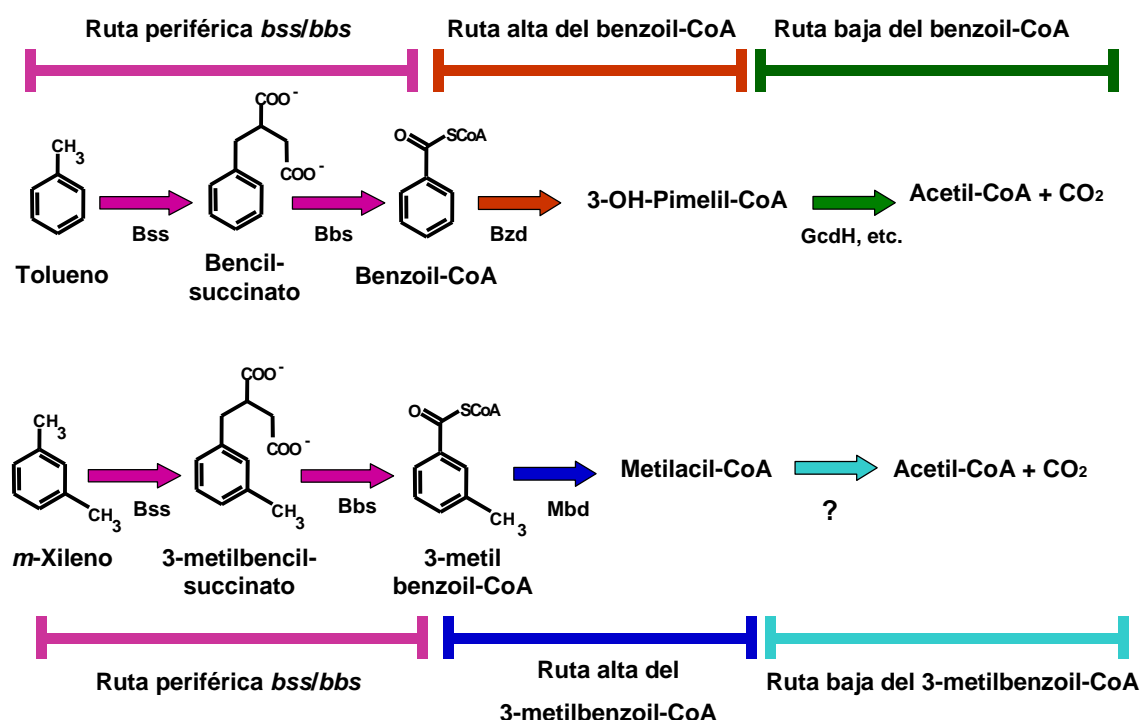


Figura 54. Representación esquemática de las rutas de degradación anaeróbica propuestas para el tolueno y *m*-xileno en *Azoarcus* sp. CIB.

2. Análisis del gen *tol*.

De todos los genes presentes en el *cluster bss/bbs* de *Azoarcus* sp. CIB, el único que no posee ortólogo en ninguno de los *clusters* equivalentes descritos hasta la fecha es el gen *tol*.

El gen *tol* codifica una proteína de 1.049 aminoácidos y función desconocida. La proteína Tol presenta dos regiones bien diferenciadas, la región N-terminal presuntamente sensora, y la región C-terminal presuntamente catalítica, que se organizan en diferentes dominios (Fig. 12). La peculiar arquitectura modular de la proteína Tol no ha sido descrita hasta la fecha en ninguna proteína. En la región N-terminal de Tol se localizan un dominio PAS, un dominio HK y un dominio REC (Fig. 55) que posee una organización modular similar y una identidad significativa con los dominios ortólogos presentes en las proteínas TutC, StyS y TodS (Fig. 12). La superfamilia PAS incluye diferentes dominios con un núcleo estructural común, y con capacidad de reconocer y responder a distintas señales medioambientales tales como

potenciales rédox, luz, niveles de oxígeno y pequeñas moléculas (Mascher y col., 2006). El dominio HK posee el residuo His-191 conservado y susceptible de fosforilación así como el motivo catalítico (HATPasa) que contiene las cajas N,D,F y G (aminoácidos 304-405) responsables de la autofosforilación del dímero histidín-quinasa (Wolanin y col., 2002). El dominio REC posee los residuos conservados que caracterizan a este tipo de dominios y que incluyen, entre otros, al residuo Asp-499 que se fosforila utilizando como donador de fosfato el motivo HK fosforilado (Pao y Saeier., 1995; Mascher y col., 2006). La región C-terminal de la proteína Tol presenta dos dominios, DGC y EAL (Fig. 55 y Fig. 12), que se han visto implicados en el metabolismo celular del dinucleótido cíclico di-GMPc. El dominio DGC se ha observado en proteínas con actividad diguanilato ciclasa, implicadas en la síntesis de di-GMPc a partir de dos moléculas de GTP. El dominio EAL se encuentra en diversas proteínas con actividad fosfodiesterasa y contiene los residuos catalíticos implicados en la hidrólisis de di-GMPc a pGpG (Römling y col., 2005; Römling y Amikan, 2006). En los últimos años, se ha visto que el di-GMPc tiene un papel de segundo mensajero dentro de la célula, interviniendo en procesos de formación de biopelículas, movilidad, virulencia, supervivencia, etc. (Römling y col., 2005; Römling y Amikan, 2006).

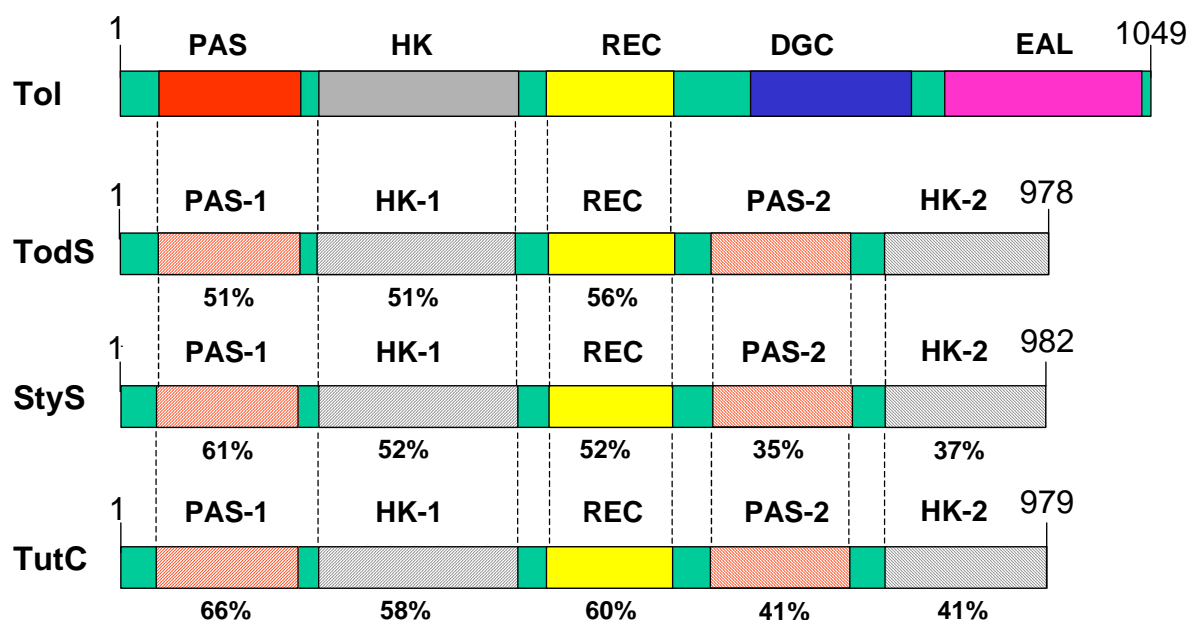


Figura 55. Esquema de la arquitectura modular de las proteínas Tol, StyS, TodS y TutC. Se indican los dominios presentes en las diferentes proteínas: dominios sensores (PAS-1 y PAS-2), dominios histidín-quinasa (HK-1 y HK-2), dominio *receiver* (REC), dominio diguanilato ciclasa (DGC) y dominio fosfodiesterasa (EAL). El diferente rayado de los dominios PAS-1/HK-1 y PAS-2/HK-2 indica que éstos se diferencian significativamente en su estructura primaria. Debajo de los dominios PAS-1, HK-1 y REC se indica el % de identidad con los dominios ortólogos de la proteína TOL. Debajo de los dominios PAS-2 e HK-2 se indica el % de identidad de secuencia con los dominios equivalentes de la proteína TodS. Las proteínas StyS, TodS y TutC indicadas en la figura pertenecen a las cepas *Pseudomonas* sp. Y2 (CAA03998), *P. putida* DOT-T1E (AAC45438) y *T. aromatica* T1 (AAD12184), respectivamente.

Las histidín-quinasas TutC, TodS y StyS constituyen el componente sensor del sistema regulador de dos componentes que controla la expresión de los genes *tut*, *tod* y *sty*, respectivamente, implicados en el catabolismo aeróbico de tolueno (*tut*, *tod*) y estireno (*sty*) en distintas cepas de los géneros *Pseudomonas* y *Thauera* (Tropel y van der Meer, 2004; Carmona y col., 2008). Estas tres histidín-quinasas híbridas poseen una arquitectura molecular similar y consistente en un supradominio N-terminal que incluye los dominios PAS-1/HK-1/REC, y un supradominio C-terminal que comparte el dominio REC e incluye también a los dominios PAS-2/HK-2 (Tropel y van der Meer, 2004; Carmona y col., 2008). Dentro de este grupo de proteínas, la proteína TodS ha sido la más estudiada. El dominio PAS-1 de TodS es el responsable de la interacción con el tolueno, y también con sus análogos tales como benceno, xilenos y estireno, siendo los residuos Phe-46, Ile-74, Phe-79, y Ile-114 los encargados de dicha interacción en el modelo de estructura propuesto para dicho dominio, y tal y como se ha confirmado experimentalmente mediante su mutagénesis (Busch y col., 2007). Estos cuatro residuos se encuentran también conservados en los dominios PAS-1 de TutC y StyS, así como en el dominio PAS de la proteína Tol (Phe-46, Ile-74, Phe-79, y Val-114), sugiriendo que dichos dominios PAS están también implicados en la interacción con hidrocarburos aromáticos (e.g., tolueno, xilenos, estireno), y que dicha interacción es la que induce la autofosforilación del correspondiente dominio HK en el dímero de la proteína sensora. El dominio REC de las proteínas TodS, TutC, StyS y Tol sería el encargado de recibir el grupo fosfato del dominio HK-fosforilado, lo que, a su vez, se traducirá en un cambio conformacional que permitirá transmitir la señal detectada a otros dominios efectores de la respuesta.

La región C-terminal de la proteína Tol posee los dominios DGC y EAL (Fig. 55 y Fig. 12). Al analizar con más detalle el dominio DGC de la proteína Tol, se puede observar que la secuencia consenso del centro activo GG(D/E)F se encuentra sustituida por la secuencia de aminoácidos 696-GDREF-702 (Fig. 12). Existen en la literatura otros ejemplos de dominios DGC con una secuencia del centro activo alterada y que no presentan actividad diguanilato ciclasa, aunque son capaces de unir GTP contribuyendo a la regulación positiva del dominio EAL adyacente (Christen y col., 2005). Por otro lado, el análisis de la secuencia del dominio EAL de la proteína Tol revela la conservación de todos los residuos clave en la interacción con el di-GMPc y en su hidrólisis al dinucleótido lineal pGpG (actividad fosfodiesterasa A) (Schmidt y col., 2005). Por lo tanto, los análisis de la estructura primaria de la región C-terminal de la

proteína Tol sugieren que ésta no posee actividad diguanilato ciclasa pero sí actividad fosfodiesterasa A que hidroliza el di-GMPc a pGpG. El pGpG será posteriormente hidrolizado a dos moléculas de GMP por otras fosfodiesterasas B de la célula (Tamayo y col., 2005). La actividad fosfodiesterasa de di-GMPc de la proteína Tol se ha confirmado en esta Tesis mediante la complementación del mutante *E. coli* TRMG que posee niveles elevados de di-GMPc, observándose una reducción de la formación de biopelículas en la cepa que expresa el gen *tol* (Fig. 17). Un resultado similar se obtuvo cuando la cepa *E. coli* TRMG se complementó con el gen *PA2133* de *P. aeruginosa*, el cual codifica una fosfodiesterasa de di-GMPc previamente caracterizada (Kulesekara y col., 2006) (Fig. 17). La reversión de la capacidad incrementada de la cepa TRMG para formar biopelículas, resultado del incremento en los niveles de di-GMPc en la célula al tener mutado el gen *csrA* que controla negativamente proteínas con actividad diguanilato ciclasa. Mediante la clonación de genes que codifican fosfodiesterasas de di-GMPc y disminuyen, por lo tanto, los niveles celulares de éste, ya se ha descrito en la literatura (Jonas y col., 2008) y, por ello, es un ensayo válido para detectar actividad fosfodiesterasa de di-GMPc *in vivo*.

En la literatura se han descrito varios ejemplos de mutaciones en dominios EAL que implican la inactivación de proteínas con actividad di-GMPc fosfodiesterasa en distintos géneros bacterianos tales como *Pseudomonas*, *Salmonella* o *Vibrio* (Jenal y Malone, 2006). Los mutantes correspondientes muestran niveles más elevados de di-GMPc, lo que se traduce en un significativo cambio en el estilo de vida de las células que pasan de ser esencialmente móviles a formar biopelículas, un fenotipo similar al que se obtiene cuando se hiperexpresan proteínas con actividad diguanilato ciclasa (Tamayo y col., 2007). En este sentido, y como cabría esperar de un mutante en una proteína con actividad di-GMPc fosfodiesterasa, *Azoarcus* sp. CIB*dtol* mostró un claro defecto en la movilidad y taxis frente a compuestos aromáticos, tales como el tolueno, en comparación con el fenotipo observado en la cepa parental *Azoarcus* sp. CIB (Figs. 14 y 16). La quimiotaxis hacia fuentes de carbono es un mecanismo de adaptación celular a la utilización de determinados compuestos como fuente de carbono y energía, por lo que muy posiblemente el gen *tol* en *Azoarcus* sp. CIB esté implicado en la adaptación de esta bacteria al catabolismo anaeróbico de ciertos hidrocarburos aromáticos, lo que explicaría la localización de este gen en el *cluster bss/bbs* y la inducción específica cuando las células se cultivan anaeróticamente en presencia de tolueno/*m*-xileno (Figs. 22 y 23). Cabría destacar aquí que la implicación del gen *tol* en la adaptación de la cepa

CIB a los hidrocarburos aromáticos muy probablemente implique también cambios en las envueltas celulares, e.g., composición de los lípidos y proteínas de membrana, expresión de bombas de expulsión de solventes, etc., que explicarían el hecho de que la cepa parental sea más resistente que la cepa mutante en el gen *tol* a la exposición súbita a una concentración tóxica de tolueno/*m*-xileno (J. Ramos Vivas, comunicación personal). La adaptación mediada por el gen *tol* también podría implicar cambios en el metabolismo general de la célula, lo que explicaría que la cepa mutante *Azoarcus* sp. CIB*dtol* crezca más lentamente que la cepa parental en distintas fuentes de carbono, incluyendo el tolueno y *m*-xileno, aunque la presencia del gen *tol* no sea un requisito indispensable para el crecimiento en estas fuentes de carbono.

Si bien se había observado un cambio en la expresión de genes que codifican proteínas con motivos DGC-EAL cuando *P. putida* se cultiva en presencia de compuestos aromáticos (Domínguez-Cuevas y col., 2006), así como un defecto general en la motilidad de *P. putida* cuando se inactiva la proteína PP2258, cuya arquitectura modular consta de los dominios PAS-DGC-EAL, que afecta a la quimiotaxis dependiente de metabolismo hacia el fenol (Sarand y col., 2008), la implicación directa de una actividad di-GMPc fosfodiesterasa en la respuesta al catabolismo de hidrocarburos aromáticos no se había descrito hasta la fecha y, por ello, constituye un importante resultado de esta Tesis.

Respecto al mecanismo de acción del di-GMPc en la célula, se ha sugerido que controla distintas funciones celulares a lo largo del tiempo, a través de la expresión diferencial de proteínas con motivos DGC/EAL, y del espacio, mediante cambios en la concentración local de este dinucleótido en puntos concretos de la célula (Jenal y Malone, 2006). A pesar de que el di-GMPc controla numerosos procesos vitales en la célula, hasta la fecha se han descrito sólo algunos motivos que reconocen e interaccionan específicamente con este segundo mensajero: i) el motivo I de algunas DGC que unen di-GMPc para inhibir alostéricamente su actividad ciclasa; ii) la proteína PelD de *P. aeruginosa* que controla la producción de exopolisacárido (Lee y col., 2007). iii) el regulador transcripcional FleQ que controla la expresión de los genes de biosíntesis de exopolisacárido en *P. aeruginosa* (Hickman y Harwood, 2008); iv) el “*riboswitch*” que detecta di-GMPc y que controla la expresión de algunos ARNm que codifican proteínas DGC/EAL o proteínas implicadas en virulencia, biosíntesis de flagelos y pili, etc (Sudarsan y col., 2008); v) el dominio PilZ presente en un gran número de bacterias (Ryjenkov y col., 2006). Así, en el genoma de *Azoarcus* sp. EbN1

y *Azoarcus* sp. BH72 se detectan al menos 1 y 4 proteínas con dominios PilZ, respectivamente. En algunas bacterias, como *E. coli*, las proteínas con dominios PilZ controlan procesos de motilidad y la síntesis de celulosa, en otras bacteria, como *Vibrio cholerae*, controlan la formación de biopelículas y la virulencia (Pratt y col., 2007).

El posible receptor de los niveles de di-GMPc controlados por la proteína Tol en *Azoarcus* sp. CIB, y la cascada de señalización que transduce el cambio en los niveles de di-GMPc a la respuesta celular que determina la adaptación al consumo de hidrocarburos aromáticos, son todavía desconocidos y serán objeto de estudio en los próximos años

La compleja organización molecular de la proteína Tol sugiere que se ha originado en el curso de la evolución como resultado de la fusión de un supradominio sensor N-terminal y otro catalítico C-terminal (Fig. 56). El supradominio sensor podría derivar de las histidín-quinasas híbridas que controlan la regulación de los genes implicados en el catabolismo aeróbico de hidrocarburos aromáticos en distintas bacterias. La detección de estos hidrocarburos (e.g., tolueno y *m*-xileno) por parte del dominio PAS implicaría la autofosforilación del dominio HK y la correspondiente activación (fosforilación) del dominio REC. El cambio conformacional inducido en el dominio REC implicaría, a su vez, un cambio conformacional del supradominio catalítico con la consiguiente inducción de la actividad di-GMPc fosfodiesterasa por parte del dominio EAL (Fig. 56). El papel del dominio DGC mediando la transferencia de la señal desde el supradominio sensor y/o modulando la actividad del dominio EAL es una hipótesis plausible que necesita ser confirmada experimentalmente.

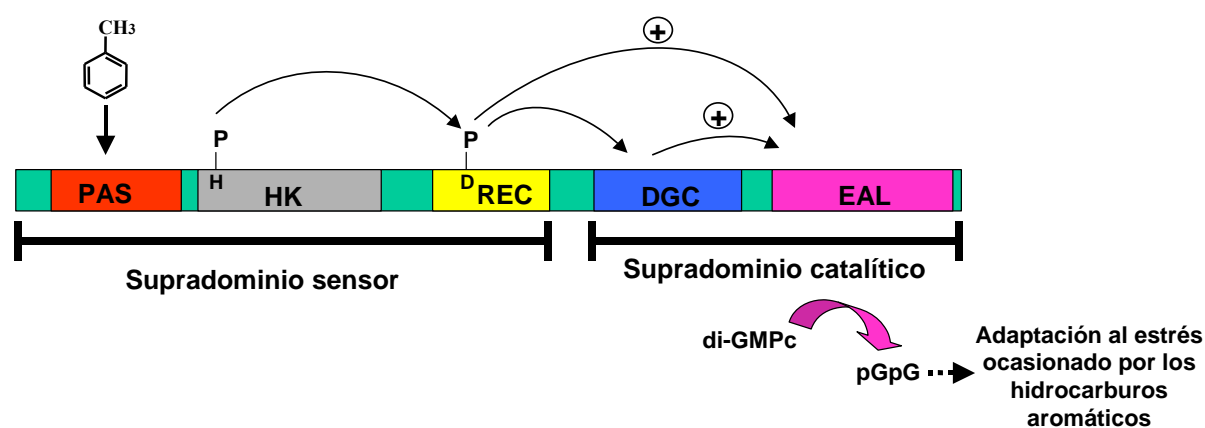


Figura 56. Modelo propuesto para la activación de la proteína Tol. Se representan los supradominios de la proteína. El dominio PAS al reconocer un hidrocarburo aromático (e.g., tolueno, *m*-xileno) dispara una cascada de fosforilación intramolecular que activa finalmente al dominio EAL e implica la hidrólisis del di-GMPc. La bajada de los niveles de di-GMPc dispara, a su vez, una cascada de señalización que facilita la adaptación de la bacteria a los hidrocarburos aromáticos.

La hidrólisis del di-GMPc por parte de la proteína Tol activa supondría una disminución en los niveles celulares de éste, lo que disparará una cascada de señalización celular que finalmente permitirá a la célula adaptarse al estrés ocasionado por los hidrocarburos aromáticos (Fig. 56). La respuesta celular a dicho estrés implicará, entre otros procesos, cambios en la movilidad y quimiotaxis, así como muy probablemente cambios en las envueltas celulares y en el metabolismo general de la célula.

Aunque las hipótesis formuladas en esta Tesis sobre la activación intramolecular de la proteína Tol, así como sobre la consiguiente cascada de señalización inducida por los cambios en los niveles de di-GMPc celulares, necesitan ser confirmadas y los mecanismos moleculares implicados estudiados en detalle, el trabajo que aquí se presenta apunta a la existencia de un nuevo sistema de transducción de señal no descrito hasta la fecha. Así, la proteína Tol constituiría el primer ejemplo de un sistema de transducción de señal monocomponente que, sin embargo, utiliza el mecanismo clásico de las histidín-quinasa que caracteriza a los sistemas de dos componentes (Leuthner y Heider, 1998), siendo por ello, de gran interés para el estudio de la evolución de estos sistemas de transducción de señales en procariotas.

3. Genómica comparada y relaciones evolutivas de los *clusters bss/bbs* en diferentes microorganismos.

Al comparar la organización génica del *cluster bss-bbs* de degradación anaeróbica de tolueno y *m*-xileno de *Azoarcus* sp. CIB con la de los *clusters* equivalentes de otras bacterias como *Azoarcus* sp. EbN1, *Azoarcus* sp. T, *T. aromatica* K172, *T. aromatica* T1, *Thaura* sp. DNT-1, *Magnetospirillum* sp. TS-6 y *G. metallireducens* GS-15 (Fig. 53), se pueden observar algunos patrones de organización comunes a todos los *clusters* y otros que difieren entre distintos grupos de *clusters* génicos.

El orden de los genes en el *cluster bssDCABE* se conserva en todos los organismos degradadores de tolueno identificados hasta la fecha (Fig. 53). Sin embargo al analizar la distancia intergénica entre los genes *bssB* y *bssE* de los siete organismos anaerobios facultativos se distinguen dos subgrupos de *clusters bss*. Así, en un subgrupo estarían incluidos los *clusters bss* de las cepas *Azoarcus* sp. EbN1, *T. aromatica* K172 y *Magnetospirillum* sp. TS-6 que presentan una distancia de 100-122 pb entre los genes

bssB y *bssE* y en la que se han detectado elementos de estructura secundaria (bucles) implicados en el control transcripcional (Hermuth y col., 2002; Kube y col., 2004). En el otro subgrupo estarían incluidos los *clusters* de las cepas *Azoarcus* sp. CIB, *Azoarcus* sp. T, *T. aromatica* T1, y *Thauera* sp. DNT-1, con una distancia intergénica entre el gen *bssB* y *bssE* de 47 pb. Estos dos mismos subgrupos también surgen al comparar la estructura primaria de las proteínas BssDCABE (Fig. 57), lo que sugiere una historia evolutiva diferente para cada uno de ellos y, a su vez, diferente a la de la bencilsuccinato sintasa de la bacteria anaerobia estricta *G. metallireducens*.

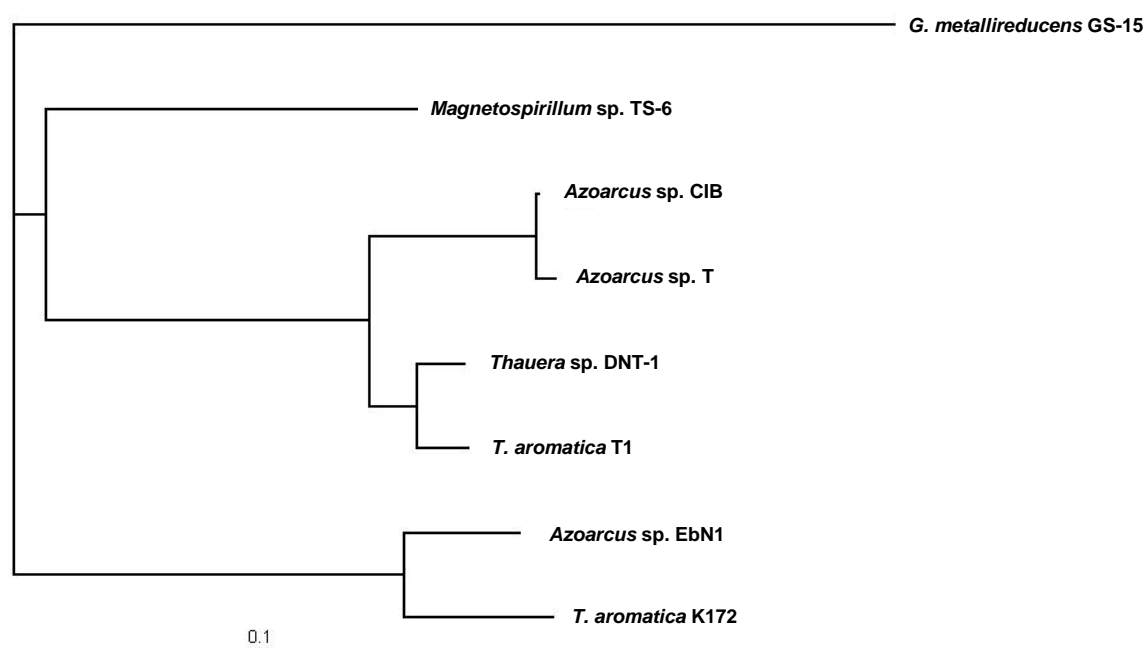


Figura 57. Árbol filogenético de las proteínas BssA de distintas bencilsuccinato sintasas. El alineamiento múltiple de las proteínas BssA de *Azoarcus* sp. CIB, *Azoarcus* sp. EbN1 (YP_158060), *Azoarcus* sp. T (AAK50372), *T. aromatica* K172 (CAA05052), *T. aromatica* T1 (AAC38454), *Thaura* sp. DNT-1 (BAC05501), *Magnetospirillum* sp. TS-6 (BAD42366) y *G. metallireducens* GS-15 (YP_384498) se realizó con el programa CLUSTALW y la visualización del árbol filogenético se realizó con el programa TreeView.

Al comparar la organización génica del *cluster bss* de *Azoarcus* sp. CIB (Fig. 53) con la del *cluster* equivalente de *Azoarcus* sp. EbN1, se puede observar que en la cepa CIB no están presentes los genes *bssH* (posible transportador) y *bssM* (función desconocida). En la posición correspondiente al gen *bssH* de la cepa EbN1 se localiza una región intergénica *bssG-I* (0,8 kb) en la cepa CIB. Esta región intergénica es precisamente la que divide el *cluster bss/bbs* de la cepa CIB en dos regiones atendiendo a su contenido G+C: una región 5' (*tdiRS-bssDCABEFG*) de bajo contenido G+C ($\leq 58\%$) y una región 3' (*bssIJKLtolbbsI-A*) cuyo contenido G+C ($\approx 65\%$) (Tabla 5) es

similar al contenido G+C promedio del genero *Azoarcus* ($\approx 65\%$). Esta diferencia en el contenido G+C sugiere que estas dos regiones del *cluster bss-bbs* de *Azoarcus* sp. CIB han tenido orígenes diferentes y se han fusionado en el transcurso de la evolución. Merece la pena resaltar el hecho de que los genes *bssDCABE* del mismo subgrupo al que pertenecen los de la cepa CIB también poseen un contenido G+C ($<57\%$) significativamente inferior al de los genes *bbs* (64% en el caso de *T. aromatica* T1) y al contenido G+C promedio de los genomas correspondientes (aprox. 65% G+C para las cepas de *Azoarcus* y *Thauera*). Por el contrario, el contenido G+C de los genes *bss* del subgrupo que forman *Azoarcus* sp. EbN1, *T. aromatica* K172 y *Magnetospirillum* sp. TS-6 es superior al 61% y más próximo al contenido G+C de los genes *bbs* y del contenido G+C promedio del genoma. Estos análisis apuntan, una vez más, a que los genes *bssDCABEFG* de *Azoarcus* sp. cepas T y CIB, *T. aromatica* T1 y *Thauera* sp. DNT-1 tienen un origen evolutivo común diferente al de sus equivalentes en *Magnetospirillum*, *Azoarcus* sp. EbN1 y *T. aromatica* K172, y muy probablemente derivan de un microorganismo cuyo contenido G+C es claramente inferior al de las bacterias de los géneros *Thauera* y *Azoarcus*.

Los genes *bssGIJKL* (de función desconocida) presentes en la cepa *Azoarcus* sp. CIB no se encuentran presentes en *G. metallireducens* donde están sustituidos por los genes *gmet1535* (codifica una proteína de membrana externa posiblemente implicada en el transporte de tolueno), *gmet1534* (función desconocida) y *gmet1533* (codifica una posible succinato-deshidrogenasa implicada en la regeneración del fumarato) (Kube y col., 2004).

El ordenamiento génico del *cluster bbs* (*bbsABCDEFJGH*) se mantiene conservado en todos los organismos anaerobios facultativos, pero en *G. metallireducens* los genes *bbsABCD* están localizados en posición 3' a los genes *bbsEFGH* y separados de éstos por tres genes adicionales (*etfA*, *etfB* y *gmet1527*) presuntamente implicados en un sistema aceptor de electrones que podría estar acoplado a la deshidrogenasa de la ruta catabólica de β -oxidación del bencilsuccinato (Kube y col., 2004).

Al comparar la secuencia del *cluster* de genes *bbs* de *Azoarcus* sp. CIB se observa que el gen *bbsI* de función desconocida, ausente en la cepa EbN1, también está presente en las cepas *Azoarcus* sp. T y *T. aromatica* K172. Por otro lado, el gen *bbsJ* de función desconocida e inducible cuando la cepa EbN1 crece en tolueno (Kühner y col., 2005), está presente en todas las cepas analizadas a excepción de *T. aromatica* K172. Sin embargo, el *cluster bbs* de *T. aromatica* K172 posee una región no codificante que

presenta una elevada identidad de secuencia de nucleótidos con la región que codifica *bbsJ* en otros organismos degradadores de tolueno. Este dato apoya la hipótesis de que el *cluster bbs* de *T. aromatica* K172 ha evolucionado de operones *bbs* que contenían el gen *bbsJ* y ha deletado parte de dicho gen, que consiguientemente no será imprescindible para el crecimiento en tolueno, y modificado la consiguiente región intergénica (Kube y col., 2004).

Gracias a la reciente secuenciación del genoma de *Azoarcus* sp. CIB se han podido conocer las regiones flanqueantes del *cluster bss-bbs*. En el extremo 5' de los genes *tdi/bss* de *Azoarcus* sp. CIB se localiza un pequeño fragmento de una secuencia de inserción IS4 y el *cluster* de genes *mbd*, involucrado en la degradación anaeróbica del 3-metilbenzoato vía 3-metilbenzoil-CoA (J.F. Juárez, Tesis doctoral). Estos genes están ausentes en *Azoarcus* sp. EbN1, una cepa incapaz de degradar *m*-xileno. En esta bacteria, en posición 5' de los genes *tdi/bss* se localizan genes presuntamente implicados en rutas de β -oxidación. En *T. aromatica* K172, los genes *orf1* y *orf2*, que podrían estar implicados en un sistema de extrusión de tolueno, se localizan adyacentes a los genes *tdiSR* en el extremo 5' del *cluster bss/bbs* (Leuthner y Heider, 1998). Los productos génicos Orf1 y Orf2 son similares a las proteínas EbA1932 y EbA1936, respectivamente, presentes en la cepa EbN1 y cuyos genes se localizan en la proximidad del extremo 3' del *cluster bbs* de dicha cepa (Kühner y col., 2005). Por otro lado, en el extremo 3' del *cluster bbs* de *Azoarcus* sp. CIB se localizan ortólogos de los genes *cIA68*, *cIA90* y *cIA92* presentes en el *cluster* de degradación anaeróbica de etilbenceno en *Azoarcus* sp. EbN1 y cuya función es desconocida hasta la fecha (Wöhlbrand y col., 2008). Por el contrario, en el extremo 3' del *cluster bbs* de la cepa EbN1 se localizan los genes *istA/istB* del elemento de inserción ISE1 perteneciente a la familia de secuencias de inserción IS21 (Fig. 53) (Kube y col., 2004). Estos elementos de inserción parecen estar ausentes en las regiones flanqueantes de los *clusters bbs* de *T. aromatica* K172 y *G. metallireducens*.

En resumen, los análisis de comparación de los *clusters bss/bbs* de diferentes organismos revelan que si bien la estrategia bioquímica del catabolismo anaeróbico del tolueno es común a todos ellos, los genes correspondientes están organizados en *clusters* que son el resultado de, al menos, tres líneas evolutivas diferentes, dos de ellas en anaerobios facultativos y una en el anaerobio estricto *G. metallireducens*. A su vez, los *clusters bss/bbs* se localizan en distintos contextos genómicos en cada una de las cepas analizadas, y en algunos casos asociados a elementos genéticos móviles

(transposasas, IS) o fragmentos de éstos, que sugieren su adquisición por parte de las correspondientes bacterias huésped mediante sucesos de transferencia génica horizontal y como respuesta adaptativa a la colonización de ambientes anaeróbicos contaminados con tolueno.

4. Regulación del cluster de degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB.

Si bien la bioquímica de la degradación anaeróbica de tolueno se ha estudiado con cierto detalle en algunas cepas de *Azoarcus* y *Thauera*, como se ha expuesto anteriormente, (Widdel y Rabus., 2001; Boll y col., 2002; Chakraborty y Coates., 2004), la regulación de los genes implicados no ha sido estudiada con tanta profundidad. Con el objetivo de incrementar los conocimientos sobre el control de la expresión génica en rutas de catabolismo anaeróbico de tolueno y *m*-xileno, se procedió al estudio de la regulación de dicho *cluster* catabólico en *Azoarcus* sp. CIB.

Al analizar la expresión de los genes *bss* y *bbs* de *Azoarcus* sp. CIB mediante experimentos de RT-PCR se observó que existía una inducción específica de estos cuando las bacterias habían sido cultivadas en tolueno o *m*-xileno (Fig. 21), lo que indica una regulación génica a nivel transcripcional. Una inducción específica similar de los genes *bss-bbs* ya se había observado en las cepas *Azoarcus* sp. T, *Azoarcus* sp. EbN1, *T. aromatica* K172, *T. aromatica* T1 y *Magnetospirillum* sp. TS-6 cultivadas en tolueno anaeróbicamente (Heider y col., 1998; Leuthner y Heider, 2000; Achong y col., 2001; Kühner y col., 2005; Shinoda y col., 2005). La inducción de los genes implicados en el catabolismo anaeróbico de tolueno sólo cuando la célula se encuentra en presencia de dicho sustrato de crecimiento es un fenómeno general en el catabolismo (aeróbico y anaeróbico) de compuestos aromáticos, y facilita la correcta utilización de los recursos energéticos de la célula permitiendo su uso sólo cuando las fuentes de carbono correspondientes están disponibles (Cases y col., 2005; Carmona y col., 2009).

La organización transcripcional del *cluster* de degradación anaeróbica de tolueno y *m*-xileno en *Azoarcus* sp. CIB se determinó mediante amplificación por RT-PCR de regiones intergénicas de longitud superior a 400 pb (Fig. 22). También se realizaron experimentos de RT-PCR utilizando el ARNm obtenido de la cepa *Azoarcus* sp. CIB*bbsB* portadora de una mutación polar en el gen *bbsB* (Fig. 23). Estos estudios revelaron que el *cluster* que codifica la ruta periférica de degradación anaeróbica de

tolueno y *m*-xileno en *Azoarcus* sp. CIB está constituido al menos por cuatro operones: i) un operón *tdi* que incluye el presunto sistema regulador de dos componentes *tdiS* y *tdiR* controlado por el promotor *PtdiS*; ii) un operón catabólico *bssDCABEFGIJKL* que incluye, entre otros, los genes que codifican la bencilsuccinato sintasa y que está controlado por el promotor *PbssD*; iii) un operón catabólico *bbsABCDEFGHIJHltol* que incluye a los genes responsables de la β -oxidación del bencilsuccinato/3-metilbencilsuccinato y al gen *tol*, y que está controlado por el promotor *PbbsA*; iv) un operón que incluye sólo al gen *tol* bajo el control del promotor *Ptol* (Fig. 58).

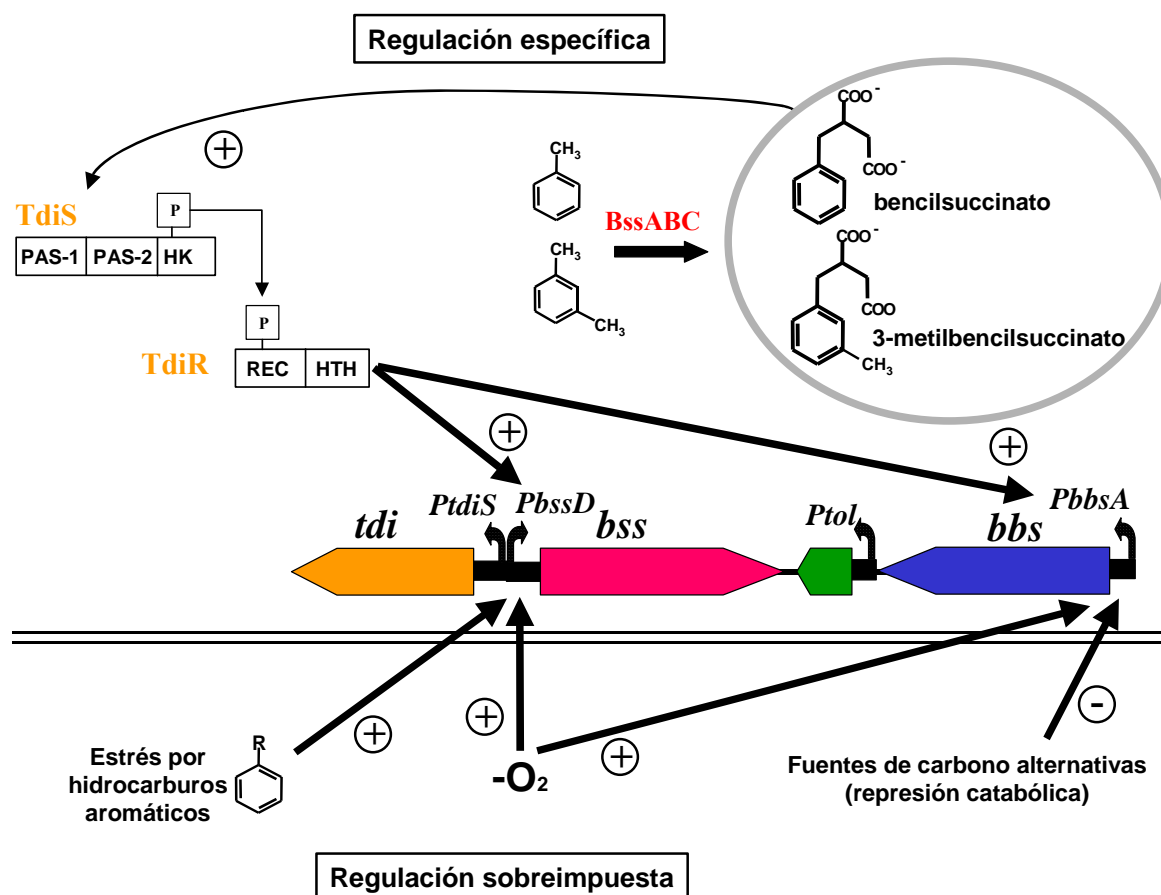


Figura 58. Modelo propuesto para la regulación del cluster de degradación anaeróbica de tolueno y *m*-xileno de *Azoarcus* sp. CIB Las flechas gruesas indican la orientación de los genes *tdi*, *bss*, *tol* y *bbs*. La proteína TdiR se activa por fosforilación mediada por la proteína TdiS en presencia de los compuestos inductores bencilsuccinato/3-metilbencilsuccinato, e interviene en la activación de los promotores catalíticos *PbssD* y *PbbsA*. A esta regulación específica se superpone una regulación dependiente de oxígeno, de la presencia de fuentes de carbono alternativas o del estrés causado por los hidrocarburos aromáticos.

La existencia de un operón catabólico *bssDCABEFGH* se ha descrito en *Azoarcus* sp. EbN1 (Kube y col., 2004), pero en esta bacteria no se conoce si los genes *bssIJKLM* se cotranscriben con los anteriores. Por el contrario, en *Azoarcus* sp. T se ha descrito la existencia de un inicio de la transcripción en el gen *bssF* (Achong y col.,

2001) y en *T. aromatica* T1 el gen *tutE* (*bssD*) forma un operón distinto al constituido por los genes *tutFDGH* (*bssCABE*) (Coschigano, 2000) Todos estos datos revelan que si bien el ordenamiento génico y la identidad de secuencia están conservados en los *clusters bss* de distintas bacterias desnitrificantes, la organización transcripcional de dichos *clusters* difiere significativamente entre distintas especies e incluso entre distintas cepas de la misma especie.

Al contrario de lo que ocurre con el *cluster bss*, hasta la fecha no se había realizado un análisis transcripcional del *cluster bbs* en ningún microorganismo. Los datos presentados en esta Tesis constituyen la primera demostración experimental de que los genes *bbs* se organizan en un único operón en *Azoarcus* sp. CIB, lo que está de acuerdo con el hecho de que la función propuesta para la mayoría de estos genes sea la de estar implicados en alguna de las etapas de la β -oxidación del bencilsuccinato/3-metilbencilsuccinato.

En posición 5' del *cluster* se dispone un hipotético sistema regulador de dos componentes *tdiS* y *tdiR* cuya expresión se induce específicamente en presencia de tolueno o *m*-xileno (Fig. 22). En *Azoarcus* sp. T estos genes también constituyen un operón (Achong y col., 2001). El gen *tdiR* de *T. aromatica* K172 se ha clonado y expresado en *E. coli* y un extracto de esta cepa recombinante mostró la capacidad de retrasar la migración de un fragmento de ADN que contenía al promotor *PbssD* de la cepa K172 en un ensayo de retardo en gel (Leuthner y Heider, 1998), lo que está de acuerdo con el hecho de que la proteína TdiR actúe como un regulador transcripcional de los genes *bss*. Por otro lado, una mutación del gen *tutB1* (*tdiR*) de *T. aromatica* T1 inhibe tanto el crecimiento anaeróbico en tolueno como la expresión de los genes *bss* (Coschigano y Bishop, 2004), indicando de nuevo su implicación en la transcripción de estos genes y su papel regulador en el metabolismo anaeróbico del tolueno. Los genes *tdiSR* de las cepas de *Azoarcus* analizadas (cepas EbN1, CIB y T) se disponen en orientación divergente a los genes catabólicos *bss*, y sus productos génicos poseen mayor identidad de secuencia entre si que con los productos de los genes ortólogos *tdiSR* de las cepas del género *Thauera*, los cuales se organizan en la misma orientación que los genes catabólicos *bss* (Fig. 53). En la cepa *T. aromatica* T1, los genes *tutC1* (*tdiS*) y *tutB1* (*tdiR*), se encuentran separados de los genes catabólicos *tutEFDGH* por los genes *tutB* y *tutC* los cuales se organizan en un operón que se transcribe de forma divergente al operón catabólico *tut* (Fig. 53). Al realizar un análisis de comparación de secuencias se observa que los productos de los genes *tutBC* muestran similitud con

sistemas reguladores de dos componentes que controlan el catabolismo aeróbico de tolueno y estireno en distintas cepas del género *Pseudomonas*, lo que sugiere que las proteínas TutBC también están implicadas en la regulación del catabolismo aeróbico del tolueno en *T. aromatica* T1 (Leuthner y Heider, 1998).

Para confirmar si en *Azoarcus* sp. CIB el gen *tdiR* codifica el regulador de los genes catabólicos *bss* y para demostrar por primera vez si esta proteína controla también los genes *bbs*, se monitorizó la actividad β -galactosidasa en la cepa silvestre *Azoarcus* sp. CIB y en la cepa mutante *Azoarcus* sp. CIB $tdiR$ que contenían las fusiones traduccionales *PbssD::lacZ* y *PbbsA::lacZ*. Los resultados obtenidos mostraron que la actividad de los promotores *PbssD* y *PbbsA* era inducible en la cepa silvestre cuando ésta crece en presencia de tolueno (Fig. 28). En cambio, en la cepa mutante *Azoarcus* sp. CIB $tdiR$ no se observa actividad del promotor *PbbsA*, ni siquiera en presencia de tolueno (Fig. 28). Estos datos indican que la proteína TdiR es un activador imprescindible para la actividad del promotor *PbbsA*. Por el contrario, la actividad del promotor *PbssD* no es estrictamente dependiente del activador TdiR ya que se observa actividad β -galactosidasa en el mutante *Azoarcus* sp. CIB $tdiR$ (*PbssD::lacZ*) cultivado en presencia de tolueno, si bien los valores obtenidos no alcanzan los de la cepa silvestre (Fig. 29). Por tanto, en *Azoarcus* sp. CIB los operones *bss* y *bbs* presentan una regulación diferencial en presencia de tolueno. Así, mientras que la expresión de los genes *bbs* es absolutamente dependiente del activador TdiR, la expresión de los genes *bss* está controlada por el activador TdiR y por otros elementos reguladores todavía no identificados y que muy probablemente forman parte de una respuesta al estrés ocasionado por los hidrocarburos aromáticos en la cepa CIB (Fig. 31 y 58).

Para identificar la molécula efectora del sistema regulador TdiSR en *Azoarcus* sp. CIB, se transfirió la fusión traduccional *PbbsA::lacZ* a las cepas mutantes *Azoarcus* sp. CIB $dbssD$, incapaz de formar bencilsuccinato (a partir de tolueno)/ 3-metilbencilsuccinato (a partir de *m*-xileno), y a la cepa *Azoarcus* sp. CIB $dbbsB$, incapaz de degradar el bencilsuccinato/3-metilbencilsuccinato a benzoil-CoA/3-metilbenzoil-CoA, respectivamente. Al analizar la actividad del promotor *PbbsA* en células de la cepa *Azoarcus* sp. CIB $dbssD$ cultivadas en piruvato en presencia de tolueno o *m*-xileno no se observó actividad β -galactosidasa, indicando que el tolueno o *m*-xileno no son las moléculas inductoras de los genes *bbs* en *Azoarcus* sp. CIB. Sin embargo, al analizar la actividad del promotor *PbbsA* en células de la cepa *Azoarcus* sp. CIB $dbbsB$ se observó

inducción cuando las células habían sido cultivadas en presencia de tolueno o *m*-xileno, indicando que la molécula inductora de los genes *bbs* es el bencilsuccinato o 3-metilbencilsuccinato (figura 30). No obstante, dado que los valores de actividad β -galactosidasa son inferiores a los observados en la cepa silvestre, no se puede descartar que alguno de los compuestos intermediarios de la β -oxidación del bencilsuccinato/3-metilbencilsuccinato pudiesen actuar también como inductores. Aunque hasta la fecha no se había estudiado la actividad del promotor *PbbsA* en ningún organismo, el bencilsuccinato había sido descrito como molécula inductora de los genes *bss* en *T. aromatica* T1 (Coschigano y Bishop., 2004). Por otro lado, la inducción de los genes *bss* dependiente de tolueno en *Azoarcus* sp. EbN1 (Kürhner y col., 2005) se podría corresponder con la regulación TdiR independiente observada en *Azoarcus* sp. CIB, aunque no se puede descartar que las moléculas inductoras de la ruta catabólica puedan variar entre distintas cepas.

La considerable actividad basal (en ausencia del inductor) del promotor *PbssD*, similar a la actividad inducida del promotor *PbbsA*, podría constituir una estrategia para mantener unos niveles basales de actividad bencilsuccinato sintasa en la célula que serían necesarios para iniciar la formación de la molécula inductora (bencilsuccinato/3-metilbencilsuccinato) del sistema regulador TdiSR y que permitiría una expresión óptima de la ruta *bss-bbs* cuando *Azoarcus* sp. CIB coloniza ambientes anaeróbicos ricos en tolueno/*m*-xileno.

Se han descrito varios sistemas reguladores de dos componentes implicados en el control de rutas de catabolismo de compuestos aromáticos. Las dos histidín-quinasa BphS (1598 aa) y BpdS (1576 aa) descritas en organismos Gram-positivos, *Rhodococcus* sp. RHA1 y *Rhodococcus* sp. M5, respectivamente, están implicadas en la detección de compuestos aromáticos policíclicos (bifenilos) y poseen un dominio N-terminal similar al de proteínas con actividad serina/treonina quinasa y un dominio C-terminal HK. El resto de las histidín-quinasa identificadas pertenecen a bacterias Gram-negativas y, atendiendo a su arquitectura molecular y similitud de secuencia, se pueden clasificar en tres grandes grupos (Fig. 59): i) las histidín-quinasa híbridas de rutas aeróbicas de degradación de tolueno en *P. putida* (TodS, 978 aa), *P. mendocina* (TmoS, 973 aa) y *T. aromatica* (TutC, 979 aa) (Busch y col., 2007), y de estireno en distintas cepas de *Pseudomonas* (StyS, 982 aa) (Velasco y col., 1998), que poseen dos supradominios (PAS-1/HK-1/REC y REC/PAS-2/HK-2) en su estructura primaria (Fig. 55), ii) las histidín-quinasa de rutas de degradación anaeróbica de tolueno, en distintas

cepas de *Azoarcus* (TdiS, 545 aa) y *T. aromatica* (TdiS (548 aa en K172)/TutC1 (546 aa en T1)) (Leuthner y Heider, 1998), y etilbenceno en *Azoarcus* sp. EbN1 (Tcs2, 568 aa) (Kühner y col., 2005), que poseen dos dominios PAS adyacentes (PAS-1 y PAS-2) en su extremo N-terminal y un dominio HK en su extremo C-terminal (Fig. 12), y iii) la histidín-quinasa Tcs1 (421 aa) presuntamente implicada en la degradación anaeróbica de acetofenona (ruta del etilbenceno) en *Azoarcus* sp. EbN1 (Kühner y col., 2005), y que posee un dominio PAS N-terminal y un dominio HK C-terminal.

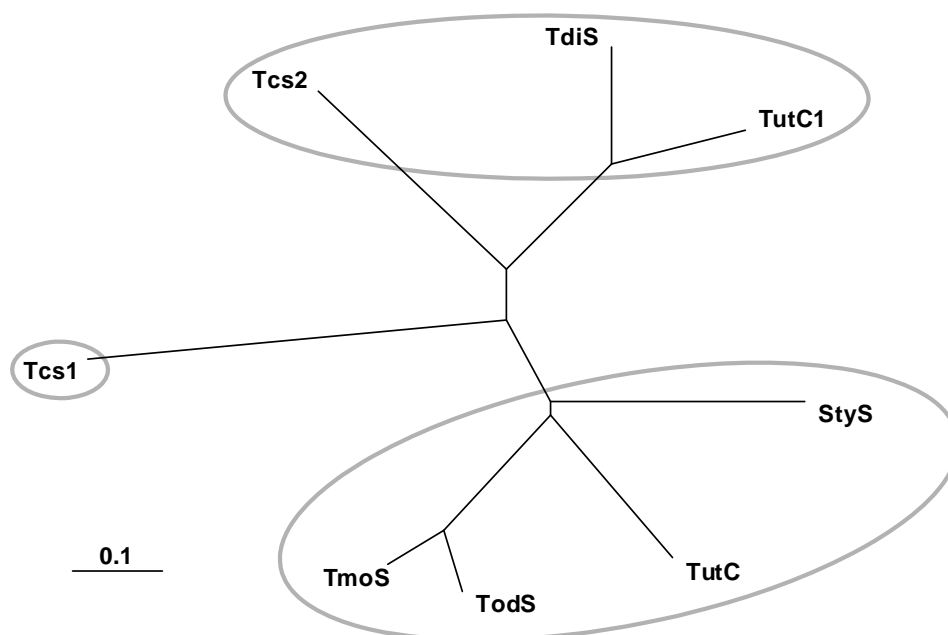


Figura 59. Árbol filogenético de diferentes histidín-quinasas sensoras de sistemas reguladores de dos componentes. El alineamiento múltiple de las proteínas TdiS de *Azoarcus* sp. CIB, TutC1 de *T. aromatica* T1 (AAD12187), Tcs2 de *Azoarcus* sp. EbN1 (YP_158339), StyS de *Pseudomonas* sp. Y2 (CAA03998), TutC de *T. aromatica* T1 (AAD12184), TodS de *P. putida* DOT-T1 (AAC45438), TmoS de *P. mendocina* (AAL13332) y Tcs1 de *Azoarcus* sp. EbN1 (YP_158337), se realizó con el programa CLUSTALW y la visualización del árbol filogenético se realizó con el programa TreeView program.

La existencia de varios dominios PAS y/o HK en la misma molécula sugiere que la mayoría de estas proteínas sensoras son capaces de reconocer e integrar en el mismo circuito regulador distintas señales ambientales tales como la presencia de hidrocarburos aromáticos y, posiblemente, el potencial redox de la célula, lo que permitiría una mayor versatilidad reguladora y explicaría la selección de este tipo de reguladores en el control de rutas catabólicas de compuestos aromáticos tóxicos que inducen varias respuestas celulares, tales como metabolismo y estrés, simultáneamente. La histidín-quinasa TodS ha sido la más estudiada hasta la fecha y es la que define a la subfamilia TodS de histidín-quinasas híbridas (Lacal y col., 2006; Busch y col.,

2007). TodS es una proteína soluble que no posee ningún motivo transmembrana en su estructura primaria, por lo que se clasifica dentro de la clase de las histidín-quinasa citoplásmicas (Mascher y col., 2006). La ausencia de posibles dominios transmembrana en la estructura primaria de otras histidín-quinasa, tales como StyS, TmoS, TutC, TdiS, TutC1, Tcs2 y Tcs1, sugiere que se trata también de sensores citoplásmicos.

Los alineamientos múltiples de las secuencias de los dominios PAS e HK de las proteínas sensoras descritas anteriormente revelan que los dominios PAS-1 e HK-1 del supradominio N-terminal de las histidín-quinasa aeróbicas de tolueno/estireno, junto con los dominios PAS y HK equivalentes de la proteína Tol (ver más arriba) constituyen ramas evolutivas claramente diferenciadas del resto de las ramas del correspondiente árbol filogenético, y sugieren un origen evolutivo común para el supradominio N-terminal de las histidín-quinasa aeróbicas y la región N-terminal sensora de la proteína Tol. Por otro lado, los dominios PAS-2 e HK-2 de las histidín-quinasa aeróbicas podrían, como ya se ha comentado más arriba (Fig. 55), tener un origen evolutivo común con los dominios PAS e HK, respectivamente, de las histidín-quinasa anaeróbicas TdiS, TutC1 y Tcs2. Finalmente, el dominio HK de la histidín-quinasa anaeróbica Tcs1 parece estar más próximo filogenéticamente a los correspondientes dominios HK de los sensores BphS y BpdS de las rutas aeróbicas de bifenilos.

Como se ha indicado anteriormente, el dominio PAS-1 de TodS es el responsable de la interacción con el tolueno y otros hidrocarburos aromáticos (Busch y col., 2007). Los cuatro residuos del dominio PAS-1 de TodS que se han visto implicados en la interacción con el tolueno se encuentran también conservados en los dominios PAS-1 del resto de los sensores aeróbicos de tolueno/estireno y en el dominio PAS de la proteína Tol (ver más arriba) pero no están conservados en los dominios PAS-2 de dichas proteínas ni tampoco en ninguno de los dominios PAS de las histidín-quinasa anaeróbicas. La posibilidad de que el tolueno y otros hidrocarburos aromáticos no interaccionen con las histidín-quinasa anaeróbicas está de acuerdo con resultados previos en *T. aromatica* T1 (Coschigano y Bishop, 2004) y con la observación realizada en esta tesis en *Azoarcus* sp. CIB de que el verdadero inductor de la regulación mediada por TdiS(TutC1)/TdiR(TutB1) es el bencilsuccinato/3-metilbencilsuccinato y no el tolueno/*m*-xileno. Por otro lado, resulta interesante destacar que el dominio PAS-2 de TdiS posee la secuencia 236-His-Asp-Gly-240 que incluye el residuo His que interacciona con el grupo hemo en la histidín-quinasa sensora de oxígeno FixL (Gilles-

Gonzalez y col., 2008), por lo que se podría asumir una función de detección de niveles de oxígeno para dicho dominio en TdiS, y quizás en otras histidín-quinasas anaeróbicas. Una función similar también se ha propuesto para el dominio PAS-2 de las histidín-quinasas aeróbicas (Lau y col., 1997), si bien no se ha detectado la presencia de ningún grupo hemo en la proteína TodS purificada (Lacal y col., 2006).

A diferencia del componente sensor, el componente regulador de los distintos sistemas de dos componentes que controlan el catabolismo de compuestos aromáticos posee una arquitectura modular más conservada y consistente en un dominio REC N-terminal y un dominio HTH C-terminal separados por una región *linker* rica en residuos Gln (*Q linker*). La estructura tridimensional del regulador StyR es conocida (Milani y col., 2005) y permite identificar los residuos conservados esenciales en otros reguladores equivalentes. Así, el dominio REC de TdiR contiene los residuos conservados Asp-15, Asp-58 (aceptor final del fosfato), Arg-60, Thr-83, Phe-102 y Lys-105, implicados en la recepción dependiente de Mg^{2+} del grupo fosfato procedente de la correspondiente histidín-quinasa fosforilada, y en la consiguiente transmisión de la señal (cambio conformacional) al dominio efector HTH. Mientras que en algunos reguladores de respuesta, e. g. StyR y FixJ, la fosforilación implica una mayor afinidad por el ADN diana debido a la dimerización de la proteína, en otros, e.g., TodT, la fosforilación no supone un cambio en la afinidad por el ADN diana y la activación viene mediada por cambios conformacionales del dominio HTH que favorecen los contactos con la ARNP en el promotor correspondiente. Resulta interesante destacar que el árbol filogenético generado con las secuencias de los distintos reguladores de respuesta (Figs. 60) se superpone al obtenido con los dominios HK de las correspondientes proteínas sensoras, lo que sugiere una coevolución del dominio HK con el consiguiente dominio REC del regulador de respuesta, reflejando así la relación funcional entre ambos dominios. En este sentido, sería el supradominio C-terminal (PAS-2/HK-2) de las histidín-quinasas aeróbicas el encargado de interactuar con el correspondiente regulador de respuesta. Por otro lado, las parejas HK/regulador de respuesta de sistemas aeróbicos y anaeróbicos parecen constituir grupos evolutivos diferentes, si bien las parejas HK/regulador de respuesta de rutas de degradación aeróbica de bifenilos en organismos Gram-positivos (BphT, BpdT) están próximos filogenéticamente a la pareja HK/Tcr1 de la ruta anaeróbica de la acetofenona en *Azoarcus* sp. EbN1 (Figs. 60).

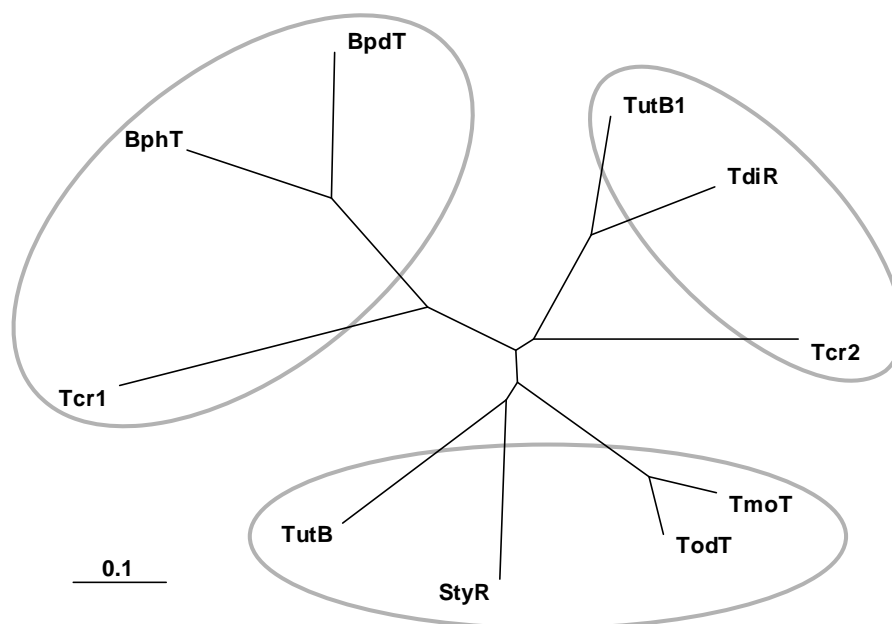


Figura 60. **Árbol filogenético de diferentes reguladores de respuesta de sistemas reguladores de dos componentes que controlan el catabolismo de compuestos aromáticos.** El alineamiento múltiple de las proteínas TdiR de *Azoarcus* sp. CIB, TutB1 de *T. aromatica* T1 (AAD12186), Tcr2 de *Azoarcus* sp. EbN1 (YP_158340), TutB de *T. aromatica* T1 (AAD12185), StyR de *Pseudomonas* sp. Y2 (CAA03999), TodT de *P. putida* DOT-T1 (CAB43736), TmoT de *P. mendocina* (AAL13333), Tcr1 de *Azoarcus* sp. EbN1 (YP_158338), BphT de *Rhodococcus* sp. RHA1 (BAC75412), y BpdT de *Rhodococcus* sp. M5 (AAB52544), se realizó con el programa CLUSTALW y la visualización del árbol filogenético se realizó con el programa TreeView program.

Los dominios HTH de todos los reguladores de respuesta identificados hasta la fecha en catabolismo de compuestos aromáticos pertenecen a la familia NarL/FixJ de reguladores transcripcionales procariontes y, por ello, poseen conservados los residuos que caracterizan a dicha familia. El dominio HTH de StyR está estructurado en 4 hélices α , de las cuales la penúltima es la hélice de reconocimiento del ADN y contiene tres residuos (Lys-175, Val-176 e His-179) presuntamente implicados en la interacción con la caja operadora (caja StyR) del promotor diana (Milani y col., 2005). Estos mismos residuos están conservados en TodT (TutB y TmoT), lo que está de acuerdo con el hecho de que las cajas operadoras Sty (ATAAACCATGGTTTAT) (Santos y col., 2002) y Tod (ATAAAC-N₄-GTTTAT) (Lacal y col., 2006) sean prácticamente idénticas. Por el contrario, estos residuos no están conservados en los dominios HTH de otros reguladores de respuesta, lo que sugiere la existencia de cajas operadoras distintas para estos reguladores transcripcionales. La comparación de los promotores *PbbsA* y *PbssD* de distintos organismos permite detectar una región palindrómica conservada (GGTGTNCGCACC) que se localiza centrada en la posición -43/-44 y que muy

probablemente se corresponde con la caja operadora del regulador TdiR (Figs. 19 y 20). Dado que los residuos Lys-175, Val-176 e His-179 de StyR (TodT, TutB, TmoT) están sustituidos por Glu-178, Ala-179 y Gly-181 en TdiR, se puede hipotetizar que son éstos los encargados de interactuar específicamente con la caja operadora TdiR. La localización de la caja TdiR en la proximidad de la región -35 de los promotores *PbssD* y *PbbsA* sugiere un mecanismo de activación de promotores dependientes de la subunidad σ^{70} de la ARNP típico de reguladores transcripcionales de clase II (Wing y col., 2000), si bien el mecanismo molecular de dicha activación por la proteína TdiR fosforilada necesita ser estudiado en detalle.

Además de la regulación específica mediada por TdiR, los genes *bss/bbs* de *Azoarcus* sp. CIB están sometidos a una regulación sobreimpuesta que depende de distintos factores ambientales como la disponibilidad de oxígeno, la presencia de fuentes de carbono preferentes, etc (Fig. 58). La propia presencia de los hidrocarburos aromáticos es una fuente de estrés para la célula y puede condicionar la expresión de los genes catabólicos, tal y como se ha visto con el control de la actividad del promotor *PbssD* independiente de TdiR en *Azoarcus* sp. CIB (Fig. 31).

En esta Tesis se ha demostrado que no existe expresión de los genes *bss/bbs* cuando las células se cultivan aeróbicamente en tolueno (Fig. 33). Un resultado se ha descrito para los genes *bss* de la cepa *Magnetospirillum* sp. TS-6 que sólo se expresan cuando las células crecen en tolueno anaeróbicamente (Shinoda y col., 2005). Por el contrario, los genes *bss* de *Thauera* sp. DNT-1 se transcriben tanto en condiciones de crecimiento anaeróbicas como aeróbicas en presencia de tolueno (Shinoda y col., 2004).

La cepa *Azoarcus* sp. CIB presenta en su genoma el gen *acpR* (ortólogo del gen *fnr* de *E. coli*) que es imprescindible para la expresión de los genes del cluster *bzd* actuando como un activador transcripcional del promotor P_N en ausencia de oxígeno (Durante-Rodríguez y col., 2006). La proteína AcpR se une a la secuencia consenso (TTGA-N₆-TCAA) que caracteriza a los reguladores de la familia FNR/CRP y que está situada en la posición -41,5 del promotor P_N , probablemente facilitando tanto la interacción de la ARNP con el promotor como la formación del complejo abierto de inicio de la transcripción (Durante-Rodríguez y col., 2006). Al analizar las regiones promotoras *PbssD* y *PbbsA* no se observa la presencia de la secuencia consenso de interacción con AcpR. Trabajos futuros permitirán confirmar experimentalmente si la proteína AcpR está involucrada en el control de la actividad de los promotores *PbssD* y *PbbsA* en condiciones de anaerobiosis, y si este control se realiza de forma directa sobre

los componentes promotores o es un efecto indirecto mediado por otras proteínas reguladoras. La existencia de otros reguladores transcripcionales que detectan niveles de oxígeno en el control de los genes *bss/bbs* de *Azoarcus* sp. CIB tampoco puede ser descartado y será objeto de estudios futuros.

Como se ha indicado en el apartado 3.2 de la introducción, la represión de los promotores catabólicos en presencia de fuentes de carbono adicionales es un fenómeno común en la degradación de compuestos aromáticos (Shingler, 2003; Carmona y col., 2008). La represión catabólica resulta frecuentemente de la integración de varias señales que hacen que la célula tenga preferencia por determinadas fuentes de carbono frente a otras presentes en el medio. La mayoría de los mecanismos moleculares que determinan la represión catabólica en degradación de compuestos aromáticos se han estudiado en rutas aeróbicas de enterobacterias y bacterias del género *Pseudomonas* (Galán y col., 2001; Morales y col., 2004; Rojo y Dinamarca, 2004; Moreno y Rojo, 2008; Pflüger y de Lorenzo, 2008). Sin embargo, el fenómeno de represión catabólica también se ha observado en la degradación anaeróbica de compuestos aromáticos y al menos se han descrito un par de ejemplos en *T. aromatica* (Heider y col., 1998) y *Azoarcus* sp. CIB (López-Barragán y col., 2004a). En esta última, se ha descrito que la represión catabólica que ejercen los ácidos orgánicos no aromáticos (succinato, acetato, etc.) sobre la ruta central del benzoil-CoA se produce a nivel del promotor P_N . Por el contrario, el promotor P_R que controla la expresión del gen regulador *bzdR* no está sometido a represión catabólica por los ácidos orgánicos (López-Barragán y col., 2004a; Durante-Rodríguez y col., 2008).

El estudio de la expresión de la fusión traduccional *PbbsA::lacZ* en células de *Azoarcus* sp. CIB cultivadas simultáneamente en tolueno y en otras fuentes de carbono adicionales, reveló que la expresión de los genes *bbs* sufre una considerable represión en presencia de diferentes fuentes de carbono (Fig. 32) y permitió identificar al promotor *PbssA* como uno de los elementos reguladores involucrados en el mecanismo de represión, indicando que se trataba de un efecto ejercido a nivel transcripcional. Aunque la represión catabólica de ácidos orgánicos no aromáticos sobre las rutas de catabolismo de tolueno ya se ha descrito en organismos aerobios (Duetz y col., 1996; Ruíz y col., 2004), el trabajo presentado en esta Tesis es el primer ejemplo de represión en el catabolismo anaeróbico de tolueno y constituye el punto de partida para iniciar los estudios que determinen la base molecular de dicha represión catabólica. En este sentido, en el genoma de *Azoarcus* sp. CIB se han identificado posibles genes

reguladores globales, i.e. *ptsN* y *ptsO*, *cyoB*, *bphQ*, similares a los ya caracterizado en organismos aerobios y que se han visto implicados en procesos de represión catabólica (Petruschka y col., 2001; Pflüger y de Lorenzo, 2008). La participación de estos reguladores globales en la represión catabólica de los genes *bbs* en *Azoarcus* sp. CIB será objeto de estudio en el futuro próximo.

Los resultados presentados en esta Tesis sobre la regulación de los genes implicados en la ruta periférica del catabolismo anaeróbico de tolueno/*m*-xileno en *Azoarcus* sp. CIB permiten considerar a este sistema como el más completo y mejor estudiado hasta la fecha, constituyendo un modelo de trabajo idóneo para, en el futuro, profundizar en los mecanismos moleculares implicados en la regulación específica y sobreimpuesta de esta ruta catabólica. No obstante, estrategias alternativas de regulación son previsibles en otros organismos anaerobios tales como *G. metallireducens*, en donde el sistema regulador de dos componentes TdiSR está sustituido por un presunto regulador transcripcional (Gmet1542) de la familia XylR/DmpR (Fig. 53), al igual que ocurre en rutas de degradación aeróbicas de tolueno muy bien caracterizadas como la ruta Tol de *P. putida* (Tropel y van der Meer, 2004; Ruíz y col., 2004).

La utilización del sistema regulador TdiSR/*PbbsA* en el desarrollo de biosensores anaeróbicos para la detección específica de tolueno/*m*-xileno en ambientes contaminados con dichos hidrocarburos es una de las posibles aplicaciones biotecnológicas de este trabajo y debería ser abordada en próximos estudios. La construcción de una casete catabólica movilizable con los genes *bss-bbs* sería también una herramienta de gran valor biotecnológico ya que permitiría transferir la capacidad de convertir el tolueno en benzoil-CoA y así ampliar las habilidades catabólicas de ciertos organismos anaerobios, e. g. *A. Evansii*, *R. palustris*, que son capaces de degradar ácidos aromáticos pero no hidrocarburos tóxicos.

5. Caracterización del gen *gcdH* de la ruta baja de degradación anaeróbica de tolueno en *Azoarcus* sp. CIB

El benzoil-CoA que se genera como producto final de la ruta periférica de degradación anaeróbica de tolueno es degradado a través de la ruta alta del benzoil-CoA, codificada por los genes *bzd* de *Azoarcus* sp. CIB, con la consiguiente formación del compuesto alifático 3-hidroxi-pimelil-CoA (Fig. 4) (López-Barragán y col., 2004a). La oxidación del 3-hidroxi-pimelil-CoA hasta intermediarios del metabolismo central de la célula implica tres series de reacciones enzimáticas: i) una ruta de β -oxidación de

ácidos grasos dicarboxílicos de cadena media-larga con la formación final de glutaril-CoA; ii) la descarboxilación oxidativa de glutaril-CoA a crotonil-CoA por una glutaril-CoA deshidrogenasa; y iii) la β -oxidación del crotonil-CoA a acetil-CoA a través de la ruta de β -oxidación de ácidos grasos de cadena corta (Fig. 4) (Harrison y Harwood, 2005). Estas tres etapas del catabolismo del 3-hidroxi-pimelil-CoA en la ruta central del benzoil-CoA se conocen con el nombre genérico de "ruta baja" del benzoil-CoA (Carmona y col., 2008). El glutaril-CoA es un intermediario que también se genera en el catabolismo de otras fuentes de carbono tales como algunos compuestos alicíclicos, e. g. ciclohexanocarboxilato, y los ácidos grasos dicarboxílicos de número impar igual o superior a 5 de átomos de carbono. Algunas rutas del catabolismo de la lisina, triptófano y compuestos heterocíclicos, tales como el ácido nicotínico, también generan glutaril-CoA (Fig. 6), siendo éste, por tanto, un intermediario central en un gran número de rutas metabólicas (Numa y col., 1964; Besrat y col., 1969; Gomes y col., 1981; Husain y Steenkamp, 1985; Kitts y col., 1992; Parke y col., 2001; Revelles y col., 2005) y la reacción catalizada por la glutaril-CoA deshidrogenasa un paso clave en dichas rutas. A pesar de que numerosos genes bacterianos han sido anotados en las bases de datos como responsables de una actividad glutaril-CoA deshidrogenasa, hasta la fecha no se había demostrado experimentalmente dicha función para ninguno de ellos. Con el objetivo de identificar y caracterizar el primer gen responsable de la actividad glutaril-CoA deshidrogenasa bacteriana, en esta Tesis se estudió el gen *gcdH* de *Azoarcus* sp. CIB.

La construcción de la cepa mutante *Azoarcus* sp. CIB*gcdH* y la observación de que dicha cepa es incapaz de degradar anaeróbicamente tolueno, benzoato, 3-hidroxibenzoato, ciclohexanocarboxilato, pimelato o glutarato, y tampoco muestra actividad glutaril-CoA deshidrogenasa (Fig. 36), demuestra por primera vez que el gen *gcdH* es esencial para el crecimiento en estas fuentes de carbono y que su función no puede ser compensada por ningún otro gen del cromosoma bacteriano. La confirmación de que el gen *gcdH* codifica la proteína GcdH con actividad glutaril-CoA deshidrogenasa se obtuvo clonando dicho gen en *E. coli* y valorando la actividad deshidrogenasa de su producto génico en extractos proteicos. Es importante resaltar que la cepa mutante *Azoarcus* sp. CIB*gcdH* sigue siendo capaz de metabolizar anaeróbicamente *m*-xileno. Este dato sugiere que en la degradación anaeróbica del *m*-xileno la ruta central del 3-metilbenzoil-CoA genera un compuesto alifático derivado de CoA (hidroxi-metilacil-CoA) cuya β -oxidación no produce glutaril-CoA como

intermediario y por lo tanto, requiere una enzima distinta a la codificada por el gen *gcdH* todavía sin identificar. En este sentido, la secuenciación del genoma de la cepa CIB ha revelado la existencia de varios genes localizados en la proximidad del *cluster mbd* implicado en la ruta central del 3-metilbenzoil-CoA, y que podrían codificar la correspondiente ruta baja incluyendo la enzima que cataliza la reacción equivalente a la catalizada por GcdH en la ruta baja del benzoil-CoA (Fig. 54).

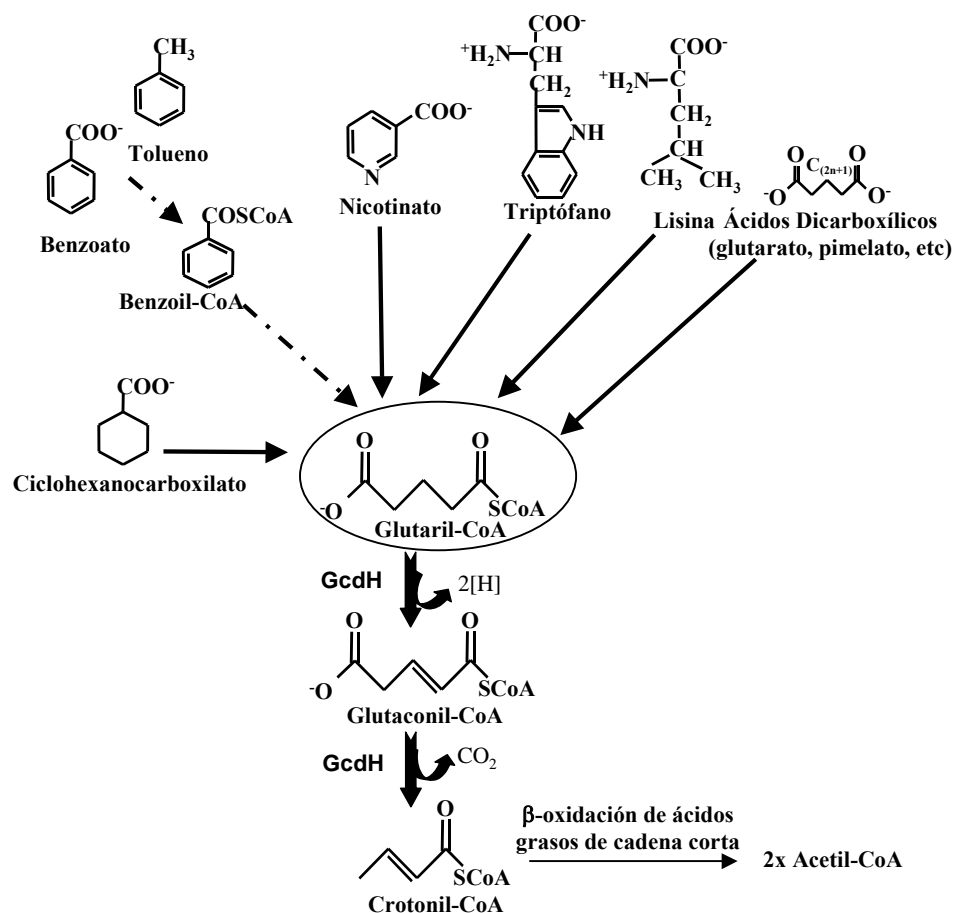


Figura 61. Esquema del catabolismo de diferentes fuentes de carbono que convergen en el intermediario glutaril-CoA y posterior metabolismo de éste. Se representan con líneas discontinuas las rutas anaeróbicas y con líneas continuas las rutas aeróbicas que en algunos casos, e. g. ciclohexanocarboxilato y ácidos dicarboxílicos, son también funcionales en anaerobiosis.

La proteína GcdH es una enzima bifuncional que cataliza inicialmente la deshidrogenación del glutaril-CoA con la formación de glutaconil-CoA como intermediario, y después la descarboxilación de este último generando crotonil-CoA (Fig. 61) (Fu y col., 2004). La enzima GcdH ha sido caracterizada bioquímicamente en algunas bacterias, tales como *A. evansii* y *T. aromatica* (Härtel y col., 1993), así como en células de organismos superiores (Fu y col., 2004). En humanos, la enzima GcdH está involucrada en el metabolismo mitocondrial de aminoácidos, tales como el

triptófano, lisina e hidroxilisina. En la enzima GcdH humana han sido descritas más de 150 mutaciones, algunas de las cuales causan graves enfermedades tales como la aciduria glutárica cerebral tipo I producida por la acumulación de los ácidos glutárico y 3-hidroxiglutárico (Zschocke y col., 2000; Gordon, 2006). La estructura tridimensional de la enzima GcdH humana (homotetrámero) es conocida, tratándose de un miembro único dentro de la familia de las flavoproteínas acil-CoA deshidrogenasas debido a que posee una actividad descarboxilasa adicional no descrita por otros miembros de la familia (Fu et al., 2004). La enzima GcdH en bacterias desnitrificantes es también un homotetrámero compuesto por monómeros de 41 kDa asociados cada uno de ellos a una molécula de FAD (Härtel y col., 1993). Dada la elevada identidad de secuencia (66% identidad con 78% de similitud) entre la proteína GcdH de *Azoarcus* sp. CIB y la proteína GcdH humana, ha sido posible generar un modelo de la estructura tridimensional predicha de la proteína bacteriana (Fig. 62). El monómero de la proteína GcdH consiste en tres dominios claramente diferenciados: un dominio N-terminal formado por 6 hélices α (α_1 - α_6), un dominio de cadenas β (β_1 - β_7) en la región central de la proteína y un dominio C-terminal formado por 5 hélices α (α_7 - α_{11}) (Fig. 62).

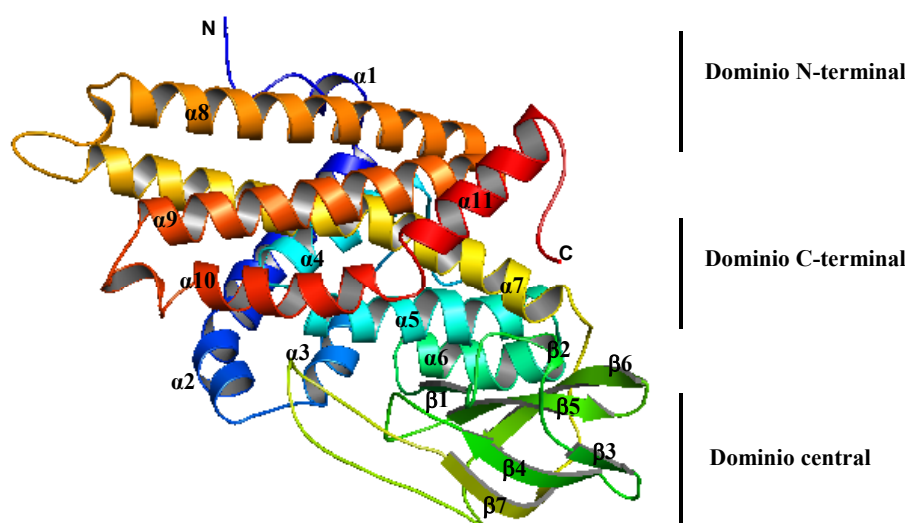


Figura 62. Modelo de la estructura tridimensional propuesta para el monómero de la proteína GcdH de *Azoarcus* sp. CIB. Diagrama de *ribbon* donde se muestran las hélices α y las cadenas β de los 3 dominios de la proteína. La figura fue generada mediante el programa PyMOL (DeLano, 2002).

El mecanismo catalítico de la proteína GcdH implica al cofactor FAD, que se localizaría entre las cadenas β del dominio central y las hélices α del dominio C-terminal (Fu y col., 2004). Al igual que ocurre en la proteína GcdH humana (Fu y col., 2004), la proteína GcdH de *Azoarcus* sp. CIB presenta un residuo Glu-373 conservado y

localizado en el dominio C-terminal de la proteína el cual actuará como base en la reacción de deshidrogenación, así como varios residuos tales como Glu-92, Arg-99, Ser-100, Thr-173 y Tyr-372 localizados a lo largo de la secuencia proteica y posiblemente implicados en la reacción de descarboxilación. Todos estos residuos también se mantienen conservados en las glutaril-CoA deshidrogenasas anotadas en las bases de datos.

El análisis de comparación de secuencias entre la proteína GcdH de la cepa CIB y sus ortólogos presentes en los genomas de otros organismos reveló una elevada identidad (91-93%) con los ortólogos de cepas del género *Azoarcus*, y una identidad alrededor del 80% con los ortólogos presentes en proteobacterias de los subgrupos β y γ . La identidad con proteínas GcdH de otros grupos de bacterias y con GcdH eucariotas ronda el 66%. Estos análisis revelan, por lo tanto, que el producto del gen *gcdH* mantiene un elevado nivel de conservación a lo largo de la escala evolutiva.

Utilizando *Azoarcus* sp. CIB*gcdH* como cepa receptora y valorando el crecimiento anaeróbico en benzoato de los clones recombinantes así como su actividad glutaril-CoA deshidrogenasa, se ha confirmado experimentalmente en esta Tesis la actividad glutaril-CoA deshidrogenasa del hipotético gen *gcdH* de *P. putida* (γ -proteobacteria) (Figs. 49 y 50). Este resultado revela, además, la utilidad de la cepa *Azoarcus* sp. CIB*gcdH* como huésped para la identificación y clonación de genes *gcdH* bacterianos presentes en librerías metagenómicas, algunos de los cuales podrían ser de gran interés biotecnológico para la síntesis de crotonil-CoA, un intermediario empleado para la producción de bioplásticos (Benner y Sismour, 2005)

6. Caracterización de la regulación del gen *gcdH*.

Experimentos de RT-PCR muestran que la transcripción del gen *gcdH* de *Azoarcus* sp. CIB se induce cuando las células se cultivan en presencia de sustratos, tales como benzoato o tolueno, que generan glutaril-CoA (Fig. 40), lo que está de acuerdo con los ensayos de actividad glutaril-CoA deshidrogenasa que indican la presencia de esta enzima cuando las células utilizan como fuente de carbono benzoato o pimelato (Fig. 34), y sugieren el control transcripcional como mecanismo básico de regulación del gen *gcdH*. Dado que no se ha estudiado la regulación transcripcional de ningún gen *gcdH* bacteriano, se procedió a dicho análisis utilizando como modelo el gen *gcdH* de *Azoarcus* sp. CIB. Un estudio de la región flanqueante al gen *gcdH* de la

cepa CIB reveló la existencia de un gen divergente, *gcdR* (Tabla 6), cuyo producto muestra similitud significativa con reguladores transcripcionales de la familia LysR (Tabla 6) (Maddocks y Oyston, 2008). Así, la proteína GcdR presenta en su región N-terminal (aminoácidos 1-66) un dominio HTH alado de unión a ADN y en su región C-terminal (aminoácidos 67-303) los residuos presuntamente implicados en la interacción con los inductores y la dimerización de la proteína. Dado que los reguladores de la familia LysR están frecuentemente implicados en la regulación (activación) de rutas catabolismo de compuestos aromáticos (Díaz y Prieto, 2000; Tropel y van der Meer, 2004; Carmona y col., 2009), se propuso la participación de GcdR en la regulación del gen *gcdH*.

Para comprobar si el gen *gcdR* estaba implicado en la regulación del gen *gcdH*, se construyó la cepa mutante *Azoarcus* sp. CIB*gcdR*. Esta cepa mutante se comportaba de forma similar a la cepa *Azoarcus* sp. CIB*gcdH*, ya que era incapaz de crecer anaeróbicamente en fuentes de carbono tales como el pimelato, benzoato o glutarato (Fig. 42). Al complementar el mutante portador de la disrupción insercional del gen *gcdR* con el plásmido pGCDR que expresa dicho gen, la cepa complementada volvía a recuperar la capacidad de metabolizar pimelato, benzoato o glutarato. Estos resultados sugerían la implicación de la proteína GcdR en la expresión del gen *gcdH*. La construcción de una fusión entre el promotor del gen *gcdH* (*Pg*) y el gen reportero *lacZ* permitió confirmar que la proteína GcdR es necesaria para la activación de dicho promotor, siendo los ácidos dicarboxílicos de 5 carbonos, glutarato y glutaconato, las moléculas inductoras (Fig. 45). Puesto que los derivados CoA del glutarato y glutaconato son los sustratos de las dos reacciones catalizadas por la enzima GcdH, cabe pensar que estos derivados (glutaril-CoA y glutaconil-CoA) puedan ser también inductores del regulador transcripcional GcdR lo que podría facilitar la expresión del gen *gcdH* cuando las células crecen en fuentes de carbono, tales como compuestos aromáticos, ciclohexanocarboxilato o pimelato, que producen glutaril-CoA, y no glutarato, como intermediario en su degradación.

Mediante experimentos de *primer extension* (figura 46) se determinó el inicio de la transcripción (+1) del gen *gcdH* que se corresponde con una adenina localizada a 59 nucleótidos del codón de inicio de la traducción (ATG) del gen *gcdH*. Un análisis detallado de la secuencia del promotor *Pg* reveló unas posibles cajas -10 (TACAGT) y -35 (TTGTCG) de interacción con la ARNP separadas por 17 nucleótidos, siendo esta estructura típica de los promotores dependientes de la subunidad σ^{70} de la ARNP. Es

sabido que los reguladores de tipo LysR interaccionan con dos regiones operadoras distintas del promotor. Inicialmente, reconocen e interaccionan con una región llamada RBS (*recognition binding site*) y caracterizada por presentar una secuencia consenso T-N₁₁-A. Además, interaccionan con una región llamada ABS (*activation binding site*), y que normalmente se encuentra muy cerca de la caja -35 de unión de la ARNP. A través de estas dos regiones de interacción generalmente el regulador es capaz de reprimir su propia expresión y activar el promotor catabólico divergente (Maddocks y Oyston, 2008). En el promotor *Pg*, la región RBS se encontraría localizada a -57 nucleótidos del sitio de inicio de la transcripción, y presenta una repetición invertida de 6 pb, i.e, 5'-GTGCGT-N₃-ACGCAC-3' que se ajusta a la secuencia consenso T-N₁₁-A. La región ABS, que también estaría formada por una repetición invertida de 6 pb, 5'ACGAAATTTTCGT-3', se localiza en posición 5' inmediatamente adyacente a la caja -35 del promotor *Pg* (Fig. 47). La comparación del promotor *Pg* con las regiones promotoras de otros genes *gcdH* bacterianos revela que los motivos de secuencia que más se conservan se corresponden a las regiones ABS, RBS, -10 y -35 de *Pg*, lo cual está de acuerdo con el hecho de que dichas regiones cumplen un papel esencial en la función de estos promotores.

El hecho de que la organización divergente de los genes *gcdH* y *gcdR* se mantiene conservada en numerosos géneros bacterianos pertenecientes a distintos subgrupos de Proteobacterias, y que las cajas reguladoras del promotor *Pg* también parecen estar conservadas en distintos organismos, sugiere un patrón común de regulación transcripcional del gen *gcdH* en numerosos microorganismos. Para confirmar esta hipótesis, se procedió a estudiar el control del gen *gcdH* de *P. putida* clonado en *Azoarcus* sp. CIB. Mientras que el gen *gcdH_{KT}* era capaz de inducirse y permitir el crecimiento de *Azoarcus* sp. CIBd*gcdH* en presencia de benzoato (Figs. 40 y 50), no existía crecimiento en benzoato en células de *Azoarcus* sp. CIBd*gcdR* que contenían el gen *gcdH_{KT}*. Estos resultados indicaban que el gen *gcdR* de *Azoarcus* controla al promotor *Pg* del gen *gcdH* de *P. putida* de forma eficaz y similar a como lo hace sobre el promotor *Pg* del propio *Azoarcus*. La existencia de regulación cruzada entre el gen *gcdR* de *Azoarcus* sp. CIB y el promotor *Pg* del gen *gcdH* de *P. putida* pone en evidencia que la estrategia de regulación del gen *gcdH* se ha conservado en la evolución de bacterias pertenecientes a distintos grupos filogenéticos, lo que está de acuerdo con el hecho de que el gen *gcdH* se encuentre muy conservado y codifique una enzima clave en el metabolismo de numerosos compuestos.

Como conclusión, en esta Tesis se ha caracterizado el primer gen *gcdH* bacteriano y se propone un modelo para su regulación que se esquematiza en la figura 63. Los mecanismos moleculares implicados en el modelo propuesto abren una nueva línea de trabajo y será objeto de estudios futuros. El sistema reportero *GcdR/Pg::lacZ* caracterizado en este trabajo podría ser utilizado también como trampa genética (Galvaio y col., 2005) para identificar en librerías metagenómicas nuevas actividades enzimáticas/rutas metabólicas que generen glutarato/glutaconato.

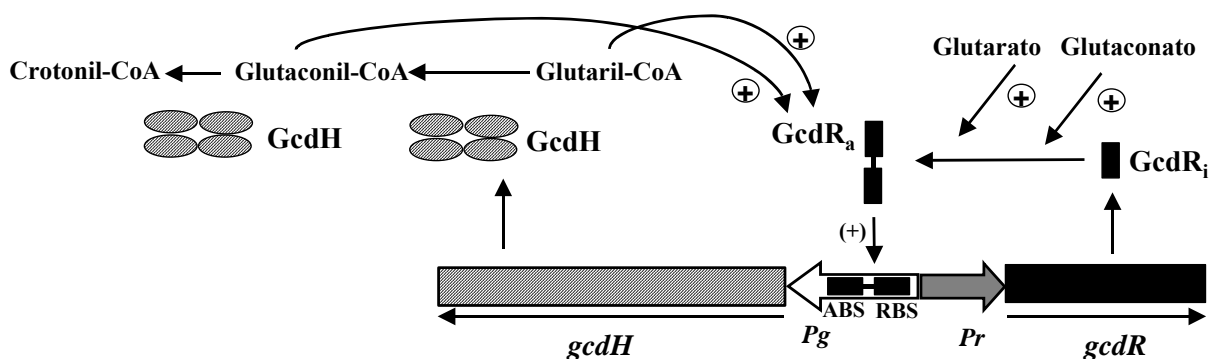


Figura 63. Modelo propuesto para la regulación del gen *gcdH* de *Azoarcus* sp. CIB Se representan los genes *gcdH* (líneas verticales negras) y *gcdR* (en negro) con sus promotores respectivos, *Pg* y *Pr*. La proteína inactiva GcdR_i se activa en la presencia de glutarato o glutaconato produciéndose su oligomerización y la activación del promotor *Pg* mediante la interacción con las cajas operadoras RBS y ABS. El glutaril-CoA y glutaconil-CoA podrían participar también como inductores alternativos.

VIII. CONCLUSIONES

El trabajo descrito a lo largo de esta tesis ha dado lugar a las siguientes conclusiones:

1. Se ha identificado y caracterizado el cluster génico *bss/bbs* responsable de la ruta periférica que convierte el tolueno en benzoil-CoA en *Azoarcus* sp. CIB. Por primera vez se demuestra que los genes *bbs*, junto con los *bss*, están también implicados en la degradación anaeróbica del *m*-xileno y la formación de 3-metilbenzoil-CoA.
2. Por primera vez se ha estudiado en un mismo microorganismo la organización transcripcional de un cluster *bss/bbs* completo. El cluster génico *bss/bbs* de *Azoarcus* sp. CIB se organiza en al menos cuatro operones: un operón regulador (*tdiSR*), dos operones catabólicos (*bss* y *bbs*), y un cuarto operón (*tol*) que no es indispensable para el crecimiento anaeróbico en tolueno/*m*-xileno.
3. Los genes *tdiSR* codifican un sistema regulador de dos componentes compuesto por la histidín-quinasa sensora TdiS y el regulador transcripcional TdiR, que controla la expresión de los operones *bss* y *bbs* activando los promotores *PbssD* y *PbbsA*, respectivamente, con la participación del bencilsuccinato/3-metilbencilsuccinato como inductores.
4. A la regulación específica de los genes *bss/bbs* mediada por TdiSR se superpone una regulación que depende de factores ambientales tales como la anaerobiosis y la presencia de fuentes de carbono preferentes (represión catabólica) tales como los ácidos orgánicos no aromáticos.
5. El gen *tol* codifica una proteína con una arquitectura modular no descrita hasta la fecha, y que se predice como el resultado de la fusión de un supradominio N-terminal sensor, similar al de la subfamilia TodS de histidín-quinzas híbridas que reconocen hidrocarburos aromáticos, y un supradominio C-terminal catalítico, que contiene un dominio EAL con actividad fosfodiesterasa sobre el segundo mensajero di-GMPc. Por ello, la proteína Tol constituye el primer ejemplo de un sistema de transducción de señales monocomponente que, sin embargo, utiliza el mecanismo clásico de las histidín-quinzas de los sistemas de dos componentes.

6. La disrupción del gen *tol* no impide el crecimiento anaeróbico en tolueno y *m*-xileno, pero afecta significativamente a la quimiotaxis y movilidad de la cepa mutante hacia el tolueno, lo que sugiere su implicación en los mecanismos de adaptación de la célula al consumo de hidrocarburos aromáticos, y predice una nueva función para el di-GMPc en bacterias controlando la respuesta al estrés causado por estos compuestos tóxicos.
7. Se ha caracterizado el primer gen bacteriano (*gcdH*) que codifica la actividad glutaril-CoA deshidrogenasa, una etapa clave de la ruta baja del benzoil-CoA en la degradación anaeróbica del tolueno y otros compuestos aromáticos. La disrupción del gen *gcdH* no impide, sin embargo, el catabolismo anaeróbico del *m*-xileno, lo que sugiere la existencia de una ruta baja alternativa para el metabolismo del 3-metilbenzoil-CoA en *Azoarcus* sp. CIB.
8. El gen *gcdR* codifica un regulador transcripcional de la familia LysR que controla la expresión del gen divergente *gcdH* activando al promotor de éste (*Pg*) en presencia de los inductores glutarato y glutaconato. El análisis del promotor *Pg* revela la existencia de las cajas reguladoras RBS y ABS que caracterizan a los promotores regulados por miembros de la familia LysR.
9. Se ha demostrado la existencia de una regulación cruzada entre el gen *gcdR* de *Azoarcus* sp. CIB y el promotor *Pg* del gen *gcdH* de *P. putida*, lo que sugiere que la estrategia de regulación de un gen tan conservado como *gcdH* también se ha conservado en bacterias pertenecientes a distintos grupos filogenéticos.

VIII. BIBLIOGRAFÍA

- Achong, G. R., A. M. Rodriguez, y A. M. Spormann.** 2001. Benzylsuccinate synthase of *Azoarcus* sp. strain T: cloning, sequencing, transcriptional organization, and its role in anaerobic toluene and *m*-mylene mineralization. *J. Bacteriol.* **183**:6763-6770.
- Adler, E.** 1977. Lignin chemistry past, present and future. *Wood Sci. Technol.* **11**:169-218.
- Alhapel, A., D. J. Darley, N. Wagener, E. Eckel, N. Elsner, y A. J. Pierik.** 2006. Molecular and functional analysis of nicotinate catabolism in *Eubacterium barkeri*. *Proc. Natl. Acad. Sci. U S A.* **103**:12341-12346.
- Alonso, S., D. Bartolomé-Martína, M. del Álamo, E. Díaz, J. L. García, y J. Perera.** 2003. Genetic characterization of the styrene lower catabolic pathway of *Pseudomonas* sp. strain Y2 *Gene* **319**:71-83
- Altschul, S., W. Gish, W. Miller, E. Myers, y D. Lipman.** 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
- Anders, H.-J., A. Kaetzke, P. Kampfer, W. Ludwig, y G. Fuchs.** 1995. Taxonomic position of aromatic-degrading denitrifying *Pseudomonad* Strains K172 and KB740 and their description as new members of the genera *Thauera*, as *Thauera aromatica* sp. nov., and *Azoarcus*, as *Azoarcus evansii* sp. nov., respectively, members of the beta subclass of the proteobacteria. *Int. J. Syst. Bacteriol.* **45**:327-333.
- Armengaud, J., K. N. Timmis, y R.-M. Wittich.** 1999. A functional 4-hydroxysalicylate/hydroxyquinol degradative pathway gene cluster is linked to the initial dibenzo-p-dioxin pathway genes in *Sphingomonas* sp. strain RW1. *J. Bacteriol.* **181**:3452-3461.
- Barragán, M. J. L., B. Blázquez, M. T. Zamarro, J. M. Mancheno, J. L. García, E. Díaz, y M. Carmona.** 2005. BzdR, a repressor that controls the anaerobic catabolism of benzoate in *Azoarcus* sp. CIB, is the first member of a new subfamily of transcriptional regulators. *J. Biol. Chem.* **280**:10683-10694.
- Bartolomé-Martín, D., E. Martínez-García, V. Mascaraque, J. Rubio, J. Perera, y S. Alonso.** 2004. Characterization of a second functional gene cluster for the catabolism of phenylacetic acid in *Pseudomonas* sp. strain Y2. *Gene* **341**:167-179.
- Bartolomé-Martín, D.** 2006. "*Estudio de los genes implicados en la degradación del ácido fenilacético en Pseudomonas sp. Y2.*" Universidad Complutense de Madrid.
- Beller, H. R., y A. M. Spormann.** 1999. Substrate range of benzylsuccinate synthase from *Azoarcus* sp. strain T. *FEMS Microbiol. Lett.* **178**:147-153
- Benner, S. A., y A. M. Sismour.** 2005. Synthetic biology. *Nat Rev Genet* **6**:533-543.
- Besrat, A., C. E. Polan, y L. M. Henderson.** 1969. Mammalian Metabolism of Glutaric Acid. *J. Biol. Chem.* **244**:1461-1467.

- Bhandare, R., M. Calabroa, y P. W. Coschigano.** 2006. Site-directed mutagenesis of the *Thauera aromatica* strain T1 *tutE tutFDGH* gene cluster. *Biochem. Biophys. Res. Commun.* **346**:992-998.
- Biegert, T., G. Fuchs, y J. Heider.** 1996. Evidence That Anaerobic Oxidation of Toluene in the Denitrifying Bacterium *Thauera aromatica* is Initiated by Formation of Benzylsuccinate from Toluene and Fumarate. *Eur. J. Biochem.* **238**:661-668.
- Bilwes, A. M., L. A. Alex, B. R. Crane, y M. I. Simon.** 1999. Structure of CheA, a signal-transducing histidine kinase. *Cell* **96**:131-141.
- Blakley, E. R.** 1978. The microbial degradation of cyclohexanecarboxylic acid by a β -oxidation pathway with simultaneous induction to the utilization of benzoate. *Can. J. Microbiol.* **24**:847-855.
- Blázquez, B., M. Carmona, J. L. García, y E. Díaz.** 2008. Identification and analysis of a glutaryl-CoA dehydrogenase-encoding gene and its cognate transcriptional regulator from *Azoarcus* sp. *CIB. Environ. Microbiol.* **10**:474-482.
- Boll, M.** 2005. Key enzymes in the anaerobic aromatic metabolism catalysing Birch-like reductions. *Biochim. Biophys. Acta.* **1707**:34-50.
- Boll, M., y G. Fuchs.** 1995. Benzoyl-coenzyme A reductase (dearomatizing), a key enzyme of anaerobic aromatic metabolism. ATP dependence of the reaction, purification and some properties of the enzyme from *Thauera aromatica* strain K172. *Eur. J. Biochem.* **234**:921-933.
- Boll, M., G. Fuchs, y J. Heider.** 2002. Anaerobic oxidation of aromatic compounds and hydrocarbons. *Curr. Opin. Chem. Biol.* **6**:604-611
- Bradford, M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-54.
- Breese, K., M. Boll, J. Alt-Mörbe, H. Schägger, y G. Fuchs.** 1998. Genes coding for the benzoyl-CoA pathway of anaerobic aromatic metabolism in the bacterium *Thauera aromatica*. *Eur. J. Biochem.* **256**:148-154.
- Browning, D. F., y S. J. W. Busby.** 2004. The regulation of bacterial transcription initiation. *Nat. Rev. Micro.* **2**:57-65.
- Brzostowicz, P. C., A. B. Reams, T. J. Clark, y E. L. Neidle.** 2003. Transcriptional cross-regulation of the catechol and protocatechuate branches of the β -ketoadipate pathway contributes to carbon source-dependent expression of the *Acinetobacter* sp. strain ADP1 *pobA* gene. *Appl. Environ. Microbiol.* **69**:1598-1606.
- Busch, A., J. Lacal, A. Martos, J. L. Ramos, y T. Krell.** 2007. Bacterial sensor kinase TodS interacts with agonistic and antagonistic signals. *Proc. Natl. Acad. Sci. U S A.* **104**:13774-13779.

- Butler, J., Q. He, K. Nevin, Z. He, J. Zhou, y D. Lovley.** 2007. Genomic and microarray analysis of aromatics degradation in *Geobacter metallireducens* and comparison to a *Geobacter* isolate from a contaminated field site. *BMC Genomics* **8**:180.
- Carmona, M., M. A. Prieto, B. Galán, J. L. García, y E. Díaz.** 2008. Signaling networks and design of pollutants biosensors. p. 97-143. In E. Díaz (ed.), *Microbial Biodegradation: genomics and molecular biology*. Caister Academic Press, Norfolk, UK.
- Carmona, M., M. T. Zamarro, B. Blázquez, G. Durante-Rodríguez, J. F. Juárez, J. A. Valderrama, M. J. L. Barragán, J. L. García, y E. Díaz.** 2009. Anaerobic catabolism of aromatic compounds: a genetic and genomic view. *Microbiol. Mol. Biol. Rev.* **73**:71-133.
- Casadaban, M. J.** 1997. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J. Mol. Biol.* **104**:541-555.
- Cases, I., y V. de Lorenzo.** 1998. Expression systems and physiological control of promoter activity in bacteria. *Curr. Opin. Microbiol.* **1**:303-310.
- Cases, I., y V. de Lorenzo.** 2001. The black cat/white cat principle of signal integration in bacterial promoters. *EmboJ.* **20**:1-11.
- Cases, I., y d V. de Lorenzo.** 2005a. Promoters in the environment: transcriptional regulation in its natural context. *Nat. Rev. Microbiol.* **3**:105-118.
- Cases, I., y V. de Lorenzo.** 2005b. Genetically modified organisms for the environment: stories of success and failure and what we have learned from them. *Int. Microbiol.* **8**:213-222.
- Chakraborty, R., y J. D. Coates.** 2004. Anaerobic degradation of monoaromatic hydrocarbons. *Appl. Microbiol. Biotechnol.* **64**:437-446.
- Chang, A. L., J. R. Tuckerman, G. González, R. Mayer, H. Weinhouse, G. Volman, D. Amikam, M. Benziman, y M.-A. Gilles-Gonzalez.** 2001. Phosphodiesterase A1, a regulator of cellulose synthesis in *Acetobacter xylinum*, is a heme-based sensor. *Biochemistry* **40**:3420-3426.
- Cheng, J., A. Z. Randall, M. J. Sweredoski, y P. Baldi.** 2005. SCRATCH: a protein structure and structural feature prediction server. *Nucl. Acids Res.* **33**:W72-76.
- Christen, M., B. Christen, M. Folcher, A. Schauerte, y U. Jenal.** 2005. Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *J. Biol. Chem.* **280**:30829-30837.

- Coates, J. D., V. K. Bhupathiraju, L. A. Achenbach, M. J. McInerney, y D. R. Lovley.** 2001a. *Geobacter hydrogenophilus*, *Geobacter chappellei* and *Geobacter grbiciae*, three new, strictly anaerobic, dissimilatory Fe(III)-reducers. *Int. J. Syst. Evol. Microbiol.* **51**:581-588.
- Coates, J. D., R. Chakraborty, J. G. Lack, S. M. O'Connor, K. A. Cole, K. S. Bender, y L. A. Achenbach.** 2001b. Anaerobic benzene oxidation coupled to nitrate reduction in pure culture by two strains of *Dechloromonas*. *Nature* **411**:1039-1043.
- Coates, J. D., R. Chakrabortya, y M. J. McInerney.** 2002. Anaerobic benzene biodegradation--a new era. *Res. Microbiol.* **153**:621-628
- Coschigano, P. W.** 2000. Transcriptional analysis of the *tutE tutFDGH* gene cluster from *Thauera aromatica* strain T1. *Appl. Environ. Microbiol.* **66**:1147-1151.
- Coschigano, P. W.** 2002. Construction and characterization of insertion/deletion mutations of the *tutF*, *tutD*, and *tutG* genes of *Thauera aromatica* strain T1. *FEMS Microbiol. Lett.* **217**:37-42.
- Coschigano, P. W., y B. J. Bishop.** 2004. Role of benzylsuccinate in the induction of the *tutE tutFDGH* gene complex of *T. aromatica* strain T1. *FEMS Microbiol. Lett.* **231**:261-266.
- Coschigano, P. W., T. S. Wehrman, y L. Y. Young.** 1998. Identification and analysis of genes involved in anaerobic toluene metabolism by strain T1: putative role of a glycine free radical. *Appl. Environ. Microbiol.* **64**:1650-1656.
- Coschigano, P. W., y L. Y. Young.** 1997. Identification and sequence analysis of two regulatory genes involved in anaerobic toluene metabolism by strain T1. *Appl. Environ. Microbiol.* **63**:652-660.
- Crooks, G., y S. Copley.** 1994. Purification and characterization of 4-chlorobenzoyl CoA dehalogenase from *Arthrobacter* sp. strain 4-CB1. *Biochemistry* **33**:11645-116459.
- Cripps, R. E.** 1973. The microbial metabolism of thiophen-2-carboxylate. *Biochem. J.* **134**:353-366.
- DeLano, W. L.** 2002. The PyMOL Molecular graphics system. DeLano Scientific, San Carlos, CA, USA.
- de Lorenzo, V., y K. Timmis.** 1994. Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods. Enzymol.* **235**:386-405.
- Denef, V. J., J. Park, T. V. Tsoi, J. M. Rouillard, H. Zhang, J. A. Wibbenmeyer, W. Verstraete, E. Gulari, S. A. Hashsham, y J. M. Tiedje.** 2004. Biphenyl and benzoate metabolism in a genomic context: outlining genome-wide metabolic networks in *Burkholderia xenovorans* LB400. *Appl. Environ. Microbiol.* **70**:4961-4970.

- Díaz, E.** 2004. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. *Int. Microbiol.* **7**:173-180.
- Díaz, E., A. Ferrández, M. A. Prieto, y J. L. García.** 2001. Biodegradation of aromatic compounds by *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **65**:523-569.
- Díaz, E., y M. A. Prieto.** 2000. Bacterial promoters triggering biodegradation of aromatic pollutants. *Curr. Opin. Biotechnol.* **11**:467-475.
- Dimroth, P., y B. Schink.** 1998. Energy conservation in the decarboxylation of dicarboxylic acids by fermenting bacteria. *Arch. Microbiol.* **170**:69-77.
- Dispensa, M., C. T. Thomas, M. K. Kim, J. A. Perrotta, J. Gibson, y C. S. Harwood.** 1992. Anaerobic growth of *Rhodospseudomonas palustris* on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators. *J. Bacteriol.* **174**:5803-5813.
- Domínguez-Cuevas, P., J. E. González-Pastor, S. Marqués, J. L. Ramos, y V. de Lorenzo.** 2006. Transcriptional tradeoff between metabolic and stress-response programs in *Pseudomonas putida* KT2440 cells exposed to toluene. *J. Biol. Chem.* **281**:11981-11991.
- Duetz, W. A., S. Marqués, B. Wind, J. L. Ramos, y J. G. van An del.** 1996. Catabolite repression of the toluene degradation pathway in *Pseudomonas putida* harboring pWW0 under various conditions of nutrient limitation in chemostat culture. *Appl. Environ. Microbiol.* **62**:601-606.
- Ducros, V. M.-A., R. J. Lewis, C. S. Verma, E. J. Dodson, G. Leonard, J. P. Turkenburg, G. N. Murshudov, A. J. Wilkinson, y J. A. Brannigan.** 2001. Crystal structure of GerE, the ultimate transcriptional regulator of spore formation in *Bacillus subtilis*. *J. Mol. Biol.* **306**:759-771.
- Duldhardt, I., I. Nijenhuis, F. Schauer, y H. J. Heipieper.** 2007. Anaerobically grown *Thauera aromatica*, *Desulfococcus multivorans*, *Geobacter sulfurreducens* are more sensitive towards organic solvents than aerobic bacteria. *Appl. Microbiol. Biotechnol.* **77**:705-711.
- Dunaway-Mariano, D., y P. Babbitt.** 1994. On the origins and functions of the enzymes of the 4-chlorobenzoate to 4-hydroxybenzoate converting pathway. *Biodegradation* **5**:259-276.
- Durante-Rodríguez, G., M. T. Zamarro, J. L. García, E. Díaz, y M. Carmona.** 2008. New insights into the BzdR-mediated transcriptional regulation of the anaerobic catabolism of benzoate in *Azoarcus* sp. *CIB. Microbiology* **154**:306-316.
- Durante-Rodríguez, G., M. T. Zamarro, J. L. García, E. Díaz, y M. Carmona.** 2006. Oxygen-dependent regulation of the central pathway for the anaerobic catabolism of aromatic compounds in *Azoarcus* sp. strain CIB. *J. Bacteriol.* **188**:2343-2354.

- Egland, P. G., J. Gibson, y C. S. Harwood.** 2001. Reductive, coenzyme A-mediated pathway for 3-chlorobenzoate degradation in the phototrophic bacterium *Rhodospseudomonas palustris*. *Appl. Environ. Microbiol.* **67**:1396-1399.
- Egland, P. G., y C. S. Harwood.** 1999. BadR, a new MarR family member, regulates anaerobic benzoate degradation by *Rhodospseudomonas palustris* in concert with AadR, an Fnr family member. *J. Bacteriol.* **181**:2102-2109.
- Egland, P. G., D. A. Pelletier, M. Dispensa, J. Gibson, y C. S. Harwood.** 1997. A cluster of bacterial genes for anaerobic benzene ring biodegradation. *Proc. Natl. Acad. Sci. U S A.* **94**:6484-6489.
- Elshahed, M. S., y M. J. McInerney.** 2001. Benzoate fermentation by the anaerobic bacterium *Syntrophus aciditrophicus* in the absence of hydrogen-using microorganisms. *Appl. Environ. Microbiol.* **67**:5520-5525.
- Elshahed, M. S., V. K. Bhupathiraju, N. Q. Wofford, M. A. Nanny, y M. J. McInerney.** 2001. Metabolism of benzoate, cyclohex-1-ene carboxylate, and cyclohexane carboxylate by *Syntrophus aciditrophicus* strain SB in syntrophic association with H₂-using microorganisms. *Appl. Environ. Microbiol.* **67**:1728-1738.
- Fernández, S., V. de Lorenzo, y J. Pérez-Martin.** 1995. Activation of the transcriptional regulator XylR of *Pseudomonas putida* by release of repression between functional domains. *Mol. Microbiol.* **16**:205-213.
- Ferrández, A., J. L. García, y E. Díaz.** 2000. Transcriptional regulation of the divergent *paa* catabolic operons for phenylacetic acid degradation in *Escherichia coli*. *J. Biol. Chem.* **275**:12214-12222.
- Ferrández, A., B. Miñambres, B. García, E. R. Olivera, J. M. Luengo, J. L. García, y E. Díaz.** 1998. Catabolism of phenylacetic acid in *Escherichia coli*. Characterization of a new aerobic hybrid pathway. *J. Biol. Chem.* **273**:25974-25986.
- Foght, J.** 2008. Anaerobic biodegradation of aromatic hydrocarbons: pathways and prospects. *J. Mol. Microbiol. Biotechnol.* **15**:93-120.
- Franklin, F. C., M. Bagdasarian, M. M. Bagdasarian, y K. Timmis.** 1981. Molecular and functional analysis of the TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring meta cleavage pathway. *Proc. Natl. Acad. Sci. U S A.* **78**:7458-7462.
- Fu, Z., M. Wang, R. Paschke, K. S. Rao, F. E. Frerman, y J.-J. P. Kim.** 2004. Crystal structures of human glutaryl-CoA dehydrogenase with and without an alternate substrate: structural bases of dehydrogenation and decarboxylation reactions. *Biochemistry* **43**:9674-9684.
- Fuchs, G.** 2008. Anaerobic Metabolism of Aromatic Compounds. *Ann. N. Y. Acad. Sci.* **1125**:82-99.

- Galán, B., E. Díaz, y J. L. García.** 2000. Enhancing desulphurization by engineering a flavin reductase-encoding gene cassette in recombinant biocatalysts. *Environ. Microbiol.* **2**:687-669.
- Gallus, C., y B. Schink.** 1994. Anaerobic degradation of pimelate by newly isolated denitrifying bacteria. *Microbiology* **140**:409-16.
- Galperin, M. Y., A. N. Nikolskaya, y E. V. Koonin.** 2001. Novel domains of the prokaryotic two-component signal transduction systems FEMS Microbiol. Lett. **203**:11-21.
- Galperin, M. Y.** 2004. Bacterial signal transduction network in a genomic perspective. *Environ. Microbiol.* **6**:552-567.
- Galperin, M. Y.** 2006. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.* **188**:4169-4182.
- Galvão, T. C., W. W. Mohn, y V. de Lorenzo.** 2005. Exploring the microbial biodegradation and biotransformation gene pool Trends. Biotechnol. **23**:497-506
- Gasson, M. J., Y. Kitamura, W. R. McLauchlan, A. Narbad, A. J. Parr, E. L. H. Parsons, J. Payne, M. J. C. Rhodes, y N. J. Walton.** 1998. Metabolism of ferulic acid to vanillin. A bacterial gene of the enoyl-SCoA hydratase/isomerase superfamily encodes an enzyme for the hydration and cleavage of a hydroxycinnamic acid SCoA thioester. *J. Biol. Chem.* **273**:4163-4170.
- Gescher, J., W. Eisenreich, J. Wörth, A. Bacher, y G. Fuchs.** 2005. Aerobic benzoyl-CoA catabolic pathway in *Azoarcus evansii*: studies on the non-oxygenolytic ring cleavage enzyme. *Mol. Microbiol.* **56**:1586-1600.
- Gescher, J., W. Ismail, E. Ölgeschläger, W. Eisenreich, J. Wörth, y G. Fuchs.** 2006. Aerobic benzoyl-coenzyme A (CoA) catabolic pathway in *Azoarcus evansii*: conversion of ring cleavage product by 3,4-dehydroadipyl-CoA semialdehyde dehydrogenase. *J. Bacteriol.* **188**:2919-2927.
- Gerhardt, P., Morroy, R.G.E., Woody, W.A. y Krieg, N.R. (1994) *Methods for General and Molecular Bacteriology*. Washington D.C.
- Gjermansen, M., P. Ragas, C. Sternberg, S. Molin, y T. Tolker-Nielsen.** 2005. Characterization of starvation-induced dispersion in *Pseudomonas putida* biofilms. *Environ. Microbiol.* **7**:894-904.
- Gibson, J., y C. S. Harwood.** 2002. Metabolic diversity in aromatic compound utilization by anaerobic microbes. *Annu. Rev. Microbiol.* **56**:345-369.
- Gibson, K. J., y J. Gibson.** 1992. Potential early intermediates in anaerobic benzoate degradation by *Rhodospseudomonas palustris*. *Appl. Environ. Microbiol.* **58**:696-698.
- Gibson, D. T., y R. E. Parales.** 2000. Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr. Opin. Biotechnol.* **11**:236-243.

- Gilles-Gonzalez, M.-A., G. Gonzalez, H. Eduardo, S. Sousa, y J. Tuckerman.** 2008. Oxygen-sensing histidine-protein kinases: assays of ligand binding and turnover of response-regulator substrates. *Methods. Enzymol.* **437**:173-189.
- Gomes, B., G. Fendrich, y R. Abeles.** 1981. Mechanism of action of glutaryl-CoA and butyryl-CoA dehydrogenases. Purification of glutaryl-CoA dehydrogenase. *Biochemistry* **20**:1481-1490.
- Gordon, N.** 2006. Glutaric aciduria types I and II. *Brain Dev.* **28**:136 -140.
- Guex, N., y M. Peitsch.** 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**:2714-23.
- Haddock, J. D., y J. G. Ferry.** 1989. Purification and properties of phloroglucinol reductase from *Eubacterium oxidoreducens* G-41. *J. Biol. Chem.* **264**:4423-4427.
- Harayama, S., y M. Rejik.** 1989. Bacterial aromatic ring-cleavage enzymes are classified into two different gene families. *J. Biol. Chem.* **264**:15328-15333.
- Harayama, S. y Timmis, K.N.** 1989 Catabolism of aromatic hydrocarbons by *Pseudomonas*. In *Genetics of bacterial diversity*. Hopwood, A. and Chater, K.F. (eds). London: Academic Press.
- Harayama, S. y Timmis, K.N.** 1992 Aerobic biodegradation of aromatic hydrocarbons by bacteria. En: *Metal Ions in Biological Systems*. Sigel, H. y Sigel, A. (eds.). New York: Marcell Dekker, Inc., pp. 99-156.
- Harms, G., K. Zengler, R. Rabus, F. Aeckersberg, D. Minz, R. Rossello-Mora, y F. Widdel.** 1999. Anaerobic oxidation of *o*-xylene, *m*-xylene, and homologous alkylbenzenes by new types of sulfate-reducing bacteria. *Appl. Environ. Microbiol.* **65**:999-1004.
- Härtel, U., E. Eckel, J. Koch, G. Fuchs, D. Linder, y W. Buckel.** 1993. Purification of glutaryl-CoA dehydrogenase from *Pseudomonas* sp., an enzyme involved in the anaerobic degradation of benzoate. *Arch. Microbiol.* **159**:174-181.
- Harrison, F. H., y C. S. Harwood.** 2005. The *pimFABCDE* operon from *Rhodospseudomonas palustris* mediates dicarboxylic acid degradation and participates in anaerobic benzoate degradation. *Microbiology* **151**:727-736.
- Harwood, C. S., G. Burchhardt, H. Herrmann, y G. Fuchs.** 1999. Anaerobic metabolism of aromatic compounds via the benzoyl-CoA pathway. *FEMS Microbiol. Rev.* **22**:439-458.
- Harwood, C. S., y J. Gibson.** 1997. Shedding light on anaerobic benzene ring degradation: a process unique to prokaryotes? *J. Bacteriol.* **179**:301-309.
- Harwood, C. S., y R. E. Parales.** 1996. The β -keto adipate pathway and the biology of self-identity. *Annu. Rev. Microbiol.* **50**:553-590.

- Heider, J.** 2007. Adding handles to unhandy substrates: anaerobic hydrocarbon activation mechanisms. *Curr. Opin. Chem. Biol.* **11**:188-194.
- Heider, J., y G. Fuchs.** 1997. Anaerobic metabolism of aromatic compounds. *Eur. J. Biochem.* **243**:577-596.
- Heider, J., M. Boll, K. Breese, S. Breinig, C. Ebenau-Jehle, U. Feil, N. Gad'on, D. Laempe, B. Leuthner, M. E.-S. Mohamed, S. Schneider, G. Burchhardt, y G. Fuchs.** 1998. Differential induction of enzymes involved in anaerobic metabolism of aromatic compounds in the denitrifying bacterium *Thauera aromatica*. *Arch. Microbiol.* **170**:120-131.
- Hermuth, K., B. Leuthner, y J. Heider.** 2002. Operon structure and expression of the genes for benzylsuccinate synthase in *Thauera aromatica* strain K172. *Arch. Microbiol.* **177**:132-138.
- Hickman, J. W., y C. S. Harwood.** 2008. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol. Microbiol.* **69**:376-389.
- Hickman, J. W., D. F. Tifrea, y C. S. Harwood.** 2005. A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proc. Natl. Acad. Sci. U S A.* **102**:14422-14427.
- Hirsch, W., H. Schagger, y G. Fuchs.** 1998. Phenylglyoxylate:NAD⁺ oxidoreductase (CoA benzoylating), a new enzyme of anaerobic phenylalanine metabolism in the denitrifying bacterium *Azoarcus evansii*. *Eur. J. Biochem.* **251**:907-915.
- Hoffman, L. R., D. A. D'Argenio, M. J. MacCoss, Z. Zhang, R. A. Jones, y S. I. Miller.** 2005. Aminoglycoside antibiotics induce bacterial biofilm formation. *Nature* **436**:1171-1175.
- Husain, M., y D. J. Steenkamp.** 1985. Partial purification and characterization of glutaryl-coenzyme A dehydrogenase, electron transfer flavoprotein, and electron transfer flavoprotein-Q oxidoreductase from *Paracoccus denitrificans*. *J. Bacteriol.* **163**:709-715.
- Ismail, W., M. E.-S. Mohamed, B. L. Wanner, K. A. Datsenko, W. Eisenreich, F. Rohdich, A. Bacher, y G. Fuchs.** 2003. Functional genomics by NMR spectroscopy. Phenylacetate catabolism in *Escherichia coli*. *Eur. J. Biochem.* **270**:3047-3054.
- Iwabuchi, T., y S. Harayama.** 1998. Biochemical and molecular characterization of 1-hydroxy-2-naphthoate dioxygenase from *Nocardioides* sp. KP7. *J. Biol. Chem.* **273**:8332-8336.
- Jackson, D. W., K. Suzuki, L. Oakford, J. W. Simecka, M. E. Hart, y T. Romeo.** 2002. Biofilm formation and dispersal under the influence of the global regulator CsrA of *Escherichia coli*. *J. Bacteriol.* **184**:290-301.

- Jenal, U.** 2004. Cyclic di-guanosine-monophosphate comes of age: a novel secondary messenger involved in modulating cell surface structures in bacteria?. *Curr. Opin. Microbiol.* **7**:185-191.
- Jenal, U., y J. Malone.** 2006. Mechanisms of Cyclic-di-GMP signaling in bacteria. *Annu. Rev. Genet.* **40**:385-407.
- Jiménez, J. I., B. Miñambres, J. L. García, y E. Díaz.** 2002. Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. *Environ. Microbiol.* **4**:824-841.
- Jonas, K., A. N. Edwards, R. Simm, T. Romeo, U. Römling, y Ö. Melefors.** 2008. The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins. *Mol. Microbiol.* **70**:236-257.
- Kader, A., R. Simm, U. Gerstel, M. Morr, y U. Römling.** 2006. Hierarchical involvement of various GGDEF domain proteins in rdar morphotype development of *Salmonella enterica* serovar Typhimurium. *Mol. Microbiol.* **60**:602-616.
- Kane, S. R., H. R. Beller, T. C. Legler, y R. T. Anderson.** 2002. Biochemical and genetic evidence of benzylsuccinate synthase intoluene-degrading, ferric iron-reducing *Geobacter metallireducens*. *Biodegradation* **13**:149 -154.
- Kluge, C., A. Tschech, y G. Fuchs.** 1990. Anaerobic metabolism of resorcylic acids (*m*-dihydroxybenzoic acids) and resorcinol (1,3-benzenediol) in a fermenting and in a denitrifying bacterium. *Arch. Microbiol.* **155**:68-74.
- Koenig, K., y J. R. Andreesen.** 1990. Xanthine dehydrogenase and 2-furoyl-coenzyme A dehydrogenase from *Pseudomonas putida* Fu1: two molybdenum-containing dehydrogenases of novel structural composition. *J. Bacteriol.* **172**:5999-6009.
- Komano, T., R. Utsumi, y M. Kawamukai.** 1991. Functional analysis of the *fic* gene involved in regulation of cell division. *Res. Microbiol.* **142**:269-277.
- Kovach, M., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. I. Roop, y K. M. Peterson.** 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**:175-176.
- Kraulis, P. J.** 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**:946-950
- Kretzer, A., K. Frunzke, y J. Andreesen.** 1993. Catabolism of isonicotinate by *Mycobacterium* sp. INA1: extended description of the pathway and purification of the molybdoenzyme isonicotinate dehydrogenase. *J. Gen. Microbiol.* **139**:2763-2772.
- Krieger, C. J., H. R. Beller, M. Reinhard, y A. M. Spormann.** 1999. Initial reactions in anaerobic oxidation of *m*-xylene by the denitrifying bacterium *Azoarcus* sp. strain T. *J. Bacteriol.* **181**:6403-6410.

- Krieger, C. J., W. Roseboom, S. P. J. Albracht, and A. M. Spormann.** 2001. A stable organic free radical in anaerobic benzylsuccinate synthase of *Azoarcus* sp. strain T. J. Biol. Chem. **276**:12924-12927.
- Kube, M., J. Heider, J. Amann, P. Hufnagel, S. Kühner, A. Beck, R. Reinhardt, y R. Rabus.** 2004. Genes involved in the anaerobic degradation of toluene in a denitrifying bacterium, strain EbN1. Arch. Microbiol. **181**:182-194.
- Kühner, S., L. Wöhlbrand, I. Fritz, W. Wruck, C. Hultschig, P. Hufnagel, M. Kube, R. Reinhardt, y R. Rabus.** 2005. Substrate-dependent regulation of anaerobic degradation pathways for toluene and ethylbenzene in a denitrifying bacterium, strain EbN1. J. Bacteriol. **187**:1493-1503.
- Kulesekara, H., V. Lee, A. Brencic, N. Liberati, J. Urbach, S. Miyata, D. G. Lee, A. N. Neely, M. Hyodo, Y. Hayakawa, F. M. Ausubel, y S. Lory.** 2006. Analysis of *Pseudomonas aeruginosa* diguanylate cyclases and phosphodiesterases reveals a role for bis-(3'-5')-cyclic-GMP in virulence. Proc. Natl. Acad. Sci. U S A. **103**:2839-2844.
- Lacal, J., A. Busch, M. E. Guazzaroni, T. Krell, y J. L. Ramos.** 2006. The TodS–TodT two-component regulatory system recognizes a wide range of effectors and works with DNA-bending proteins. Proc. Natl. Acad. Sci. U S A. **103**:8191-8196.
- Lack, J. G., S. K. Chaudhuri, R. Chakraborty, L. A. Achenbach, y J. D. Coates.** 2002. Anaerobic biooxidation of Fe(II) by *Dechlorosoma suillum*. Microb. Ecol. **43**:424-431.
- Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.
- Laempe, D., W. Eisenreich, A. Bacher, y G. Fuchs.** 1998. Cyclohexa-1,5-diene-1-carboxyl-CoA hydratase, an enzyme involved in anaerobic metabolism of benzoyl-CoA in the denitrifying bacterium *Thauera aromatica*. Eur. J. Biochem. **255**:618-627.
- Laempe, D., M. Jahn, y G. Fuchs.** 1999. 6-Hydroxycyclohex-1-ene-1-carboxyl-CoA dehydrogenase and 6-oxocyclohex-1-ene-1-carboxyl-CoA hydrolase, enzymes of the benzoyl-CoA pathway of anaerobic aromatic metabolism in the denitrifying bacterium *Thauera aromatica*. Eur. J. Biochem. **263**:420-429.
- Lau, P. C., Y. Wang, A. Patel, D. Labbé, H. Bergeron, R. Brousseau, Y. Konishi, y M. Rawlings.** 1997. A bacterial basic region leucine zipper histidine kinase regulating toluene degradation. Proc. Natl. Acad. Sci. U S A. **94**:1453-1458.
- Laurie, A. D., L. M. Bernardo, C. C. Sze, E. Skarfstad, A. Szalewska-Palasz, T. Nystrom, y V. Shingler.** 2003. The role of the alarmone (p)ppGpp in sigma N competition for core RNA polymerase J. Biol. Chem. **278**:1494-1503.
- Lee, V. T., J. M. Matewish, J. L. Kessler, M. Hyodo, Y. Hayakawa, y S. Lory.** 2007. A cyclic-di-GMP receptor required for bacterial exopolysaccharide production. Mol. Microbiol. **65**:1474-1484.

- Lehman, T. C., D. E. Hale, A. Bhala, y C. Thorpe.** 1990. An acyl-coenzyme A dehydrogenase assay utilizing the ferricenium ion. *Anal. Biochem.* **186**:280-284.
- Leuthner, B., y J. Heider.** 1998. A two-component system involved in regulation of anaerobic toluene metabolism in *Thauera aromatica*. *FEMS Microbiol. Lett.* **166**:35-41.
- Leuthner, B., y J. Heider.** 2000. Anaerobic toluene catabolism of *Thauera aromatica*: the *bbs* operon codes for enzymes of β oxidation of the intermediate benzylsuccinate. *J. Bacteriol.* **182**:272-277.
- Leuthner, B., C. Leutwein, H. Schulz, P. Hörth, W. Haehnel, E. Schiltz, H. Schägger, y J. Heider.** 1998. Biochemical and genetic characterization of benzylsuccinate synthase from *Thauera aromatica*: a new glyceryl radical enzyme catalysing the first step in anaerobic toluene metabolism. *Mol. Microbiol.* **28**:615-628.
- Leutwein, C., y J. Heider.** 2001. Succinyl-CoA:(*R*)-benzylsuccinate CoA-transferase: an enzyme of the anaerobic toluene catabolic pathway in denitrifying bacteria. *J. Bacteriol.* **183**:4288-4295.
- Leutwein, C., y J. Heider.** 2002. (*R*)-Benzylsuccinyl-CoA dehydrogenase of *Thauera aromatica*, an enzyme of the anaerobic toluene catabolic pathway. *Arch. Microbiol.* **178**:517-524.
- Levi, A., y U. Jenal.** 2006. Holdfast Formation in Motile Swarmer Cells Optimizes Surface Attachment during *Caulobacter crescentus* Development. *J. Bacteriol.* **188**:5315-5318.
- Li, L., D. P. Patterson, C. C. Fox, B. Lin, P. W. Coschigano, y E. N. G. Marsh.** 2009. Subunit structure of benzylsuccinate synthase. *Biochemistry* **48**:1284-1292.
- Lin, B., H. W. Van Verseveld, y W. F. Roling.** 2002. Microbial aspects of anaerobic BTEX degradation. *Biomed. Environ. Sci.* **15**:130-144.
- López-Barragán, M. J., M. Carmona, M. T. Zamarro, B. Thiele, M. Boll, G. Fuchs, J. L. García, y E. Díaz.** 2004a. The *bzd* gene cluster, coding for anaerobic benzoate catabolism, in *Azoarcus* sp. strain CIB. *J. Bacteriol.* **186**:5762-5774.
- López-Barragán, M. J., E. Díaz, J. L. García, y M. Carmona.** 2004b. Genetic clues on the evolution of anaerobic catabolism of aromatic compounds. *Microbiology* **150**:2018-2021.
- Lovley, D. R.** 2003. Cleaning up with genomics: applying molecular biology to bioremediation. *Nat. Rev. Microbiol.* **1**:35-44.
- Lovley, D. R., y E. J. P. Phillips.** 1989. Requirement for a microbial consortium to completely oxidize glucose in Fe(III)-reducing sediments. *Appl. Environ. Microbiol.* **55**:3234-3236.

- Luengo, J. M., J. L. García, y E. R. Olivera.** 2001. The phenylacetyl-CoA catabolon: a complex catabolic unit with broad biotechnological applications. *Mol. Microbiol.* **39**:1434-1442.
- Maddocks, S. E., y P. C. F. Oyston.** 2008. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**:3609-3623.
- McLeod, M. P., y L. D. Eltis.** 2008. Genomic insights into the aerobic pathways for degradation of organic pollutants. p. 1-23. *In* E. Díaz (ed.), *Microbial Biodegradation: genomics and molecular biology*. Caister Academic Press, Norfolk, UK.
- Magasanik, B.** 1970. Glucose effects: inducer exclusion and repression. p.189-220. *En*: J. Beckwith, *The lactose operon*. Cold Spring Harbor, N. Y.
- Marchler-Bauer, A., y S. H. Bryant.** 2004. CD-Search: protein domain annotations on the fly. *Nucl. Acids Res.* **32**:W327-331.
- Marqués, S., I. Aranda-Olmedo, y J. L. Ramos.** 2006. Controlling bacterial physiology for optimal expression of gene reporter constructs. *Curr. Opin. Biotechnol.* **17**:50-56.
- Martín, A. C., R. López, y P. García.** 1995. Nucleotide sequence and transcription of the left early region of *Streptococcus pneumoniae* Cp-1 coding for the terminal protein and the DNA polymerase. *Virology* **211**:21-32.
- Martínez-Blanco, H., A. Reglero, L. B. Rodríguez-Aparicio, y J. M. Luengo.** 1990. Purification and biochemical characterization of phenylacetyl-CoA ligase from *Pseudomonas putida*. A specific enzyme for the catabolism of phenylacetic acid. *J. Biol. Chem.* **265**:7084-7090.
- Mascher, T., J. D. Helmann, y G. Udden.** 2006. Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol. Mol. Biol. Rev.* **70**:910-938.
- Merritt, E. A., y M. E. Murphy.** 1994. Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **1**:869-873.
- Milani, M., L. Leoni, G. Rampioni, E. Zennaro, P. Ascenzi, y M. Bolognesi.** 2005. An active-like structure in the unphosphorylated StyR response regulator suggests a phosphorylation- dependent allosteric activation mechanism. *Structure* **13**:1289-1297.
- Miller, J. H.** 1972. *Experiments in molecular genetics*. Cold Spring Harbord Laboratory, Cold Sprin Harbor, N.Y.
- Mohamed, M. E.** 2000. Biochemical and molecular characterization of phenylacetate-coenzyme A ligase, an enzyme catalyzing the first step in aerobic metabolism of phenylacetic acid in *Azoarcus evansii*. *J. Bacteriol.* **182**:286-294.
- Mohamed, M. E., W. Ismail, J. Heider, y G. Fuchs.** 2002. Aerobic metabolism of phenylacetic acids in *Azoarcus evansii*. *Arch. Microbiol.* **178**:180-92.

- Morales, G., J. F. Linares, A. Beloso, J. P. Albar, J. L. Martínez, y F. Rojo.** 2004. The *Pseudomonas putida* Crc global regulator controls the expression of genes from several chromosomal catabolic pathways for aromatic compounds. *J. Bacteriol.* **186**:1337-1344.
- Morasch, B., y R. U. Meckenstock.** 2005. Anaerobic degradation of *p*-xylene by a sulfate-reducing enrichment culture. *Curr. Microbiol.* **51**:127-130.
- Morasch, B., B. Schink, C. C. Tebbe, y R. U. Meckenstock.** 2004. Degradation of *o*-xylene and *m*-xylene by a novel sulfate-reducer belonging to the genus *Desulfotomaculum*. *Arch. Microbiol.* **181**:407-417.
- Moreno, R., y F. Rojo.** 2008. The target for the *Pseudomonas putida* Crc global regulator in the benzoate degradation pathway Is the BenR transcriptional regulator. *J. Bacteriol.* **190**:1539-1545.
- Moreno-Ruiz, E., M. J. Hernáez, O. Martínez-Perez, y E. Santero.** 2003. Identification and functional characterization of *Sphingomonas macrogolita* strain TFA genes involved in the first two steps of the tetralin catabolic pathway. *J. Bacteriol.* **185**:2026-2030.
- Morgenstern, B., A. Dress, y T. Werner.** 1996. Multiple DNA and protein sequence alignment based on segment-to-segment comparison. *Proc. Natl. Acad. Sci. U S A.* **93**:12098-12103.
- Nogales, J., A. Canales, J. Jiménez-Barbero, J. L. García, y E. Díaz.** 2005. Molecular characterization of the gallate dioxygenase from *Pseudomonas putida* KT2440: the prototype of a new subgroup of extradiol dioxygenases. *J. Biol. Chem.* **280**:35382-35390.
- Neuwald, A. F., L. Aravind, J. L. Spouge, y E. V. Koonin.** 1999. AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**:27-43.
- Niemetz, R., U. Altenschmidt, S. Brucker, y G. Fuchs.** 1995. Benzoyl-coenzyme-A 3-monooxygenase, a flavin-dependent hydroxylase. Purification, some properties and its role in aerobic benzoate oxidation via gentisate in a denitrifying bacterium. *Eur. J. Biochem.* **227**:161-168.
- OCDE** (1994) *Biothecnology for a clean environments: prevention, detection and remediation.* OCDE. Press, Paris.
- Ohtsubo, Y., M. Delawary, K. Kimbara, M. Takagi, A. Ohta, y Y. Nagata.** 2001. BphS, a key transcriptional regulator of *bph* genes involved in polychlorinated biphenyl/biphenyl degradation in *Pseudomonas* sp. KKS102. *J. Biol. Chem.* **276**:36146-36154.
- Ohtsubo, Y., H. Goto, Y. Nagata, T. Kudo, y M. Tsuda.** 2006. Identification of a response regulator gene for catabolite control from a PCB-degrading β -proteobacteria, *Acidovorax* sp. KKS102. *Mol. Microbiol.* **60**:1563-1575.

- Page, R. D. M.** 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**:357-358.
- Pao, G., y M. J. Saier.** 1995. Response regulators of bacterial signal transduction systems: selective domain shuffling during evolution. *J. Mol. Evol.* **40**:136-154.
- Parales, R. E., y C. S. Harwood.** 2002. Bacterial chemotaxis to pollutants and plant-derived aromatic molecules. *Curr. Opin. Microbiol.* **5**:266-273.
- Parke, D., M. A. García, y L. N. Ornston.** 2001. Cloning and genetic characterization of *dca* genes required for β -oxidation of straight-chain dicarboxylic acids in *Acinetobacter* sp. strain ADP1. *Appl. Environ. Microbiol.* **67**:4817-4827.
- Paul, R., S. Weiser, N. C. Amiot, C. Chan, T. Schirmer, B. Giese, y U. Jenal.** 2004. Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes Dev.* **18**:715-727.
- Pelletier, D. A., y C. S. Harwood.** 1998. 2-Ketocyclohexanecarboxyl coenzyme A hydrolase, the ring cleavage enzyme required for anaerobic benzoate degradation by *Rhodospseudomonas palustris*. *J. Bacteriol.* **180**:2330-2336.
- Pelletier, D. A., y C. S. Harwood.** 2000. 2-Hydroxycyclohexanecarboxyl Coenzyme A dehydrogenase, an enzyme characteristic of the anaerobic benzoate degradation pathway used by *Rhodospseudomonas palustris*. *J. Bacteriol.* **182**:2753-2760.
- Pérez-Martin, J., y V. de Lorenzo.** 1996. VTR expression cassettes for engineering conditional phenotypes in *Pseudomonas*: activity of the *Pu* promoter of the TOL plasmid under limiting concentrations of the XylR activator protein. *Gene* **172**:81-86
- Peters, F., M. Rother, y M. Boll.** 2004. Selenocysteine-containing proteins in anaerobic benzoate metabolism of *Desulfococcus multivorans*. *J. Bacteriol.* **186**:2156-2163.
- Peters, F., Y. Shinoda, M. J. McInerney, y M. Boll.** 2007. Cyclohexa-1,5-diene-1-carbonyl-coenzyme A (CoA) hydratases of *Geobacter metallireducens* and *Syntrophus aciditrophicus*: evidence for a common benzoyl-CoA degradation pathway in facultative and strict anaerobes. *J. Bacteriol.* **189**:1055-1060.
- Petruschka, L., G. Burchhardt, C. Müller, C. Weihe, y H. Herrmann.** 2001. The *cyo* operon of *Pseudomonas putida* is involved in carbon catabolite repression of phenol degradation. *Mol. Genet. Genomics* **266**:199-206.
- Pflüger, K., y V. de Lorenzo.** 2008. Evidence of in vivo cross talk between the nitrogen-related and fructose-related branches of the carbohydrate phosphotransferase system of *Pseudomonas putida*. *J. Bacteriol.* **190**:3374-3380.
- Philipp, B., y B. Schink.** 2000. Two distinct pathways for anaerobic degradation of aromatic compounds in the denitrifying bacterium *Thauera aromatica* strain AR-1. *Arch. Microbiol.* **173**:91-96.

- Pieper, D. H., y W. Reineke.** 2000. Engineering bacteria for bioremediation. *Curr. Opin. Biotechnol.* **11**:262-270.
- Pieper, D. H., V. A. Martins dos Santos, y P. N. Golyschin.** 2004. Genomic and mechanistic insights into the biodegradation of organic pollutants. *Curr. Opin. Biotechnol.* **15**:215-224.
- Ponting, C., y L. Aravind.** 1997. PAS: a multifunctional domain family comes to light. *Curr. Biol.* **7**:R674-677.
- Pratt, J. T., R. Tamayo, A. D. Tischler, and A. Camilli.** 2007. PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*. *J. Biol. Chem.* **282**:12860-12870.
- Prieto, M., B. Galán, B. Torres, A. Ferrández, C. Fernández, B. Miñambres, J. L. García, y E. Díaz.** 2004. Aromatic metabolism versus carbon availability: the regulatory network that controls catabolism of less-preferred carbon sources in *Escherichia coli*. *FEMS Microbiol. Rev.* **28**:503-518.
- Rabus, R.** 2005. Functional genomics of an anaerobic aromatic-degrading denitrifying bacterium, strain EbN1. *Appl. Microbiol. Biotechnol.* **68**:580-587.
- Rabus, R., M. Kube, J. Heider, A. Beck, K. Heitmann, F. Widdel, y R. Reinhardt.** 2005. The genome sequence of an anaerobic aromatic-degrading denitrifying bacterium, strain EbN1. *Arch. Microbiol.* **183**:27-36.
- Rabus, R., y F. Widdel.** 1995. Anaerobic degradation of ethylbenzene and other aromatic hydrocarbons by new denitrifying bacteria. *Arch. Microbiol.* **163**:96-103.
- Rojo, F.** 2001. Mechanisms of transcriptional repression. *Curr. Opin. Microbiol.* **4**:145-151.
- Rojo, F., y A. Dinamarca.** 2004. Catabolite repression and physiological control. p. 365-387. In J. L. Ramos (ed.), *Pseudomonas vol.2: Virulence and Gene Regulation*. Kluwer Academic, New York.
- Romanov, V., y R. P. Hausinger.** 1996. NADPH-dependent reductive ortho dehalogenation of 2,4-dichlorobenzoic acid in *Corynebacterium sepedonicum* KZ-4 and *Coryneform bacterium* strain NTB-1 via 2,4-dichlorobenzoyl coenzyme A. *J. Bacteriol.* **178**:2656-2661.
- Romeo, T., M. Gong, M. Y. Liu, y A. M. Brun-Zinkernagel.** 1993. Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J. Bacteriol.* **175**:4744-4755.

- Ross, P., R. Mayer, H. Weinhouse, D. Amikam, Y. Huggirat, M. Benziman, E. de Vroom, A. Fidder, P. de Paus, y L. A. Sliedregt.** 1990. The cyclic diguanylic acid regulatory system of cellulose synthesis in *Acetobacter xylinum*. Chemical synthesis and biological activity of cyclic nucleotide dimer, trimer, and phosphothioate derivatives. *J. Biol. Chem.* **265**:18933-18943.
- Römling, U., M. Gomelsky, y M. Y. Galperin.** 2005. C-di-GMP: the dawning of a novel bacterial signalling system. *Mol. Microbiol.* **57**:629-639.
- Römling, U., y D. Amikam.** 2006. Cyclic di-GMP as a second messenger. *Curr. Opin. Microbiol.* **9**:218-228.
- Rost, R., S. Haas, E. Hammer, H. Herrmann, y G. Burchhardt.** 2002. Molecular analysis of aerobic phenylacetate degradation in *Azoarcus evansii*. *Mol. Genet. Genomics* **267**:656-663.
- Ruíz, R., M. Isabel Aranda-Olmedo, P. Domínguez-Cuevas, M. I. Ramos-González, y S. Marqués.** 2004. Transcriptional regulation of the toluene catabolic pathways. p. 509-537. *In* J. L. Ramos (ed.), *Pseudomonas vol.2: Virulence and Gene Regulation*. Kluwer Academic, New York.
- Ryjenkov, D. A., R. Simm, U. Römling, y M. Gomelsky.** 2006. The PilZ domain is a receptor for the second messenger c-di-GMP: the PilZ domain protein YcgR controls motility in enterobacteria. *J. Biol. Chem.* **281**:30310-30314.
- Saier, M. H., Jr.** 2000. Families of transmembrane transporters selective for amino acids and their derivatives. *Microbiology* **146**:1775-1795.
- Sambrook, J., y D. W. Russell.** 2001. *Molecular cloning: A laboratory manual*. Volume 1-3. Joseph Sambrook and David W Russell. 3rd Edition. New York: Cold Spring Harbor Laboratory Press.
- Sanger, F., S. Nicklen, y A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U S A.* **74**:5463-5467.
- Santos, P. M., L. Leoni, I. Di Bartolo, and E. Zennaro.** 2002. Integration host factor is essential for the optimal expression of *styABCD* operon in *Pseudomonas fluorescens* ST. *Res. Microbiol.* **153**:527-536.
- Sarand, I., S. Österberg, S. Holmqvist, P. Holmfeldt, E. Skärfstad, R. E. Parales, y V. Shingler.** 2008. Metabolism-dependent taxis towards (methyl)phenols is coupled through the most abundant of three polar localized Aer-like proteins of *Pseudomonas putida*. *Environ. Microbiol.* **10**:1320-1334.
- Schäfer, A., A. Tauch, W. Jäger, J. Kalinowski, G. Thierbach, y A. Pühler.** 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**:69-73.

- Schell, M. A.** 1993. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **47**:597-626.
- Schink, B., B. Philipp, y J. Muller.** 2000. Anaerobic degradation of phenolic compounds. *Naturwissenschaften* **87**:12-23.
- Schmidt, A. J., D. A. Ryjenkov, y M. Gomelsky.** 2005. The ubiquitous protein domain EAL Is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. *J. Bacteriol.* **187**:4774-4781.
- Schöcke, L., y B. Schink.** 1999. Energetics and biochemistry of fermentative benzoate degradation by *Syntrophus gentianae*. *Arch. Microbiol.* **171**:331-337.
- Schühle, K., J. Gescher, U. Feil, M. Paul, M. Jahn, H. Schägger, y G. Fuchs.** 2003. Benzoate-coenzyme A ligase from *Thauera aromatica*: an enzyme acting in anaerobic and aerobic pathways. *J. Bacteriol.* **185**:4920-4929.
- Schühle, K., M. Jahn, S. Ghisla, y G. Fuchs.** 2001. Two similar gene clusters coding for enzymes of a new type of aerobic 2-aminobenzoate (anthranilate) metabolism in the bacterium *Azoarcus evansii*. *J. Bacteriol.* **183**:5268-5278.
- Selmer, T., A. J. Pierik, y J. Heider.** 2005. New glycyl radical enzymes catalysing key metabolic steps in anaerobic bacteria. *Biol. Chem.* **386**:981-988.
- Serre, L., B. Vallée, N. Bureaud, F. Schoentgen, y C. Zelwer.** 1998. Crystal structure of the phosphatidylethanolamine-binding protein from bovine brain: a novel structural class of phospholipid-binding proteins. *Structure* **6**:1255-1265.
- Shingler, V.** 2003. Integrated regulation in response to aromatic compounds: from signal sensing to attractive behaviour. *Environ. Microbiol.* **5**:1226-1241.
- Shinoda, Y., J. Akagi, Y. Uchihashi, A. Hiraishi, H. Yukawa, H. Yurimoto, Y. Sakai, y N. Kato.** 2005. Anaerobic degradation of aromatic compounds by *Magnetospirillum* strains: isolation and degradation genes. *Biosci. Biotechnol. Biochem.* **69**:1483-1491.
- Shinoda, Y., Y. Sakai, H. Uenishi, Y. Uchihashi, A. Hiraishi, H. Yukawa, H. Yurimoto, y N. Kato.** 2004. Aerobic and anaerobic toluene degradation by a newly isolated denitrifying bacterium, *Thauera* sp. strain DNT-1. *Appl. Environ. Microbiol.* **70**:1385-1392.
- Sofia, H. J., G. Chen, B. G. Hetzler, J. F. Reyes-Spindola, y N. E. Miller.** 2001. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucl. Acids Res.* **29**:1097-1106.
- Spormann, A. M., y F. Widdel.** 2000. Metabolism of alkylbenzenes, alkanes, and other hydrocarbons in anaerobic bacteria. *Biodegradation* **11**:85-105.

- Srivatsana, A., y J. D. Wang.** 2008. Control of bacterial transcription, translation and replication by (p)ppGpp. *Curr. Opin. Microbiol.* **11**:100-105.
- Stülke, J., y W. Hillen.** 1999. Carbon catabolite repression in bacteria. *Curr. Opin. Microbiol.* **2**:195-201.
- Sudarsan, N., E. R. Lee, Z. Weinberg, R. H. Moy, J. N. Kim, K. H. Link, y R. R. Breaker.** 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. *Science* **321**:411-413.
- Swank, R. T., y K. D. Munkres.** 1971. Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium dodecyl sulfate. *Anal. Biochem.* **39**:462-477.
- Swoboda-Colberg, N.G.** 1995 "Chemical contamination of the environment: sources, types, and fates of synthetic organic chemicals." p.27-74. En: L.Y. Young y C.E. Cerniglia, *Microbial transformations and degradation of toxic organic chemicals.* Wiley-Liss, Inc., New York.
- Takenaka, S., S. Murakami, R. Shinke, K. Hatakeyama, H. Yukawa, y K. Aoki.** 1997. Novel genes encoding 2-aminophenol 1,6-dioxygenase from *Pseudomonas* species AP-3 growing on 2-aminophenol and catalytic properties of the purified enzyme. *J. Biol. Chem.* **272**:14727-14732.
- Tamayo, R., J. T. Pratt, y A. Camilli.** 2007. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annu. Rev. Microbiol.* **61**:131-148.
- Tamayo, R., A. D. Tischler, y A. Camilli.** 2005. The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. *J. Biol. Chem.* **280**:33324-33330.
- Taylor, B. L.** 2007. Aer on the inside looking out: paradigm for a PAS-HAMP role in sensing oxygen, redox and energy. *Mol. Microbiol.* **65**:1415-1424.
- Taylor, B. L., I. B. Zhulin, y M. S. Johnson.** 1999. Aerotaxis and other energy-sensing behavior in bacteria. *Annu. Rev. Microbiol.* **53**:103-128.
- Teodorescu, O., T. Galor, J. Pillardy, y R. Elber.** 2004. Enriching the sequence substitution matrix by structural information. *Proteins* **54**:41-48.
- Thompson, J. D., D. G. Higgins, y T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**:4673-4680.
- Trautwein, K., S. Kuhner, L. Wohlbrand, T. Halder, K. Kuchta, A. Steinbuchel, y R. Rabus.** 2008. Solvent stress response of the denitrifying bacterium *Aromatoleum aromaticum* strain EbN1. *Appl. Environ. Microbiol.* **74**:2267-2274.
- Tropel, D., y J. R. van der Meer.** 2004. Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol. Mol. Biol. Rev.* **68**:474-500.

- Ullmann, A.** 1996. Catabolite repression: a story without end. *Res. Microbiol.* **147**:455-458.
- Vaillancourt, F. H., J. T. Bolin, y L. D. Eltis.** 2004. Ring-cleavage dioxygenases., p. 359-395, *Pseudomonas*, vol. 3. Kluwer Academic/Plenum Publishers, New York.
- van der Meer, J. R.** 1997. Evolution of novel metabolic pathways for the degradation of chloroaromatic compounds. *Antonie Van Leeuwenhoek* **71**:159-178.
- van der Meer, J. R., W. M. de Vos, S. Harayama, y A. J. Zehnder.** 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol. Rev.* **56**:677-694.
- Velasco, A., S. Alonso, J. L. García, J. Perera, y E. Díaz.** 1998. Genetic and functional analysis of the styrene catabolic cluster of *Pseudomonas* sp. strain Y2. *J. Bacteriol.* **180**:1063-1071.
- Venturi, V., F. Zennaro, G. Degrassi, B. C. Okeke, y C. V. Bruschi.** 1998. Genetics of ferulic acid bioconversion to protocatechuic acid in plant-growth-promoting *Pseudomonas putida* WCS358. *Microbiology* **144**:965-973.
- VerBerkmoes, N. C., M. B. Shah, P. K. Lankford, D. A. Pelletier, M. B. Strader, D. L. Tabb, W. H. McDonald, J. W. Barton, G. B. Hurst, L. Hauser, B. H. Davison, J. T. Beatty, C. S. Harwood, F. R. Tabita, R. L. Hettich, y F. W. Larimer.** 2006. Determination and comparison of the baseline proteomes of the versatile microbe *Rhodospseudomonas palustris* under its major metabolic states. *J. Proteome Res.* **5**:287-298.
- Verfürth, K., A. J. Pierik, C. Leutwein, S. Zorn, y J. Heider.** 2004. Substrate specificities and electron paramagnetic resonance properties of benzylsuccinate synthases in anaerobic toluene and *m*-xylene metabolism. *Arch. Microbiol.* **181**:155-162.
- Weber, H., C. Pesavento, A. Possling, G. Tischendorf, y R. Hengge.** 2006. Cyclic-di-GMP-mediated signalling within the sigmaS network of *Escherichia coli*. *Mol. Microbiol.* **62**:1014-1034.
- Widdel, F., y R. Rabus.** 2001. Anaerobic biodegradation of saturated and aromatic hydrocarbons. *Curr. Opin. Biotechnol.* **12**:259-276
- Wing, H. J., J. Green, J. R. Guest, y S. J. W. Busby.** 2000. Role of activating region 1 of *Escherichia coli* FNR protein in transcription activation at class II promoters. *J. Biol. Chem.* **275**:29061-29065.
- Wirth, R., A. Friesenegger, y S. Fiedler.** 1989. Transformation of various species of gram-negative bacteria belonging to 11 different genera by electroporation. *Mol. Gen. Genet.* **216**:175-177.

Wischgoll, S., D. Heintz, F. Peters, A. Erxleben, E. Sarnighausen, R. Reski, A. Van Dorsselaer, y M. Boll. 2005. Gene clusters involved in anaerobic benzoate degradation of *Geobacter metallireducens*. *Mol. Microbiol.* **58**:1238-1252.

Wöhlbrand, L. Wilkes, H. , T. Halder, y R. Rabus. 2008. Anaerobic degradation of *p*-ethylphenol by *Aromatoleum aromaticum* strain EbN1: pathway, regulation, and involved proteins. *J. Bacteriol.* **190**:5699-5709.

Wolanin, P., P. Thomason, y J. Stock. 2002. Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol.* **3**:reviews3013.1-8.

Yang, Y., S. Yuan, T. Chen, P. Ma, G. Shang, y Y. Dai. 2009. Cloning, heterologous expression, and functional characterization of the nicotinate dehydrogenase gene from *Pseudomonas putida* KT2440. *Biodegradation.*

Zaar, A., W. Eisenreich, A. Bacher, y G. Fuchs. 2001. A novel pathway of aerobic benzoate catabolism in the bacteria *Azoarcus evansii* and *Bacillus stearothermophilus*. *J. Biol. Chem.* **276**:24997-25004.

Zaar, A., J. Gescher, W. Eisenreich, A. Bacher, y G. Fuchs. 2004. New enzymes involved in aerobic benzoate metabolism in *Azoarcus evansii*. *Mol. Microbiol.* **54**:223-238.

Zhuang, Z., K.-H. Gartemann, R. Eichenlaub, y D. Dunaway-Mariano. 2003. Characterization of the 4-hydroxybenzoyl-Coenzyme A thioesterase from *Arthrobacter* sp. strain SU. *Appl. Environ. Microbiol.* **69**:2707-2711.

Zhulin, I. B., B. L. Taylor, y R. Dixon. 1997. PAS domain S-boxes in archaea, bacteria and sensors for oxygen and redox. *Trends Biochem. Sci.* **22**:331-333

Zschocke, J., E. Quak, P. Guldborg, y G. F. Hoffmann. 2000. Mutation analysis in glutaric aciduria type I. *J. Med. Genet.* **37**:177-181.