

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Farmacia

Departamento de Microbiología II

**CARACTERIZACIÓN MOLECULAR DE LA RUTA CATABÓLICA
DEL 4-HIDROXIFENILACETATO DE *Escherichia coli* W**

Memoria que para optar al Grado de Doctor en Farmacia presenta:

M^a AUXILIADORA PRIETO JIMENEZ

Director de la Tesis:

Dr. José Luis García López

Investigador Científico

Consejo Superior de Investigaciones Científicas

Madrid, 1995

A Raza

A mis padres

Quiero expresar mi agradecimiento a aquellas personas que me han ayudado en la realización de esta Tesis.

Especialmente, quiero agradecer a mi director, José Luis García López, la formación científica que de él he recibido, siendo para mí un punto de referencia no sólo por su alto nivel científico sino también por la comprensión y el respeto con el que trata a las personas que le rodean. Realmente, me siento afortunada por haber tenido la oportunidad de trabajar con él durante estos años.

Por otra parte, quisiera mostrar mi más profundo agradecimiento al Dr. Rubens López por haberme aceptado en su laboratorio, por sus valiosos consejos y por el inmejorable trato humano que siempre me ha dispensado. También quiero agradecer a los doctores Pedro García y Ernesto García la gran ayuda profesional y moral que me han brindado en todo momento, demostrando que son personas con las que siempre se puede contar. Además, quiero hacer mención a la Dra. Concepción Ronda, que en los difíciles momentos iniciales, supo crear un ambiente de amistad y compañerismo a su alrededor, facilitándome la integración en el grupo.

De manera especial, agradezco sinceramente a todos mis compañeros de laboratorio con los que he compartido mis buenos y malos momentos, la amistad, comprensión y a veces paciencia que siempre han mostrado hacia mí. Igualmente, agradezco la valiosa colaboración de Isabel Jado y la excelente asistencia técnica de Eloisa Cano, Manuel Carrasco, Paquita Morante, Aurelio Hurtado y Ricardo Uña que ha facilitado considerablemente la realización de este trabajo.

También quisiera agradecer al Dr. Kenneth Timmis del G.B.F. (Braunschweig, Alemania) y a todo su grupo, el haberme permitido trabajar durante un tiempo en su laboratorio, poniendo a mi disposición todo lo necesario para llevar a cabo mi proyecto con éxito. Quisiera mencionar especialmente al Dr. Eduardo Díaz porque sin su ayuda moral y profesional mi estancia en este centro no habría sido tan fructífera.

A los doctores Víctor de Lorenzo, Juan Luis Ramos, Miguel Vicente, Miguel Angel Peñalva, Alicia Gibello y Amando Garrido Pertierra les agradezco su ayuda y colaboración en diversos aspectos de este trabajo.

Mi agradecimiento al director del departamento de Microbiología de la Facultad de Farmacia, el Dr. César Nombela, por haber facilitado la lectura de esta Tesis y haber aceptado ser ponente de la misma.

También quisiera agradecer a mis amigos, en particular a los miembros de la N.B.A y M.B.A., el ánimo que en todo momento han sabido transmitirme.

*Por último, quisiera agradecer a toda mi familia el constante apoyo y comprensión que siempre me han ofrecido y particularmente a Rafa, que ha compartido con respeto y cariño todas mis alegrías y penas durante este tiempo, hasta tal punto que en la actualidad es el único abogado capaz de sorprender a propios y extraños con "amenas" charlas sobre el papel de *Escherichia coli* en la descontaminación medioambiental.*

Este trabajo ha sido realizado gracias a la concesión de una Beca de Formación de Personal Investigador de la Comunidad de Madrid.

ABREVIATURAS

ϵ :	coeficiente de extinción molar
Δ :	deleción
2-HPA:	ácido 2-hidroxifenilacético
3-HPA:	ácido 3-hidroxifenilacético
4-HPA:	ácido 4-hidroxifenilacético
Ap ^r :	resistencia a ampicilina
C12O:	catecol 1,2-dioxigenasa
C23O:	catecol 2,3-dioxigenasa
cAMP:	adenosín-monofosfato cíclico
CAP:	proteína receptora de cAMP
Ci:	curio
Cm ^r :	resistencia a cloranfenicol
CoA:	coenzima A
CHM:	5-carboximetil-2-hidroxi-muconato
CHMS:	5-carboximetil-2-hidroxi-mucónico semialdehido
Da	dalton
dATP:	desoxiadenosín 5'-trifosfato
dCTP:	desoxicitosín 5'-trifosfato
EDTA:	etilendiaminotetraacetato
FAD:	flavín-adenin-dinucleótido
FMN:	flavín-mononucleótido
g:	aceleración de la gravedad
HCA:	ácido fenilpropiónico
HHDD:	2-hidroxi-hept-2,4-dien-1,7-dioato
HHED:	2,4-dihidroxi-hept-2-en-1,7-dioato
HMS:	2-hidroxi-mucónico semialdehido
HPC:	homoprotocatecuato
HPLC:	cromatografía líquida de alta resolución
IPTG:	isopropil- β -D-tiogalactopiranosido
kb:	1000 pares de bases

Km ^r :	resistencia a kanamicina
LB:	medio de cultivo de Luria y Bertani
MHP:	ácido 3-(3-hidroxifenil)propiónico
MMGs:	microorganismos modificados genéticamente
NADH:	nicotinamida-adenín-dinucleótido
NADPH:	fosfato de nicotinamida-adenín-dinucleótido
Nal ^r :	resistencia a ácido nalidíxico
OHED:	2-oxo-hept-3-en-1,7-dioato
ONPG:	<i>o</i> -nitrofenil-galactósido
OPET:	5-oxo-pent-3-en-1,2,5-tricarboxilato
ORF:	marco de lectura abierta
p/v:	relación peso/volumen
P34O:	protocatecuato 3,4-dioxigenasa
PA:	ácido fenilacético
PGA:	penicilina G acilasa
Py:	piruvato
RBS:	sitio de unión al ribosoma
Rif ^r :	resistencia a rifampicina
SDS:	dodecil sulfato sódico
SS:	succinato semialdehido
Tc ^r :	resistencia a tetraciclina
Tris:	2-amino-2-hidroximetilpropano-1,3-diol
v/v:	relación volumen/volumen
X-Gal:	5-bromo-4-cloro-3-indolil- β -D-galactopiranosido

ÍNDICE

<u>I. INTRODUCCIÓN</u>	13
1. ACUMULACIÓN DE COMPUESTOS AROMATICOS ENLA BIOSFERA. UN PROBLEMA MEDIOAMBIENTAL	14
2. ADAPTACIÓN INDUCIDA DE LOS MICROORGANISMOS A LOS COMPUESTOS AROMÁTICOS CONTAMINANTES	15
3. RUTAS CATABÓLICAS DE COMPUESTOS AROMÁTICOS EN BACTERIAS	17
3.1. Metabolismo anaerobio	17
3.1.1. Bacterias fotosintéticas	18
3.1.2. Bacterias desnitrificantes	19
3.1.3. Bacterias sulfato-reductoras	21
3.1.4. Fermentación	22
3.2. Metabolismo aerobio	22
3.2.1. Rutas periféricas	26
3.2.1.1. Oxigenasas	26
3.2.1.2. Compuestos aromáticos halogenados	32
3.2.1.3. Compuestos nitroaromáticos	35
3.2.1.4. Compuestos aromáticos azufrados	36
3.2.2. Ruta del catecol	37
3.2.2.1. Ruta meta	37
3.2.2.2. Ruta orto	40
3.2.3. Ruta del gentisato	42
3.2.4. Regulación de las rutas catabólicas de compuestos aromáticos	42
4. AMPLIACIÓN DIRIGIDA DE LA CAPACIDAD BIODEGRADATIVA BACTERIANA	46
5. LAS ENTEROBACTERIAS Y LA BIODEGRADACIÓN DE COMPUESTOS AROMÁTICOS	48
5.1. La capacidad biodegradativa de <i>E. coli</i>	50
5.1.1. Rutas catabólicas de compuestos aromáticos de <i>E.coli</i>	52
5.2. Fuentes naturales de los compuestos aromáticos metabolizables por <i>E. coli</i>	56

6. OBJETIVOS	60
<u>II. MATERIALES Y MÉTODOS.</u>	62
1. ESTIRPES BACTERIANAS Y PLÁSMIDOS	63
2. MEDIOS Y CONDICIONES DE CULTIVO	64
3. PROCEDIMIENTOS DE TRANSFERENCIA GENÉTICA	64
3.1 Transformación	64
3.2. Conjugación.....	65
4. MANIPULACIÓN DEL DNA CON ENZIMAS DE USO COMÚN EN BIOLOGÍA MOLECULAR.....	65
5. PREPARACIÓN DE PLÁSMIDOS Y DNA CROMOSÓMICO.....	65
6. ELECTROFORESIS DEL DNA EN GELES DE AGAROSA	65
7. AISLAMIENTO Y PURIFICACIÓN DE LOS FRAGMENTOS DE DNA.....	66
7.1. Geles de poliacrilamida	66
7.2. Técnica de GeneClean.....	66
7.3. Geles de agarosa de bajo punto de fusión.....	67
7.4. Técnica de la β-agarasa	67
8. SELECCIÓN DE LOS CLONES PRODUCTORES DE PGA.....	67
9. TÉCNICAS DE HIBRIDACIÓN DE DNA	68
9.1. Técnica de Southern-blot.....	68
9.2. Hibridación de DNA sobre filtros de colonias celulares.....	68
10. SECUENCIACIÓN DEL DNA	69
10.1. Secuenciación manual	69
10.2. Secuenciación automática	69
10.3. Análisis de la secuencia de DNA	69

11. OBTENCIÓN DE EXTRACTOS DE CULTIVOS CELULARES	70
12. ENSAYOS DE ACTIVIDAD ENZIMÁTICA	
12.1 Penicilina G acilasa	70
12.2. 4-HPA-hidroxilasa	71
12.2.1. Cuantificación de la oxidación del NADH	71
12.2.2. Reacción acoplada con la catecol 2,3-dioxigenasa I (C23O I) de <i>P. putida</i>	71
12.3. HPC-dioxigenasa	72
12.4. β -lactamasa	72
12.5. β -galactosidasa	72
13. IDENTIFICACIÓN DE LOS PRODUCTOS DE LA REACCIÓN DE LA 4-HPA-HIDROXILASA	73
14. DETECCIÓN DE PROTEÍNAS CODIFICADAS POR FRAGMENTOS CLONADOS DE DNA	73
15. PURIFICACIÓN DE LA 4-HPA-HIDROXILASA	74
16. CUANTIFICACIÓN DE PROTEÍNAS	74
17. ELECTROFORESIS DE PROTEÍNAS EN GELES DE POLIACRIAMIDA- SDS	74
18. DETERMINACIÓN DE LA SECUENCIA N-TERMINAL DE HpaB	75
19. OBTENCIÓN DE ANTICUERPOS ANTI-HpaB	75
20. TÉCNICA DE WESTERN-BLOT	76
21. AISLAMIENTO DE RNA	76
22. DETERMINACIÓN DE LOS SITIOS DE INICIACIÓN DE LA TRANSCRIPCIÓN MEDIANTE EL EMPLEO DE LA TÉCNICA DE EXTENSIÓN CON UN INICIADOR	77
22.1. Marcaje del oligonucleótido con $[\gamma\text{-}^{32}\text{P}]$ ATP.	77
22.2. Incorporación de $[\alpha\text{-}^{32}\text{P}]$ dCTP.	78

III. RESULTADOS	79
1. CLONACIÓN DEL GEN <i>pac</i> DE <i>E. coli</i> ATCC 11105	80
1.1. Localización del gen responsable del fenotipo negro	80
1.2. Detección de las proteínas clonadas en el plásmido pAJ19	83
2. CARACTERIZACIÓN DE LA HIDROXILASA	85
2.1. Identificación de los productos de la reacción de hidroxilación	85
2.2. Estudio de los cofactores requeridos por la hidroxilasa	88
2.3. Especificidad de sustrato	88
2.4. Secuenciación del operón de la 4-HPA-hidroxilasa	90
2.5. Implicación de la proteína HpaC en la actividad hidroxilasa	93
2.6. Purificación de la hidroxilasa HpaB	97
2.7. Presencia de genes homólogos a <i>hpaB</i> en otros microorganismos...	100
3. CLONACIÓN Y SECUENCIACIÓN DEL OPERÓN <i>meta</i> DE LA RUTA DEL 4-HPA	102
3.1. Clonación del operón <i>meta</i> y construcción del mapa físico de la ruta del 4-HPA	102
3.2. Secuenciación del <i>cluster</i> de genes que componen la ruta del 4-HPA de <i>E. coli</i> ATCC 11105	104
3.3. Análisis de las regiones flanqueantes a la ruta del 4-HPA y su localización en el cromosoma de <i>E. coli</i>	128
3.4. Las regiones intergénicas	133
4. REGULACIÓN DEL OPERÓN <i>hpaBC</i>	136
4.1. Localización del promotor P_{BC}	136
4.2. Determinación de los inductores del operón <i>hpaBC</i>	139
4.3. Localización de la región reguladora del operón <i>hpaBC</i>	139
4.4. Implicación de las proteínas HpaA y HpaX en la regulación del operón <i>hpaBC</i>	140
4.5. Represión por catabolito de la 4-HPA-hidroxilasa	141
4.6. Determinación de los sitios de iniciación de la transcripción de los promotores P_A y P_{BC}	142
4.7. El promotor P_X	143
5. EXPRESIÓN DE LOS GENES DE LA RUTA DEL 4-HPA EN ORGANISMOS HETERÓLOGOS	153

5.1 Construcción de módulos transponibles con los genes de la ruta del 4-HPA	153
5.2 Integración de la ruta del 4-HPA en el cromosoma de <i>E. coli</i> K12 ..	154
5.3 Producción de la 4-HPA-hidroxilasa en <i>P. putida</i> KT2442	155

IV. DISCUSIÓN..... 163

1. CARACTERIZACIÓN DE LA 4-HPA-HIDROXILASA DE *E. coli* W 164

2. COMPARACIÓN DE LA RUTA DEL 4-HPA DE *E. coli* W CON OTRAS RUTAS CATABÓLICAS 168

2.1. Comparación de los genes *hpa* con los genes estructurales de otras rutas metabólicas de derivados aromáticos..... 168

2.2. Los genes reguladores de la ruta del 4-HPA 173

2.3. La organización física de los genes..... 176

3. PERSPECTIVAS DE FUTURO EN LA UTILIZACIÓN DE LOS GENES DE LA RUTA DEL 4-HPA EN LA DESCONTAMINACIÓN MEDIO-AMBIENTAL 179

V. CONCLUSIONES 181

VI. BIBLIOGRAFÍA 184

Probablemente no va en contra de la ciencia sugerir que en un lugar u otro existe algún organismo que, bajo las condiciones adecuadas, pueda oxidar cualquier sustancia...
(Gale, 1952)

I. INTRODUCCIÓN

1. ACUMULACIÓN DE COMPUESTOS AROMÁTICOS EN LA BIOSFERA. UN PROBLEMA MEDIOAMBIENTAL

La gran capacidad biodegradativa de los microorganismos se conoce desde hace décadas. De hecho, el *principio de infalibilidad microbiana* enunciado por Gale en 1952, proponía que, en las condiciones adecuadas, todas las sustancias naturales podían ser degradadas por algún microorganismo (Stanier, 1986). Ciertamente, la mayoría de los compuestos naturales se degradan con facilidad, aunque existen claras excepciones como la lignina, presente en la madera de los árboles, que es muy resistente a la biodegradación (Kuhad y Singh, 1993). Sin embargo, este principio no contemplaba la masiva introducción de una amplia variedad de compuestos aromáticos en la Biosfera como consecuencia de la actividad humana y el gran progreso industrial desarrollado desde mediados del pasado siglo (Smith, 1990). En líneas generales, podemos definir el término *xenobiótico* como un compuesto sintético cuya estructura química no ha sido expuesta a los microorganismos a lo largo de la evolución (Leisinger y Brunner, 1986). La aparición reciente de estos compuestos plantea el problema de su integración en los grandes ciclos de recambio de la materia, que conduce a una acumulación de estas especies químicas debido a la resistencia que presentan frente al ataque microbiano. El impacto que la acumulación de este tipo de contaminantes puede provocar en un ecosistema viene dado por su concentración y su grado de toxicidad, pudiendo crear graves problemas de contaminación ambiental.

La etapa limitante en la biodegradación o detoxificación de estos compuestos es la eficiencia de las rutas catabólicas microbianas. Aunque a veces existen las enzimas adecuadas para llevar a cabo la completa mineralización de ciertos compuestos xenobióticos, estos procesos no son lo suficientemente rápidos como para mantener un equilibrio entre la liberación y eliminación de éstos, siendo la causa principal de este desfase la gran cantidad de vertidos que de forma incontrolada se liberan en la Biosfera (Smith, 1990). La eficacia y cinética de estos procesos catabólicos puede verse afectada por diversos factores como por ejemplo la accesibilidad del sustrato, la presencia de sustancias adsorbentes en el suelo, o factores fisico-químicos como la temperatura del

hábitat, pH, potencial redox, concentración de oxígeno, etc. (Swindol y cols., 1988; Spain y van Veld, 1983; Bottomley, 1993).

2. ADAPTACIÓN INDUCIDA DE LOS MICROORGANISMOS A LOS COMPUESTOS AROMÁTICOS CONTAMINANTES

La selección natural actúa favorablemente sobre una población cuando las variaciones introducidas permiten a dicha población obtener un crecimiento más rápido. La continua y rápida evolución genética característica de las bacterias ha facilitado la aparición y diseminación de rutas metabólicas para la degradación de una gran variedad de hidrocarburos aromáticos (Leisinger y Brunner, 1986). Sin embargo, la masiva liberación de agentes xenobióticos durante un corto espacio de tiempo, desde el punto de vista evolutivo, no ha permitido a los microorganismos desarrollar nuevas propiedades catabólicas con la rapidez suficiente para paliar el problema medioambiental originado por la acumulación de estos compuestos en distintos ecosistemas.

Para acelerar los procesos evolutivos implicados en la adquisición de nuevas rutas catabólicas se han diseñado en el laboratorio, mediante el uso de quimiostatos, diversos experimentos consistentes en la aplicación controlada de una presión selectiva sobre una estirpe, o una comunidad microbiana, que metabolice un compuesto parecido al que se quiere degradar (Dorn y cols., 1974). Estos estudios indican que es posible conseguir una completa mineralización de ciertos agentes xenobióticos sustituyendo lentamente la fuente de carbono metabolizable por el compuesto en estudio. El periodo de exposición al nuevo sustrato puede consistir en días o meses, denominándose *proceso de adaptación microbiana* al cambio que provoca la ampliación de la capacidad biodegradativa (Spain y van Veld, 1983). Estas estrategias se han utilizado con éxito en varias ocasiones y tienen la ventaja de ser sencillas y de no ser necesaria la caracterización *a priori* de las rutas catabólicas que intervienen en el proceso. La desventaja principal que presentan es que el resultado final no es totalmente predecible.

Los mecanismos que han desarrollado los microorganismos para llevar a cabo estos procesos consisten, por una parte, en la inducción de enzimas específicas para la

degradación del sustrato, y por otra, en la alteración de la actividad enzimática originada por procesos de adaptación genética como la recombinación, transposición o mutación. También pueden estar implicados procesos de transferencia genética, de forma que, por transformación, transducción o conjugación, el microorganismo adquiera genes que le permitan ampliar su capacidad catabólica (Reineke y Knackmuss, 1979; van der Meer y cols., 1992). Estos procesos se ven facilitados por el hecho de que muchas rutas catabólicas están localizadas en plásmidos transmisibles o transposones. En la tabla 1 se muestran algunos ejemplos de plásmidos que contienen genes implicados en rutas catabólicas de compuestos aromáticos (Frantz y Chakrabarty, 1986).

Otro procedimiento para ampliar la capacidad biodegradativa bacteriana consiste en la utilización de técnicas de ingeniería genética para la transferencia *in vitro* de genes clonados y/o de sistemas reguladores conocidos (de Lorenzo y Rojo, 1994). Esto da lugar a microorganismos modificados genéticamente (MMGs) que, por su importancia, se tratarán en otro apartado. La evolución *in vitro* requiere una caracterización exhaustiva de las rutas metabólicas que se quieren expandir, lo que justifica el gran avance que se ha producido en este campo durante las últimas décadas.

Tabla 1. Rutas catabólicas de compuestos aromáticos localizadas en plásmidos

Plásmido	Microorganismo	Ruta degradativa	Referencia
TOL	<i>Pseudomonas putida</i>	xileno, tolueno	(Duggleby y cols., 1977)
SAL1	<i>P. putida</i>	salicilato	(Chakrabarty, 1972) (Yen y cols., 1983)
NAH	<i>P. putida</i>	naftaleno	(Yen y cols., 1983)
NIC	<i>P. convexa</i>	nicotina, nicotinato	(Thacker y Gunsalus, 1979)
pRA500	<i>P. putida</i>	3,5-xilenol	(Hopper y Kemp, 1980) (Jain y cols., 1984)
pEG	<i>P. fluorescens</i>	estireno	(Bestetti y cols., 1984)
PCIT1	<i>Pseudomonas</i> sp.	anilina	(Anson y Mackinnon, 1984)
pJP4	<i>Alcaligenes eutrophus</i>	2,4-diclorofenoxiacetato	(Don y Pemberton, 1981)
pWR1	<i>Pseudomonas</i> sp.	3-clorobenzoato	(Reineke y Knackmuss, 1979)
pAC25	<i>P. putida</i>	3-clorobenzoato	(Chatterjee y Chakrabarty, 1983)
pVI150	<i>Pseudomonas</i> CF600	fenol	(Nordlund y cols., 1990)

3. RUTAS CATABÓLICAS DE COMPUESTOS AROMÁTICOS EN BACTERIAS

Junto con los residuos glucosídicos, el anillo bencénico constituye la unidad de estructura química más difundida en la Naturaleza, por lo que no es extraño que los microorganismos hayan desarrollado rutas metabólicas que les permitan mineralizar este tipo de compuestos (Smith, 1990). Sin embargo, la introducción artificial de grupos electronegativos, cloro o nitro, en el anillo aromático disminuye considerablemente su biodegradabilidad (Leisinger y Brunner, 1986).

El uso moderno del vocablo *aromático* no tiene ninguna relación con el aroma. Aunque el término se originó en los albores de la química orgánica cuando se descubrió que muchos compuestos que contenían anillos de benceno se caracterizaban por olores fuertes, en la actualidad se asocia con compuestos que tienen una estabilidad de índole *aromática* (Streiwieser y Heathcock, 1979). Los hidrocarburos aromáticos son muy estables debido a su configuración electrónica por lo que requieren una activación previa para ser metabolizados. El mecanismo de activación es diferente dependiendo de si se lleva a cabo en un microorganismo anaerobio o aerobio. En el primer caso, las reacciones que conllevan a la desestabilización del anillo aromático consisten, generalmente, en una tioesterificación del anillo bencénico mediante la incorporación enzimática de coenzima A (CoA). Sin embargo, los microorganismos aerobios, normalmente, desestabilizan la estructura resonante mediante una serie de reacciones de oxidación. En ambos casos, estas reacciones enzimáticas conducen a la formación de metabolitos intermediarios de rutas centrales como el ciclo de Krebs (Leisinger y Brunner, 1986).

3.1. Metabolismo anaerobio

Los ecosistemas anóxicos se originan cuando el consumo de oxígeno supera al suministro de éste, como sucede en aguas estancadas, sedimentos de lagos, plantas industriales que producen metano a partir de desechos orgánicos, plantas de depuración de aguas residuales, tracto digestivo de los animales, sedimento de los océanos, etc. El metabolismo anaerobio de los compuestos orgánicos y su mineralización a dióxido de

carbono y metano depende de la disponibilidad de la luz y de aceptores de electrones inorgánicos como nitrato, sulfato o dióxido de carbono.

En 1934, Tarvin y Buswell demostraron, mediante evidencias químicas, que al incubar un compuesto aromático en condiciones anaerobias, utilizando como medio fango y aguas residuales, este compuesto era completamente transformado en metano y dióxido de carbono. Posteriormente, y mediante el uso de benzoato marcado con ^{14}C , Clark y Fina (1952) demostraron que, en condiciones similares, el 50% del carbono procedente de benzoato era convertido en metano, con lo que definitivamente quedó demostrado que los microorganismos anaerobios eran capaces de catabolizar este tipo de compuestos. Actualmente ha surgido un gran interés por el estudio de este tipo de rutas dado que muchos ecosistemas anóxicos contienen grandes cantidades de residuos aromáticos procedentes de desechos industriales.

La clasificación de los microorganismos anaerobios capaces de metabolizar compuestos aromáticos se ha diseñado atendiendo al sistema que la bacteria utiliza para obtener energía (Evans y Fuchs, 1988) (tabla 2).

3.1.1. Bacterias fotosintéticas

Desde el punto de vista biodegradativo, las bacterias rojas no sulfúreas son las más interesantes dentro del grupo de bacterias fotosintéticas. Son microorganismos principalmente fotoheterótrofos, es decir, que utilizan la luz como fuente de energía y compuestos orgánicos como fuente de carbono. Las bacterias rojas no sulfúreas se presentan de modo típico en lagos y charcas de agua dulce, donde hay materia orgánica pero no hay sulfuro, o está en muy baja concentración. Dentro de este grupo de bacterias, se han descrito algunas especies pertenecientes al género *Rhodospseudomonas* como *R. palustris* que pueden metabolizar benzoato y ω -alcanofenilcarboxilatos en condiciones anaerobias y en presencia de luz (Elder y cols., 1992). El benzoato entra en la célula mediante un sistema de transporte dependiente de energía y es convertido en benzoil-CoA por una benzoil-CoA ligasa inducible. Posteriormente, mediante una serie de reacciones enzimáticas de reducción e hidratación, da lugar a 3-hidroxipimeil-CoA.

Este compuesto produciría acetil-CoA mediante una ruta metabólica similar a la β -oxidación de los ácidos grasos (Evans y Fuchs, 1988; Koch y cols., 1993; Gibson y Gibson, 1992).

Otra especie interesante es *R. gelatinosa* que es capaz de mineralizar floroglucinol. Se cree que el primer paso consiste en una deshidrogenación ya que el dihidrofloroglucinol se ha identificado en el medio de cultivo pero el resto de la ruta metabólica está aún por determinar (Whittle y cols., 1976).

3.1.2. Bacterias desnitrificantes

Muchas bacterias aerobias pueden utilizar nitrato, en lugar de oxígeno, como aceptor de electrones cuando las condiciones son de anaerobiosis. Siempre que la materia orgánica se descompone en el suelo o en el agua y se agota el oxígeno como resultado de la respiración aeróbica microbiana, algunos de estos microorganismos continuarán utilizando la materia orgánica si hay nitrato presente, es decir, mediante respiración anaerobia. En este proceso el nitrógeno combinado es eliminado del suelo y del agua con liberación de N_2 a la atmósfera.

En 1970, Taylor y cols. aislaron un microorganismo Gram negativo que fue clasificado como *Pseudomonas* estirpe PN-1, que era capaz de metabolizar el benzoato en condiciones aerobias mediante una ruta que conllevaba una hidroxilación que daba lugar al ácido protocatéquico y posterior ruptura del anillo en posición *meta*. En anaerobiosis y en presencia de nitrato este microorganismo también era capaz de utilizar el benzoato como única fuente de carbono. Se demostró que el primer paso de esta ruta consiste en la formación de benzoil-CoA. Posteriormente, Blake y Hegeman (1987) reclasificaron esta estirpe denominándola *Alcaligenes xylooxidans* subespecie *desnitrificans* y demostraron, mediante experimentos genéticos, que la ruta anaeróbica de degradación de benzoato estaba localizada en el plásmido conjugativo pCB1. Este microorganismo también puede mineralizar 2- y 4-fluorobenzoato mediante una reacción de deshalogenación.

Tabla 2. Metabolismo anaerobio de compuestos aromáticos

Obtención de energía	Microorganismo	Sustrato	Referencia
<u>Fotosíntesis</u>	<i>Rhodospseudomonas palustris</i>	benzoato	(Proctor y Scher, 1960)
	<i>R. gelatinosa</i>	<i>m,p</i> -hidroxibenzoato floroglucinol	(Dutton y Evans, 1967) (Whittle y cols., 1976)
<u>Desnitrificación</u>	<i>Alcaligenes xylooxidans</i>	benzoato hidroxibenzoatos fluorobenzoato	(Blake y Hegeman, 1987)
	<i>Paracoccus denitrificans</i>	protocatecuato vanilato	(Williams y Evans, 1975)
	<i>Parococcus</i> sp. <i>Bacillus</i> sp.	<i>o, p</i> -cresol <i>o, m, p</i> -ftalato	(Rudolphi y cols., 1991) (Afring y cols., 1981)
	<i>Pseudomonas</i>	2-aminobenzoato fenol <i>o, m, p</i> -cresol fenilacetato, 4-hidroxifenilacetato, fenol <i>p</i> -cresol benzoato	(Ziegler y cols., 1987) (Tschsch y Fuchs, 1987) (Bossert y Young, 1986) (Dangel y cols., 1991)
<u>Reducción de sulfato</u>	<i>Desulfovibrio</i> sp.	benzoato	(Evans y Fuchs, 1988)
	<i>Desulfococcus</i>	hidroxibenzoatos	(Evans y Fuchs, 1988)
	<i>Desulfonema</i>	fenilacetato	(Evans y Fuchs, 1988)
	<i>Desulfosarcina</i>	fenol indol	(Bak y Widdel, 1986a) (Bak y Widdel, 1986b)
<u>Fermentación</u>	<i>Coprococcus</i> sp.	floroglucinol	(Tsai y Jones, 1975)
	<i>Streptococcus</i>	resorcinol	(Tschsch y Schink, 1985)
	<i>Pelobacter acidigallici</i>	galato, pirogalol	(Schink y Pfennig, 1982)
	<i>Eubacterium oxidoreducens</i>	polifenoles, quercetina	(Krumholz y Bryant, 1986)
<u>Fermentación metanogénica</u>	Asociación microbiana: bacteria fermentativa + bacteria acetogénica y metanogénica	lignina benzoato tirosina cinamato fenilpropionato fenilacetato, benzoato fenol, catecol hidroquinona ferulato vanilato siringinato fenilalanina benceno, tolueno triptófano, indol clorobenceno clorofenoles clorobenzoatos clorofenoxiacetatos nitrofenoles cloroguaiacol	(Boruff y Buswell, 1934) (Tarvin y Buswell, 1934) (Mountfort y Bryant, 1982) (Balba y Evans, 1980a) (Balba y Evans, 1980b) (Healy y Young, 1978) (Balba y Evans, 1980c) (Healy y cols., 1980) (Kaiser y Haselmann, 1982) (Sleat y Robinson, 1983) (Balba y Evans, 1980a) (Grbic-Galic y Vogel, 1987) (Berry y cols., 1987) (Reineke y Knackmuss, 1984) (Boyd y cols., 1983) (Boyd y Shelton, 1984) (Suffita y cols., 1982) (Suffita y Gibson, 1985) (Neilson y cols., 1987)

Actualmente se sabe que existen otras estirpes del género *Pseudomonas* que, en condiciones desnitrificantes, degradan el benzoato y otros compuestos aromáticos como el hidroxibenzoato, el 2-aminobenzoato, el fenol, el ácido fenilacético (PA) y el ácido 4-hidroxifenilacético (4-HPA) (Altenschmidt y cols., 1991; Biegert y cols., 1993; Dangel y cols., 1991; Mohamed y cols., 1993; Mohamed y Fuchs, 1993; Elder y Kelly, 1994). En todos los casos interviene una acil-CoA ligasa que da lugar al tioéster correspondiente.

Mediante la intervención de diversas reacciones enzimáticas de α -oxidación, reducción o carboxilación estos compuestos darán lugar a benzoil-CoA como metabolito intermediario central que posteriormente es reducido e hidratado enzimáticamente por un mecanismo similar al propuesto en *R. palustris* (Koch y Fuchs, 1992; Koch y cols., 1993). Las acil-CoA ligasas son de vital importancia en los primeros pasos de las rutas catabólicas anaeróbicas de derivados aromáticos ya que desestabilizan el anillo aromático mediante tioesterificación.

3.1.3. Bacterias sulfato-reductoras

El hábitat típico de estos microorganismos anaerobios estrictos desasimiladores de sulfato son sedimentos anaeróbicos que contienen materia orgánica y sulfato. Las actividades de estos organismos dan como resultado una generación masiva de ácido sulfhídrico. Esto conduce con frecuencia al desarrollo de bacterias rojas y verdes sulfúreas, que utilizan el ácido sulfhídrico como donador fotosintético de electrones, oxidándolo a sulfato bajo condiciones anaeróbicas en presencia de luz. Esta acción combinada determina un ciclo anaeróbico típico del azufre. Los géneros *Desulfovibrio*, *Desulfococcus*, *Desulfonema* y *Desulfosarcina*, aislados de sedimentos marinos y de agua salada, son capaces de metabolizar algunos derivados aromáticos como el benzoato, el PA, el 3-fenilpropionato y el 2-, 3-, y 4-hidroxibenzoato, pero hasta ahora se desconoce el mecanismo por el cual este tipo de bacterias mineralizan estos compuestos (Evans y Fuchs, 1988).

3.1.4. Fermentación

La fermentación puede definirse como un proceso metabólico generador de ATP en el que los compuestos orgánicos sirven tanto de donadores de electrones (oxidándose) como de aceptores de electrones (reduciéndose).

En 1982, Schink y Pfennig aislaron de fango marino un microorganismo perteneciente a la familia *Bacteroidaceae* al que llamaron *Pelobacter acidigallici*. Este microorganismo es capaz de convertir ácido gálico, floroglucinol, pirogalol y 2,4,6-trihidroxibenzoato en acetato y CO₂. Durante las fermentaciones no se reducía nitrato ni sulfato. La asociación de este microorganismo con *Acetobacterium woodii* durante el proceso fermentativo permitía la utilización de un nuevo sustrato ya que esta comunidad bacteriana convertía el ácido síngico en acetato. El acoplamiento de diferentes rutas metabólicas permite a las comunidades microbianas utilizar ciertos compuestos aromáticos que no serían degradables por un único microorganismo. Este tipo de asociaciones co-metabólicas existen de forma natural en la Biosfera aumentando la capacidad biodegradativa de los microorganismos.

3.2. Metabolismo aerobio

Como se mencionó anteriormente, el mecanismo general de desestabilización del anillo aromático desarrollado por los microorganismos aerobios implica una oxidación progresiva de la estructura resonante. Actualmente se tiene un gran conocimiento, a nivel molecular, de algunas rutas catabólicas microbianas de compuestos aromáticos, especialmente de las pertenecientes al género *Pseudomonas* (van der Meer y cols., 1992). Una comparación detallada de estas rutas metabólicas indica que el primer paso consiste en la incorporación de dos grupos hidroxilo en el anillo bencénico para lo cual los microorganismos han desarrollado una serie de *rutas periféricas* de oxidación, deshalogenación, desnitración o desulfuración, que dan lugar a un intermediario dihidroxilado susceptible a la acción de dioxigenasas específicas que provocarán la apertura del anillo. Estos derivados dihidroxilados serán canalizados por dos rutas generales: por *la ruta del catecol*, en la que los intermediarios catecólicos pueden ser

metabolizados a su vez mediante la ruta *meta* o la ruta *orto* dependiendo de la posición en la que se efectúe la apertura del anillo, o por *la ruta del gentisato*. Los metabolitos finales de ambas rutas entran directamente en el metabolismo intermediario de la célula (figura 1).

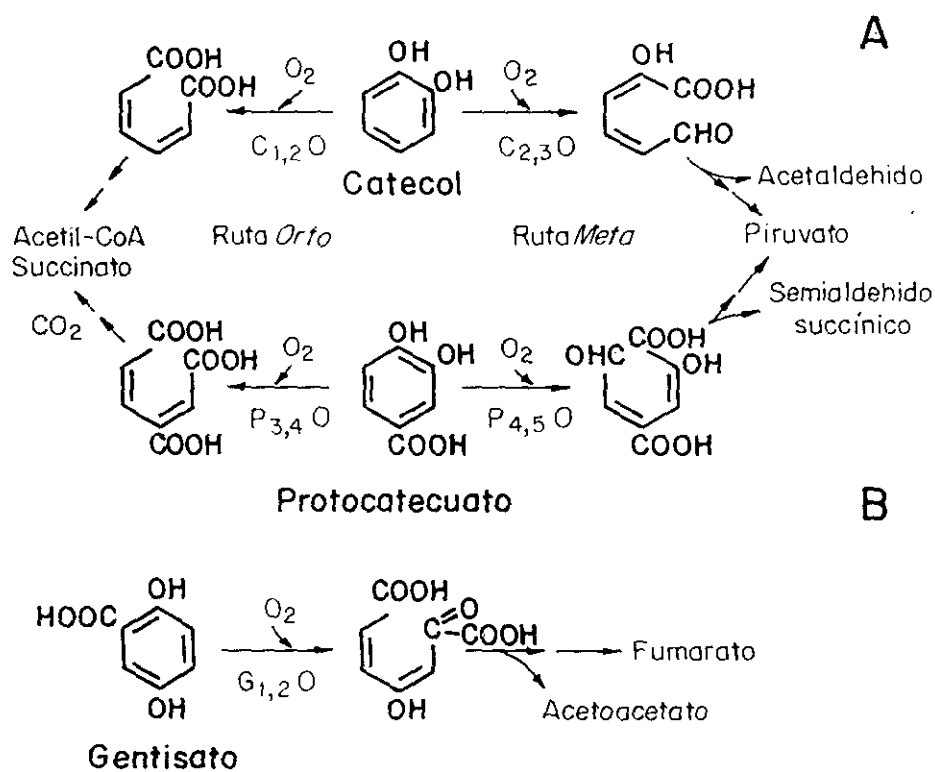


Figura 1. Rutas catabólicas centrales de compuestos aromáticos de microorganismos aerobios. Los hidrocarburos aromáticos siguen dos rutas generales: A, la ruta del catecol y B, la ruta del gentisato. En el primer caso, la ruptura del anillo aromático puede ser en posición 1,2 (*orto*), o en 2,3 (*meta*), según la enzima utilizada. Abreviaturas: C_{1,2}O, catecol 1,2-dioxigenasa; C_{2,3}O, catecol 2,3-dioxigenasa; P_{3,4}O, protocatecuato 3,4-dioxigenasa; P_{4,5}O, protocatecuato 4,5-dioxigenasa; G_{1,2}O, gentisato 1,2-dioxigenasa. (de Lorenzo y Rojo, 1994).

En general, los genes que codifican para las enzimas responsables de la degradación de compuestos aromáticos están estructurados en operones organizados en agrupaciones o *clusters* de genes que, en muchas ocasiones, están localizados en plásmidos transmisibles o en transposones (tabla 1), lo que permite su propagación. En la figura 2 se muestra la organización de algunas de estas rutas catabólicas.

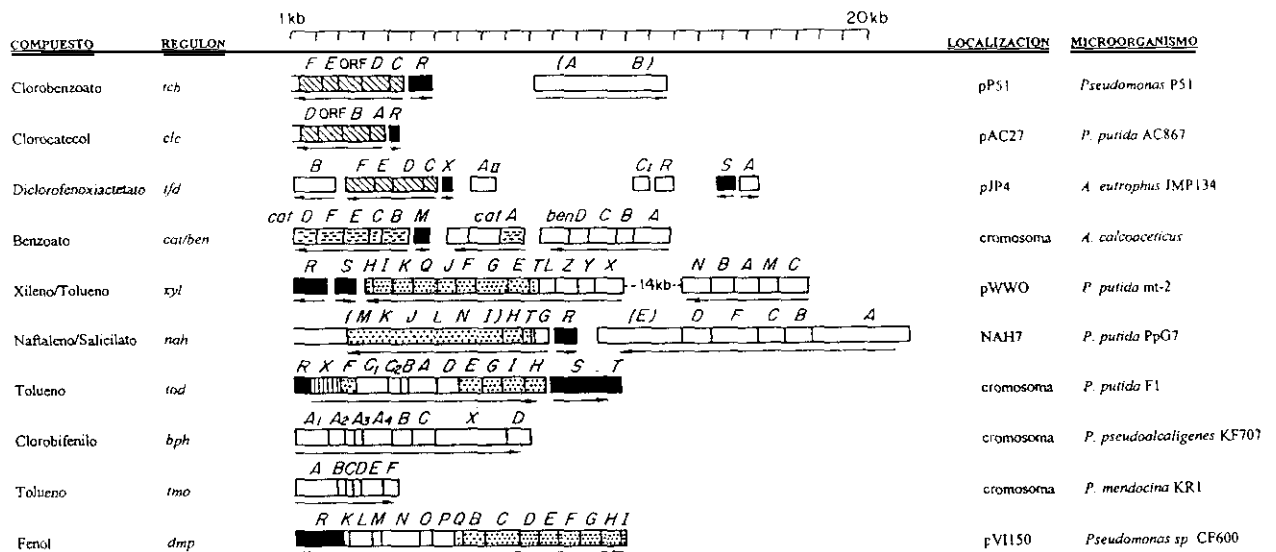


Figura 2. Organización física de algunas rutas catabólicas de compuestos aromáticos. □ genes que forman parte de las rutas periféricas; ■, genes reguladores; □, genes implicados en transporte; ▨, ruta *meta*; ▩, ruta *orto*; ▧, ruta *orto modificada*. Las flechas indican la dirección de transcripción de los genes. Los genes cuya secuencia de nucleótidos no se conoce en su totalidad están indicados entre paréntesis.

La ruta catabólica mejor caracterizada es la ruta TOL o ruta de degradación de xileno y tolueno de *Pseudomonas putida* (Marqués y Ramos, 1993). Esta ruta está localizada en los plásmidos TOL, de los cuales, el prototipo es el plásmido pWWO que fue descrito por primera vez en el año 1974 por Williams y Murray aunque sus propiedades catabólicas ya habían sido observadas en 1973 por Nakazawa y Yokota. Este plásmido pertenece al grupo de incompatibilidad P-9 y tiene un tamaño de 117 kb de las que, aproximadamente 40 kb están implicadas en la ruta catabólica (Downing y Broda, 1979). El plásmido pWWO es autotransmisible y su rango de huésped se restringe al género *Pseudomonas* y algunas *Enterobacteriaceae* (Nakazawa, 1978; Ramos-González y cols., 1991). Los genes implicados en esta ruta catabólica del xileno, o genes *xyl*, están localizados en el transposón Tn4651 de 56 kb que a su vez está incluido en el transposón Tn4653 de 70 kb, ambos pertenecientes a la familia de transposones Tn3 (Tsuda y cols., 1989). El plásmido TOL es termosensible ya que en *P. aeruginosa* no puede mantenerse a 42°C. Hasta ahora no han sido localizados con exactitud los genes responsables de la replicación y transferencia conjugativa pero se sabe que la expresión y formación del *pilus* dependiente de pWWO es constitutiva, lo que justifica el alto grado de transferencia de este plásmido a las estirpes receptoras (Ramos-González y cols., 1991).

Los genes catabólicos de la ruta TOL están organizados en dos operones llamados operón “*upper*” (*xylCMABN*) y operón “*meta*” (*xylXYZLTEGFJQKIH*). Los genes *xylTEGFJQKIH* codifican para proteínas que forman parte de la ruta *meta* para la degradación del catecol. El resto de los genes *xyl* constituyen una ruta periférica mediante la cual el grupo metilo del tolueno es secuencialmente oxidado hasta benzoato, o metilbenzoato si el producto inicial es xileno, por los productos de los genes del operón *upper*; posteriormente estos compuestos son oxidados y descarboxilados para producir catecol mediante la toluato 1,2-dioxigenasa (*xylXYZ*) y la dihidroxiciclohexadieno descarboxilasa deshidrogenasa (*xylL*), respectivamente.

3.2.1. Rutas periféricas

A lo largo de la evolución, los microorganismos han ampliado su capacidad biodegradativa mediante variaciones en las rutas periféricas como la adquisición de nuevas enzimas y la aparición espontánea de mutaciones que provocan un cambio en la especificidad de sustrato de las enzimas existentes. Normalmente la formación del intermediario dihidroxilado en las diferentes rutas periféricas está mediada por oxigenasas que, en muchos casos, presentan una gran similitud en su estructura primaria, lo cual indica que podrían derivar de un gen ancestral común (van der Meer y cols., 1992).

En este apartado se expondrán, de forma general, los tipos de enzimas implicadas en las rutas periféricas.

3.2.1.1. Oxigenasas

Las *oxigenasas* son enzimas que catalizan la inserción de oxígeno molecular en sustratos orgánicos. Estas enzimas se clasifican en *monooxigenasas* y *dioxigenasas*, dependiendo del número de átomos de oxígeno que incorporan en la molécula de sustrato. Las monooxigenasas incorporan sólo un átomo de oxígeno en el sustrato, reduciendo el otro hasta agua mediante una reacción de oxidación. Debido al acoplamiento entre las reacciones de oxigenación y oxidación estas enzimas también se denominan *oxidasas de función mixta*. Las dioxigenasas introducen dos átomos de oxígeno en el sustrato y se subdividen en *dioxigenasas hidroxilantes*, que incorporan dos residuos hidroxilo en el anillo aromático, y *dioxigenasas implicadas en la ruptura del anillo* que, a su vez, se subdividen en *extradiol dioxigenasas* o *intradiol dioxigenasas* dependiendo de si la apertura del anillo aromático dihidroxilado se lleva a cabo mediante un proceso de *meta*-escisión (entre los carbonos 2 y 3 en el caso del catecol) u *orto*-escisión (entre los carbonos 1 y 2 en el caso del catecol) respectivamente (Harayama y cols., 1992) (figura 1). Las dioxigenasas implicadas en la ruptura del anillo aromático se tratarán en el apartado 3.2.2. ya que normalmente no forman parte de las rutas periféricas.

Tabla 3. Oxigenasas implicadas en la hidroxilación de compuestos aromáticos

MONOOXIGENASAS			
Familia	Enzima	Microorganismo	Referencia
<u>MONOCOMPONENTES</u>			
<u>4-hidroxibenzoato 3-hidroxilasa</u>	4-hidroxibenzoato 3-hidroxilasa	<i>P. fluorescens</i>	(Berkel y cols., 1992)
	salicilato hidroxilasa	<i>P. putida</i>	(You y cols., 1991)
<u>Fenol 2-hidroxilasa</u>	2,4-diclorofenol 6-monooxigenasa	<i>Alcaligenes eutrophus</i>	(Perkins y cols., 1990)
	fenol 2-hidroxilasa	<i>Pseudomonas</i> EST1001	(Nurk y cols., 1991)
<u>Flavoproteínas no clasificadas</u>	2-nitrofenol hidroxilasa	<i>P. putida</i> B2	(Zeyer y Kocher, 1988)
	3-hidroxibenzoato 6-hidroxilasa	<i>P. cepacia</i> <i>Micrococcus</i> sp.	(Wang y cols., 1987) (Rajasekharan y cols., 1990)
<u>MULTICOMPONENTES</u>			
<u>Fenol 2-hidroxilasa</u>	fenol 2-hidroxilasa	<i>Pseudomonas</i> CF600	(Nordlund y cols., 1990)
	Tolueno 4-monooxigenasa	<i>P. mendocina</i>	(Yen y cols., 1991)
<u>No clasificadas</u>	4-hidroxifenilacetato 3-hidroxilasa	<i>P. putida</i>	(Arunachalam y cols., 1992)
<u>Alquil hidroxilasas</u>	xileno monooxigenasa	<i>P. putida</i>	(Suzuki y cols., 1991)
DIOXIGENASAS			
Clase	Enzima	Microorganismo	Referencia
I A	Ftalato dioxigenasa	<i>P. cepacia</i>	(Batie y cols., 1987)
	4-clorofenilacetato 3,4-dioxigenasa	<i>Pseudomonas</i> sp. CBS3	(Markus y cols., 1986)
	4-sulfobenzoato 3,4-dioxigenasa	<i>Comamonas testosteroni</i>	(Locher y cols., 1991)
I B	benzoato dioxigenasa	<i>P. arvilla</i> <i>Acinetobacter calcoaceticus</i>	(Yamaguchi y Fujisawa, 1982) (Neidle y cols., 1991)
	4-metoxibenzoato O-desmetilasa	<i>P. putida</i>	(Bernhardt y cols., 1975)
II A	Pirazón dioxigenasa	<i>Pseudomonas</i> sp.	(Sauber y cols., 1977)
II B	Benceno 1,2-dioxigenasa	<i>P. putida</i>	(Irie y cols., 1987)
	Tolueno dioxigenasa	<i>P. putida</i>	(Gibson y cols., 1982)
III	Naftaleno dioxigenasa	<i>P. putida</i>	(Ensley y Gibson, 1983)

Las monooxigenasas y dioxigenasas hidroxilantes requieren donadores de electrones externos como el NADH o NADPH y un centro redox, como mínimo, que permita la activación de la molécula de oxígeno. Estos cofactores suelen ser metales de transición como hierro, manganeso, cobre y cobalto, o una molécula de flavina o pteridina. La clasificación actual de las hidroxilasas de derivados aromáticos se ha establecido en función del peso molecular de las subunidades y de la similitud que presentan estas enzimas en su estructura primaria (tabla 3). A diferencia de las dioxigenasas, la mayoría de las monooxigenasas son flavoproteínas monocomponentes, aunque se han descrito algunos sistemas multicomponentes como la fenol 2-hidroxilasa de *Pseudomonas* CF600 que cataliza la conversión de fenol en catecol y está codificada por los genes *dmpKLMNOP* (Nordlund y cols., 1990), y la tolueno 4-monooxigenasa de *P. mendocina* KR1 codificada por los genes *tmoABCDEF*. Los componentes de estas dos monooxigenasas están relacionados entre sí y con otros sistemas enzimáticos con cadena de transporte electrónico como la metano monooxigenasa y algunas dioxigenasas hidroxilantes de compuestos aromáticos (Yen y cols., 1991).

Una monooxigenasa de interés para la discusión posterior de este trabajo es la 4-HPA-hidroxilasa de *P. putida* (Arunachalam y cols., 1992). Es un sistema enzimático que consta de dos componentes proteicos; una flavoproteína dimérica con dos subunidades idénticas de 30,7 kDa y una proteína cooperadora de 38,5 kDa. Los genes que codifican para este sistema todavía no han sido clonados, pero la purificación de los dos componentes enzimáticos ha permitido estudiar parcialmente su mecanismo de reacción (Arunachalam y cols., 1994; Arunachalam y Massey, 1994). La flavoproteína puede, por sí sola, catalizar la oxidación de NADH en presencia de 4-HPA independientemente de la incorporación del grupo hidroxilo en la molécula de sustrato, siendo necesaria la presencia de la proteína cooperadora para acoplar las reacciones de oxidación e hidroxilación (figura 3). La oxidación de NADH mediada por la flavoproteína también se produce en presencia de otros compuestos estructuralmente relacionados con el 4-HPA, como son el *p*-clorofenilacetato y el *p*-fluorofenilacetato, aunque sólo el 4-HPA y en menor medida el *p*-hidroxifenilpropionato, son susceptibles a la hidroxilación mediada por la proteína cooperadora. La proteína cooperadora no presenta ningún centro redox, lo que descarta la posibilidad de que actúe como un

receptor de electrones equivalente al componente oxigenasa terminal de la cadena de transporte electrónico característico de los sistemas multicomponentes.

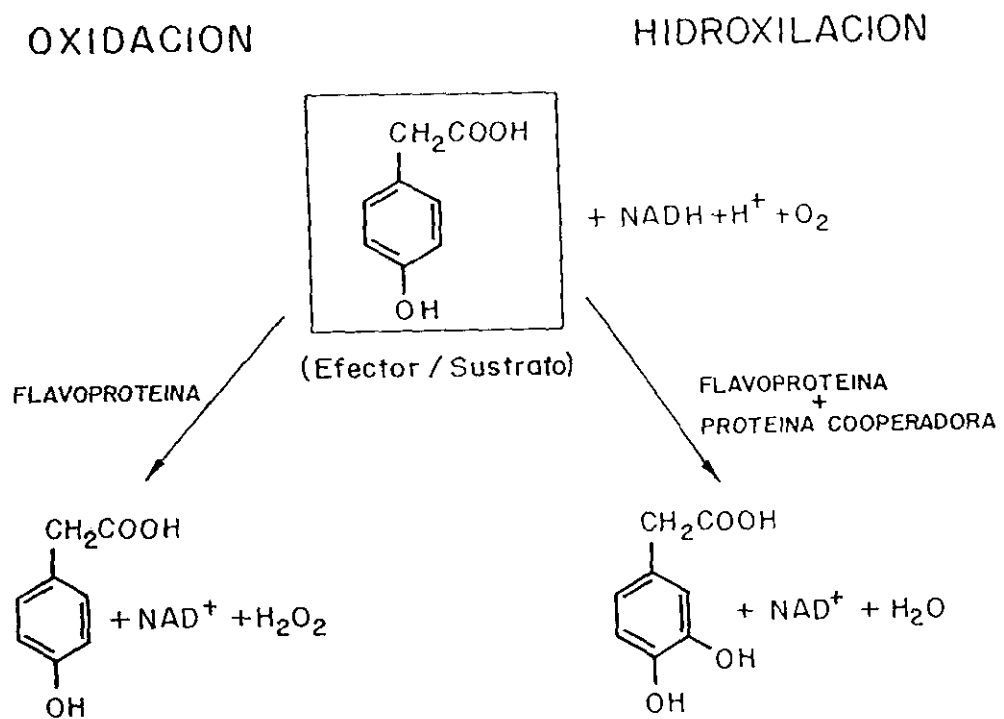


Figura 3. Reacción catalizada por el componente flavoproteico de la 4-HPA-hidroxilasa de *P. putida* en presencia y ausencia de la proteína cooperadora.

Las dioxigenasas hidroxilantes son sistemas enzimáticos multicomponentes. Los componentes proteicos de este tipo de dioxigenasas se dividen, desde el punto de vista funcional, en dos clases: oxigenasa terminal (o hidroxilasa) y componentes transportadores de electrones. Generalmente, los genes que codifican para las dioxigenasas hidroxilantes están agrupados con un gen que codifica para una dihidrodiol deshidrogenasa que pertenece a la familia de las alcohol deshidrogenasas (Manson y Cammack, 1992).

Las oxigenasas terminales son proteínas oligoméricas formadas por una o dos subunidades diferentes que presentan una configuración α_n ó $(\alpha\beta)_n$. Contienen al menos un núcleo redox de tipo "Rieske" y un átomo de hierro no hematínico responsable de la activación de oxígeno. El núcleo redox de tipo "Rieske" está formado por dos átomos de azufre y dos de hierro en el que, a diferencia de los centros redox de las ferredoxinas de las plantas donde el hierro y el azufre están coordinados a cuatro cisteínas, un átomo de hierro está coordinado a dos cisteínas y el otro a dos histidinas. Estos aminoácidos están conservados en el N-terminal de la subunidad α lo que ha permitido idear la siguiente secuencia consenso donde X indica cualquier aminoácido (Manson y Cammack, 1992):

C-X-H-15 a 17 aminoácidos-C-X-X-H

En el sitio de unión del átomo de hierro no hematínico están implicadas dos histidinas y dos tirosinas conservadas en la parte central de la subunidad α . Otras oxigenasas dependientes de hierro como las monooxigenasas de alcanos y xileno (Suzuki y cols., 1991) contienen regiones ricas en histidina y tirosina que están implicadas en la unión del átomo de hierro, pero su disposición es diferente a la de las oxigenasas terminales.

Además de la oxigenasa terminal, en la cadena de transporte electrónico intervienen una reductasa flavínica que reduce el NADH y otros componentes con núcleos redox de tipo ferredoxina o "Rieske". Atendiendo al número de componentes que formen el sistema y al grupo prostético que utilicen, las dioxigenasas se clasifican en (tabla 3):

-Clase I: Dioxigenasas de dos componentes en los que el núcleo redox azufre-hierro y el flavín nucleótido están combinados en una sola proteína. Dependiendo de si el cofactor requerido es FMN ó FAD, se subdividen en Clase IA y IB respectivamente.

-Clase II: Dioxigenasas de tres componentes en las que en el transporte electrónico están implicados una flavoproteína y una proteína con un núcleo redox de tipo ferredoxina (Clase IIA) o de tipo "Rieske" (Clase IIB).

-Clase III: Dioxigenasas de tres componentes que constan de una flavoproteína con un núcleo redox azufre-hierro y una ferredoxina.

A pesar de la gran diversidad que caracteriza a este tipo de enzimas, se piensa que algunas provienen de un mismo gen ancestral debido a la relación estructural y funcional que presentan. Como ejemplo podemos citar la benzoato 1,2-dioxigenasa de *A. calcoaceticus* (Neidle y cols., 1991) y la toluato 1,2-dioxigenasa de *P. putida*. Estas dos enzimas pertenecen a la Clase IB de las dioxigenasas hidroxilantes. La benzoato 1,2-dioxigenasa está codificada por los genes *benABC*, que pertenecen a la ruta *orto* de degradación de benzoato de *A. calcoaceticus*. Esta enzima está codificada en el cromosoma al igual que la benzoato 1,2-dioxigenasa de *P. arvilla* C-1, que es una enzima equivalente perteneciente a la ruta *orto* del benzoato en este microorganismo (Yamaguchi y cols., 1982). La toluato 1,2-dioxigenasa, codificada por los genes *xyXYZ* de la ruta TOL, es decir de la ruta *meta* de degradación de benzoato, está localizada en el plásmido pWWO a diferencia de las dos anteriores. La comparación de la secuencia de aminoácidos entre estas enzimas demuestra que existen grandes zonas de homología en todos sus componentes enzimáticos, fundamentalmente en los sitios de unión de los centros redox. Estas enzimas presentan la misma función pero difieren en la especificidad de sustrato ya que, aunque las tres son capaces de oxidar el benzoato, la toluato 1,2-dioxigenasa además de toluato, hidroxila otros alquilbenzoatos sustituidos en las posiciones 3 y 4 del anillo aromático. Este es un claro ejemplo de cómo los microorganismos amplían su capacidad metabólica mediante variaciones en sus rutas periféricas.

3.2.1.2. Compuestos aromáticos halogenados

Los hidrocarburos aromáticos halogenados constituyen uno de los grupos más amplios de compuestos xenobióticos. El uso de herbicidas, insecticidas, fungicidas y material dieléctrico en diversas industrias, que contienen mezclas de hasta 10^5 isómeros diferentes de compuestos polihalogenados como en el caso de los policlorobifenilos (PCBs) (Furukawa, 1994), contribuye a incrementar la resistencia a la biodegradación que normalmente presentan los compuestos aromáticos.

Los microorganismos capaces de metabolizar compuestos halogenados requieren la intervención de enzimas, denominadas comúnmente deshalogenasas, que específicamente catabolizan la escisión del enlace carbono-halógeno. Las reacciones de deshalogenación pueden formar parte de las rutas periféricas cuando se producen en un paso previo a la ruptura del anillo aromático, mientras que si se producen en un paso posterior a la apertura del anillo dan lugar a las rutas catabólicas modificadas.

Se han descrito siete mecanismos de deshalogenación que permiten clasificar las deshalogenasas en siete grupos (tabla 4):

- a) *Deshalogenasas reductasas*. La deshalogenación reductiva está catalizada por reductasas dependientes de NADH que catalizan la sustitución del halógeno por un átomo de hidrógeno mediante una reacción de oxidoreducción. Generalmente, este tipo de deshalogenación se lleva a cabo en un paso posterior a la *orto*-fisión del anillo aromático. Las enzimas mejor caracterizadas de este grupo son las cloromaleilacetato reductasas de las rutas de degradación de derivados aromáticos clorados de *A. eutrophus* JMP134 (Ghosal y You, 1988) y *Pseudomonas* P51 (van der Meer y cols., 1991b) codificadas por los genes *tfdF* y *tcbF* respectivamente. Ambos genes están localizados en plásmidos.

- b) *Deshalogenasas oxigenasas*. La deshalogenación oxigenolítica está mediada por monooxigenasas como la pentaclorofenol monooxigenasa de *Flavobacterium*

ATCC 33790, o por dioxigenasas como la 4-clorofenilacetato dioxigenasa de *Pseudomonas* sp. CBS3. A diferencia de las reductasas, las oxigenasas intervienen antes de la ruptura del anillo aromático dando lugar a un derivado dihidroxilado que seguirá metabolizándose por la ruta del catecol.

c) *Halurohidrolasas*. La deshalogenación hidrolítica conlleva la incorporación de un grupo hidroxilo procedente del agua mediante una sustitución nucleofílica catalizada por las halurohidrolasas. Estas enzimas utilizan el agua como único cosustrato. Las halurohidrolasas se han clasificado en función de la especificidad de sustrato y de su estructura primaria (Janssen y cols., 1994).

d) *Glutation-S-transferasas*. La deshalogenación tiolítica consiste en el desplazamiento nucleofílico de un átomo de halógeno por una molécula de glutatión, que posteriormente será a su vez desplazada por una reacción mediada por el mismo sistema enzimático. En algunas rutas de degradación de compuestos haloaromáticos intervienen conjuntamente dos sistemas enzimáticos de deshalogenación como en el caso de la ruta de degradación de pentaclorofenol de *Flavobacterium* ATCC 33790 (Orser y cols. 1993a y 1993b). En primer lugar, el pentaclorofenol sufre una deshalogenación oxigenolítica catalizada por una monooxigenasa codificada por el gen *pcpB*. Durante este proceso se elimina un átomo de cloro. Posteriormente, el derivado dihidroxilado es deshalogenado mediante una glutatión-S-transferasa codificada por el gen *pcpC*. Las glutatión-S-transferasas están implicadas no sólo en las rutas degradativas microbianas sino que también intervienen en mecanismos de detoxificación de compuestos aromáticos de organismos superiores.

e) *Deshalogenasas isomerasas*. Uno de los mecanismos desarrollados por los microorganismos para degradar compuestos haloaromáticos es la deshalogenación por sustitución intramolecular. El clorocatecol se produce como metabolito intermediario de muchas rutas degradativas de compuestos cloroaromáticos. Este intermediario se metaboliza en *Pseudomonas* B13 y en *A. eutrophus* JMP134 por una ruta de tipo *orto* denominada ruta *orto modificada* (ver apartado 3.2.2.2.). El

3-clorocatecol se transforma en el ácido 3-cloromucónico que, mediante la 3-cloromuconato cicloisomerasa, da lugar a un intermediario inestable que se descompone espontáneamente provocando la liberación del cloruro y la formación del ácido oxomucónico. Este compuesto se metaboliza directamente por la ruta *orto* de degradación de benzoato.

f) *Deshidrodeshalogenasas*. La deshidrohalogenación conlleva la formación de un doble enlace y la liberación de una molécula de haluro de hidrógeno. La γ -pentaclorociclohexano deshidrohalogenasa de *P. paucimobilis* UT26, codificada por el gen *linA*, transforma este producto en tetraclorociclohexadieno que a su vez puede ser transformado en triclorobenceno. Este último paso no es muy frecuente *in vivo* ya que el intermediario tetraclorado es metabolizado por un mecanismo hidrolítico.

g) *Deshalogenasas hidratasas*. Dentro de este grupo, la enzima mejor caracterizada es la 4-clorobenzoato deshalogenasa de *Pseudomonas* CBS3 (Babbitt y cols., 1992). Este microorganismo está especialmente capacitado para la mineralización de compuestos haloaromáticos ya que produce dos 4-clorobenzoato halurohidrolasas, una 4-clorofenilacetato oxigenasa, y una 4-clorobenzoato hidratasa que cataliza la conversión de 4-clorobenzoato en 4-hidroxibenzoato. La 4-clorobenzoato hidratasa es un sistema enzimático que consta de tres componentes proteicos: una CoA-ligasa que activa el anillo aromático mediante la incorporación de una molécula de CoA, una hidratasa que lleva a cabo la deshalogenación del anillo aromático activado incorporando un grupo hidroxilo, y una tioesterasa que moviliza el CoA dando lugar a 4-hidroxibenzoato. Este mecanismo de deshalogenación se ha detectado en otros microorganismos como *Arthrobacter* sp. US (Crooks y Copley, 1993).

Tabla 4. Deshalogenasas

Enzimas	Ruta	Microorganismo	Gen	Referencia
<u>Cloromaleilacetoreductasas</u>	2,4-diclorofenoxiacetato 1,2,4-triclorobenceno	<i>A. eutrophus</i> JMP134 (pJP4) <i>Pseudomonas</i> P51 (pP51)	<i>tfdF</i> <i>tcbF</i>	(Ghosal y You, 1988) (v. d. Meer y cols., 1991b)
<u>Monooxigenasas</u>	pentaclorofenol	<i>Flavobacterium</i> ATCC 33790	<i>pcpB</i>	(Orser y cols., 1993b)
<u>Hidrolasas</u>	4-clorobenzoato 4-clorobenzoato	<i>Pseudomonas</i> CBS3 <i>Pseudomonas</i> CBS3	<i>dehCI</i> <i>dehCI</i>	(Schneider y cols., 1991)
<u>Glutation-S-transferasas</u>	pentaclorofenol	<i>Flavobacterium</i> ATCC 39723	<i>pcpC</i>	(Orser y cols., 1993a)
<u>Cloromuconatocicloisomerasa</u>	3-clorobenzoato 2,4-diclorofenoxiacetato 1,2,4-triclorobenceno	<i>P. putida</i> AC867 (pAC27) <i>A. eutrophus</i> JMP134 (pJP4) <i>Pseudomonas</i> P51 (pP51)	<i>clcB</i> <i>tfdD</i> <i>tcbD</i>	(Ghosal y You, 1988) (Perkins y cols., 1990) (v. d. Meer y cols., 1991b)
<u>Deshidrodeshalogenasas</u>	γ -hexaclorociclohexano	<i>P. paucimobilis</i> UT26	<i>linA</i>	(Imai y cols., 1991)
<u>Hidratasas</u>	4-clorobenzoato 4-clorobenzoato	<i>Pseudomonas</i> CBS3 <i>Arthrobacter</i> SU	ORF1 ORF1	(Babbitt y cols., 1992) (Crooks y Copley, 1993)

3.2.1.3. Compuestos nitroaromáticos

Los compuestos nitroaromáticos se producen a gran escala en los procesos de fabricación de colorantes, plásticos y explosivos. Son compuestos altamente tóxicos y su liberación provoca grandes daños no sólo al medio ambiente, sino que también puede afectar directamente al hombre ya que, incluso a muy bajas dosis, producen cianosis, hepatotoxicidad, anemia e ictericia.

Se han descrito algunos microorganismos capaces de mineralizar nitrobencono, nitrofenoles, cloronitrofenoles, nitroanilinas, nitrobenzoatos, dinitrobenzoatos e incluso 2,4,6-trinitrotolueno (TNT) (Higson, 1992). El TNT es un compuesto derivado de la fabricación de explosivos que se caracteriza por su notable estabilidad química que le confiere una gran resistencia al ataque microbiano. Los microorganismos han

desarrollado dos mecanismos generales para mineralizar este tipo de compuestos. Por una parte, una reducción progresiva del grupo nitro generalmente llevada a cabo por nitroreductasas pertenecientes a rutas centrales, que conduce a la formación de aminas y amoniaco. La otra alternativa conlleva la liberación de nitrito mediante una reacción de oxidación. Estos dos mecanismos coexisten en una estirpe de *P. putida* aislada por Zeyer y Kearney (1984) ya que es capaz de mineralizar 2-nitrofenol, mediante la formación de nitrito y catecol, y por otra parte es capaz de metabolizar 3-nitrofenol liberándose amonio en el proceso.

Actualmente el conocimiento a nivel molecular de este tipo de rutas catabólicas es muy escaso, aunque el aislamiento de microorganismos degradadores de compuestos nitroaromáticos permitirá el desarrollo de este tipo de estudios.

3.2.1.4. Compuestos aromáticos azufrados

El azufre está presente en combustibles fósiles en forma de azufre pirítico, tioles, sulfuros, tiofenos, dibenzotiofenos y ácidos sulfónicos. La combustión de carbón y petróleo da lugar, además de dióxido de carbono y agua, a óxidos de azufre y nitrógeno que son muy perjudiciales ya que generan *lluvias ácidas* que ocasionan grandes daños en los ecosistemas vegetales. La reducción del contenido en azufre y nitrógeno en este tipo de combustibles por métodos químicos y biológicos contribuiría a paliar este problema. Existen varios procedimientos químicos y microbiológicos para eliminar el azufre inorgánico del carbón y petróleo (Monticello y Finnerty, 1985), pero el verdadero problema lo plantea el azufre que forma parte de los hidrocarburos aromáticos que componen estos combustibles. Se han descrito microorganismos y comunidades microbianas capaces de oxidar selectivamente el enlace C-S facilitando la liberación del azufre en forma de sulfato, pero hasta ahora, las rutas catabólicas implicadas sólo se han estudiado con compuestos modelo como el dibenzotiofeno y el ácido naftalen-sulfónico (de Lorenzo y Rojo, 1994).

3.2.2. Ruta del catecol

Mediante las rutas periféricas muchos derivados aromáticos se transforman en catecol como metabolito intermediario, lo que justifica que comúnmente se le denomine “embudo catabólico” o “cuello de botella” de las rutas catabólicas de compuestos aromáticos. Como ya se ha mencionado, dependiendo de la posición en la que se efectúe la ruptura del anillo, el catecol será mineralizado por la ruta *meta* o la ruta *orto* (figura 1).

3.2.2.1. Ruta *meta*

Un compuesto aromático sigue una ruta de degradación de tipo *meta* cuando la ruptura del anillo se lleva a cabo por una extradiol dioxigenasa (tabla 5). En la figura 2 están indicadas algunas rutas catabólicas de tipo *meta*.

Tabla 5. Dioxigenasas implicadas en la ruptura del anillo aromático

Familia	Enzima	Microorganismo	Referencia
<u>EXTRADIOL DIOXIGENASAS</u>			
<u>Catecol 2,3-dioxigenasa I</u>	Catecol 2,3-dioxigenasa I	<i>P. putida</i> (TOL)	(Nakai y cols., 1983)
	2,3-Dihidroxibifenil dioxigenasa	<i>P. paucimobilis</i>	(Taira y cols., 1988)
	1,2-Dihidroxinaftaleno dioxigenasa	<i>P. putida</i>	(Harayama y Rekik, 1989)
<u>Protocatecuato 4,5-dioxigenasa</u>	Protocatecuato 4,5-dioxigenasa	<i>P. paucimobilis</i>	(Nebert y González, 1987)
	Catecol 2,3-dioxigenasa II	<i>A. eutrophus</i>	(Kabisch y Fortnagel, 1990)
<u>Homoprotocatecuato dioxigenasa</u>	Homoprotocatecuato dioxigenasa	<i>Escherichia coli</i> C	(Roper y Cooper, 1990a)
<u>INTRADIOL DIOXIGENASAS</u>			
<u>Catecol 1,2-dioxigenasa I</u>	Catecol 1,2-dioxigenasa I	<i>P. aeruginosa</i>	(Kukor y cols., 1988)
	Catecol 1,2-dioxigenasa II (Clorocatecol 1,2-dioxigenasa)	<i>Pseudomonas</i> sp. P51	(van der Meer y cols., 1991b)
	Protocatecuato 3,4-dioxigenasa	<i>A. calcoaceticus</i>	(Hartnett y cols. 1990)
<u>GENTISATO 1,2-DIOXIGENASA</u>			
	Gentisato 1,2 dioxigenasa	<i>P. testosteroni</i> <i>P. acidovorans</i>	(Harpel y Lipscomb, 1990a y b)

La extradiol dioxigenasa mejor caracterizada es la catecol 2,3-dioxigenasa I (C23O I) que está codificada por el gen *xylE* y forma parte de la ruta TOL de *P. putida*. La C23O I está formada por cuatro subunidades idénticas de 32 kDa y contiene un catión ferroso en cada subunidad. El producto de la reacción es el 2-hidroxi-mucónico semialdehído o su alquilderivado, dependiendo de si el sustrato inicial es catecol o alquilatecol, que es fácilmente detectable por ser de color amarillo. El rango de sustrato de esta enzima es bastante amplio, oxida 3-metil-, 3-etil-, 4-metil- y 4-cloro-catecol. La comparación de la secuencia de aminoácidos de la C23O I de *P. putida* con la catecol 2,3-dioxigenasa II de *A. eutrophus* sugiere que estas enzimas no derivan del mismo origen por lo que ambas proteínas deben pertenecer a dos superfamilias diferentes.

Los pasos siguientes a la *meta*-escisión en la ruta TOL pueden seguir dos ramas diferentes que convergen en el intermediario 2-oxopent-4-enoato (figura 4). La vía utilizada dependerá de si el semialdehído, producto de la dioxigenasa, proviene del *m*-toluato o si proviene del benzoato o del *p*-metilbenzoato. En el primer caso, el semialdehído correspondiente es metabolizado mediante una hidrolasa codificada por *xylF* mientras que, en el segundo caso, la ruta sigue la rama del oxalocrotonato en la que intervienen tres enzimas: la hidroximucónico semialdehído deshidrogenasa (XylG), la 4-oxalocrotonato tautomerasa (XylH) y la 4-oxalocrotonato descarboxilasa (XylI). El 2-oxopent-4-enoato es posteriormente hidratado por XylJ y mediante la acción de una aldolasa (XylK) da lugar a acetaldehído y piruvato que entran directamente en el metabolismo intermediario de la célula (Marqués y Ramos, 1993).

Otras rutas catabólicas de tipo *meta* como, por ejemplo, la ruta de degradación de fenol de *Pseudomonas* sp CF600 (Powlowski y Shingler, 1994) y la ruta de degradación de naftaleno de *P. putida* (Assinder y Williams, 1988), siguen un esquema similar al de la ruta TOL y se ha demostrado, mediante técnicas de hibridación de DNA y comparación de secuencias de aminoácidos, que algunas de las enzimas con función equivalente pertenecientes a diferentes rutas proceden de un gen ancestral común (Williams y Sayers, 1994).

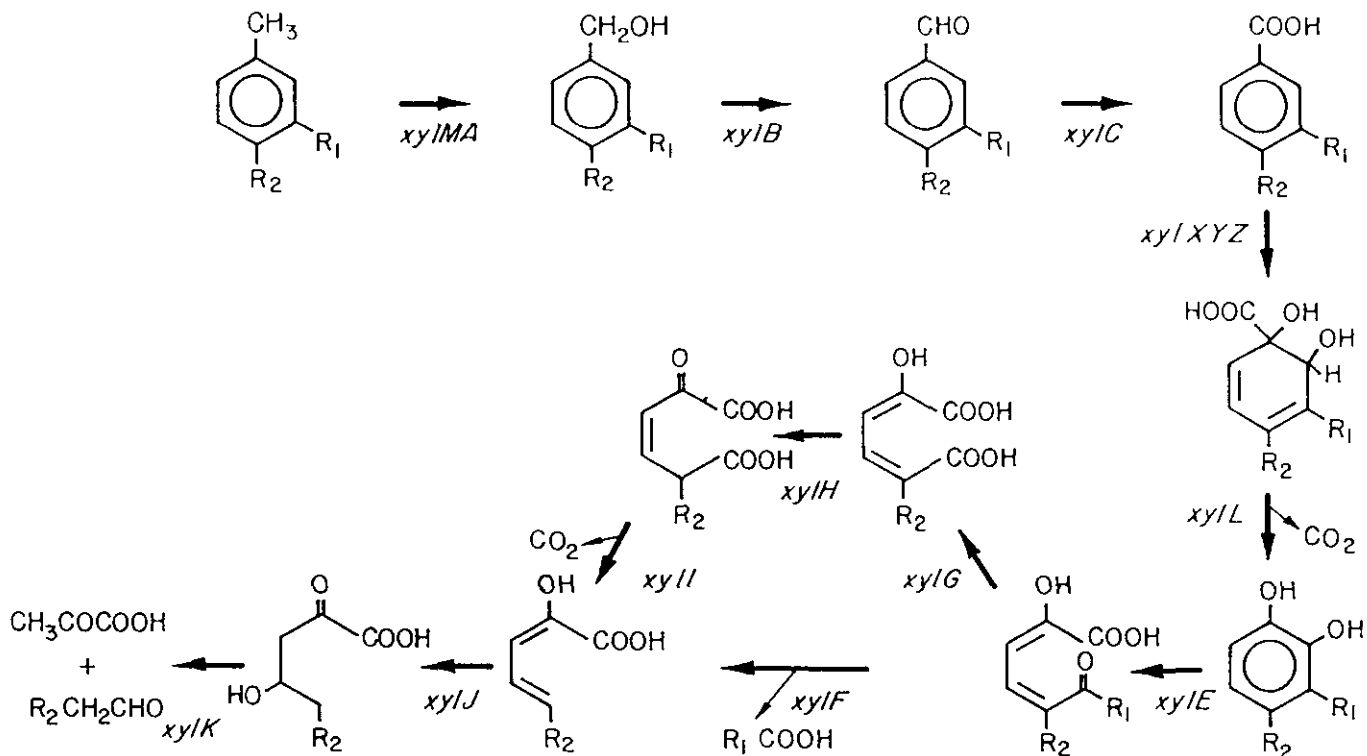


Figura 4. Pasos enzimáticos de la ruta TOL. Los sustituyentes del anillo aromático pueden ser: R₁ = R₂ = -H; R₁ = -CH₃, R₂ = -H; R₁ = H, R₂ = -CH₃; R₁ = -C₂H₅, R₂ = -H. Los genes *xyl* del operón *upper* (*xylCMABN*) y del operón *meta* (*xylXYZLEGFJKIH*) codifican para las siguientes enzimas: XylMA, xileno oxidasa; XylB, bencil alcohol deshidrogenasa; XylC, benzaldehído deshidrogenasa; XylXYZ, toluato 1,2-dioxigenasa; XylL, dihidroxiciclohexadieno descarboxilasa deshidrogenasa; XylE, catecol 2,3-dioxigenasa; XylF, hidroximucónico semialdehído hidrolasa; XylG, hidroximucónico semialdehído deshidrogenasa; XylH, 4-oxalocrotonato tautomerasa; XylI, 4-oxalocrotonato descarboxilasa; XylJ, 2-oxopent-4-3-enoato hidratasa; XylK, 2-oxo-4-hidropentanoato aldolasa (Marqués y Ramos, 1993).

3.2.2.2. Ruta orto

Un compuesto aromático sigue una ruta de degradación de tipo *orto* o ruta del β -cetoadipato cuando la ruptura del anillo es llevada a cabo por una intradiol dioxigenasa. En la figura 2 están indicadas algunas rutas catabólicas de tipo *orto*.

La familia de las intradiol dioxigenasas incluye tres tipos diferentes de enzimas (tabla 5): catecol 1,2-dioxigenasa I (C12O I), protocatecuato 3,4-dioxigenasa (P340) y catecol 1,2-dioxigenasa II (C12O II). Las dos primeras se han identificado en muchos microorganismos que contienen la ruta *orto* propiamente dicha como *P. putida*, *P. aeruginosa*, *P. cepacia* y *A. calcoaceticus* (Aldrich y cols., 1987; Doten y cols., 1987; Hartnett y cols., 1990; Neidle y cols., 1988). Estas enzimas están codificadas en el cromosoma a diferencia de la C12O II perteneciente a la *ruta orto modificada* que está codificada en un plásmido.

Las intradiol dioxigenasas utilizan Fe(III) como cofactor y se ha demostrado, mediante la determinación de la estructura tridimensional de la P340 de *P. putida*, que el Fe(III) se une a la proteína por dos residuos de tirosina y dos residuos de histidina conservados en todas las intradiol dioxigenasas. Los genes que codifican para la C12O I (*catA*) y las dos subunidades diferentes de la P340 (*pcaH* y *pcaG*) se transcriben separadamente del resto de los genes de la ruta *orto*. En la figura 5 se muestran los pasos enzimáticos que conducen a la formación de acetil-CoA y succinil-CoA.

El aislamiento y caracterización de algunas bacterias capaces de mineralizar compuestos aromáticos clorados derivados de benceno, benzoato y el ácido 2,4-diclorofenoxiacético demostró la existencia de un grupo nuevo de enzimas responsables de la metabolización de estos compuestos mediante una vía *orto* que se denomina *ruta orto modificada* (Ghosal y You 1988; Don y Pemberton, 1981). La mayoría de ellas convierten los derivados aromáticos clorados en clorocatecoles que son susceptibles a la acción de las dioxigenasas pertenecientes al grupo de la de la C12O II. Las enzimas implicadas en la mineralización de clorocatecoles presentan una especificidad de sustrato más amplia que las enzimas equivalentes pertenecientes a la ruta *orto* no modificada y

3.2.3. Ruta del gentisato

El gentisato es un metabolito dihidroxilado intermediario de algunas rutas de degradación microbiana de compuestos aromáticos como el antranilato (Ladd, 1962), β -naftol (Walker y Lippert, 1965), 3- y 4-hidroxibenzoato (Crawford, 1975), salicilato (Crawford y cols., 1975), flavonas (Tomasek y Crawford, 1986), naftalenodisulfonato (Whittich y cols., 1988), *m*-cresol, 2,5- y 3,5-xilenol (Hopper y Chapman, 1970). En el primer paso de esta ruta catabólica (figura 1) interviene la gentisato 1,2-dioxigenasa que provoca la apertura del anillo aromático dando lugar a maleilpiruvato que puede degradarse directamente por el metabolismo intermediario microbiano o transformarse previamente en fumarilpiruvato mediante una reacción de isomerización. La gentisato 1,2-dioxigenasa se ha purificado de *P. acidovorans* y *P. testosteroni* (Harpel y Lipscomb, 1990a) y se ha demostrado que contiene cuatro subunidades idénticas y utiliza Fe (II) como cofactor igual que las extradiol dioxigenasas, aunque el mecanismo de acción de ambos grupos de enzimas es diferente (Harpel y Lipscomb, 1990b).

3.2.4. Regulación de las rutas catabólicas de compuestos aromáticos

La expresión de los genes que componen las rutas catabólicas está controlada por proteínas reguladoras específicas. De esta forma, los microorganismos producen enzimas para utilizar diferentes fuentes de carbono, nitrógeno y fósforo degradándolas, solamente, cuando no existe en el medio otro compuesto más fácilmente metabolizable. Este mecanismo supone un ahorro de energía en el metabolismo microbiano.

La mayoría de las rutas catabólicas de compuestos aromáticos están reguladas positivamente. Generalmente las proteínas reguladoras se unen a un efector o coinductor e interaccionan con la región promotora de los operones catabólicos a los que regulan favoreciendo su transcripción. Actualmente se ha caracterizado un gran número de proteínas reguladoras de este tipo de rutas. El sistema más estudiado es el formado por las proteínas XylS/XylR que regulan coordinadamente la expresión de los operones *upper* y *meta* de la ruta TOL de *P. putida* (Marqués y Ramos, 1993). La figura 6 muestra el modelo de la regulación de esta ruta. Los dos genes reguladores están

situados en una posición adyacente uno del otro en el extremo 3' del operón *meta* y se transcriben divergentemente mediante los promotores *Pr* en el caso de *xylR* y *Ps* en el caso de *xylS*. XylR es una proteína de 566 aa (67 kDa) que pertenece a la familia de reguladores NtrC (Inouye y cols., 1988). El gen *xylR* se expresa constitutivamente y responde a la presencia de xileno activando la transcripción del operón *upper* y *xylS*, en presencia del factor σ^{54} (RpoN) de la RNA polimerasa y, para el caso del operón *upper*, la proteína reguladora IHF (Köhler y cols., 1989, Holtel y cols., 1990, de Lorenzo y cols., 1991).

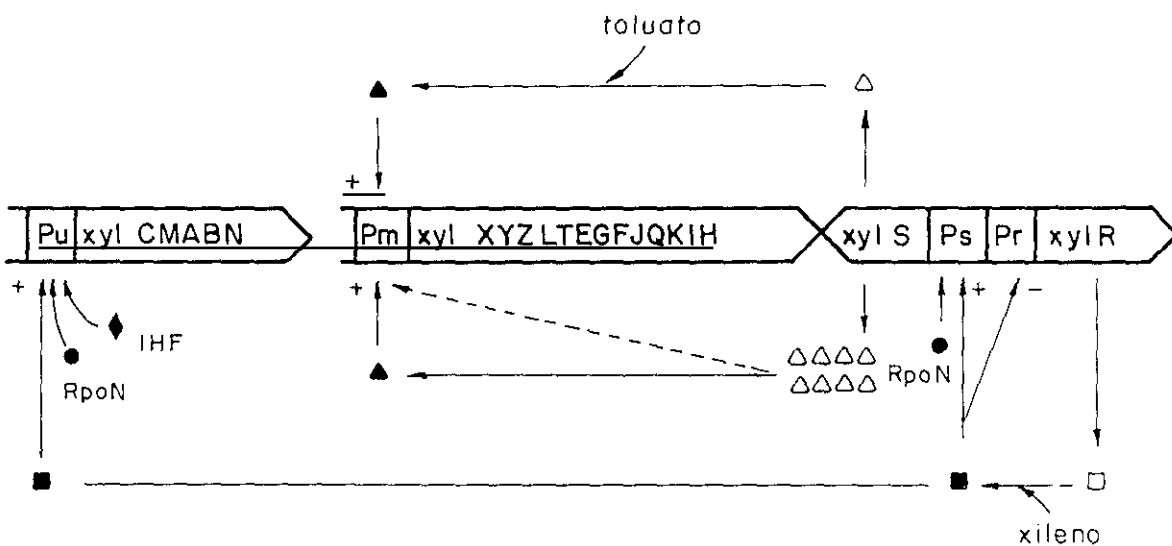


Figura 6. Regulación de la ruta TOL de *P. putida*. Símbolos: □, forma inactiva de XylR incapaz de estimular eficientemente la transcripción a partir de *Pu* y *Ps*; ■, forma activa de XylR que i) autorreprime su propia síntesis, ii) en presencia del factor de transcripción σ^{54} ó proteína RpoN (●) estimula la transcripción del gen *xylS*, y iii) en presencia de RpoN y de la proteína IHF (◆) estimula la transcripción del operón *upper*; △, forma inactiva de XylS que a baja concentración no es capaz de activar al promotor *Pm*; ▲, forma activa de XylS que estimula la transcripción a partir de *Pm*; +, estimulación de la transcripción; -, inhibición de la transcripción (Marqués y Ramos, 1993).

La familia de reguladores NtrC se caracteriza por estar estructurada en cuatro dominios (A, B, C y D), tres de los cuales están altamente conservados entre los miembros de esta familia. El dominio D está situado en el extremo C-terminal de estas proteínas y presenta un motivo de hélice-giro-hélice responsable de la unión de la proteína reguladora al DNA. El dominio C es la región más conservada y se cree que es la responsable de la interacción de estas proteínas con la RNA polimerasa permitiendo la formación de un complejo abierto de transcripción (Kustu y cols., 1991). El dominio B comprende una región hidrofílica muy corta que actuaría como sistema de unión entre los dominios C y A. Por último, el dominio A parece estar implicado en la recepción de la señal activadora, bien sea una señal mediada por una proteína como en el sistema NtrC/NtrB (Ninfa y cols., 1988) o bien una señal mediada por un compuesto químico como en el caso de XylR. El dominio A constituye el N-terminal de estas proteínas y es el menos conservado en esta familia.

Además de XylR, existen otros miembros de esta familia implicados en la regulación de las rutas catabólicas de derivados aromáticos como, por ejemplo, las proteínas DmpR y PhhR pertenecientes a la rutas de degradación de fenol de *Pseudomonas* CF600 (Shingler y cols., 1993) y de *P. putida* P35X, respectivamente (Ching y cols., 1995). El porcentaje de identidad que presentan estas tres proteínas es superior al 60% (Ching y cols., 1995).

La activación de la transcripción del operón *meta* está mediada por la proteína XylS que se hiperproduce como consecuencia de la activación provocada por XylR o se activa independientemente de XylR en presencia de benzoato o determinados alquilbenzoatos que actúan como efectores (Inouye y cols., 1987; Mermoud y cols., 1987; Ramos y cols., 1986). XylS es una proteína de 321 aa (36 kDa) que pertenece a la familia de reguladores AraC (Gallegos y cols., 1993). Algunos de estos reguladores (RhaR, RhaS, MelR, AraC y XylS) están implicados en el catabolismo de ciertas fuentes de carbono como la ramnosa, la melobiosa, la arabinosa y los alquilbenzoatos (Tobin y Schleif, 1987 y 1990; Webster y cols., 1989; Miyada y cols., 1990; Stoner y Schleif, 1982; Lei y cols., 1985). Otros regulan la expresión de factores de virulencia en *E. coli* (Rns) (Caron y cols., 1989) y *Yersinia enterocolitica* (VirF) (Cornelis y cols., 1989). La

mayoría de ellos son reguladores positivos excepto CelD que actúa como represor del operón *celABCF* de *E. coli* (Parker y Hall, 1990). Estas proteínas reguladoras están estructuradas en dos dominios el C-terminal y el N-terminal. En el C-terminal existe un motivo característico de hélice-giro-hélice responsable de la unión de la proteína al DNA. Este dominio está muy conservado en toda la familia lo que ha permitido diseñar una secuencia consenso que caracteriza a la familia (Gallegos y cols., 1993). El dominio N-terminal no está conservado y se cree que, en el caso de los reguladores implicados en el metabolismo de fuentes de carbono, es el responsable de la unión al efector (Ramos y cols., 1990).

Otra familia de reguladores muy frecuentes en este tipo de rutas catabólicas son los pertenecientes a la familia de reguladores LysR (Henikoff y cols., 1988; Schell, 1993). Normalmente codifican para proteínas activadoras de la transcripción excepto *catM* (Neidle y cols., 1989) que actúa como represor. La mayoría de estos genes tienen un tamaño aproximado de 1 kb y se transcriben en sentido contrario al gen u operón al que regulan (Schell, 1993). En la región N-terminal de los miembros de esta familia existe un motivo de hélice-giro-hélice, característico de proteínas que se unen a DNA, que está conservado en la mayoría de los miembros de esta familia. Estas proteínas necesitan un coinductor para activar la transcripción pero se unen al DNA diana independientemente de la presencia del coinductor. La tabla 6 muestra algunos ejemplos de proteínas reguladoras implicadas en el catabolismo de derivados aromáticos pertenecientes a esta familia.

Tabla 6. Familia de reguladores LysR.

Proteínas	Origen	Ruta catabólica	Coinductor	Referencia
NahR	plásmido NAH7 de <i>P. putida</i>	naftaleno/salicilato	salicilato	(Schell, 1985)
CatR	<i>P. putida</i>	benzoato	<i>cis-cis</i> -muconato	(Rothmel y cols., 1990)
CatM	<i>A. calcoaceticus</i>	benzoato	<i>cis-cis</i> -muconato	(Neidle y cols., 1989)
TfdS	plásmido pJP4 de <i>A. eutrophus</i>	diclorofenoxiacetato clorodienlactona	cloromaleilacetato	(Kaphammer y Olsen, 1990)
TcbR	<i>Pseudomonas</i> sp. estirpe P51	clorobenzoato	3-clorobenzoato	(van der Meer y cols., 1991a)
ClcR	plásmido pAC25 de <i>P. putida</i>	clorocatecol	clorocatecol	(Coco y cols., 1993)

4. AMPLIACIÓN DIRIGIDA DE LA CAPACIDAD BIODEGRADATIVA BACTERIANA

Como se comentó en el apartado 2., la caracterización exhaustiva de las rutas metabólicas microbianas para la mineralización de compuestos aromáticos, permite la utilización de técnicas de ingeniería genética para la ampliación de la capacidad biodegradativa de los microorganismos. Esta aproximación experimental puede abordarse por distintas vías; por una parte, desde el punto de vista cuantitativo, puede mejorarse la actividad catalítica de ciertos sistemas enzimáticos incrementando la transcripción a partir de los promotores nativos mediante el uso de proteínas reguladoras mutadas, aumentando el nivel de traducción o actuando globalmente sobre los circuitos reguladores de la ruta (Ramos y cols., 1988; Timmis y cols., 1994). Desde el punto de vista cualitativo, las rutas metabólicas pueden mejorarse mediante la variación controlada de la estabilidad y la actividad catalítica de las enzimas que componen la ruta, la construcción de enzimas híbridas con amplio rango de sustrato o el aislamiento de mutantes con enzimas reguladoras modificadas que permitan aumentar el rango de efectores que activen la ruta (Timmis y cols., 1994).

Otra forma de ampliar la capacidad metabólica de una bacteria es la construcción de nuevas rutas mediante el ensamblaje de genes provenientes de diversas rutas catabólicas de otras bacterias de forma que, en conjunto, se comporten como una ruta coordinada (de Lorenzo y Rojo, 1994). Generalmente, las rutas metabólicas de hidrocarburos aromáticos convergen, mediante rutas periféricas, en un intermediario metabólico como el catecol o el gentisato (ver apartado 3.2.). Esta estructura modular permite la combinación de diferentes módulos bioquímicamente compatibles que posibiliten la canalización de nuevos sustratos hacia rutas metabólicas centrales. Dependiendo de si los nuevos compuestos metabolizables son o no análogos estructurales del sustrato que originalmente el microorganismo podía metabolizar, la expansión de la capacidad biodegradativa de una ruta metabólica puede ser *lateral* o *vertical* (Timmis y cols., 1994). Un ejemplo clásico de este tipo de estrategias es la expansión de la ruta de degradación de 3-clorobenzoato de *Pseudomonas* sp. B13. Este microorganismo contiene una ruta de tipo *orto* para la degradación de benzoato y una

ruta *orto modificada* para la degradación de 3-clorobenzoato, pero no es capaz de mineralizar 4-clorobenzoato o mezclas de alquil- y halobenzoatos (Dorn y cols., 1974). Sin embargo, *Pseudomonas* sp. FR1 (pFRC20P), una cepa derivada de *Pseudomonas* sp.B13 que fue construida mediante el ensamblaje de genes de diferentes microorganismos, es capaz de mineralizar estos compuestos (Rojo y cols., 1987).

La liberación al medio ambiente de microorganismos modificados genéticamente (MMGs) para su utilización en la descontaminación de suelo, aguas subterráneas, o en plantas de tratamiento de aguas residuales, plantea el problema de la predictibilidad del impacto ecológico que tendrán las bacterias recombinantes en el medio en que se introducen así como los posibles riesgos para la salud humana y animal (de Lorenzo y Rojo, 1994). La evaluación de riesgos requiere el estudio de sistemas modelo o microcosmos, que simulen el nicho ecológico en el que las cepas recombinantes están destinadas a actuar. El microcosmos debe ser diseñado de forma que se puedan controlar diferentes parámetros como la capacidad de los MMGs para sobrevivir y multiplicarse, la estabilidad del material genético introducido, la capacidad de los MMGs para llevar a cabo la función para la cual habían sido construidos o la transferencia genética del material introducido en los MMGs a otros microorganismos que comparten el microcosmos (Ramos y cols., 1994). Por tanto, los MMGs destinados a aplicaciones medioambientales deben presentar ciertas propiedades que complican significativamente su manipulación genética en comparación con los MMGs de uso exclusivo de laboratorio. La mayoría de las bacterias recombinantes contienen marcadores de resistencia a antibióticos, por lo que la liberación de este tipo de bacterias está prohibida por la mayoría de las normativas vigentes en los países de nuestro entorno (de Lorenzo y Rojo, 1994). Por ello se han desarrollado nuevos sistemas de selección inocuos desde el punto de vista medioambiental como, por ejemplo, resistencias a herbicidas y metales pesados (Herrero y cols., 1990; de Lorenzo y cols., 1990). Estos marcadores están incluidos en vectores-transposones, siendo los más utilizados los mini-transposones derivados de Tn5 (Herrero y cols., 1990). Otra alternativa consiste en utilizar anticuerpos monoclonales como marcadores de selección que reconozcan epítomos de la superficie celular de la bacteria recombinante en estudio. Este sistema es muy sensible y permite la identificación de los MMGs *in situ* (Ramos y cols., 1994).

Por último, es importante resaltar los trabajos que actualmente se están desarrollando en sistemas de contención biológica, consistentes en el diseño de circuitos genéticos que permiten la expresión de un gen tóxico cuando el MMG no se encuentra bajo las condiciones de crecimiento preestablecidas (Molin y cols., 1987; Knudsen y Karlström, 1991; Ramos y cols., 1994). Con estos sistemas se pretende evitar la proliferación incontrolada de los MMGs fuera del nicho ecológico para los que fueron diseñados o cuando el compuesto tóxico ya ha sido eliminado. Asimismo, también han sido diseñados recientemente sistemas de contención genética que reducen la posible transferencia genética horizontal del material recombinante a la población nativa de microorganismos (Díaz y cols., 1994).

5. LAS ENTEROBACTERIAS Y LA BIODEGRADACIÓN DE COMPUESTOS AROMÁTICOS

Las Enterobacterias constituyen uno de los grupos mayores y mejor definidos entre las Eubacterias Gram-negativas no fotosintéticas. Son bacterias de forma bacilar recta o curva cuyas dimensiones oscilan entre 0,3-1,0 x 1,0-6,0 μm (Brenner, 1984). Las especies pertenecientes a ciertos géneros son inmóviles mientras que otras se mueven mediante flagelos peritricos. Se distinguen del resto de las Eubacterias Gram-negativas de estructura similar por ser anaerobios facultativos. En condiciones anaeróbicas obtienen energía por fermentación de carbohidratos mientras que, en condiciones aeróbicas, pueden utilizar una amplia gama de compuestos orgánicos como, por ejemplo, ácidos orgánicos, aminoácidos y carbohidratos (Stanier y cols., 1986). La consideración de la inserción flagelar como un carácter taxonómico ha impedido la inclusión en el grupo entérico de algunas bacterias acuáticas muy relacionadas con las enterobacterias, como las pertenecientes a los géneros *Vibrio*, *Aeromonas*, *Beneckea* y *Photobacterium*. Estas bacterias presentan flagelos polares o inserción flagelar mixta y son oxidasa-positivos a diferencia de los miembros del grupo entérico (Stanier y cols., 1986). En la figura 7 se muestra un esquema simplificado del porcentaje de similitud entre el DNA de los géneros de la familia *Enterobacteriaceae* (Brenner, 1984).

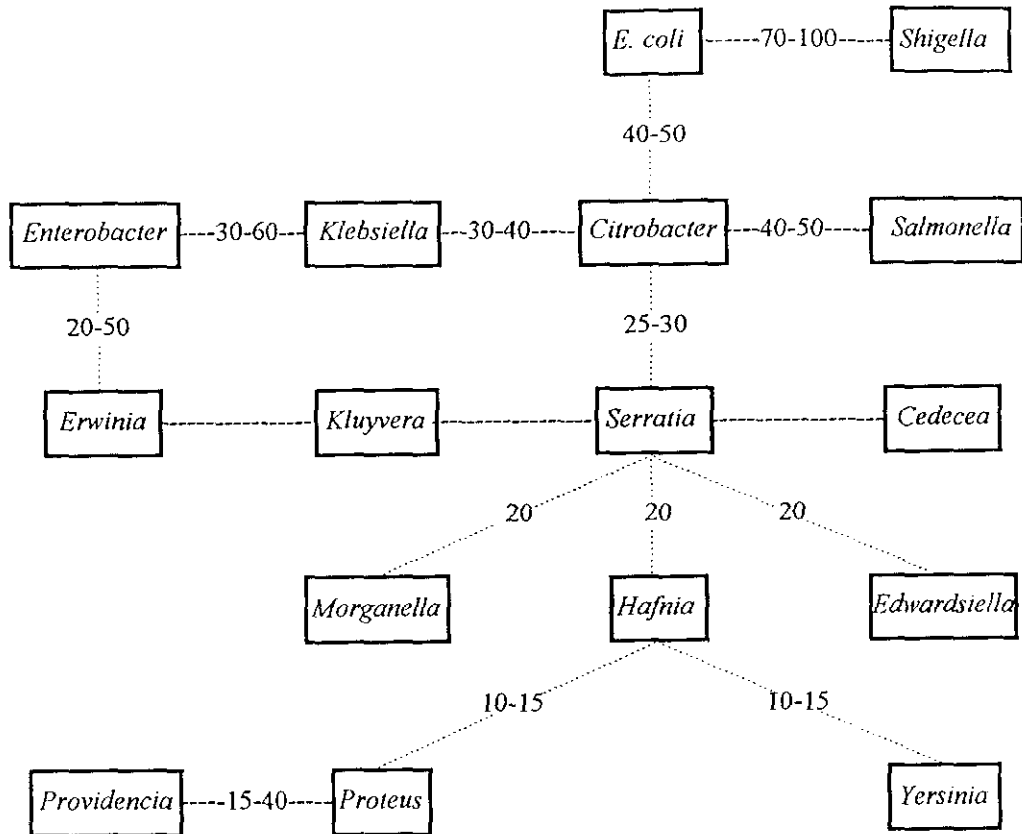


Figura 7. Porcentaje de similitud a nivel de nucleótidos entre los miembros de la familia Enterobacteriaceae.

Escherichia coli es el representante clásico de las enterobacterias (Brenner, 1984). Este microorganismo forma parte de la flora intestinal habitual de los mamíferos, aunque algunas estirpes patógenas causan infecciones intestinales y del tracto urinario. Además, esta bacteria está presente en el suelo y en ciertos ambientes acuáticos como consecuencia del tratamiento del suelo con estiércol y del vertido de aguas residuales urbanas (Cieslak y cols., 1993; Tsai y Olson, 1992; Manja y cols., 1992). Los géneros *Salmonella* y *Shigella* están constituidos por bacterias patógenas responsables de infecciones intestinales como la disentería bacilar, la fiebre tifoidea y la intoxicación alimentaria bacteriana (Schaechter y Neidhardt, 1987). Dada la importancia clínica de estos géneros se han hecho subdivisiones intraespecíficas muy minuciosas en base al

estudio inmunológico de las superficies celulares (Brenner, 1984). Los géneros *Enterobacter*, *Serratia*, *Erwinia*, y *Proteus* están muy relacionados con los anteriores pero su hábitat es diferente, ya que principalmente están presentes en el suelo y en algunos ambientes acuáticos. Las especies pertenecientes al género *Erwinia* son patógenos de plantas. Algunas bacterias patógenas de animales clasificadas antiguamente dentro del género *Pasteurella* se incluyen actualmente dentro del género *Yersinia*. Entre las especies que componen este género la más conocida es *Y. pestis*, el agente causal de la peste bubónica que se caracteriza por ser una infección muy diferente a las infecciones entéricas desde el punto de vista de la sintomatología que provoca y del mecanismo de transmisión (Stanier y cols., 1986).

No cabe duda de que *E. coli* es un microorganismo paradigmático que ha servido como herramienta para el estudio y desarrollo de la Biología Molecular. Ya desde principios de siglo fue elegido por los fisiólogos para sus investigaciones, debido a su capacidad de crecer rápidamente en medios simples. A finales de 1930, las investigaciones realizadas por Wollmans y Bronfenbrenner sobre los bacteriófagos de *E. coli*, determinaron el establecimiento definitivo de esta bacteria como modelo de estudio. Actualmente el conocimiento de esta bacteria a nivel molecular es mayor que el de cualquier otro organismo unicelular. De hecho, en el año 1987 aproximadamente un tercio de los productos génicos de esta bacteria habían sido estudiados en detalle desde el punto de vista bioquímico y sus genes habían sido identificados (Schaechter y Neidhardt, 1987; Bachmann, 1990).

5.1. La capacidad biodegradativa de *E. coli*

E. coli presenta un alto grado de adaptabilidad al medio ambiente, propio de las bacterias que compiten por los nutrientes con otros microorganismos en un mismo nicho ecológico. Por ejemplo, ante ciertos estímulos como el aumento en la concentración de nutrientes en el medio de cultivo, este microorganismo puede variar en segundos la fase de crecimiento desde fase estacionaria hasta fase exponencial (Schaechter y Neidhardt, 1987). Por otra parte, *E. coli* puede degradar una amplia gama de compuestos orgánicos como diferentes azúcares, polioles y ácidos orgánicos utilizándolos como fuente de

carbono, lo que favorece la adaptación de este microorganismo a nichos ecológicos con diferente composición química (Lin, 1987). Sin embargo, resulta sorprendente que siendo *E. coli* la bacteria mejor estudiada bioquímica y genéticamente no se haya analizado, hasta hace relativamente poco tiempo, su capacidad para degradar compuestos aromáticos. Aunque el número de trabajos en este sentido es aún escaso, se sabe que algunas cepas de *E. coli* poseen rutas tan complejas como las descritas en otros microorganismos para mineralizar distintos derivados bencénicos como los ácidos fenilacético (PA), 3- y 4-hidroxifenilacético (3- y 4-HPA), homoprotocatéuico (HPC), 3-fenilpropiónico (HCA), 3-(3-hidroxifenil)propiónico (MHP), 3-hidroxicinámico y 2-feniletilamina (Cooper y Skinner, 1980; Burlingame y Chapman, 1983; Cooper y cols., 1985). Otras enterobacterias que poseen rutas catabólicas de este tipo son *K. pneumoniae*, que puede mineralizar 3- y 4-hidroxibenzoato así como 3-y 4-HPA (Suárez y cols., 1991; Martín y cols., 1991), y algunas especies del género *Serratia* que pueden utilizar como única fuente de carbono benzoato, benzilformato, 3- y 4-hidroxibenzoato y PA (Brenner, 1984).

No todas las estirpes de *E. coli* pueden catabolizar los mismos sustratos (tabla 7). Por ejemplo, la estirpe K-12 puede metabolizar el PA pero no el 3- y 4-HPA, a diferencia de la estirpes B y C que sólo pueden mineralizar los hidroxiderivados del PA. La estirpe W puede degradar tanto el PA como el 3- y 4-HPA así como todos los derivados aromáticos que han sido estudiados hasta ahora, por lo que parece estar especialmente capacitada para la mineralización de este tipo de compuestos (Burlingame y Chapman 1983). Esta estirpe posee una enzima denominada penicilina G acilasa (PGA) capaz de hidrolizar el resto PA de la molécula de penicilina G, liberando ácido 6-amino penicilánico que es la base de las penicilinas semisintéticas (Cole, 1964). La PGA ha sido muy estudiada en la industria para la producción de antibióticos β -lactámicos siendo una de las pocas enzimas que actualmente se emplean en procesos industriales de biotransformación (Valle y cols., 1991). Además de la penicilina G esta enzima puede hidrolizar otras amidas y ésteres del PA, 4-HPA y de otros compuestos aromáticos (Margolin y cols., 1980; Roa y cols., 1994; Duggleby y cols., 1995) por lo que se ha especulado que su función debe estar relacionada con la posibilidad de hidrolizar compuestos que puedan ser posteriormente utilizados como fuente de carbono (Valle y

cols. 1991). El mecanismo de regulación de la acilasa apoya esta hipótesis ya que su producción está sujeta a represión catabólica y se induce por PA y otros derivados aromáticos relacionados con éste (Merino y cols., 1992; Vandame, 1985). Existen otras enterobacterias que poseen PGA como *Kluyvera citrophila* ATCC 21285 y *Proteus rettgeri* ATCC 31052 (Barbero y cols., 1986; Daumy y cols., 1985) pero, sin duda alguna, la acilasa mejor caracterizada es la de la estirpe W de *E. coli* ATCC 11105.

Tabla 7. Estirpes de *E. coli* que metabolizan compuestos aromáticos

	B	C	W	Estirpe		Aislados clínicos
				K-12	NTCTC5928	
Número examinado	1	1	2	2	1	19 5 2 8 1 2 3 27
Fenilacetato	-	-	+	+	+	- + - - - + + +
3-Hidroxifenilacetato	+	+	+	-	-	- - - - + + - +
4-Hidroxifenilacetato	+	+	+	-	-	- - - - + + - +
3-Fenilpropionato	+	+	+	+	+	- - - + - - + +
3-(3-Hidroxifenil)propionato	+	+	+	+	+	- - + + - - + +
3-Hidroxifenilcinamato	+	+	+	+	+	- - + + - - + +

5.1.1. Rutas catabólicas de compuestos aromáticos de *E. coli*

Los primeros trabajos sobre rutas catabólicas de aromáticos de *E. coli* fueron realizados en 1980 por Cooper y Skinner. Mediante la identificación de intermediarios catabólicos en extractos crudos de células de la estirpe C de *E. coli* incubadas en presencia de los ácidos 3- y/o 4-HPA y la caracterización bioquímica de las enzimas implicadas, propusieron una ruta metabólica de ruptura *meta* para la degradación de 3- y 4-HPA. El primer paso de la ruta consistiría en una hidroxilación del sustrato para dar

lugar en, ambos casos, al HPC, por lo que demostraron que ambas rutas están conectadas en esta estirpe. Durante los últimos quince años se ha caracterizado bioquímicamente y clonado los genes (*hpc*) de la ruta de degradación *meta* del HPC de *E. coli* C, aunque el mapa de restricción así como el orden de transcripción de los genes de la ruta no han sido determinados hasta 1993 por Roper y cols., (figura 8) (Cooper y Skinner, 1980; Garrido-Pertierra y Cooper, 1981; Skinner y Cooper 1982; Jenkins y Cooper 1988; Ferrer y Cooper, 1988; Fawcett y cols., 1989; Roper y Cooper, 1990a, b; Roper y Cooper, 1993; Roper y cols., 1995)

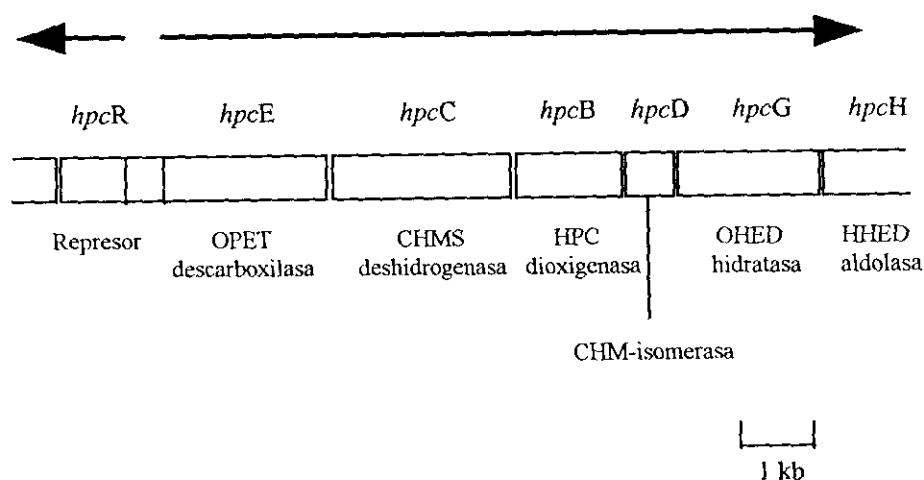


Figura 8. Orden de transcripción de los genes de la ruta catabólica del homoprotocatecuato de *E. coli* C. Esta ruta está compuesta por el gen regulador *hpcR* y el operón *meta* (*hpcECBDGH*). Las flechas indican la dirección de transcripción de los genes. Los genes *hpc* codifican para las enzimas que aparecen en la parte inferior de la figura. Abreviaturas: HPC, homoprotocatecuato; CHMS, 5-carboximetil-2-hidroxi-mucónico semialdehído; CHM, 5-carboximetil-2-hidroxi-muconato; OPET, 5-oxo-pent-3-en-1,2,5-tricarboxilato; OHED, 2-oxo-hept-3-en-1,7-dioato; HHED, 2,4-dihidroxi-hept-2-en-1,7-dioato (Roper y col., 1993).

La ruta catabólica del PA no ha sido prácticamente estudiada. En 1985, Cooper y cols. aislaron una serie de mutantes de *E. coli* K-12 incapaces de metabolizar el PA y localizaron los genes implicados en esta ruta metabólica en el minuto 30,4 del cromosoma de *E. coli* K-12. Se ha demostrado bioquímicamente que la ruta de degradación de 2-feniletilamina está relacionada con la ruta del PA, ya que esta amina se transforma en fenilacetaldehído mediante una amina oxidasa dependiente de cobre y posteriormente en PA mediante una deshidrogenasa, por lo que aparentemente las rutas de degradación de la 2-feniletilamina y del PA podrían estar relacionadas (Parrot y cols., 1987; Cooper y cols., 1992). Por otra parte, se ha demostrado la existencia de una fenilacetil-CoA ligasa en *E. coli* W inducible por PA que podría estar involucrada en la ruta catabólica de este compuesto, activando el anillo aromático mediante tioesterificación, de igual forma que otras acetil-CoA ligasas implicadas en las rutas catabólicas de derivados aromáticos de microorganismos, generalmente, anaerobios (Vitovski, 1993).

La ruta catabólica del HCA fue caracterizada bioquímicamente en 1983 por Burlingame y Chapman. Posteriormente, mediante el aislamiento de mutantes incapaces de metabolizar el HCA, estos autores determinaron los pasos enzimáticos de la ruta de degradación del HCA y del MHP de *E. coli* K-12 (figura 9) (Burlingame y cols., 1986). Esta ruta sigue un mecanismo similar al de las rutas de ruptura *meta* descritas en otros microorganismos para la degradación de otros hidrocarburos aromáticos (ver apartado 3.2.2.1.). Actualmente se han localizado en el minuto 8 del cromosoma de *E. coli* K-12 los genes que codifican para las tres primeras enzimas de la ruta de degradación del MHP, la MHP-hidroxilasa (*mhpA*), la 3-(2,3-dihidroxifenil)propionato 1,2-dioxigenasa (*mhpB*) y la 2-hidroxi 6-cetononadienedionato hidrolasa (*mhpC*) (figura 9) y se han clonado e hiperexpresado los genes *mhpB* y *mhpC* (Bugg, 1993). La comparación de la secuencia del extremo N-terminal de la 1,2-dioxigenasa MhpB indica que esta enzima puede estar relacionada con la catecol 2,3-dioxigenasa II de *A. eutrophus* (Kabisch y Fortnagel, 1990).

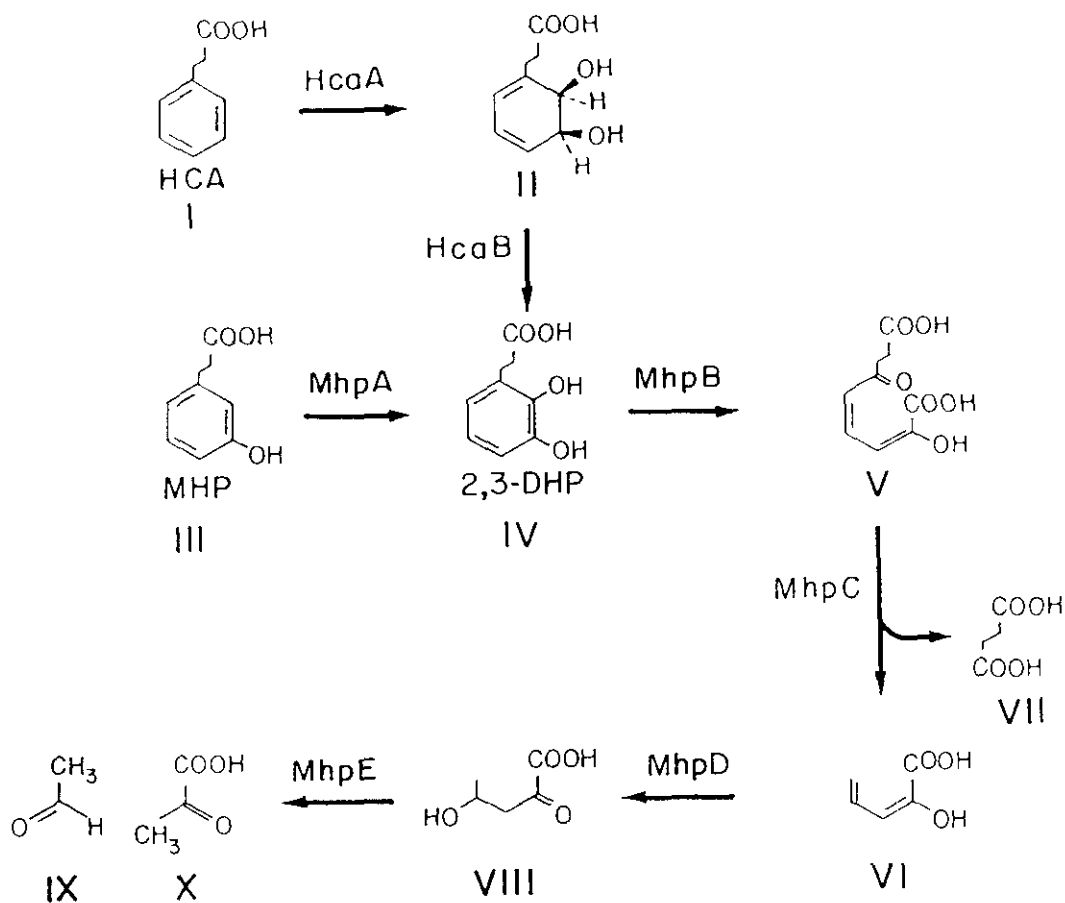


Figura 9. Ruta de degradación del ácido 3-fenilpropiónico y del ácido 3-(3-hidroxifenil)propiónico de *E. coli* K12. Los metabolitos son los siguientes: ácido 3-fenilpropiónico (HCA) (I), *cis*-3-(3-carboxietil)-3,5-ciclohexadieno-1,2-diol (II), ácido 3-(3-hidroxifenil)propiónico (MHP) (III), ácido 3-(2,3-dihidroxifenil)propiónico (2,3-DHP) (IV), ácido 2-hidroxi-6-cetnonadienedióico (V), ácido 2-ceto-4-pentenoico (enol) (VI), ácido succínico (VII), ácido 4-hidroxi-2-cetovalérico (VIII), acetaldehído (IX), ácido pirúvico (X). Denominación de las enzimas implicadas en la ruta: dioxigenasa HcaA, deshidrogenasa HcaB, hidroxilasa MhpA, dioxigenasa MhpB, hidrolasa MhpC, hidratasa MhpD, aldolasa MhpE (Bugg, 1993).

5.2. Fuentes naturales de los compuestos aromáticos metabolizables por *E. coli*

Para poder establecer una hipótesis sobre el posible papel fisiológico de estas rutas en un organismo como *E. coli*, es necesario conocer las fuentes naturales de las que provienen los compuestos aromáticos que este microorganismo puede metabolizar. Como se mencionó anteriormente, *E. coli* forma parte de la flora intestinal del hombre y otros mamíferos y, por tanto, no sería extraño que estos compuestos estuvieran presentes en el hábitat intestino/fecal. Se ha planteado la hipótesis de que *E. coli* utiliza estos compuestos cuando están presentes en el medio y no hay un sustrato más fácilmente metabolizable, lo que justifica que las rutas sean inducibles y que algunas de ellas estén sujetas a represión catabólica (Valle y cols., 1991).

Los derivados aromáticos metabolizables por *E. coli* no pueden ser considerados como agentes xenobióticos, ya que están presentes en la Biosfera como compuestos naturales y de hecho existen otros microorganismos que contienen rutas metabólicas para su completa mineralización (tabla 8). El ácido cinámico e hidroxicinámico son metabolitos centrales implicados en la síntesis de compuestos fenólicos en plantas, y se encuentran en la mayoría de los vegetales como ácidos conjugados en forma de ésteres o como ácidos libres. La reducción del doble enlace del sustituyente alifático del ácido cinámico da lugar al HCA, por lo que este compuesto también puede aislarse en extractos vegetales (Martínez y cols., 1992; Elder y Kelly 1994). Por tanto, este tipo de compuestos aparece de forma natural en los alimentos de origen vegetal como, por ejemplo, frutas y hortalizas (Li y cols., 1993). También se ha detectado la presencia de PA y MHP en champiñones, ciruelas, calabacines, pimientos verdes, tomates y berenjenas (Chen y Wu, 1984, Heimhuber y Herrmann, 1990). En muchas ocasiones, este tipo de compuestos son los responsables del sabor y el aroma de ciertos alimentos como por ejemplo el de la miel. De hecho, la presencia o ausencia del PA y del HCA en este alimento se ha utilizado como parámetro para la identificación de diferentes tipos de miel (Steeg y Montag, 1988).

Tabla 8. Bacterias que metabolizan 4-HPA

Bacteria	Referencia
<i>Acinetobacter</i> sp.	(Sparnins y cols., 1974)
<i>Bacillus brevis</i>	(Que y cols., 1981)
<i>B. stearothermophilus</i>	(Jamaluddin, 1977)
<i>E. coli</i> (30 aislados clínicos)	(Burlingame y Chapman, 1983)
<i>E. coli</i> B	(Cooper y Skinner, 1980)
<i>E. coli</i> C	(Cooper y Skinner, 1980)
<i>E. coli</i> W (ATCC 9637, ATCC 11105)	(Cooper y Skinner, 1980)
<i>Flavobacterium</i> sp.	(van den Tweel y cols., 1988)
<i>K. pneumoniae</i> M5a1	(Grant, 1967)
<i>P. acidovorans</i>	(Hareland y cols., 1975)
<i>P. ovalis</i>	(Adachi y cols., 1964)
<i>P. putida</i>	(Raju y cols., 1988)
<i>P. putida</i> T	(Sparnins y cols., 1974)
<i>P. putida</i> U (NCIB 10015)	(Sparnins y cols., 1974)
<i>Pseudomonas</i> sp.	(Blakley y cols., 1967)
Bacteria de suelo	(Blakley 1972)

Se ha demostrado que el PA, HCA, MHP, HPC, 3- y 4-HPA y algunos derivados de éstos, son productos intermediarios del metabolismo de aminoácidos aromáticos en organismos superiores y en bacterias (Samid y cols., 1994; Spoelstra, 1978; Sparnins y Chapman, 1976). Por otra parte, el 4-HPA y HPC se producen como consecuencia del metabolismo bacteriano de aminas simpaticomiméticas como la sinefrina que está presente en ciertas plantas en grandes cantidades (Devi y cols., 1975). El hecho de que estos productos se acumulen durante procesos metabólicos microbianos, justifica su presencia en alimentos cuya elaboración conlleva procesos fermentativos. Por ejemplo, se ha detectado la acumulación de PA durante la maduración de licores y vinos, siendo uno de los componentes responsables del sabor y aroma de estas bebidas (Kepner y cols.,

1968; Gianotti y Stefano, 1991; Yavas y Rapp, 1992). También se ha identificado el PA entre los componentes volátiles responsables del sabor del vinagre (Kubota y cols., 1989) y como producto de la descomposición de la fenilalanina en la elaboración del queso (Lee y Richard, 1984).

La presencia de este tipo de compuestos en el hábitat intestino/fecal no procede sólo del aporte alimenticio, sino también del metabolismo de la flora intestinal. En este sentido, se ha demostrado la presencia de PA, HCA, MHP y 4-HPA en estiércol procedente de granjas dedicadas a la cría del ganado porcino. Estos compuestos se producen como metabolitos del catabolismo de la tirosina de microorganismos que forman parte de la flora intestinal (Spoelstra, 1978).

El PA y sus derivados hidroxilados están presentes habitualmente en cerebro, tejidos periféricos y fluidos corporales de mamíferos como consecuencia del metabolismo de las aminas simpaticomiméticas tiramina y 2-feniletilamina, las cuales se originan a partir de la L-fenilalanina y la L-tirosina (Saavedra, 1989). La toxicidad del PA y sus derivados en humanos es dosis-dependiente. En condiciones normales, la concentración de estos compuestos en tejidos humanos es muy baja, sin embargo, en procesos patológicos como la fenilcetonuria provocada por una deficiencia en la enzima L-fenilalanina hidroxilasa, los niveles de estos compuestos se elevan considerablemente provocando daños en el sistema nervioso central y retraso mental en niños (Thibault, 1994). No obstante, desde el punto de vista farmacológico, a la dosis adecuada el PA, HCA y algunos derivados de éstos se comportan como agentes anticancerígenos y analgésicos (Samid y cols., 1994; Adams, 1992).

Ya se ha mencionado que, desde el punto de vista medioambiental, los sustratos aromáticos metabolizables por *E. coli* no pueden ser considerados como compuestos xenobióticos. Sin embargo, su acumulación puede ser la causa del mal olor del estiércol que se genera en explotaciones de ganado porcino ya que, como se ha demostrado, el 4-HPA y otros derivados de éste se transforman mediante reacciones metabólicas de las bacterias presentes en las heces, en los compuestos fenólicos responsables de dicho mal olor (Spoelstra, 1978).

Durante el proceso de molturación de la aceituna para la extracción del aceite de oliva se originan una serie de subproductos entre los cuales cabe destacar el *alpechín* o agua residual originada durante dicho proceso (Martínez y cols., 1992). Este residuo es muy recalcitrante debido a su elevado contenido en compuestos fenólicos de alto poder contaminante. Entre los componentes principales de carácter aromático de los alpechines se encuentra el ácido hidroxicinámico y algunos derivados del HCA y del PA como el 4-HPA, que junto con el ácido siríngico puede constituir hasta el 50% del total de los compuestos fenólicos en ciertos alpechines (Balice y Cera, 1984; Martínez y cols., 1992). Tan solo en la cuenca del río Guadalquivir se puede estimar una producción media de 1.800.000 m³ de alpechín al año, de los cuales el 25% se vierte de una forma incontrolada, lo que implica un importante deterioro de las aguas. El 75% restante, en la actualidad, se trata en balsas de evaporación que tan solo palián parcialmente el problema medioambiental (Martínez y cols., 1992). Actualmente se han puesto en práctica diversos procesos que van desde la simple evaporación, hasta procesos físico-químicos como la ósmosis inversa y la ultrafiltración. Pero estos procesos requieren altas inversiones de difícil amortización dado el carácter temporal de la industria almazarera. Distintos organismos de la Administración y empresas privadas están desarrollando planes de actuación con el fin de paliar el problema causado por el vertido de estos residuos aprovechando, además, su alto contenido en materia orgánica e inorgánica mediante procesos de reciclado que permitan su uso en la producción de combustibles sólidos, compost y biogas (de Andrés y García, 1982). Uno de los proyectos más interesantes es el diseño de biorreactores para la depuración biológica de los alpechines, que permita la biotransformación de los compuestos aromáticos con el fin de disminuir el carácter tóxico de estos residuos (Janer, 1980, Fiestas y cols., 1982; Martínez y cols., 1986; Martínez, y cols. 1992). Dado que los principales componentes del alpechín son metabolizables por algunas estirpes de *E. coli*, este microorganismo podría ser considerado como un candidato en el diseño de biorreactores para la depuración biológica del alpechín, a lo que también contribuiría su gran capacidad de adaptación a distintos ecosistemas. El estudio y caracterización a nivel molecular de las rutas catabólicas de este tipo de compuestos, nos aportaría nueva información que sería de gran utilidad a la hora de diseñar proyectos biotecnológicos como los descritos anteriormente.

6. OBJETIVOS

Durante el periodo de tiempo en el que el Dr. José Luis García formaba parte del equipo de investigación de Antibióticos S.A., desarrolló un procedimiento para clonar genes *pac*, que codifican para la penicilina G acilasa (PGA), de diferentes microorganismos. Este método está basado en la complementación auxotrófica de *E. coli* HB101 (*leuB*) incubando las células en medio mínimo en presencia de fenilacetil-L-leucina como única fuente de L-leucina (García y Buesa, 1986). En el proceso de aislamiento del gen *pac* de *E. coli* ATCC 11105, se detectó un clon productor de PGA que al incubarlo en medio Luria Bertani (LB) presentaba un extraño fenotipo negro. Este clon contenía el plásmido pAEC19, el cual es un derivado de pBR322 que contiene un fragmento *Hind*III de aproximadamente 8,5 kb del DNA cromosómico de *E. coli* ATCC 11105. El hecho de que los derivados catecólicos se oxiden espontáneamente dando lugar a productos de coloración marrón o negra y la posible implicación de la PGA en el metabolismo de derivados del PA nos indujo a pensar que, además del gen *pac*, el plásmido pAEC19 podía contener un gen que codificara para una hidroxilasa de compuestos aromáticos.

Ya se ha comentado que *E. coli*, un microorganismo que se caracteriza por crecer rápidamente en medios simples, es la bacteria más estudiada a nivel molecular. Por tanto, la caracterización y el estudio de la expresión de una ruta de derivados aromáticos de *E. coli* resulta interesante desde el punto de vista biotecnológico, ya que facilitaría el uso de esta ruta con fines medioambientales. Además, aportaría nuevos datos sobre la posibilidad de utilizar este microorganismo como vehículo de expresión de rutas catabólicas de compuestos recalcitrantes.

Por todo ello, los objetivos de esta Tesis son los siguientes:

I. Caracterización de la hidroxilasa de compuestos aromáticos de *E. coli* ATCC 11105

- i) Localización del gen que codifica para la hidroxilasa.
- ii) Construcción de un sistema que posibilite la alta producción de la enzima.

- iii) Purificación de la enzima.
- iv) Estudio de los requerimientos bioquímicos de la hidroxilasa.
- v) Determinación de la especificidad de sustrato de la enzima.
- vi) Secuenciación del gen que codifica para la 4-HPA-hidroxilasa.
- vii) Comparación de la secuencia de aminoácidos de la 4-HPA-hidroxilasa con otras oxigenasas ya secuenciadas.
- viii) Clasificación de la hidroxilasa en función de la similitud que presente con otras oxigenasas.

II. Caracterización molecular del resto de la ruta metabólica del 4-HPA de *E. coli* ATCC 11105

- i) Clonación de todos los genes implicados en la mineralización del 4-HPA.
- ii) Construcción del mapa físico de la ruta del 4-HPA de *E. coli* W.
- ii) Secuenciación de la ruta completa del 4-HPA

III. Estudio del mecanismo de regulación de la 4-HPA hidroxilasa

- i) Búsqueda de regiones promotoras.
- ii) Construcción de fusiones de las regiones promotoras a genes trazadores.
- iii) Identificación de las proteínas implicadas en la regulación de la expresión de la 4-HPA-hidroxilasa.

IV. Expresión de los genes de la ruta del 4-HPA en organismos heterólogos

- i) Obtención mediante Ingeniería Genética de módulos transponibles que permitan transferir los genes de la ruta del 4-HPA a otros microorganismos.
- ii) Expresión de los genes de la ruta catabólica del 4-HPA en organismos heterólogos con el fin de ampliar su capacidad biodegradativa.

II. MATERIALES Y MÉTODOS

1. ESTIRPES BACTERIANAS Y PLASMIDOS

Para el desarrollo de este trabajo se utilizaron las estirpes de *E. coli* que se detallan a continuación: *E. coli* ATCC 11105 (derivada de *E. coli* W ATCC 9637 auxótrofa para vitamina B₁₂, productora de PGA); B/rK (*E. coli* B resistente a UV, cedida por el Dr. M. Vicente); C (estirpe salvaje cedida por el Dr. M. Vicente); W (estirpe salvaje cedida por el Dr. A. Garrido-Pertierra); MC1116 ($\Delta(ara-leu)$); HB101 (*proA2*, *leuB*, *thi*, *recA*); TG1 (*supE*, *hsdD5*, *thi*, $\Delta(lac-proAB)$, F' (*traD36*, *proAB*⁺, *lacI*^f, *lacZDM15*)); MC4100 (*araD139*, $\Delta(lacIPOZYA)$); SBS688 (MC4100, $\Delta crp39$), W3110 (CECT 416, ATCC 27325), DH1 (*supE44*, *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi1*, *relA1*) (Sambrook y cols., 1989), SE5000 (MC4100, *recA56*) (Silhavy y cols., 1984) ET8000 (Nal^r, *rbs lacZ::IS1*, *gyrA*, *hutC*) (MacNeil, 1981); CC118- λpir ($\Delta[ara-leu]$, *araD*, $\Delta lacX74$, *galE*, *galK*, *phoA20*, *thi-1*, *rpsE*, *rpoB*, *argE* (Am), *recA1*, λpir); S17- λpir (Tp^r, Sm^r, *recA*, *thi*, *pro*, *hsdR*, M^r RP4:2-Tc:Mu:Km Tn7, λpir) (Herrero y cols., 1990).

Otras especies utilizadas: *K. citrophila* ATCC 21285 (productora de PGA) (García y Buesa, 1986) and *K. pneumoniae* M5a1 (cedida por el Dr. A. Garrido-Pertierra) (Martín y cols., 1991), *P. putida* KT2442 (*hsdR*, Rif^r) (Herrero y cols., 1990).

Los plásmidos utilizados son los siguientes: pBR322 (Ap^r, Tc^r); pBR325 (Ap^r, Cm^r, Tc^r) (Bolívar, 1978); pACYC184 (Cm^r, Tc^r) (Rose, 1988); pUC18 (Ap^r) (Yanisch-Perron y cols., 1985); pRS551 (Ap^r, Km^r; plásmido para la búsqueda de promotores) (Simons y cols., 1987); pUTmini-Tn5Km2 (Ap^r, Km^r) (de Lorenzo y cols., 1990); pUC18Not (Ap^r) (Herrero y cols., 1990); pCNB5 (Ap^r, Km^r) (de Lorenzo y cols., 1993); pAW31 (Ap^r) (derivado de pEMBL9 con un fragmento *SaI*I de 1,7 kb que contiene el gen *xylE*) (cedido por el Dr. V. de Lorenzo); pAG464 (Ap^r) (derivado de pUC18 que contiene el gen de la HPC-dioxigenasa de *K. pneumoniae* M5a1) (cedido por el Dr. A. Garrido-Pertierra).

2. MEDIOS Y CONDICIONES DE CULTIVO

Las cepas bacterianas se conservaron congeladas a -70°C en medio de cultivo al que se añadió glicerol al 10% (v/v). En el momento de inocularlas se descongelaron, incubándose en los medios correspondientes a 37°C , a menos que se indique lo contrario, con agitación. Los principales medios de cultivo utilizados han sido: LB, M9 y M63 suplementados adecuadamente en cada ensayo y añadiendo agar al 1% (p/v) para los cultivos en medio sólido (Miller, 1972; Sambrook y cols., 1989). Cuando se requiere una fuente de carbono específica se preparan soluciones estériles de éstas para suplementar el medio mínimo hasta una concentración final de 5 mM, en el caso de los derivados aromáticos, 20 mM si se trata de glicerol y 10 mM si la fuente de carbono utilizada es glucosa.

La concentración de antibiótico añadido al medio de cultivo de cepas resistentes fue de 100 $\mu\text{g/ml}$ para la ampicilina, 10 $\mu\text{g/ml}$ para la tetraciclina, 50 $\mu\text{g/ml}$ para la kanamicina, 30-60 $\mu\text{g/ml}$ para el cloranfenicol, 50 $\mu\text{g/ml}$ para la rifampicina y 25 $\mu\text{g/ml}$ para el ácido nalidíxico.

El crecimiento de las cepas se siguió por turbidimetría a 600 nm con un espectrofotómetro Shimadzu UV-260.

3. PROCEDIMIENTOS DE TRANSFERENCIA GENÉTICA

La transferencia de DNA entre las diferentes estirpes se ha llevado a cabo por transformación y conjugación.

3.1. Transformación

Las cepas de *E. coli* se transformaron con el método del RbCl (Kushner, 1978) y por electroporación (Wirth y cols., 1989) utilizando un equipo Gene Pulser de Bio-Rad.

3.2. Conjugación

Los plásmidos que contienen transposones derivados de mini-Tn5 se transfieren por conjugación según el método descrito por de Lorenzo y Timmis (1994). Para ello se utilizaron *E. coli* S17-1 λ pir como cepa donadora y *E. coli* ET8000 y *P. putida* KT2442 como bacterias receptoras

4. MANIPULACIÓN DEL DNA CON ENZIMAS DE USO COMÚN EN BIOLOGÍA MOLECULAR

Las endonucleasas de restricción se obtuvieron de Boehringer Mannheim, Amersham, Pharmacia o New England Biolabs. La fosfatasa alcalina de intestino de ternera se obtuvo de Boehringer Mannheim. Amersham suministró la DNA ligasa, el fragmento Klenow de la DNA polimerasa I de *E. coli* y la polinucleótido quinasa del fago T4. Todas las enzimas y sus tampones correspondientes se utilizaron siguiendo las indicaciones recomendadas por las casas suministradoras.

5. PREPARACIÓN DE PLÁSMIDOS Y DNA CROMOSÓMICO

La preparación de plásmidos a partir de cepas de *E. coli* se realizó utilizando el método rápido de lisis alcalina y el método de purificación de DNA por centrifugación en gradiente de CsCl (Sambrook y cols., 1989). La extracción del DNA cromosómico a partir de las bacterias utilizadas en este trabajo se realizó según el método descrito para *E. coli* por Sambrook y cols., (1989).

6. ELECTROFORESIS DEL DNA EN GELES DE AGAROSA

Se utilizaron geles de agarosa al 0,7 ó 1% en tampón TAE (Tris-HCl 40 mM, ácido acético 20 mM, EDTA 2 mM, pH 8,1), utilizando el mismo tampón como electrolito. A las muestras se les añadió 1/5-1/10 de su volumen de una solución compuesta por Ficoll 400 al 30% (p/v), azul de bromofenol al 0,2% (p/v), xilencianol al 0,2% y EDTA 40 mM pH 8,0. La electroforesis se realizó a 100-150 V durante 60-90

minutos. Una vez finalizada la electroforesis, los geles se tiñeron con bromuro de etidio (5 µg/ml) para su visualización con radiación ultravioleta. Como marcador de tamaño se utilizó el DNA del fago λ digerido con la enzima de restricción *Bst*EII (Amersham).

7. AISLAMIENTO Y PURIFICACIÓN DE LOS FRAGMENTOS DE DNA

Los fragmentos resultantes de la digestión del DNA con las enzimas de restricción se aislaron y purificaron por cuatro procedimientos distintos.

7.1. Geles de poliacrilamida

Se utilizaron geles de poliacrilamida no desnaturalizantes preparados en TBE (Tris-borato 89 mM, ácido bórico 89 mM y EDTA 2 mM, pH 8,0) a una concentración variable dependiendo del tamaño de fragmento a aislar. Se empleó como electrolito para la electroforesis el tampón TBE. A las muestras se les añadió 1/10 de su volumen de una solución compuesta por Ficoll 400 al 30% (p/v), azul de bromofenol al 0,2% (p/v), xilencianol al 0,2% y EDTA 40 mM pH 8,0. Una vez terminada la electroforesis, los geles se tiñeron con azul de metileno y se recortó la banda de DNA deseada en cada caso. El DNA se eluyó de la banda de acrilamida tal como describen Sambrook y cols., (1989).

7.2. Técnica de GeneClean

Los fragmentos que se desean aislar se separan mediante electroforesis en geles de agarosa. Posteriormente, se recortan las bandas de interés y el DNA se purifica de la agarosa mediante su adsorción a partículas de vidrio siguiendo la técnica del GeneClean, tal como recomiendan los fabricantes (Bio 101 Inc., La Jolla, CA, USA).

7.3. Geles de agarosa de bajo punto de fusión

Para algunos aislamientos de fragmentos de DNA se utilizaron geles de agarosa de bajo punto de fusión (Bio-Rad) al 1-1,4% en tampón TAE. Una vez realizada la electroforesis, se recortó del gel la banda deseada en cada caso y se resuspendió en tampón TE (Tris-HCl 10 mM, pH 8,0; EDTA 1mM) con NaCl 0,25 M. La agarosa se fundió manteniéndola a 65 °C durante 5 minutos. A continuación, la muestra se trató con fenol y posteriormente con éter (v/v), incubándola después a 65°C durante 10-15 minutos con el fin de eliminar los posibles restos de fenol. El DNA se precipitó con dos volúmenes y medio de etanol absoluto durante 45 minutos a -70°C. Tras una centrifugación a 10.000 x g durante 10 minutos, el precipitado se lavó con etanol 70% y finalmente se disolvió en un volumen adecuado de tampón TE.

7.4. Técnica de la β -agarasa

El empleo de esta técnica también requiere la separación previa de los fragmentos mediante electroforesis en geles de agarosa de bajo punto de fusión y posterior tratamiento de la banda extraída con la β -agarasa. Esta enzima hidroliza la agarosa liberando el DNA. Para ello es necesario fundir previamente la banda de agarosa recortada en el tampón de la enzima (10 mM Bis Tris-HCl pH 6,5; 10 mM EDTA) y posterior tratamiento con la β -agarasa durante una hora a 40-42°C. Después de tratar con fenol y precipitar el DNA se resuspendió en el volumen de TE adecuado. Tanto la enzima como el tampón se obtuvieron de la casa comercial Biolabs.

8. SELECCIÓN DE LOS CLONES PRODUCTORES DE PGA

La selección de los clones con actividad PGA se realizó según el método descrito por Garcia y Buesa (1986). La fenilacetil-L-leucina fue cedida por Antibióticos S.A. (León). Las células de *E. coli* HB101 transformadas con una genoteca *Hind*III de *E. coli* ATCC 11105 construida en pBR322 se sembraron en medio mínimo M9 suplementado con glucosa 10 mM, 1mg/l de tiamina-HCl, 10 mg/l de L-prolina, 100 mg/l de fenilacetil-L-leucina y ampicilina a 75 μ g/ml. Después de 72 horas de incubación a 30 °C, las

colonias que crecieron fueron aisladas y su fenotipo se estudió resemebrándolas en M9 con y sin fenilacetil-L-leucina. Se consideró que un clon era positivo cuando la colonia crecía exclusivamente en presencia de L-leucina o su fenilacetil derivado.

9. TÉCNICAS DE HIBRIDACIÓN DE DNA

Los fragmentos de DNA marcados radiactivamente (sondas) se obtuvieron mediante la técnica del *random-primer* utilizando [α - 32 P] dCTP (400 Ci/mmol, 10 μ Ci/ μ l) (Amersham), el fragmento Klenow, la solución de iniciadores (Pharmacia) y los dNTPs, según se indica en Sambrook y cols., (1989). Todas las hibridaciones y lavados se hicieron a 65 °C.

9.1. Técnica de Southern-blot

La hibridación del DNA mediante la técnica descrita por Southern (1975), se realizó siguiendo el protocolo de Sambrook y cols., (1989). Las membranas de nylon para la transferencia del DNA se obtuvieron de Schleicher and Shuell. Las bandas radiactivas se detectaron a -70°C utilizando películas HyperfilmTM-MP (Amersham) y pantallas amplificadoras Cronex Lightning Plus (Dupont).

9.2. Hibridación del DNA sobre filtros de colonias celulares

Las colonias de *E. coli* se transfirieron a los filtros de nitrocelulosa (Millipore HATF, tamaño de poro 0,45 μ m) como se indica en Sambrook y cols., (1989). Las células se lisaron sobre el filtro mediante tratamientos sucesivos, de 5 minutos de duración cada uno, con las soluciones que se indican: i) NaCl 1,5 M y NaOH 0,5 M (2 veces); ii) Tris-HCl 1 M, pH 7,0; iii) Tris-HCl 0,5 M, pH 7,0 y NaCl 1,5 M; iv) 2 x SSC (NaCl 0,3 M, citrato sódico 0,003 M). La fijación del DNA al filtro se realizó mediante incubación a 80 °C durante 2 horas. Para eliminar los restos celulares, los filtros se hirvieron en agua destilada durante 10 minutos. La hibridación con la sonda radiactiva se realizó como se describe en Sambrook y cols., (1989).

10. SECUENCIACIÓN DEL DNA

La secuenciación del DNA se realizó según el método de Sanger y cols., (1977) por dos procedimientos:

10.1. Secuenciación manual

Para la reacción de secuenciación se utilizó el *T7 DNA polymerase Kit* (Pharmacia) y como desoxinucleósido trifosfato marcado radiactivamente, [α - 35 S] dATP (Amersham), siguiendo las recomendaciones de los fabricantes.

10.2. Secuenciación automática

Parte de la secuenciación del DNA se llevó a cabo en el Centro Nacional de Biotecnología (G.B.F.) (Braunschweig, Alemania) utilizando un secuenciador automático modelo 373A (Applied Biosystems). Para la reacción de secuenciación se utilizó el *Prism ready reaction dye deoxyterminator cycle sequencing Kit* de Applied Biosystems que conlleva el uso de didesoxinucleótidos coloreados y la DNA polimerasa AmpliTaq siguiendo las recomendaciones de los fabricantes. Las reacciones de polimerización se llevaron cabo mediante la técnica de PCR (reacción en cadena de la polimerasa) para lo que se utilizó un termociclador Perkin-Elmer Cetus 480.

10.3. Análisis de la secuencia de DNA

Para el análisis de la secuencia de DNA se utilizó el grupo de programas para Biología Molecular de la Universidad de Wisconsin (WINP) presente en el VAX del Centro Nacional de Biotecnología de Madrid. La secuencia de aminoácidos deducida a partir de la secuencia de nucleótidos se comparó con las proteínas de las siguientes bases de datos: GenBank/EMBL y Swissprot.

11. OBTENCIÓN DE EXTRACTOS DE CULTIVOS CELULARES

Para obtener extractos de las cepas de *E. coli* se cultivaron éstas en 100 ml del medio indicado en cada caso a 37 °C durante 16 horas (fase estacionaria de crecimiento). A continuación, los cultivos se enfriaron a 4 °C y se centrifugaron a 6.000 x g durante 10 minutos, resuspendiéndose las células en 10 ml de tampón fosfato sódico 20 mM ó 100 mM pH 8, dependiendo del ensayo que se fuera a realizar. Una vez homogeneizada la suspensión, las células se rompieron en un sonicador MSE modelo MK2 mediante cuatro pulsos de 15 segundos cada uno, manteniendo siempre la muestra a 4 °C. Finalmente la suspensión sonicada se centrifugó a 4 °C y 30.000 x g durante 20 minutos para eliminar los restos celulares. El sobrenadante de esta centrifugación (extracto crudo) se conservó a -20°C. Cuando se cultivaron volúmenes mayores, las células se rompieron usando una *French-Press* (American Instruments Company) operando a una presión de 7,6 atmósferas.

12. ENSAYOS DE ACTIVIDAD ENZIMÁTICA

En todos los casos las medidas de absorbancia se realizaron en un espectrofotómetro Shimadzu UV-160.

12.1. Penicilina G acilasa

La amplia especificidad que presenta la PGA para el resto ácido o amino de los enlaces que hidroliza, permite el empleo de sustratos derivados del PA cuya hidrólisis supone la aparición de cromóforos. Durante este trabajo, la actividad penicilina G acilasa se determinó de acuerdo al método colorimétrico descrito por Szewczuk y cols., (1979). Los cultivos se incubaron a 30 °C con agitación en presencia de fenilacético al 0,15% durante toda la noche. La mezcla de reacción contenía 0,2 ml de una solución de 5 mM fenilacetil-*p*-aminobenzoico y una alícuota de 0,05 ml de medio de cultivo con células en estado estacionario. Esta mezcla se incubó 20 minutos a 37 °C deteniéndose al añadir 0,250 ml de una solución que contiene 10 mM de nitrito sódico en ácido acético 0,25 M. Posteriormente se añadieron 0,750 ml de la solución reactiva (ácido 1-amino 8-

hidroxinaftaleno 3,6-disulfónico al 0,01% (p/v) en carbonato sódico 0,66 M). Por centrifugación en microfuga se eliminan los restos celulares y se mide la reacción en el espectrofotómetro a 535 nm utilizando un coeficiente de extinción molar (ϵ) de 36.000 $M^{-1} \text{ cm}^{-1}$.

12.2. 4-HPA-hidroxilasa

Para determinar la actividad de la 4-HPA hidroxilasa sobre diferentes sustratos se utilizaron dos métodos que se detallan a continuación.

12.2.1. Cuantificación de la oxidación del NADH

La actividad hidroxilasa se midió añadiendo 0,01 ml de sustrato 0,1 M a 0,5 ml de una solución que contiene, NADH 0,2 mM y 0,06 ml de extracto preparado en tampón fosfato sódico 100 mM pH 8. La oxidación del NADH se determina a 30 °C midiendo la disminución de absorbancia a 340 nm usando un $\epsilon=6.220 M^{-1} \text{ cm}^{-1}$ para el NADH. Los valores se corrigieron por la oxidación del NADH en ausencia de sustrato. Una unidad de actividad hidroxilasa se define como la cantidad de enzima necesaria para la oxidación de 1 μmol de NADH por minuto.

12.2.2. Reacción acoplada con la catecol 2,3-dioxigenasa I (C230 I) de *P. putida*

Este ensayo se llevó a cabo utilizando fenol como sustrato de la 4-HPA hidroxilasa. En presencia de concentraciones saturantes de C230 I el catecol se transforma instantáneamente en 2-hidroximucónico semialdehído (HMS) que presenta un máximo de absorción a 375 nm lo que permite su cuantificación ($\epsilon=46.000 M^{-1} \text{ cm}^{-1}$ a pH 8). La reacción se inicia añadiendo 0,050 ml de extracto de *E. coli* JM101 (pAW31) en el instante en que todo el NADH se ha oxidado. Una unidad de actividad se define como la cantidad de enzima necesaria para la transformación de 1 μmol de HMS por minuto.

12.3. HPC-dioxigenasa

El ensayo de actividad de la HPC-dioxigenasa se llevó a cabo de acuerdo al procedimiento espectrofotométrico descrito por Cooper y Skinner (1980). La mezcla de reacción contenía 0,005 ml de HPC 20 mM, 0,05 ml de extracto crudo y 0,445 ml de tampón fosfato 50 mM pH 8. La formación de 5-carboximetil-2-hidroximucónico semialdehído (CHMS) se detectó a 380 nm utilizando un $\epsilon = 35.000 \text{ M}^{-1} \text{ cm}^{-1}$.

12.4. β -lactamasa

La actividad β -lactamasa se determinó mediante el método espectrofotométrico de Ross y O'Callaghan (1975). Se utilizaron 0,05 ml de extracto enzimático en 0,5 ml de tampón fosfato 100 mM pH 7 iniciándose la reacción por la adición de cefaloridina 0,1 mM. La reacción se desarrolló a 25 °C. La hidrólisis del anillo β -lactámico se siguió a una longitud de onda de 255 nm, registrando la disminución de absorbancia durante 5 min. Se calcularon los μ moles de cefaloridina hidrolizados empleando un $\epsilon = 14000 \text{ M}^{-1} \text{ cm}^{-1}$.

12.5. β -galactosidasa

Los niveles de β -galactosidasa se determinaron de acuerdo al protocolo de Miller (1972). Las células se cultivaron a 30 °C y los efectores ensayados se añadieron al medio de cultivo a una concentración de 1 mM. Se analizaron diferentes alícuotas de medio de cultivo en función de los niveles de β -galactosidasa producidos por cada clon, entre 0,02 y 0,1 ml, completando a un volumen final de 0,5 ml con tampón Z, (fosfato disódico 0,06 M, fosfato monosódico 0,04 M, KCl 0,01 M, SO_4Mg 0,001 M y β -mercaptoetanol 0,05 M). Las células se permeabilizaron añadiendo dos gotas de cloroformo (Merck) y una gota de dodecilsulfato sódico (SDS) al 0,01 %. La reacción se inicia añadiendo 0,1 ml del sustrato, 2-nitrofenil- β -D-galactopiranosido (ONPG) (4 mg/ml). Tras un periodo variable de 5 a 15 minutos a 28 °C se para la reacción con 0,25 ml de carbonato sódico 1 M. Los restos celulares se eliminaron centrifugando las muestras a 12.000 x g durante 10 minutos y posteriormente se midió la reacción en un espectrofotómetro Shimadzu UV-

160 a 420 nm. La medida es estable durante 20 min a 4 °C. Las unidades de β -galactosidasa se calcularon en base a la siguiente ecuación:

$$\text{Unidades } \beta\text{-gal} = [\text{DO}_{420} / t \text{ (min)} \times \text{vol (ml)} \times \text{DO}_{600} \text{ del medio de cultivo}] \times 1000$$

13. IDENTIFICACIÓN DE LOS PRODUCTOS DE LA REACCIÓN DE LA 4-HPA-HIDROXILASA

Los productos obtenidos por incubación de las células en medio M9 suplementado con glucosa 10 mM y el sustrato a analizar a una concentración final de 1 mM, se analizaron por cromatografía líquida de alta resolución (HPLC). Después de centrifugar los cultivos se analizaron los sobrenadantes en un equipo Gilson utilizando una columna Nucleosil 300-5C18 (250 x 4 mm) precedida de una precolumna Nucleosil 300-5C18 (11 x 4 mm). Para detectar la formación de catecol se utilizó una solución (1:1) de fosfato potásico 50 mM y metanol pH 5 como fase móvil. Para detectar L-Dopa, se utilizó como fase móvil una solución de fosfato potásico 50 mM, EDTA 0,1 mM, trifluoroacético 100 mM y acetonitrilo 8 % v/v a pH 4,8. El flujo se mantuvo a 0,5 ml/min y la detección se llevó a cabo espectrofotométricamente a 280 nm. Los metabolitos se identificaron por comparación con los tiempos de retención de las sustancias puras.

14. DETECCIÓN DE PROTEÍNAS CODIFICADAS POR FRAGMENTOS CLONADOS DE DNA

Las proteínas codificadas por los genes *hpaB*, *hpaC* y la ORF 14, presentes en diferentes plásmidos, clonados en *E. coli* SE5000 se detectaron mediante la técnica de maxicélulas descrita por Silhavy y cols., (1984) utilizando [α -³⁵S]metionina (1.500 Ci/mol) (Amershan). La electroforesis en gel de poliacrilamida se llevó a cabo según se describe en el apartado 18.

15. PURIFICACIÓN DE LA 4-HPA-HIDROXILASA

Las células de *E. coli* DH1 (pAJ221) fueron incubadas a 37 °C en 50 ml de medio LB conteniendo 34 µg/ml de cloranfenicol. Después de centrifugar, el sedimento se resuspendió en 5 ml de tampón fosfato 20 mM, pH 8. Las células se rompieron mediante sonicación y el extracto resultante se centrifugó a 30.000 x g durante 20 minutos. Los 5 ml de extracto se cargaron en una columna (3 x 1,5 cm) de Cibacrón Blue 3GA Agarosa-3000-CL (Sigma) equilibrada en tampón fosfato 20 mM, pH 8. La columna se lavó con el mismo tampón y la enzima se eluyó con tampón fosfato 20 mM conteniendo 30 mM NADH. Las fracciones que presentaban actividad hidroxilasa se mezclaron y concentraron con polietilenglicol 20.000. Para la cromatografía de filtración se utilizó una columna Superosa 12 HR 10/30 (Pharmacia) equilibrada en tampón fosfato 50 mM, pH 8 utilizando un equipo de HPLC Gilson. Los patrones utilizados para las determinaciones del masa molecular fueron: ovoalbúmina (45.000 Da), seroalbúmina bovina (67.000 Da) y aldolasa (160.000 Da).

16. CUANTIFICACIÓN DE PROTEÍNAS

La concentración de la proteína existente en preparaciones puras de HpaB y en extractos crudos celulares ha sido determinada mediante el método descrito por Bradford (1976).

17. ELECTROFORESIS DE PROTEÍNAS EN GELES DE POLIACRIAMIDA-SDS

Las electroforesis analíticas de proteínas se realizaron en todos los casos en condiciones desnaturalizantes en presencia de SDS. La técnica empleada fue descrita por Laemmli (1970) utilizando geles de poliacrilamida en placa (160 x 100 x 2 mm) preparados a concentraciones que, según los casos, oscilaron entre el 8 y el 15 %. Las muestras se hirvieron durante 5 minutos en presencia del tampón de ruptura (Tris-HCl 62,5 mM, pH 6,8; SDS al 2%, β-mercaptoetanol al 5%, glicerol al 10% y azul de bromofenol al 0,005%). Las electroforesis se realizaron a temperatura ambiente y

corriente constante (50 mA), utilizando un electrolito que contenía Tris-HCl 0,025 M, glicina 0,192 M y SDS al 0,1%. Las proteínas de los geles se visualizaron con azul brillante de Coomassie R-250, según se describe en Swank y Munkress (1971). Las proteínas empleadas como marcadores de tamaño molecular se adquirieron de Bio-Rad.

18. DETERMINACIÓN DE LA SECUENCIA N-TERMINAL DE HpaB

La secuencia de aminoácidos N-terminal de la proteína HpaB se determinó mediante el método de degradación de Edman utilizando un secuenciador de fase gaseosa conectado a un equipo de HPLC (Applied Biosystems).

19. OBTENCIÓN DE ANTICUERPOS ANTI-HpaB

Para la obtención de anticuerpos policlonales anti-HpaB se inmunizaron conejos con la enzima HpaB purificada. Se utilizaron dos dosis de 100 µg de proteína en solución salina; la primera en adjuvante completo de Freund (1:1 v/v) y la segunda en adjuvante incompleto de Freund (1:1 v/v) dejando transcurrir un mes entre ambas. La administración del antígeno se realizó por vía subcutánea. Una semana después de la administración de la segunda dosis se extrajo la sangre a los conejos para la obtención del suero. La sangre se incubó a 37 °C durante 5 horas y después se mantuvo a 4 °C 12 horas para permitir la formación del coágulo. El sobrenadante se centrifugó a 10.000 x g a 4 °C durante 5 minutos para eliminar los restos celulares y, por último, se inactivaron las proteínas del complemento durante 15 minutos a 56 °C. Se ensayó la especificidad del suero obtenido, así como la posible reacción inespecífica del suero pre-inmune, mediante análisis por Western blot utilizando diferentes diluciones del mismo. Para evitar la detección de bandas inespecíficas en el Western blot formadas por la posible reacción cruzada del suero con otras proteínas celulares diferentes a HpaB se preincubó el suero a 37 °C con un extracto celular de la cepa que posteriormente se iba a emplear para producir la proteína HpaB.

20. TÉCNICA DE WESTERN BLOT

Una vez separadas las proteínas mediante electroforesis en geles de poliacrilamida-SDS, se transfirieron a membranas de nitrocelulosa (Schleicher and Shuell) según describen Sambrook y cols., (1989). La membrana se saturó con leche en polvo desnatada al 5% (p/v) en tampón PBS (fosfato sódico 10 mM, pH 7,4; NaCl 140 mM) mediante incubación a 4°C durante 12 horas. Posteriormente, se incubó a temperatura ambiente con agitación suave durante 4 horas en presencia del suero anti-HpaB. Después de tres lavados con PBS que contenía Tween-20 al 0,1% (v/v), de 10 minutos de duración cada uno, la membrana se hizo reaccionar durante 1 hora a temperatura ambiente y agitación suave con anticuerpos de cabra anti-IgGs de conejo conjugados con peroxidasa (Jackson Immunoresearch). Finalmente se lavó con tampón PBS y las bandas de reacción con los anticuerpos se visualizaron con peróxido de hidrógeno y 4-cloro-1-naftol (Sigma).

21. AISLAMIENTO DE RNA

Para obtener el RNA total de un cultivo de *E. coli*, las células se cultivaron en 20 ml de medio de cultivo hasta una DO_{600} de 0,5 (fase exponencial de crecimiento). El sedimento celular se resuspendió en 1 ml de TE 10:10 (Tri-HCl 10 mM, pH 8; EDTA 10 mM) lavándolo a continuación con el mismo tampón. Posteriormente se resuspendió en 0,4 ml de tampón de lisis (TE 10:10 con 2 mg/ml de lisozima (Sigma)) y se provocó la lisis celular con tres ciclos sucesivos de congelación-descongelación utilizando nitrógeno líquido. Se añadió a la mezcla SDS al 0,8%, acetato sódico 166 mM, pH 4 y 0,5 ml de fenol equilibrado con agua a 60 °C. Se calentaron a 60 °C durante 5 minutos y se centrifugaron a 30.000 x g 10 minutos a 4 °C. El sobrenadante recibió un nuevo tratamiento con fenol ácido caliente. El RNA total se precipitó con etanol absoluto disolviéndose finalmente en un volumen adecuado de TE 10:10. La concentración de RNA se estimó a partir de la DO de la preparación a 260 nm. Una unidad de DO a 260 nm equivale a una concentración de 40 µg/ml. La cantidad de RNA obtenida a partir de los 20 ml de cultivo de *E. coli* osciló entre 0,3 y 0,8 mg. Para comprobar la calidad del RNA, las muestras se visualizaron en un gel de agarosa del 1 %, tras eliminar las

RNAsas de la cubeta y tampones. Una extracción correcta permite observar bandas de 1,1, 0,8 y 0,2 kb. Los RNAs se conservaron a -70 °C. Se eliminaron los posibles restos de DNA de la muestra, que pudieran interferir en el análisis posterior del RNA, mediante un tratamiento con DNAsa libre de RNAsa (DNAsa I de Pharmacia) en presencia de RNAsina (inhibidor de RNAsa) (Boehringer) durante 1 hora a 37 °C. Posteriormente se repitió el proceso de fenolización ácida y precipitación con etanol del RNA en condiciones similares a las descritas anteriormente.

22. DETERMINACIÓN DE LOS SITIOS DE INICIACIÓN DE LA TRANSCRIPCIÓN MEDIANTE EL EMPLEO DE LA TÉCNICA DE EXTENSIÓN CON UN INICIADOR.

Para localizar el sitio de iniciación de la transcripción de los genes *hpaA* y *hpaB* de *E. coli* ATCC 11105 se han seguido dos procedimientos diferentes, ambos basados en la técnica de extensión del cDNA a partir de un iniciador o *primer extension*. Los oligonucleótidos utilizados para determinar los sitios de iniciación de la transcripción de los genes *hpaA* y *hpaB* son O38A (figura 23) y OLACZ50 5'-GCGGAAGCTGGCGGCTGTGGG-3', respectivamente.

22.1. Marcaje del oligonucleótido con [γ -³²P] ATP.

Este método se llevó a cabo según el protocolo descrito por Sambrook y cols., (1989). Para el marcaje de oligonucleótidos se empleó la T4 polinucleótido quinasa utilizando como isótopo [γ -³²P] dATP (5000Ci/mml, 10 μ Ci/ μ l). Para ello se utilizaron 7 pmoles del oligonucleótido y 50 μ Ci de [γ -³²P] dATP. Posteriormente se mezclaron 1 pmol de oligonucleótido y 15 μ g RNA precipitando la mezcla según el procedimiento descrito por Sambrook y cols., (1989). El precipitado se resuspendió en 30 μ l de tampón de hibridación (40 mM PIPES pH 6.4, 1 mM EDTA pH 8, 0,4 M NaCl y formamida al 50 %), desnaturalizando a 85 °C durante 10 minutos e incubando la mezcla 12 horas a 30 °C para permitir la hibridación. El híbrido RNA-oligonucleótido se precipitó con 2,5 volúmenes de etanol absoluto resuspendiéndolo posteriormente en la mezcla de reacción de la transcriptasa reversa (1x tampón AMV (Promega), 0,9 mM dNTPs, pH 7, 0,4 U de

RNAsina, y 3 U de transcriptasa reversa AMV (avian myeloblastosis virus) de Promega. La reacción de extensión transcurrió durante una hora a 42 °C y se detuvo con NaOH a una concentración final de 0,4 M, manteniéndose 12 horas a temperatura ambiente para desnaturalizar los ácidos nucleicos. El siguiente paso consistió en la neutralización de la mezcla con ácido acético 0,4 M y posterior precipitación con etanol, resuspendiendo el precipitado en una solución 1:1 (v/v) de agua y solución colorante (azul de bromofenol al 0,3 %, xilencianol al 0,3 %, 10 mM EDTA, pH 7,5 y formamida desionizada al 97,5 %). Para analizar la longitud de los productos de la extensión, y por tanto el punto de iniciación de la transcripción, se realizó una electroforesis cargando las muestras en un gel de poliacrilamida al 6 %-urea 8 M, detectándose el resultado por autorradiografía. La secuenciación de los plásmidos que contenían las zonas promotoras subclonadas, con el mismo oligonucleótido, sirvió para localizar el punto de iniciación de la transcripción en el genoma.

22.2. Incorporación de [α -³²P]dCTP.

Este método se realizó según el ensayo descrito por Chandry y cols., (1994). Para ello, se calentaron a 90 °C durante un minuto 3 μ l de la mezcla que contenía 40 pmoles de oligonucleótido y 15 μ g de RNA y posteriormente se mantuvo la mezcla a 42 °C durante 5 minutos para realizar el anillamiento RNA-oligonucleótido. El cDNA es sintetizado en un volumen de reacción de 10 μ l que contiene 24 U de la enzima transcriptasa reversa (AMV), 15 μ Ci de [α -³²P]dCTP, dCTP 10 μ M, dATP 100 μ M, dGTP 100 μ M, dTTP 100 μ M, Tris-HCl 50 mM pH 8,5, KCl 50 mM, MgCl₂ 20 mM y ditiotreitol 10 mM. La reacción se incubó a 42 °C durante 30 minutos y se detiene con 5 μ l de formamida al 0,5 %, EDTA 20 mM, azul de bromofenol 0,05 % y xilencianol 0,05 %. El resultado se visualiza y analiza como en el método anterior.

Este procedimiento es mucho más rápido que el anterior y tiene la ventaja de conseguir una mejor desnaturalización cuando se trabaja con material genético de elevado contenido en estructura secundaria.

III. RESULTADOS

1. CLONACIÓN DEL GEN *pac* DE *E. coli* ATCC 11105

Como se ha comentado en la Introducción (apartado 6.) el clon HB101 (pAEC19), productor de PGA, presentaba un fenotipo negro al incubarlo en medio LB que había sido relacionado con la presencia de una hidroxilasa de compuestos aromáticos en esta cepa. Además, la producción del pigmento negro también se observaba al incubar las células en medio mínimo suplementado con diferentes compuestos aromáticos como L-tirosina, N-acetil-L-tirosina, L-tirosina-metil éster, 4-HPA, 3-HPA y fenol, lo cual apoyaba la hipótesis de la existencia de una hidroxilasa de aromáticos en este clon (figura 10). Sin embargo, cuando las células se incubaron en presencia de PA, L-fenilalanina o 2-HPA no se detectó la producción de ningún pigmento. Dado que por una parte, el clon HB101 (pAEC19) era propiedad de Antibióticos S.A y que, además, era necesario comprobar que este fenotipo era realmente consecuencia de la expresión de un gen de *E. coli* ATCC 11105, se construyó nuevamente la genoteca *Hind*III de esta estirpe en pBR322 y, utilizando el método desarrollado por García y Buesa (1986) para seleccionar clones productores de PGA, se seleccionó el clon HB101 (pAJ19) que presentaba un fenotipo idéntico al de las células transformadas con el plásmido pAEC19. Para confirmar que el plásmido pAJ19 contenía el gen responsable del fenotipo negro se transformaron diferentes estirpes de *E. coli* (DH1, JM101, TG1 y SE5000) obteniéndose siempre el mismo resultado.

1.1. Localización del gen responsable del fenotipo negro

El plásmido pAJ19 contiene un fragmento *Hind*III de 8,5 kb del DNA cromosómico de *E. coli* ATCC 11105. El gen *pac* fue localizado en uno de los extremos de este fragmento mediante secuenciación, utilizando oligonucleótidos sintéticos diseñados a partir de la secuencia de nucleótidos del vector pBR322. Para localizar el gen responsable del fenotipo negro, al que se denominó *hpaB*, se hicieron distintas construcciones algunas de las cuales se muestran en la figura 11. Un paso decisivo en la localización del gen *hpaB* consistió en la construcción del plásmido pAJ21 mediante una deleción *Sal*I del plásmido pAJ20. Las células transformadas con el plásmido pAJ21 no producían el pigmento negro al incubarlas en LB, lo que indicaba que el gen *hpaB* estaba

localizado en el extremo opuesto al gen *pac* del fragmento *Hind*III. El plásmido pAJ27 (figura 11) contenía un fragmento *Hind*III-*Ssp*I de 1,7 kb siendo la construcción más pequeña en la que se detectaba la producción del pigmento negro.

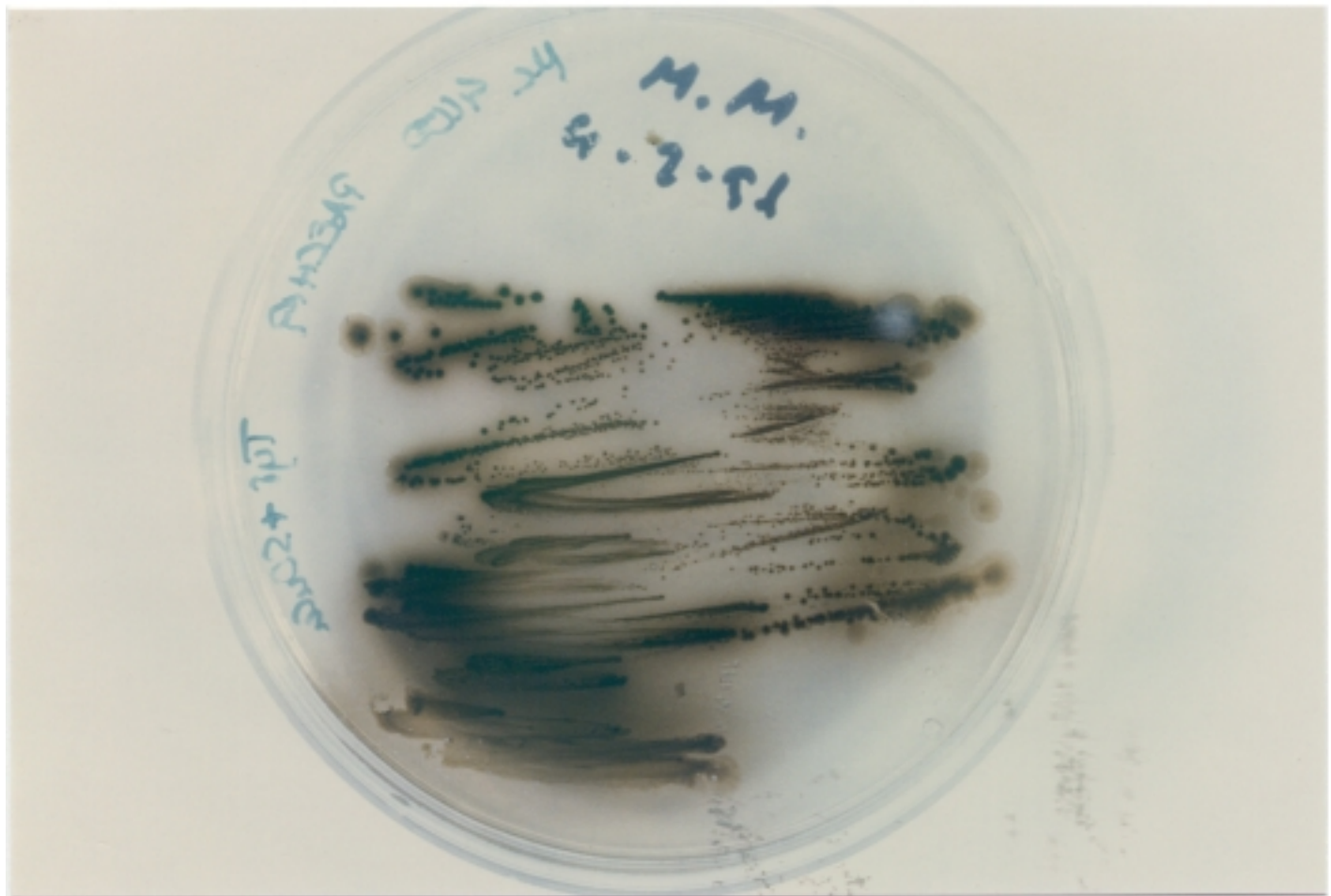


Figura 10. Fenotipo de la cepa HB101 (pAEC19). Las células se cultivaron en medio M9 suplementado con glucosa 10 mM y L-tirosina 1 mM.

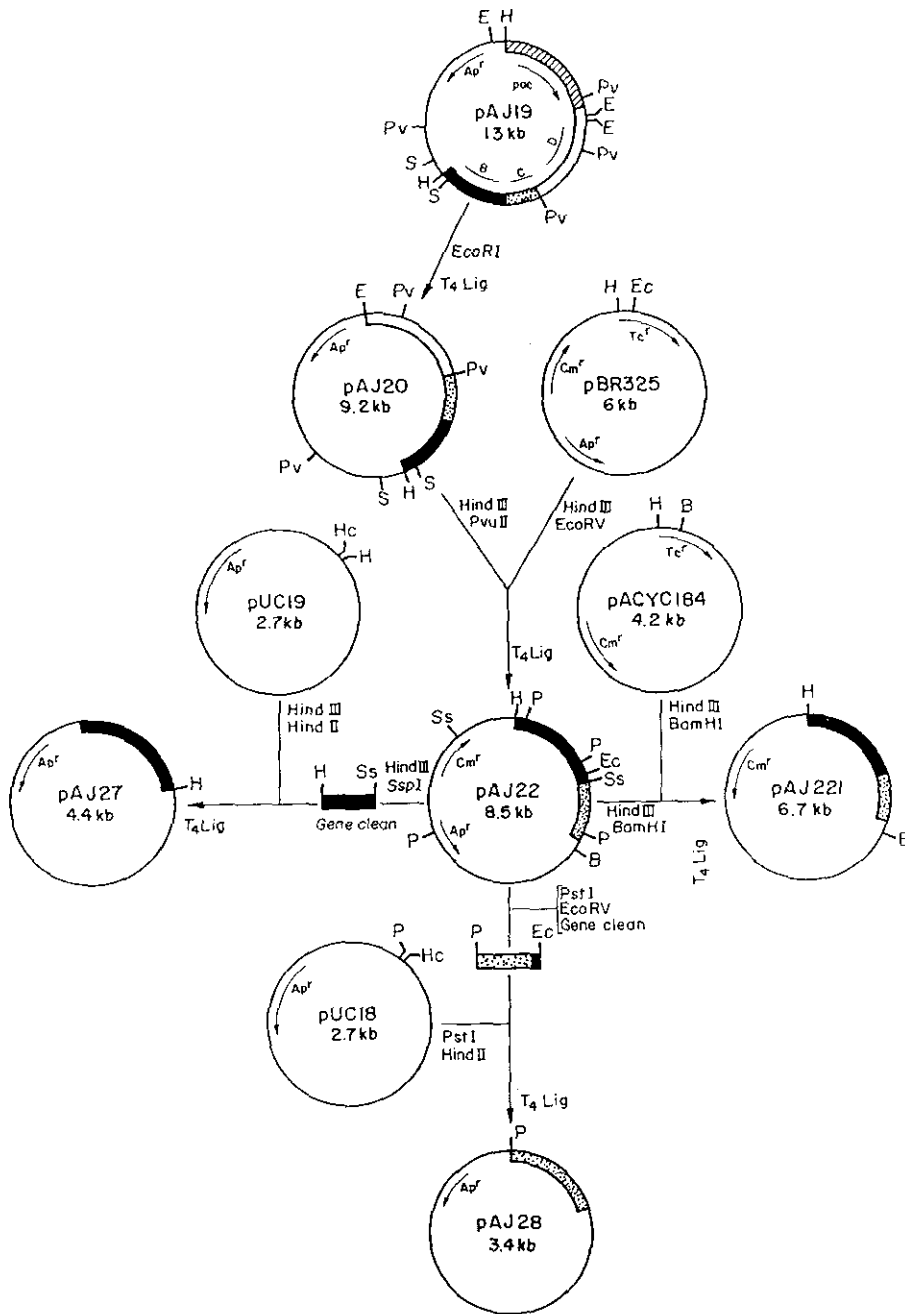


Figura 11. Subclonación del gen *hpaB*. Símbolos: ▨, gen *pac*; ■, gen *hpaB*; ▩, gen que codifica para la proteína C; □, gen que codifica para la proteína D. Las flechas indican la dirección de transcripción de los genes. Abreviaturas: Ap^r , resistencia a ampicilina; Cm^r , resistencia a cloranfenicol; Tc^r , resistencia a tetraciclina; B, *BamHI*; E, *EcoRI*; Ec, *EcoRV*; Hc, *HindII*; H, *HindIII*; P, *PstI*; Pv, *PvuII*; S, *SalI*; Ss, *SspI*.

1.2. Detección de las proteínas clonadas en el plásmido pAJ19

Una vez localizado el gen *hpaB* en el plásmido pAJ27 se analizaron los productos de los genes clonados en el plásmido pAJ19 mediante la técnica de maxicélulas (figura 12). En el carril correspondiente al clon SE5000 (pAJ19), además de la β -lactamasa (30.000 Da), se observan tres bandas 59.000, 40.000 y 19.000 Da que se indican como proteínas B, D y C, respectivamente. No se detectaron las bandas correspondientes a las dos subunidades de la PGA de 60.000 y 26.000 Da, ya que la producción de esta enzima sólo tiene lugar a 30 °C y en presencia de PA. Los plásmidos pAJ19, pAJ22 y pAJ27 contenían el gen *hpaB* a diferencia del plásmido pAJ28, dado que en las células transformadas con este último no se observó el fenotipo negro. Teniendo en cuenta que en las células que contenían el plásmido pAJ22 se producía la cloranfenicol acetil transferasa (27.000 Da) y las proteínas B y C, y en el clon SE5000 (pAJ27) sólo se observaban la β -lactamasa y la proteína B, se dedujo que el producto del gen *hpaB* podría ser la proteína B que se denominó HpaB.

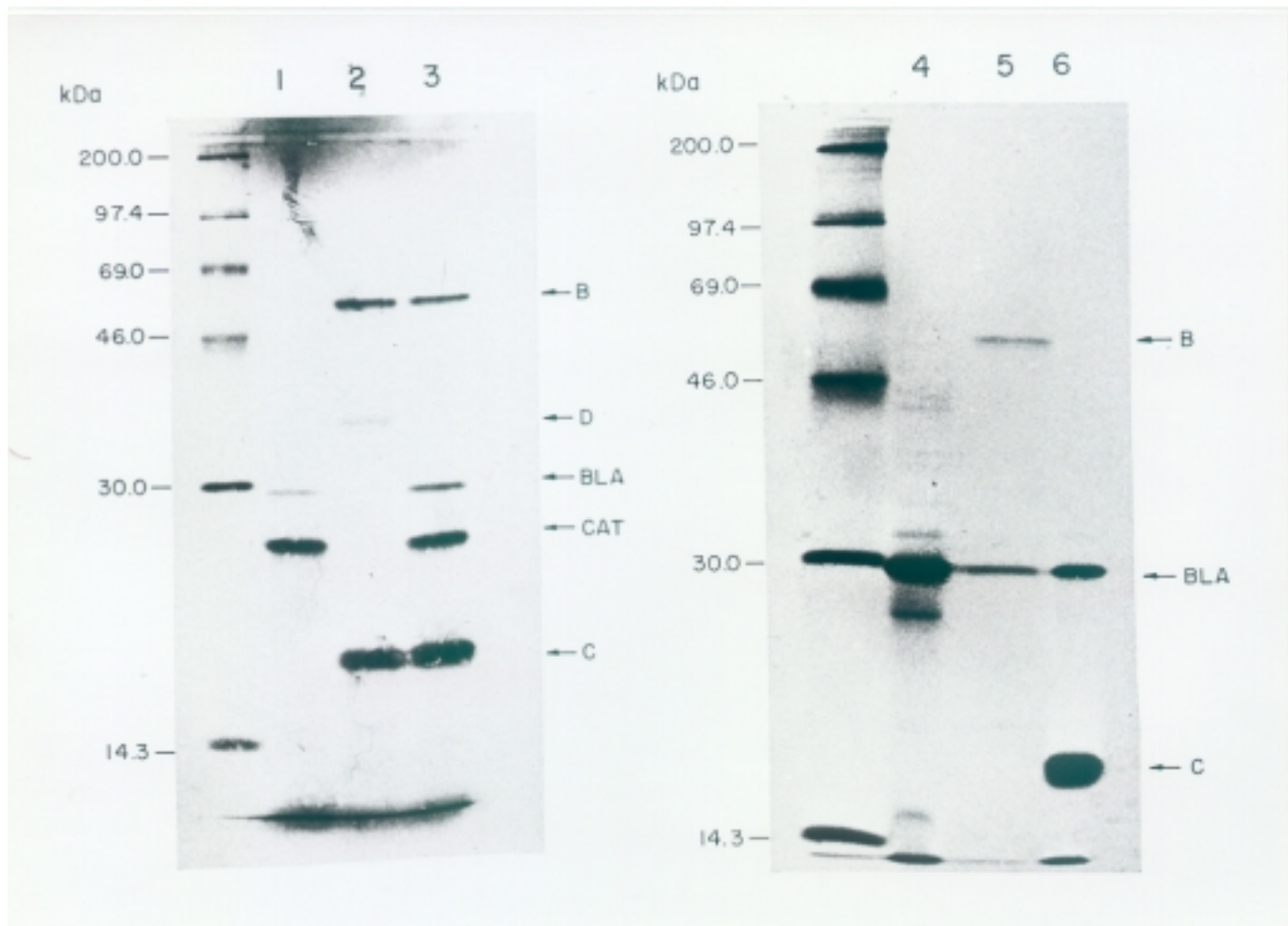


Figura 12. Detección de las proteínas codificadas por los genes flanqueantes al gen *pac*. Análisis por maxicélulas de la cepa *E. coli* SE5000 transformada con los siguientes plásmidos: línea 1, pBR325; línea 2, pAJ19; línea 3, pAJ22; línea 4, pUC18; línea 5, pAJ27; línea 6, pAJ28. Abreviaturas: B, producto del gen *hpaB*; BLA, β -lactamasa; C, proteína C; CAT, cloranfenicol acetil transferasa; D, proteína D. La masa molecular de los patrones se indica en el margen izquierdo de la figura.

2. CARACTERIZACIÓN DE LA HIDROXILASA

2.1. Identificación de los productos de la reacción de hidroxilación

Como se ha mencionado, los clones que expresan la hidroxilasa producen pigmentos de coloración marrón o negra al incubarlos en medio mínimo en presencia de compuestos aromáticos monohidroxilados como L-tirosina, fenol, y 3- ó 4-HPA. Para comprobar que los productos coloreados procedían de la oxidación espontánea de derivados dihidroxilados obtenidos mediante la acción de una hidroxilasa, se identificaron los metabolitos de la reacción mediante cromatografía líquida de alta resolución (HPLC). Para ello, la cepa *E. coli* DH1 (pAJ22) se incubó durante toda la noche a 37 °C en medio M9 con glucosa suplementado con L-tirosina o fenol 1 mM. Para comprobar que los sustratos no se transformaban por otro sistema diferente a la reacción mediada por la hidroxilasa, se incubaron en las mismas condiciones células de *E. coli* DH1 (pBR325). Posteriormente los cultivos se centrifugaron a 6.000 x g durante 10 minutos y el sobrenadante se utilizó para el análisis en HPLC empleando como patrones sustancias puras para determinar el tiempo de retención de cada compuesto. Los resultados de este experimento indicaron que la cepa *E. coli* DH1 (pAJ22) producía L-dopa y catecol a partir de L-tirosina y fenol respectivamente, demostrándose que la enzima codificada en el plásmido pAJ22 transformaba estos sustratos aromáticos monohidroxilados en los correspondientes dioles mediante una reacción de hidroxilación (figura 13). Este sistema se utilizó para comprobar la transformación de otros posibles sustratos como N-acetil-L-tirosina y L-tirosina-metil éster, detectándose en ambos casos la aparición de un metabolito más polar que no pudo ser identificado por carecer de las sustancias puras correspondientes. Mediante este método también fue analizada la transformación del PA y de sus derivados hidroxilados, observándose que los picos correspondientes al PA y al 2-HPA permanecían invariables en el cromatograma del sobrenadante del medio de cultivo de *E. coli* DH1 (pAJ22). Sin embargo, el 3-HPA y el 4-HPA se transformaban en un metabolito que, dada su inestabilidad, no era identificable mediante este método.

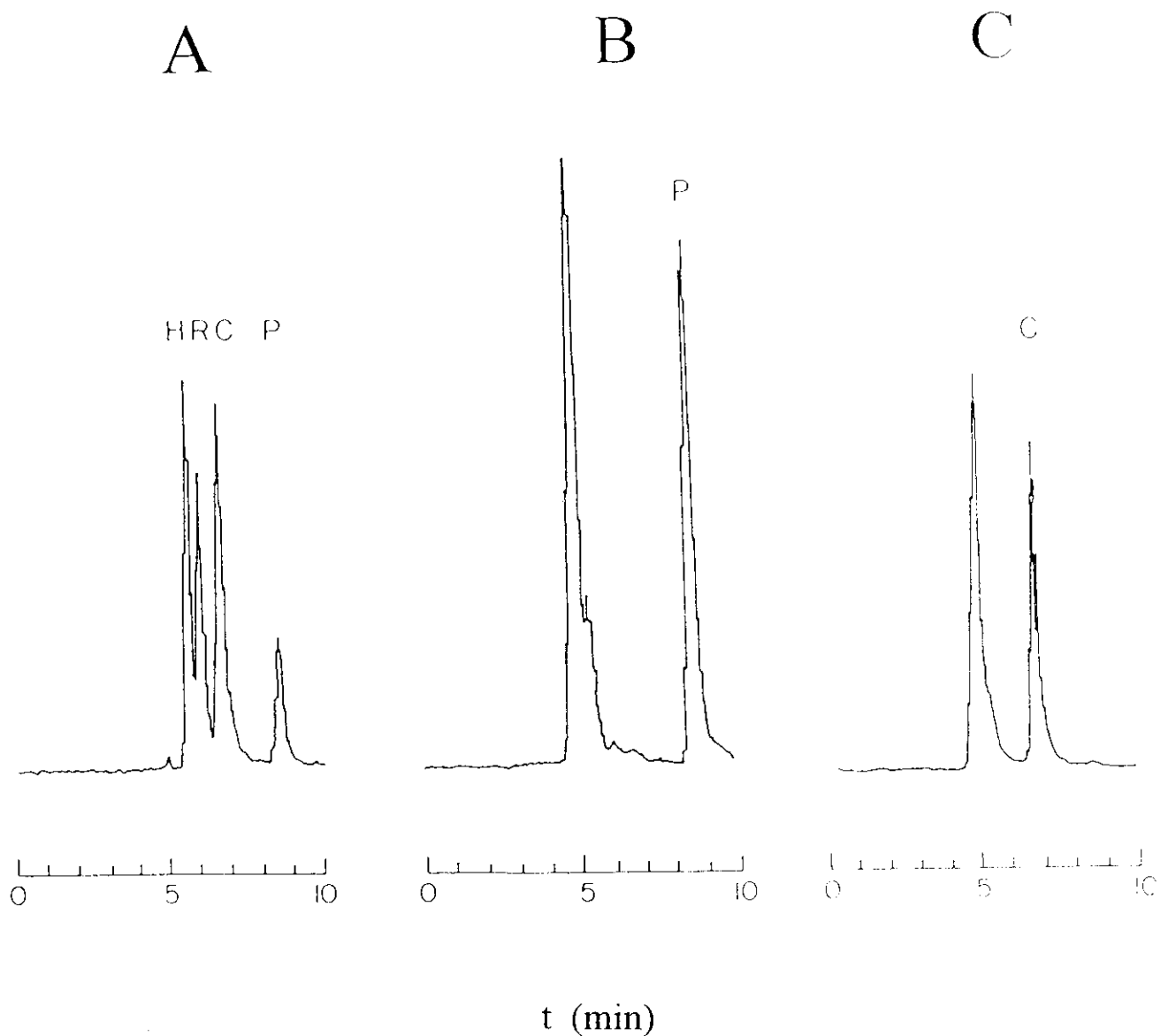


Figura 13. Identificación por HPLC del producto de la oxidación de fenol mediada por la 4-HPA-hidroxilasa. Panel A, solución estándar de hidroquinona (H), resorcinol (R), catecol (C) y fenol (P). Panel B, sobrenadante del cultivo de *E. coli* DH1 (pBR325) cultivada en medio M9 suplementado con fenol 1mM. Panel C, sobrenadante del cultivo de *E. coli* DH1 (pAJ22) cultivada en las mismas condiciones.

La transformación de 3- y 4-HPA en HPC por la acción de la hidroxilasa fue comprobada por otro procedimiento consistente en el acoplamiento *in vivo* de la reacción de hidroxilación y la reacción de oxigenación mediada por la HPC-dioxigenasa de *K. pneumoniae* M5a1 (Gibello y cols., 1994). Esta dioxigenasa lleva a cabo la ruptura en posición *meta* del anillo aromático del HPC dando lugar al 5-carboximetil-2-hidroximucónico semialdehído (CHMS), que es de color amarillo y presenta un máximo de absorbancia a 380 nm. Para expresar los genes de la hidroxilasa y de la dioxigenasa en una misma cepa se construyó el plásmido pAJ221 (figura 11), compatible con el plásmido pAG464, el cual es un derivado de pUC18 que contiene el gen de la HPC-dioxigenasa (Gibello y cols., 1994). El vector pACYC184 se digirió con las nucleasas *Hind*III y *Bam*HI ligándose posteriormente al fragmento de 2,7 kb *Hind*III-*Bam*HI de pAJ22, que previamente había sido purificado mediante la técnica de la β -agarosa. La mezcla de ligación se utilizó para transformar las células competentes de *E. coli* DH1 y los clones recombinantes fueron seleccionados en base a su resistencia a cloranfenicol y al fenotipo negro que, en este caso, era más evidente que en cualquier otra construcción analizada anteriormente ya que, como se observó posteriormente, la cepa *E. coli* DH1 (pAJ221) hiperproduce la hidroxilasa. Una vez obtenido el plásmido pAJ221 se transformaron las células competentes de *E. coli* DH1 (pAG464).

Al incubar la células de *E. coli* DH1 (pAJ221, pAG464) en medio M9 suplementado con glucosa y con 4-HPA ó 3-HPA, se producía un compuesto de color amarillo cuyo espectro de absorción coincidía con el del CHMS lo que confirmaba la presencia de este compuesto en el medio de cultivo. Al suplementar el medio con 2-HPA como sustrato aromático de la hidroxilasa no se detectaba CHMS en el medio de cultivo, ya que el 2-HPA no era un sustrato de la hidroxilasa.

Un experimento similar al anterior pero utilizando la catecol 2,3-dioxigenasa de *P. putida* codificada en el plásmido pAW31, confirmó la transformación de fenol en catecol por acción de la hidroxilasa. La cepa *E. coli* JM101 (pAW31, pAJ221), incubada en condiciones similares a las anteriormente descritas pero suplementando el medio mínimo con fenol, producía 2-hidroximucónico semialdehído (HMS) que también es amarillo y presenta un máximo de absorbancia a 375 nm (datos no mostrados).

2.2. Estudio de los cofactores requeridos por la hidroxilasa

Las oxigenasas hidroxilantes requieren donadores de electrones externos como el NADH ó NADPH y un centro redox, como mínimo, normalmente compuesto por metales de transición o una molécula de flavina (apartado 3.2.1.1. de la Introducción). Para llevar a cabo el estudio de los cofactores requeridos por la hidroxilasa se utilizó el fenol como sustrato ya que, a concentraciones saturantes de la catecol 2,3-dioxigenasa de *P. putida* (C23O), todo el catecol se convierte instantáneamente en HMS que puede ser detectado a simple vista en un análisis cualitativo. El análisis colorimétrico permite cuantificar la cantidad de sustrato transformado. El acoplamiento de estas dos reacciones *in vitro* no sólo permitió determinar los requerimientos bioquímicos de la hidroxilasa sino que también facilitó la detección de la enzima durante el proceso de purificación que se detalla en el apartado 2.6.

La formación de HMS en el extracto crudo de la cepa *E. coli* JM101 (pAW31, pAJ221) se midió espectrofotométricamente llevándose a cabo la reacción en presencia de FAD o FMN $1\mu\text{M}$ añadiendo en ambos casos NADH ó NADPH $0,2\text{ mM}$ como cosustrato. Mediante estos experimentos se llegó a la conclusión de que la hidroxilasa era una enzima dependiente de NADH. Aunque los niveles de HMS no aumentaban en presencia de FAD o FMN, esto no indicaba que no fueran requeridos en la reacción de hidroxilación ya que estos cofactores estaban presentes en el extracto crudo, probablemente, en concentraciones saturantes.

2.3. Especificidad de sustrato

La especificidad de sustrato de la hidroxilasa se determinó utilizando el extracto crudo de la cepa DH1 (pAJ221) preparado en tampón fosfato sódico 100 mM , pH 8, midiendo espectrofotométricamente la oxidación de NADH (tabla 9). Los resultados de este experimento permitieron clasificar la enzima como una 4-HPA-hidroxilasa, a pesar de que presentaba un rango de sustrato muy amplio actuando sobre una gran variedad de derivados aromáticos como la L-tirosina, el HPC, el *p*-cresol, el fenol y algunos derivados clorados de éste. La oxidación de NADH no aumentaba en presencia de PA ni

de 2-HPA, lo que confirmaba los resultados obtenidos en la identificación mediante HPLC de los productos de la reacción de hidroxilación.

Tabla 9. Sustratos de la hidroxilasa

Compuesto	Actividad (%)
4-Hidroxifenilacetato	100 ^a
3-Hidroxifenilacetato	82
2-Hidroxifenilacetato	N.D. ^b
Homoprotocatecuato	65
2,5-Difidroxifenilacetato	155
3-Cl-4-Hidroxifenilacetato	16
Fenilacetato	N.D.
4-Clorofenilacetato	5
<i>o</i> -Cresol	N.D.
<i>m</i> -Cresol	N.D.
<i>p</i> -Cresol	51
Fenol	28
2-Clorofenol	N.D.
3-Clorofenol	5
4-Clorofenol	41
Catecol	2
Resorcinol	28
Hidroquinona	32
<i>L</i> -Tirosina	5
<i>L</i> -Fenilalanina	N.D.
<i>L</i> -Dopa	7
D-4-Hidroxifenilglicina	N.D.

^a 100 % de actividad equivale a 0.2 U/mg de proteína.

^b N.D. No detectado

2.4. Secuenciación del operón de la 4-HPA-hidroxilasa

El análisis de la expresión de los genes contenidos en el plásmido pAJ22 mediante la técnica de maxicélulas indicó que el fragmento *HindIII-PvuII* de 2,5 kb contenía el gen que codifica para la 4-HPA-hidroxilasa. Este plásmido fue utilizado como punto de partida para la construcción de una serie de subclones (plásmidos pAJ24-28) (tabla 10 y figura 22). Los fragmentos clonados en estos plásmidos fueron secuenciados por ambas cadenas revelando la existencia de dos marcos de lectura abiertos (ORFs) que parecían estar formando parte de la misma unidad de transcripción. La región de DNA analizada es la comprendida entre los nucleótidos 9.185 y 11.695 de la figura 23, donde se indican los oligonucleótidos sintéticos utilizados en la secuenciación. En la figura 23 se puede observar que la primera ORF comienza en el nucleótido 9.258 y su codon de iniciación está precedido por la secuencia AGAGG correspondiente a un posible sitio de unión al ribosoma (RBS). La masa molecular deducida de esta ORF era de 58.781 Da que se correspondía con el valor de la masa molecular de la proteína HpaB determinada mediante la técnica de maxicélulas (figura 12). El codon de iniciación de la segunda ORF está localizado en la posición 10.838 de la figura 23 y, como en el caso anterior, está precedido por un posible RBS que se corresponde con la secuencia AGGAGG. Esta ORF codificaba para un polipéptido cuya masa molecular (18.679 Da) se correspondía con la observada para la proteína C (figura 12) a la que se denominó HpaC por estar codificada en el mismo operón que HpaB. En la región inmediatamente posterior al extremo 3' del gen *hpaC*, entre los nucleótidos 11.357-11383, se encontró un posible terminador de la transcripción Rho-independiente con un energía libre de -22.6 kcal/mol, lo que parecía indicar que *hpaC* era el último gen del operón.

Para determinar la presencia de otros genes que formaran parte del operón de la hidroxilasa en la región anterior al extremo 5' del gen *hpaB*, se construyó una genoteca *EcoRV* con el DNA de *E. coli* ATCC 11105 en pUC18. Así se aisló el plásmido pAJ33 (figura 14) que contenía un fragmento *EcoRV* de 2,6 kb situado entre los nucleótidos 8147-10.788 de la figura 23. El análisis de la secuencia de este fragmento reveló la existencia de otro gen al que se denominó *hpaA*, que terminaba en el nucleótido 9.005 de la figura 23. Aunque el gen *hpaA* no estaba aún completo en su extremo 5' terminal,

Tabla 10. Plásmidos construidos para la secuenciación de la ruta del 4-HPA

Plásmido (kb)	Vector	Diana del Vector	Inserto			
			Diana	(pb)	Procedencia	
pAJ19	13	pBR322	<i>Hind</i> III	<i>Hind</i> III	8.593	Genoteca <i>Hind</i> III de <i>E. coli</i> ATCC 11105
pAJ20	9,2	pBR322	<i>Hind</i> III- <i>Eco</i> RI	<i>Hind</i> III- <i>Eco</i> RI	4.890	Delección <i>Eco</i> RI de pAJ19
pAJ21	8,5	pBR322	<i>Sal</i> I- <i>Eco</i> RI	<i>Sal</i> I- <i>Eco</i> RI	4.822	Delección <i>Sal</i> I de pAJ20
pAJ22	7,8	pBR325	<i>Hind</i> III- <i>Eco</i> RV*	<i>Hind</i> III- <i>Pvu</i> II*	2.506	Fragmento <i>Hind</i> III- <i>Pvu</i> II de pAJ20
pAJ23	7,8	pACYC184	<i>Eco</i> RI	<i>Eco</i> RI- <i>Hind</i> III	3.510	Fragmento <i>Eco</i> RI de pAJ19
pAJ24	3	pUC19	<i>Pst</i> I- <i>Bam</i> HI	<i>Pst</i> I- <i>Pvu</i> II*	163	Fragmento <i>Pst</i> I- <i>Bam</i> HI de pAJ22
pAJ25	3,5	pUC19 +	<i>Pst</i> I	<i>Pst</i> I	800	Fragmento <i>Pst</i> I de pAJ22
pAJ26	4	pUC19 _	<i>Pst</i> I	<i>Pst</i> I	1.341	Fragmento <i>Pst</i> I de pAJ22
pAJ27	4,4	pUC19	<i>Hind</i> III- <i>Hind</i> II*	<i>Hind</i> III- <i>Ssp</i> I*	1.717	Fragmento <i>Hind</i> III- <i>Ssp</i> I de pAJ22
pAJ28	3,4	pUC18	<i>Pst</i> I- <i>Hind</i> II*	<i>Pst</i> I- <i>Eco</i> RV*	739	Fragmento <i>Pst</i> I- <i>Eco</i> RV de pAJ22
pAJ32	4	pUC18 _	<i>Sma</i> I*	<i>Pvu</i> II*	1.319	Fragmento <i>Pvu</i> II de pAJ20
pAJ33	5,3	pUC18 _	<i>Sma</i> I*	<i>Eco</i> RV*	2.641	Genoteca <i>Eco</i> RV de <i>E. coli</i> ATCC 11105
pAJ34	4,5	pUC18	<i>Bam</i> HI- <i>Sma</i> I*	<i>Bam</i> HI- <i>Eco</i> RV*	1.868	Delección <i>Bam</i> HI de pAJ33
pAJ35	3,4	pUC18 _	<i>Bam</i> HI	<i>Eco</i> RV*- <i>Bam</i> HI	733	Fragmento <i>Bam</i> HI de pAJ33
pAJ36	3,7	pUC18 +	<i>Hind</i> III	<i>Hind</i> III- <i>Eco</i> RV*	1.038	Fragmento <i>Hind</i> III de pAJ33
pAJ37	3,7	pUC18	<i>Hind</i> II*- <i>Eco</i> RI	<i>Pvu</i> II*- <i>Eco</i> RI	1.065	Fragmento <i>Pvu</i> II- <i>Eco</i> RI de pAJ20
pAJ38	5,2	pUC18 _	<i>Bam</i> HI	<i>Bam</i> HI	2.521	Genoteca <i>Bam</i> HI de <i>E. coli</i> ATCC 11105
pHCB1	8,7	pUC18 _	<i>Bam</i> HI	<i>Bam</i> HI	6.017	Genoteca <i>Bam</i> HI de <i>E. coli</i> ATCC 11105
pHCB2	5,3	pUC18	<i>Bam</i> HI- <i>Eco</i> RI	<i>Bam</i> HI- <i>Eco</i> RI	2.631	Delección <i>Eco</i> RI de pHCB1
pHCB3	13,7	pUC18 _	<i>Eco</i> RI	<i>Eco</i> RI	11.061	Genoteca <i>Eco</i> RI de <i>E. coli</i> ATCC 11105
pHCB4	7,2	pUC18	<i>Bam</i> HI- <i>Sma</i> I*	<i>Bam</i> HI- <i>Eco</i> RV*	4.566	Delección <i>Eco</i> RV- <i>Sma</i> I de pHCB1
pHCB5	2,9	pUC18 _	<i>Sma</i> I*	<i>Eco</i> RV*	288	Fragmento <i>Eco</i> RV de pHCB1
pHCB6	3,8	pUC18	<i>Sma</i> I*- <i>Bam</i> HI	<i>Eco</i> RV*- <i>Bam</i> HI	1.163	Fragmento <i>Eco</i> RV- <i>Bam</i> HI de pHCB1
pHCR1	8,5	pUC18 _	<i>Eco</i> RI	<i>Eco</i> RI	5.900 ^a	Genoteca <i>Eco</i> RI de <i>E. coli</i> ATCC 11105
pHCR2	10	pACYC184	<i>Eco</i> RI	<i>Eco</i> RI	5.900 ^a	Fragmento <i>Eco</i> RI de pHCR1
pHCR3	7,6	pUC18	<i>Hind</i> III- <i>Eco</i> RI	<i>Hind</i> III- <i>Eco</i> RI	5.000 ^a	Delección <i>Hind</i> III de pHCR1
pAJ40	19	pUC18	<i>Hind</i> II- <i>Eco</i> RI	<i>Hind</i> II- <i>Eco</i> RI- <i>Eco</i> RI	16.000 ^a	Digestión <i>Eco</i> RI pHCR3 ligado a fragmento <i>Eco</i> RI de pHCB3

* : Diana de restricción perdida durante el proceso de clonación.

a : Tamaño aproximado.

+ : Promotor *P_{lac}* orientado en el sentido de la transcripción de los operones catabólicos.

- : Promotor *P_{lac}* orientado en el sentido contrario al de la transcripción de los operones catabólicos.

parecía no formar parte del mismo operón que los genes *hpaB* y *hpaC* ya que en la región de 254 pb comprendida entre los genes *hpaA* y *hpaB* existía una secuencia de nucleótidos capaz de originar una posible estructura de bucle con una energía libre de -19.8 kcal/mol, que podría actuar como un terminador de la transcripción (nucleótidos 9.044-9.084 de la figura 23). Por otra parte, el promotor del operón de la 4-HPA-hidroxilasa fue localizado en la región situada entre el gen *hpaA* y el operón *hpaBC*, como se demostrará en el apartado 4.1.

La proteína truncada HpaA se parecía (21,4% de identidad) a la proteína MelR de *E. coli* (figura 31), un miembro de la familia de reguladores de transcripción XylS/AraC (Gallegos y cols., 1993), lo que indicaba que HpaA podría ser una proteína reguladora. La secuencia de HpaB no mostró una similitud significativa con ninguna proteína secuenciada hasta el momento, a diferencia de la proteína HpaC que se parecía (24,7% de identidad) a la ORF6 del *cluster* de genes de la síntesis de actinorrodina de *Streptomyces coelicolor* (Fernández-Moreno y cols., 1992) (figura 15).

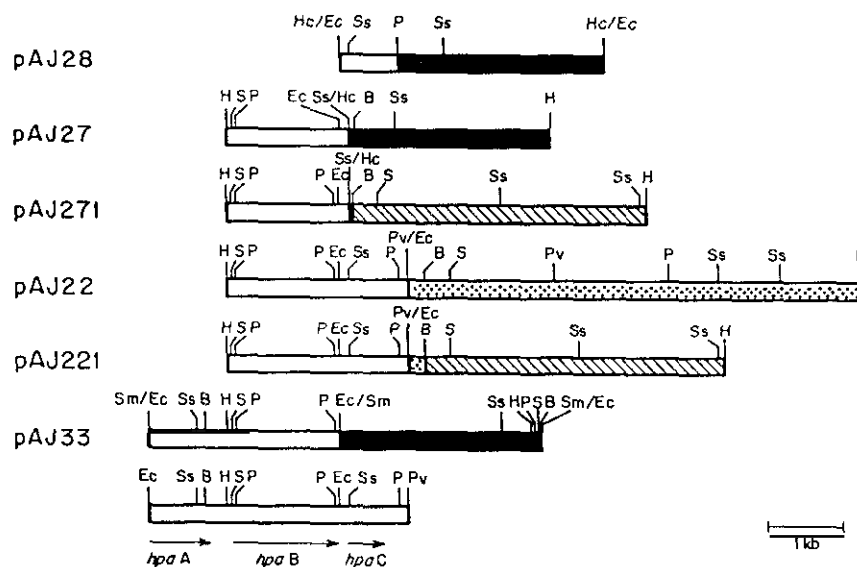


Figura 14. Mapa de restricción de los plásmidos pAJ28, pAJ27, pAJ271, pAJ22, pAJ221 y pAJ33. Símbolos: ■, pUC18 y pUC19; ▨, pBR325; ▩, pACYC184; □ región clonada. Las flechas indican la dirección de transcripción de los genes. Abreviaturas: B, *Bam*HI; Ec, *Eco*RV; Hc, *Hind*II; H, *Hind*III; P, *Pst*I; Pv, *Pvu*II; S, *Sal*I; Sm, *Sma*I; Ss, *Ssp*I.

HpaC	MQLDEQRLRFRDAMASLSAAVNIITTEGDADNA-GLRQRPSCSVTDTPPSLMVCINANSA	59
	: : : ::: : ::: ::: : : : : : : : : : : :: :::	
ORF6	MAADQGML--RDAMARVPAGVALVTAHDRGGVPHGFTASSFVSVSMEPPLALVCLARTAN	58
HpaC	MNPVFQNGKLCVNVLNHEQEVMARHFAGMTGMAMEERFSLSCWQKGPLAQPVLKGSIAS	119
	:: : : : : : : :	
ORF6	SFPVFDSCGEFAVSVLREDHTDLAMRFA----RKSADKFAGGEFVRTARGATVLDGAVAV	114
HpaC	LEGEIRDVQAIGTHLVYLVEIKNIILSAEGHGLIYFKRRFHPVMLEMEAAI	170
	: : : : : : : : : : : : : : : : : : : : :	
ORF6	VECTVHERIPAGDHIILLGEVQSVHVEEKGVPAVYVDRRFAALCSAAGACPSATGRGVPH	175
ORF6	AG	177

Figura 15. Alineamiento de la proteína HpaC y la ORF6 del *cluster* de genes de la síntesis de actinorrodina de *S. coelicolor*. | y : indican, respectivamente, aminoácidos idénticos y conservados.

2.5. Implicación de la proteína HpaC en la actividad hidroxilasa

Como se comentó en el apartado 3.2.1.1. de la Introducción, la 4-HPA-hidroxilasa de *P. putida* es un sistema enzimático formado por dos componentes proteicos; una flavoproteína homodimérica, que cataliza la oxidación de NADH dependiente de 4-HPA, y una proteína cooperadora, que acopla la reacción de oxidación de NADH a la incorporación del grupo hidroxilo en el anillo aromático (Arunachalam y cols., 1992). En este sistema, la transformación de 4-HPA en HPC depende de la presencia de ambos componentes.

Durante la subclonación del gen *hpaB* se había observado que, en las mismas condiciones de incubación, *E. coli* DH1 (pAJ27) presentaba un fenotipo negro menos acusado que las cepas que producían conjuntamente las proteínas HpaB y HpaC. Este hecho podría ser atribuido a una diferencia cuantitativa en la expresión del gen *hpaB* en estos clones pero, teniendo en cuenta el mecanismo de reacción descrito para la 4-HPA-hidroxilasa de *P. putida*, y que los genes *hpaB* y *hpaC* forman parte del mismo operón, cabía la posibilidad de que la proteína HpaC actuara como proteína cooperadora incrementando la actividad hidroxilasa de HpaB.

El plásmido pAJ271 se construyó ligando el fragmento *HindIII-BamHI* de 1,7 kb al vector pACYC184 digerido con las mismas enzimas de restricción (figura 14). La mezcla de ligación se utilizó para transformar células competentes de *E. coli* DH1 y los clones recombinantes se seleccionaron en base a su resistencia a cloranfenicol y a la producción del fenotipo negro al incubarlos en medio LB. Como en el caso del clon DH1 (pAJ27), el fenotipo negro que presentaba la cepa DH1 (pAJ271) no era tan acusado como el de DH1 (pAJ221) a pesar de que, la proteína HpaB se producía en la cepa DH1 (pAJ271) a unos niveles similares a los de el clon DH1 (pAJ221) (datos no mostrados). El plásmido pAJ271 se utilizó para transformar las células competentes de *E. coli* DH1 (pAJ28). En la figura 16 se puede observar el fenotipo negro que presenta la cepa DH1 (pAJ271, pAJ28) al incubarla en medio LB comparado con las cepas DH1 (pAJ271) y DH1 (pAJ28) incubadas en las mismas condiciones. El plásmido pAJ28 sólo contiene el gen *hpaC* por lo que el incremento de la producción del pigmento negro, y por tanto, el incremento de la actividad hidroxilasa observado en la cepa DH1 (pAJ271, pAJ28) respecto a la cepa DH1 (pAJ271), sólo puede atribuirse a la presencia de la proteína HpaC. Los resultados que se muestran en la tabla 11 indican que las células portadoras del plásmido pAJ271 presentan una actividad hidroxilasa muy baja que aumenta considerablemente al añadir el extracto crudo de la cepa DH1 (pAJ28) a la mezcla de reacción. Cuando los genes *hpaB* y *hpaC* se expresan en la misma célula en *cis* (DH1 (pAJ221) ó en *trans* (DH1 (pAJ271, pAJ28)), los extractos son capaces de transformar eficientemente el fenol en catecol. Por otro lado, se observó que la oxidación de NADH dependiente de fenol aumenta considerablemente cuando ambas proteínas están presentes en la mezcla de reacción (tabla 11). Estos experimentos permitieron demostrar que la

hidroxilación de fenol (tabla 11) ó 4-HPA (figura 17) es dependiente de FAD. Este hecho es más evidente cuando los dos genes de la 4-HPA-hidroxilasa se expresan en *trans* lo que podría sugerir que la formación del complejo enzima-FAD requiere la presencia de ambas proteínas.



Figura 16. Fenotipo de las cepas productoras de los componentes de la 4-HPA-hidroxilasa. Las células se incubaron en medio LB a 37 °C durante toda la noche. Matraz 1, sobrenadante del medio de cultivo de la cepa DH1 (pAJ271, pAJ28); matraz 2, sobrenadante del medio de cultivo de la cepa DH1 (pAJ271); matraz 3, sobrenadante del medio de cultivo de la cepa DH1 (pAJ28).

Tabla 11. Actividad hidroxilasa de los extractos de *E. coli*

Cepa	Actividad (mU) ^a			
	Oxidación de NADH		Formación de HMS	
	-FAD	+FAD	-FAD	+FAD
DH1(pAJ221)	34	47	46	58
DH1(pAJ271)	1	2	1	4
DH1(pAJ28)	ND ^b	ND	ND	ND
DH1(pAJ271,pAJ28)	18	50	23	63
DH1(pAJ271) + DH1(pAJ28)	10	42	10	51

^a Actividad (mU): La actividad hidroxilasa se ensayó utilizando fenol 2 mM como sustrato. El FAD se añadió a la mezcla de reacción a una concentración final de 0,6 μ M.

^b ND : no detectado

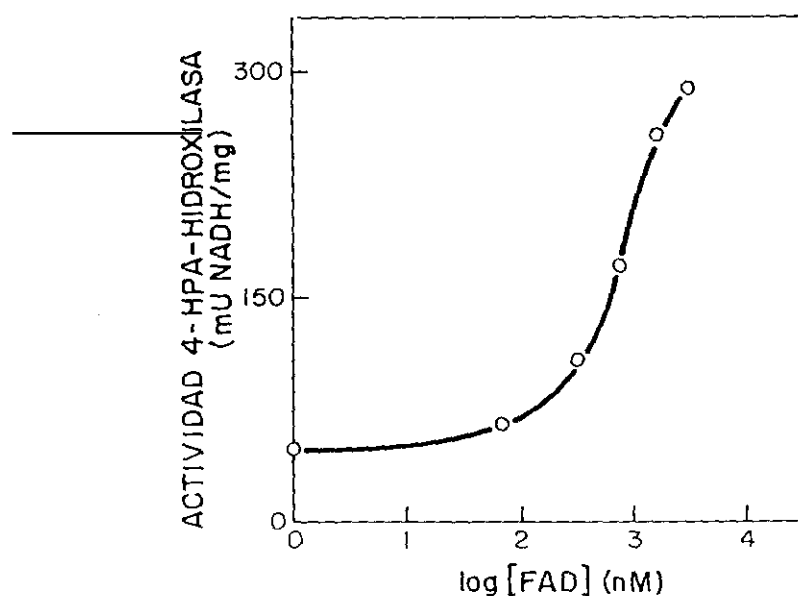


Figura 17. Dependencia de la actividad 4-HPA-hidroxilasa por el FAD. Los niveles de oxidación de NADH dependiente de 4-HPA se determinaron utilizando una mezcla 1:1 de los extractos celulares de *E. coli* DH1 (pAJ271) y *E. coli* DH1 (pAJ28).

2.6. Purificación de la hidroxilasa HpaB

Las enzimas dependientes de coenzimas derivados de la adenina interaccionan específicamente con la molécula de Cibacrón blue que presenta analogía estructural con esta base púrica (Stellwagen y Baker, 1976). Dado que la 4-HPA-hidroxilasa es una enzima dependiente de NADH y FAD cabía la posibilidad de que el Cibacrón blue pudiera actuar como ligando de adsorción de esta proteína. De este modo se podría conseguir la purificación de la enzima en un número reducido de pasos mediante cromatografía de afinidad. Para purificar la hidroxilasa HpaB se utilizó el extracto crudo de la cepa *E. coli* DH1 (pAJ221) ya que esta cepa hiperproducía la 4-HPA-hidroxilasa (figura 18A). El proceso de purificación de la enzima se llevó a cabo mediante una cromatografía de afinidad en una columna de Cibacrón blue 3GA-agarosa 3000-CL equilibrada en tampón fosfato sódico 20 mM, pH 8 y eluida con el mismo tampón conteniendo NADH 30 mM. Las fracciones que presentaban actividad hidroxilasa se concentraron con polietilenglicol 20.000 y se cargaron en una columna de filtración en gel de Superosa 12 HR. Durante todo el proceso de purificación la medida de la actividad hidroxilasa se llevó a cabo por el método acoplado con la C23O de *P. putida* dada la alta sensibilidad y reproducibilidad de este sistema.

En el análisis en geles de poliacrilamida-SDS que se muestra en la figura 18B puede observarse que la proteína HpaB queda retenida en la resina y eluye específicamente al añadir NADH al tampón de lavado. Lo mismo sucede con la cloranfenicol acetil transferasa (CAT) codificada en el plásmido pAJ221, ya que la CAT es una enzima dependiente de acetil-CoA y por tanto, susceptible de ser purificada por este tipo de resinas (Stellwagen y Baker, 1976). La proteína HpaC eluye de la columna durante el proceso de lavado lo que indica que la interacción entre los dos componentes de la 4-HPA-hidroxilasa es muy débil.

El extracto crudo presentaba una actividad específica de la hidroxilasa de 50 mU sobre el fenol. Al contrario de lo que cabría esperar, esta actividad disminuyó a lo largo del proceso de purificación obteniéndose al final un valor de actividad específica para la proteína HpaB purificada de tan solo 18 mU en presencia de FAD 0.6 μ M, pero que

aumentó hasta 200 mU al añadir el extracto crudo de la cepa DH1 (pAJ28). La actividad de la enzima purificada disminuyó gradualmente hasta inactivarse de forma irreversible a las cuatro horas de haber finalizado el proceso de purificación.

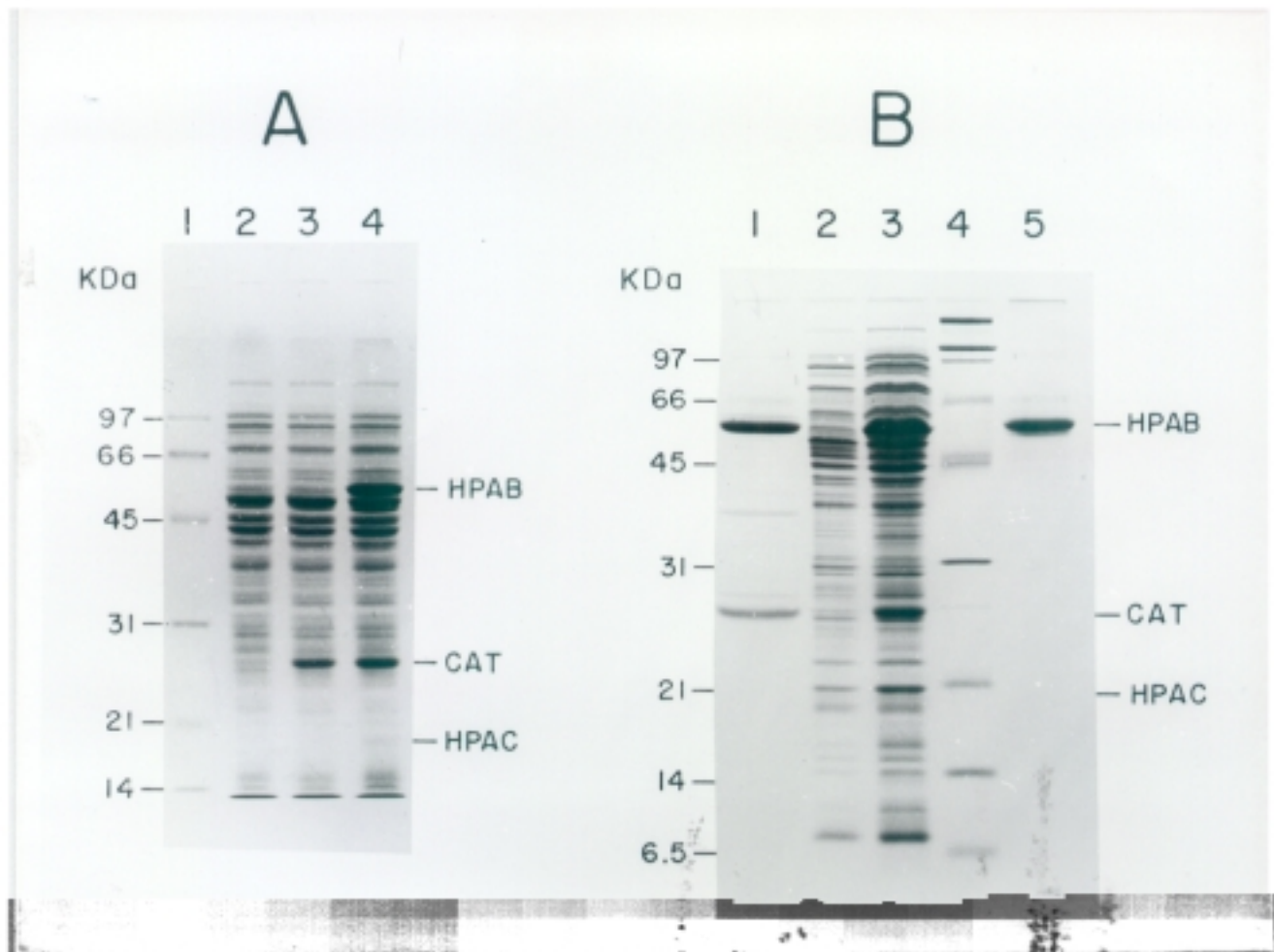


Figura 18. Electroforesis en gel de SDS-poliacrilamida de los distintos pasos de la purificación de la proteína HpaB. Panel A, extractos crudos de las cepas *E. coli* DH1 (línea 2), *E. coli* DH1 (pACYC184) (línea 3), *E. coli* DH1 (pAJ221) (línea 4). Panel B, fracción eluida con NADH de la columna de Cibacron blue (línea 1); proteínas no retenidas en la columna de Cibacron blue (línea 2); extracto crudo de *E. coli* DH1 (pAJ221) (línea 3); proteína HpaB purificada (línea 5). En la figura están indicadas las posiciones de la CAT (cloranfenicol acetil transferasa), de las proteínas HpaB y HpaC, y de los marcadores de peso molecular (panel A, línea 1 y panel B, línea 4)

En la figura 19 se muestra el cromatograma correspondiente a la elución de la enzima HpaB purificada de la columna de Superosa 12 HR con un tiempo de retención equivalente a una proteína de 120.000 Da, lo que sugiere que la proteína HpaB nativa es una enzima homodimérica.

La purificación de la hidroxilasa HpaB permitió determinar la secuencia de aminoácidos MKPEDFRASTQRPFTEEYL, correspondiente a los primeros 19 residuos del extremo N-terminal. Esta secuencia coincide exactamente con la del N-terminal de la proteína HpaB deducida a partir de la secuencia de nucleótidos.

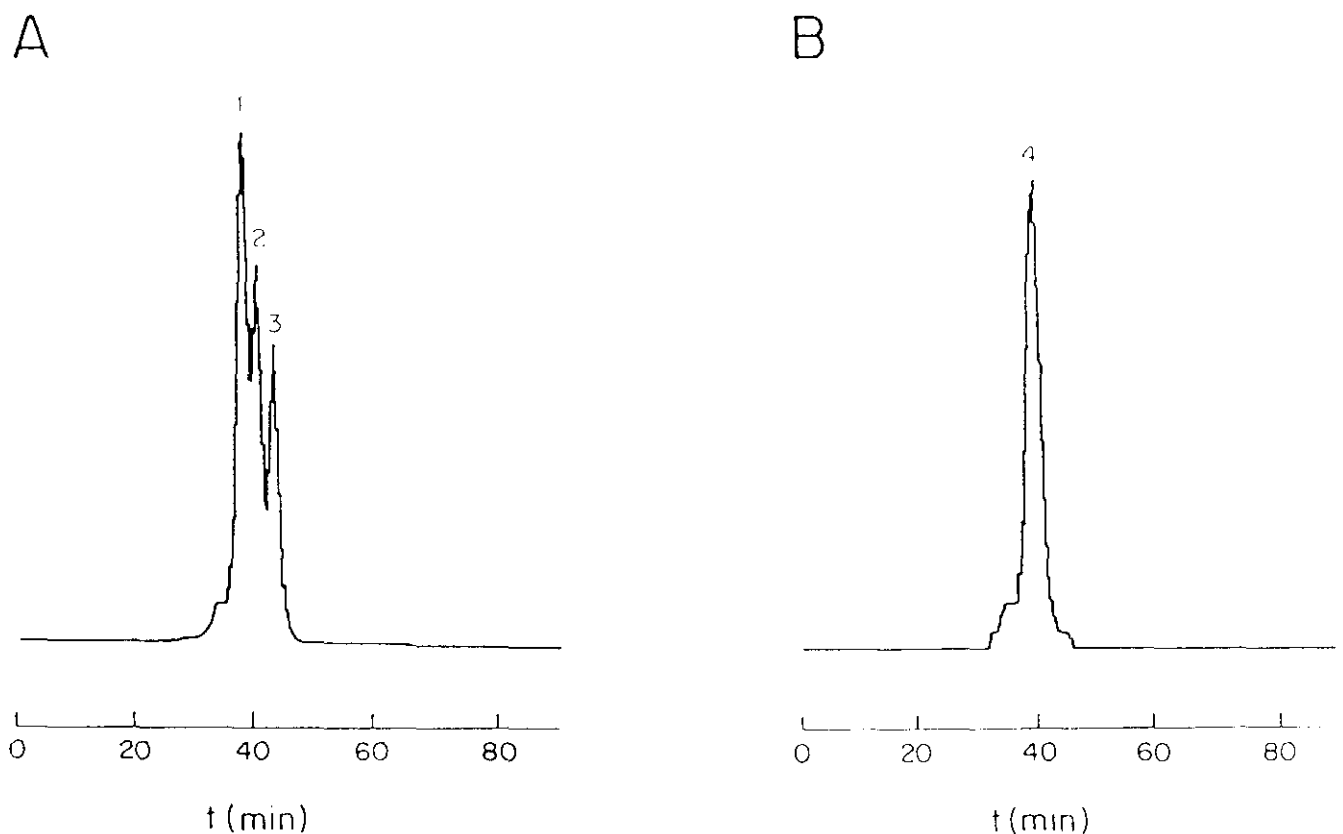


Figura 19. Determinación de la masa molecular de la proteína HpaB mediante filtración en HPLC. Panel A: 1, aldolasa (160 kDa); 2, seroalbúmina (67 kDa); 3, ovoalbúmina (45 kDa). Panel B: proteína HpaB (120 kDa).

2.7. Presencia de genes homólogos a *hpaB* en otros microorganismos

Teniendo en cuenta que el 4-HPA y 3-HPA eran buenos sustratos de la hidroxilasa (tabla 9) y que había sido descrita la presencia de una hidroxilasa de 3-HPA y 4-HPA en *E. coli* C (Cooper y Skinner, 1980), cabía pensar que el operón *hpaBC* formara parte de la ruta de degradación de estos compuestos en *E. coli* ATCC 11105 y, por tanto, otras estirpes capaces de mineralizar estos compuestos podrían tener genes homólogos a *hpaB*. Utilizando como sonda el fragmento *HindIII-EcoRV* de 1,6 kb del plásmido pAJ27 se investigó mediante análisis por Southern blot la presencia del gen *hpaB* en otras estirpes (figura 20). Estos análisis indicaron que las estirpes B, C y W de *E. coli*, así como otros microorganismos capaces de degradar el 4-HPA como *K. pneumoniae* y *K. citrophila*, contienen genes similares a *hpaB*. Los genomas de dos estirpes diferentes de *E. coli* K12, DH1 y W3110, no contienen ningún fragmento de DNA homólogo. Esta sonda fue utilizada para clonar la 4-HPA-hidroxilasa de *E. coli* C mediante la construcción de una genoteca *HindIII* en pBR322. De esta forma se aisló el clon DH1 (pHC1) que contenía un fragmento *HindIII* de 6 kb y presentaba un fenotipo negro similar al de los clones que expresan la 4-HPA-hidroxilasa de *E. coli* W (ATCC 11105). La secuenciación de los extremos de este fragmento nos permitió deducir la secuencia de los primeros 45 aminoácidos de la hidroxilasa de la estirpe C, que era idéntica a la secuencia N-terminal de la 4-HPA-hidroxilasa de la estirpe W, *E. coli* ATCC 11105.

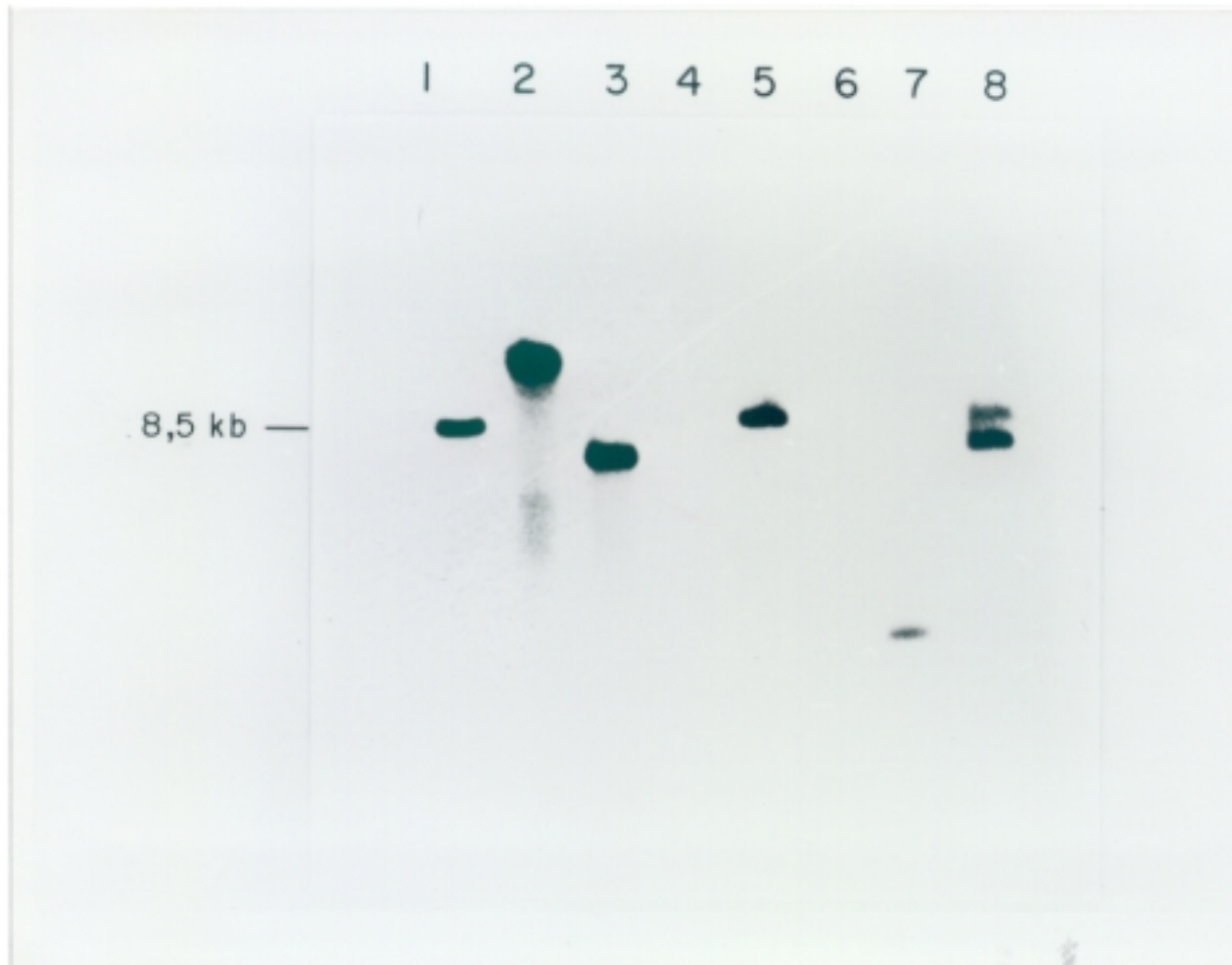


Figura 20. Análisis por Southern blot de la presencia del gen *hpab* en otras bacterias. Líneas 1 a 7: DNA cromosómico de *E. coli* ATCC 11105, *E. coli* B/rK, *E. coli* C, *E. coli* DH1, *E. coli* W, *E. coli* W3110 y *K. pneumoniae* M5a1 digerido con *Hind*III. Línea 8, DNA cromosómico de *K. citrophila* digerido con *Eco*RI. En la figura está indicado el tamaño de la banda correspondiente al fragmento *Hind*III clonado en el plásmido pAJ19.

3. CLONACIÓN Y SECUENCIACIÓN DEL OPERÓN *meta* DE LA RUTA DEL 4-HPA

3.1. Clonación del operón *meta* y construcción del mapa físico de la ruta del 4-HPA

K. pneumoniae M5a1 es un microorganismo capaz de utilizar 3-y 4-HPA como única fuente de carbono. En el apartado 2.7, se comprobó que existen secuencias homólogas al gen *hpaB* en el DNA de esta bacteria y por otra parte, se había descrito que otros genes de la ruta de degradación del 4-HPA de *K. pneumoniae* son homólogos a los genes equivalentes en *E. coli* W (Fawcett y cols., 1989). Por todo ello, se analizó mediante Southern blot la posibilidad de utilizar como sonda el gen de la HPC-dioxigenasa de *K. pneumoniae* M5a1 para clonar el operón *meta* de la ruta de degradación del 4-HPA de *E. coli* ATCC 11105. En la figura 21 se muestra la presencia de genes similares al de la HPC-dioxigenasa de *K. pneumoniae* en las bacterias capaces de mineralizar el 4-HPA como *K. citrophila* y las estirpes B, C y W de *E. coli*. Posteriormente se construyeron diferentes genotecas de *E. coli* ATCC 11105 en pUC18 utilizando la misma sonda para la selección de células recombinantes portadoras de genes homólogos a la HPC-dioxigenasa. Mediante este procedimiento se aislaron dos plásmidos solapantes, pHCB1 y pHCB3, que contenían un fragmento *Bam*HI de 6 kb y un fragmento *Eco*RI de 11 kb, respectivamente (figura 22). Al incubar las células que contenían el plásmido pHCB3 en medio LB se observó que producían un pigmento negro similar al observado en las cepas que contenían el plásmido pAJ19, lo que indicaba que el fragmento *Eco*RI clonado en el plásmido pHCB3 podía contener el operón *hpaBC*. Esto se confirmó posteriormente mediante Southern blot y análisis bioquímico. El fragmento *Bam*HI-*Eco*RI de 2.7 kb del plásmido pHCB1 se utilizó como sonda para seleccionar el plásmido pHCR1 a partir de una genoteca *Eco*RI (tabla 10, figura 22). El análisis de restricción de estos plásmidos permitió determinar el mapa físico de la ruta completa del 4-HPA (figura 22).

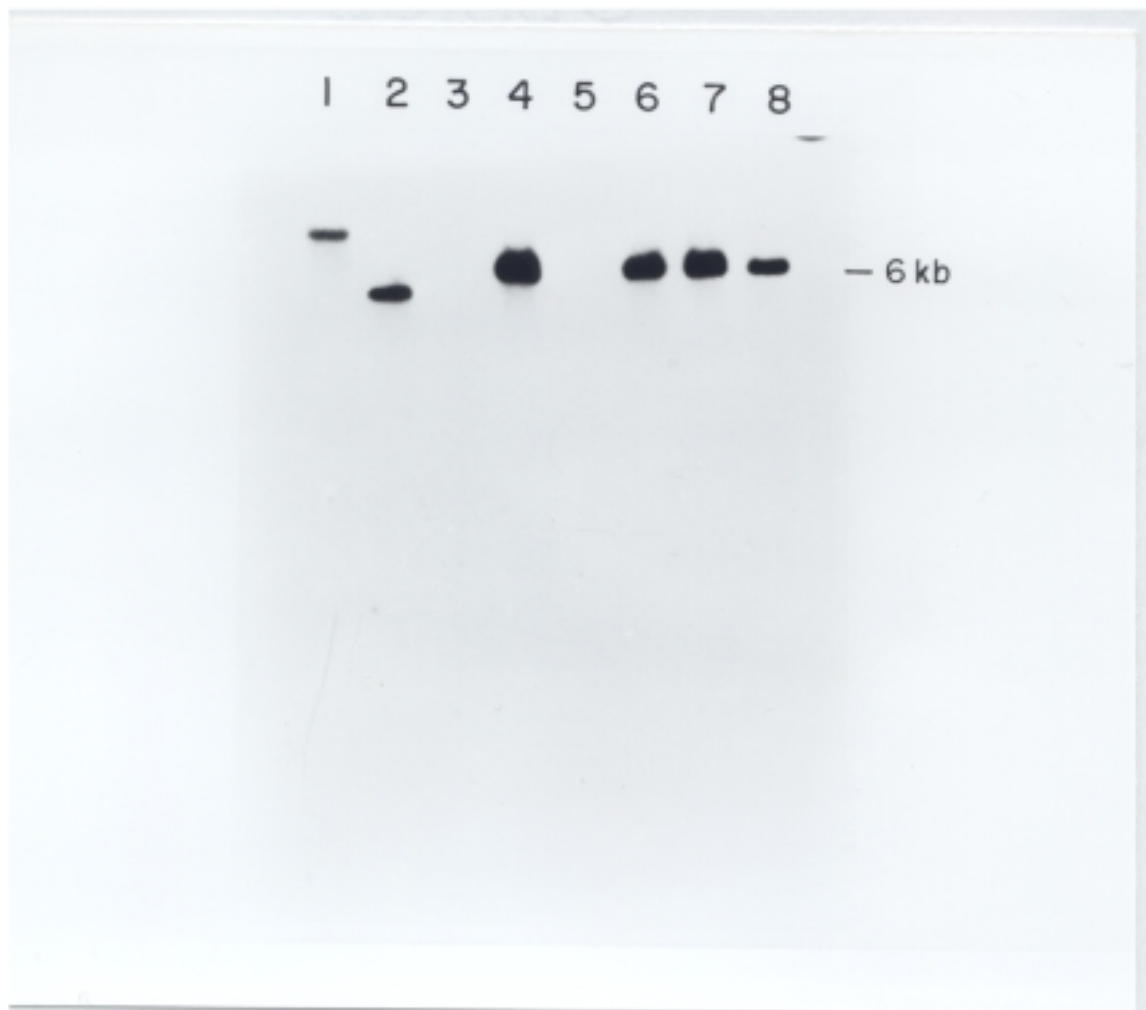


Figura 21. Análisis por Southern blot de la presencia de genes homólogos a la HPC-dioxigenasa de *K. pneumoniae* M5a1 en otras bacterias. La sonda utilizada fue un fragmento *Bbu*I de 2 kb del plásmido pAG464. Líneas 1 a 8, DNA cromosómico de *K. citrophila*, *K. pneumoniae* M5a1, *E. coli* W3110, *E. coli* W, *E. coli* DH1, *E. coli* C, *E. coli* B/rK, *E. coli* W ATCC 11105 digerido con *Bam*HI. En la figura está indicado el tamaño de la banda correspondiente al fragmento *Bam*HI clonado en el plásmido pHCB1.

3.2. Secuenciación del *cluster* de genes que componen la ruta del 4-HPA de *E. coli* ATCC 11105

La secuenciación del grupo de genes que componen la ruta de degradación del 4-HPA y de parte de las regiones flanqueantes, fue llevada a cabo mediante secuenciación manual y automática. Para ello se construyeron los plásmidos detallados en la figura 22 y en la tabla 10, y se utilizaron los oligonucleótidos sintéticos marcados en la figura 23. El análisis de los marcos de lectura abiertos (figuras 22 y 23) y de la comparación de secuencias (figuras 15, 25-35), sugirió que esta ruta estaba compuesta por once genes, ocho de los cuales estaban organizados en dos operones; el operón de la 4-HPA-hidroxilasa u operón *hpaBC*, y el operón *hpaGEDFHI* u operón *meta*. Además existen dos genes reguladores *hpaR* y *hpaA*, y el gen *hpaX*, de función desconocida, que parece formar parte de la misma unidad de transcripción que *hpaA*. Paralelamente al desarrollo de esta tesis, el grupo del Dr. Cooper ha secuenciado y caracterizado bioquímicamente las enzimas pertenecientes al operón *hpcECBDGH* u operón *meta* de la estirpe C de *E. coli* (apartado 5.1.1 de la Introducción). Dada la homología entre las proteínas que componen los operones *meta* de ambas estirpes, la función de las enzimas codificadas por el operón *hpaGEDFHI* fue atribuida de acuerdo a lo establecido por estos autores para la estirpe C de *E. coli*. En la figura 24 se detallan los pasos enzimáticos propuestos para la conversión de 4-HPA en piruvato y succinato semialdehído.

El primer gen del operón *meta* es el gen *hpaG* (nucleótidos 851-2.138 de la figura 23) que codifica para una proteína de 46.927 Da y es homólogo al gen *hpcE* de *E. coli* C (Roper y Cooper, 1993) que codifica para una enzima bifuncional que cataliza la descarboxilación del ácido 5-oxo-pent-3-en-1,2,5-tricarboxílico (OPET) y la isomerización del producto de la reacción de descarboxilación, el ácido 2-hidroxi-hept-2,4-dien-1,7-dioico (HHDD) (figura 24). La diferencia más notable entre la proteína HpaG y la proteína equivalente HpcE es que la primera tiene 24 aminoácidos más en su extremo C-terminal debido a una terminación prematura de la transcripción del gen *hpcE*, provocada por una delección de 7 pb en su extremo 3'. Se ha propuesto que HpcE podría proceder de una duplicación génica ya que el N-terminal de esta proteína es muy similar a su C-terminal (Roper y Cooper, 1993). En este sentido, el alineamiento entre el

N- y C-terminal de HpaG es más preciso debido a la contribución de los 24 aminoácidos adicionales (figura 25). También se había descrito que HpcE se parecía a la 4-oxalocrotonato descarboxilasa codificada por el gen *dmpH* de la ruta de degradación de fenol de *Pseudomonas* CF600 (Shingler y cols., 1992) y a una proteína de función desconocida que parecía formar parte de la ruta de degradación de catecol de *Alcaligenes eutrophus* (Kabisch y Fortnagel, 1990). Por otra parte, el producto del gen *hpaH* (nucleótidos 4.957-5.758 de la figura 23) también se parece a la descarboxilasa DmpH de *Pseudomonas* CF600 (figura 25.). Este gen codifica para una proteína de 29.714 Da homóloga a la hidratasa del ácido 2-oxo-hept-3-en-1,7-dioico (OHED) o proteína HpcG (Roper y cols., 1995) (figura 24). En la figura 25 se muestra la comparación de la secuencia de aminoácidos de las proteínas HpaG y HpaH con otras descarboxilasas e hidratasas de diferentes microorganismos e incluso de organismos eucariontes. Por último, hay que destacar que no se ha encontrado ninguna isomerasa que presente similitud con HpaG.

El gen *hpaE* (nucleótidos 2.137-3.601) codifica para una proteína de 53.011 Da perteneciente a la superfamilia de las aldehído deshidrogenasas (Horn y cols., 1991). La deshidrogenasa HpaE es homóloga a la CHMS-deshidrogenasa codificada por el gen *hpcC* de *E. coli* C (Roper y cols., 1995). En la figura 26 se muestra el alineamiento de estas proteínas con la deshidrogenasa DmpC de la ruta de degradación de fenol de *Pseudomonas* CF600 (Shingler y cols., 1992). Como en el caso del gen *hpcE*, una deleción en el extremo 3' de *hpcC* provoca la terminación prematura de la transcripción dando lugar a una proteína con 18 aminoácidos menos que la enzima HpaE.

El gen *hpaD* (nucleótidos 3.605-4.454 de la figura 23) codifica para una proteína de 32.018 Da homóloga a la HPC 2,3-dioxigenasa de *E. coli* C codificada por el gen *hpcB* (figura 27) (Roper y Cooper, 1990a). La diferencia más importante entre estas dos proteínas fue observada en el C-terminal donde una inserción de 2 pb en el gen *hpcB* provoca un cambio en la fase de lectura provocando una pérdida de similitud en los últimos 23 residuos de HpaD. Estas proteínas no presentan ninguna similitud con otras dioxigenasas secuenciadas lo que sugiere que puedan formar parte de una nueva familia de dioxigenasas .

La proteína codificada por el gen *hpaF* (nucleótidos 4.466-4.844) es idéntica a la 5-carboximetil-2-hidroximuconato-isomerasa (CHM-isomerasa) de *E. coli* C (figura 28) (Roper y Cooper, 1990b). Como en el caso de la HPC-dioxigenasa, esta enzima no se parece a ninguna proteína secuenciada hasta el momento.

Según el trabajo de Roper y cols. (1993), la última enzima del operón *meta* es la 2,4-dihidroxi-hept-2-en-1,7-dioato-aldolasa (HHED-aldolasa) codificada por el gen *hpcH* (GenBank/EMBL ac. Z47799). El gen equivalente *hpaI* (nucleótidos 5771-6.557) codifica para una proteína de 28.072 Da idéntica a HpcH. Además, HpaI es similar a una ORF incompleta codificada por un gen de función desconocida localizado en la región 5' de dos ORFs relacionadas con el metabolismo del gluconato en *E. coli* (Komine e Inokuchi, 1991). El grado de identidad entre HpaI y esta proteína es del 44% (figura 29).

El gen *hpaR* (nucleótidos 577-133 de la figura 23) se transcribe en sentido opuesto a los operones *hpaBC* y *hpaGEDFHI* (figura 24). Codifica para una proteína de 17.235 Da idéntica a la proteína reguladora HpcR (figura 30) que actúa como represor del operón *meta* de *E. coli* C (Roper y cols., 1993). Estas proteínas no se parecen a ninguna proteína reguladora descrita hasta el momento. El gen *hpaA* (nucleótidos 8.120-9.005) codifica para una proteína de 34.129 Da que conserva la secuencia consenso presente en los miembros de la familia de reguladores XylS/AraC (Gallegos y cols., 1993) (figura 31). Este gen parece formar parte de la misma unidad de transcripción que el gen *hpaX* (nucleótidos 6.734-8.108) ya que entre el codon de terminación de éste y el codon de iniciación del gen *hpaA* sólo hay 9 nucleótidos. La proteína HpaX tiene una masa molecular de 50.568 Da y presenta similitud con la proteína codificada por el gen *phtI* (34 % de identidad) de la ruta de degradación del ftalato en *P. putida* (Nomura y cols., 1992) (figura 32). La función de la proteína PhtI no está aún determinada, pero se ha sugerido que podría actuar como un regulador positivo de la expresión de los genes *pht* ó como un transportador de ftalato (Nomura y cols., 1992) ya que se parece a la proteína transportadora de glicerol-3-fosfato (GlpT) de *E. coli* (Eiglmeier y cols., 1987) perteneciente al grupo 4 de la superfamilia de transportadores transmembranales (MFS) (Marger y Saier, 1993). La comparación de la secuencia de aminoácidos de HpaX con la

de los miembros de este grupo como GlpT (49 % de similitud y 17 % de identidad), la proteína transportadora de fosfoglicerato de *S. typhimurium* (52 % de similitud y 21 % de identidad), la proteína transportadora de hexosa-fosfato UhpT de *E. coli* (46 % de similitud y 20 % de identidad) y su proteína reguladora UhpC (50 % de similitud y 22 % de identidad), sugirió que HpaX podía ser también un miembro de esta superfamilia. Todos los miembros de la MFS tienen aproximadamente 400 aa y presentan un motivo estructural común de 12 α -hélices transmembranales. El perfil hidrofóbico de la proteína HpaX (figura 33) es similar al que presentan los miembros de la MFS (Marger y Saier, 1993).

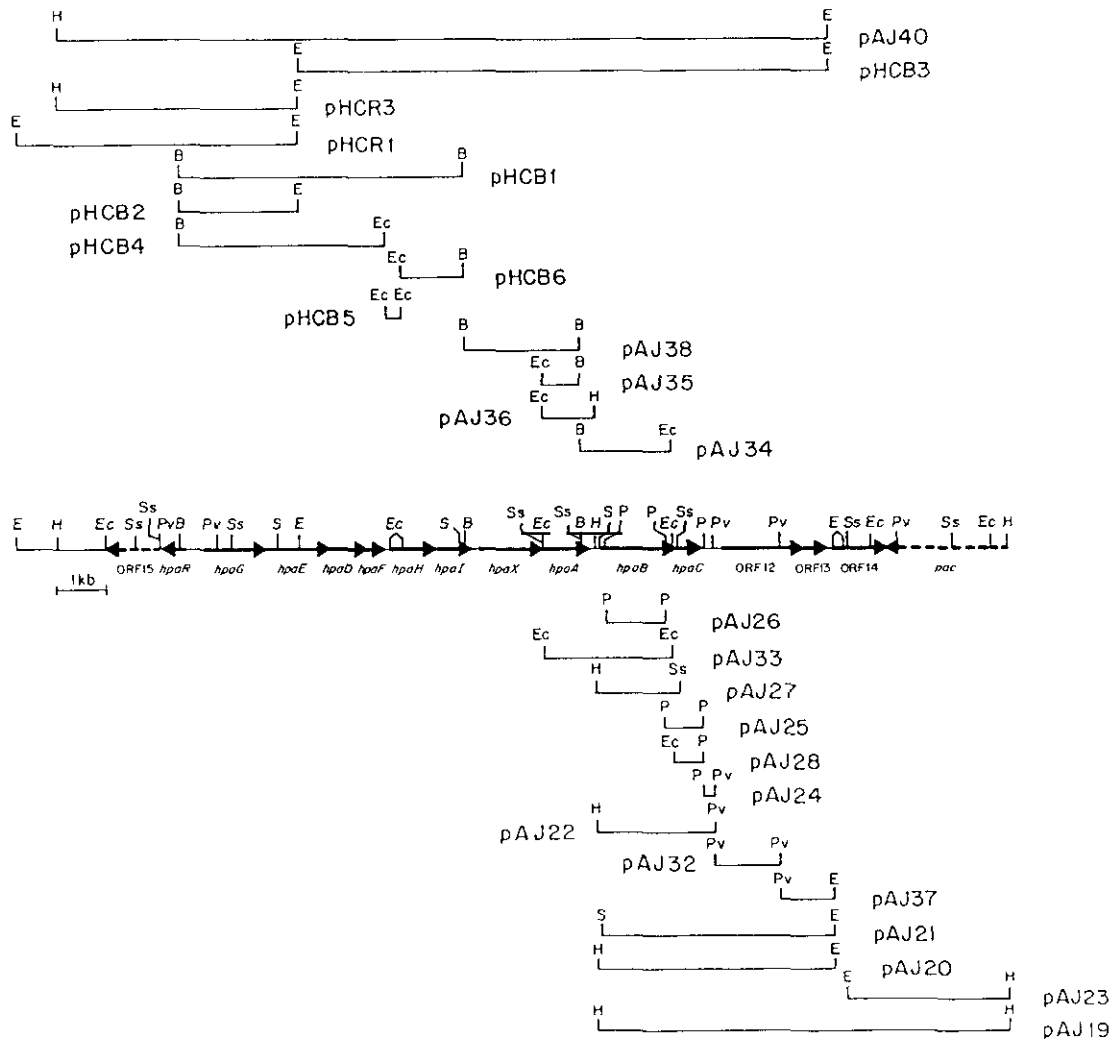


Figura 22. Organización física y mapa de restricción del cluster de genes de la ruta catabólica del 4-HPA y de las regiones flanqueantes. En la parte central de la figura están indicados los genes localizados en la región secuenciada y los sitios de restricción más relevantes. En la parte superior e inferior se indican los fragmentos contenidos en los plásmidos utilizados para la secuenciación de esta región indicando, únicamente, los sitios de restricción utilizados para la construcción de los mismos (tabla 10). Las flechas en línea continua indican el sentido de la transcripción de los genes y las flechas en línea discontinua indican las regiones secuenciadas parcialmente en este trabajo. Abreviaturas: B, *Bam*HI; E, *Eco*RI; Ec, *Eco*RV; H, *Hind*III; P, *Pst*I; Pv, *Pvu*II; S, *Sal*I; Ss, *Ssp*I.

ACGTTTTAACATGGTTTCTCTTCTCTGACCGATGCAAGATTTTCGGCCTGGAAGCA 60
 1 TGCAAAATGTACCAAGAGAAAAGACACCTGGCTACGTCTCAAAGCGCGACCTTCGT
 R K L M ← ORF15

AAAACCTTATGAAAAGGTGTAACCCGCCATCCATGGAAAAACAGGCGCTGTTATCAGA 120
 61 TTTTGAATACTTTTCCACATTGGCGGTAGTACCTTTTGGTCCGCGACCAATAGTCT

TACTAAAAAGTTATTCATTCATCCCGGGATGTCTTCTGTGGGAATTACCCAGAGCAA 180
 121 ATGATTTTCAATAAGTAATAGCGGGCCCTACAGAAGAACGCCCTTATGGTCTCGTT
 * E N D G P I D E Q R S N G L A

TAAATTCCTCTAACAGATGGGTTAACTGTTGCATTTTCTCGGCAGTAAACTGAGCTTCAA 240
 181 ATTTAAGGAGATTGTCTACCCAAATGACACAGTAAAAGAGCGCTTATGACTCGAAGTT
 I F E E L L H T L Q Q M K E A T F Q A E

TCGCGGATAGGCTTCTTCAATCTGTGTCTGAGCGGGTTATACAGCGCCTGCCCTCTT 300
 241 AGACGGGTATCCGAAGAAGTAGACACAGACTCGCGCAATATGTCGGGACGGGGAGAA
 I Q R Y A E E I Q T Q A R N Y L A Q G E

TGGTCAGCGAGATATAGAGCTTCCGCTGATCATTGATCGGCTTAAATCGCAACTAAAC 360
 301 ACCAGTCGCTCTATATCTCGAAGGCGACTAGTAACAGCGAAATAGCGTTGTGATTG
 K T L S I Y L K R Q D N I P K L R L V L

CGTCGCGTCCATCCGCGTAAGGATCCCGTCCAGGCTTGGCGCAAAATGACGGCGGAT 420
 361 GCAGCGCAAGGTAGCGGCATCTCTAGGCGAGTCCGAACCCCGTTTACCTCCCGCTA
 G D R E M R T L I G G T L S P R L I C A R

ACGCCAGATCGTGAATAATCCATGATGGACTTTCGCGCAGAAATACGCAGATGCGCACT 480
 421 TGGCGTCTAGCACTTTAGTAACTACCTGAAAGCGGCTTATGCGTCTACCGGGTGA
 Y A L D H F D M S P S E A L I R V I R W

GTTGCTCAGTCAAAATATGCGCTTAAACGATGGGCGAAATAACTCATCGCGCTTCTC 540
 481 CAACGAGTCAGTTTAAATACGGCAAATGCTAACCGCTTTTATGAGTAGGGCGAAGAG
 Q Q E T L N H R K V I P R F Y S M A A E

CGCGCTGGAGCAACGCAATGGTGTAGTGGTGGTATCTTTCCCTGAAAATTCAA 600
 541 CGCGCACTCGTTCGCTTACCAATCACTCAGCACGTAATAGAAAGGGGACTTTTAAAGTT
 R A Q L L A I T L S D H M ← hpaR

AATCATTAAATATAGAAACAGTAAATCATCAGTTACCGGACTCGGCTCTCCCTGCAATT 660
 601 TTAGTAATTATATCTTTGTCAATTAGTAGTCAATGGCCCTGACCGACAGGGACGTTAA

CCGATTTTATCAAAATGAGCATGTTTTATTAACCTCACTGATATTTAATATATCTTAT 720
 661 GGACTAAAATAGTTTTAACTCGTACAAAATAATGAGTACTATAAATATATAAGAATA

CITTAATGTAATAGTTTGTAAATAGATCACATTACATCACTTACTTTTCCAAACT 780
 721 GAAATTTACATTTATCAACAATTAATCTAGTGTAAATGAGTAAATGAAACGGTTTGA
 CAP -35

ACGCAATCGGGAATAAAGTAATCATTAAACATATTAATGATTAAAGATGATCGAATCCGAG 840
 781 TGGCTAGCCCTTATTTTTCATTAGTAATGTATAAATTAATACTACTAGCTTAGGCTC
 -10

GAGTGGTACCATGAAAGGCACTATCTTCGCGTAGCGTTGAAACCATCGCAGCCAATTGA 900
 841 CTCACCATGTAATTTCCGTGATAGAAGCGCATCGCAACTTGGTAGCGTGGTTGAACT
 hpaG → M K G T I F A V A L N H R S Q L D

TGCATGGCAGGAAGCGTCCAGCAATCCCGCTACAAAGCCCCGCTAAAAGTCGGTCTG 960
 901 ACGTACCGTCCITTCGCAAGGTGTTAGGGGGATGTTTCGGGGCGGATTTGACGCCAGAC
 A W Q E A F Q Q S P Y K A P P K T A V W

GTTTATTAACCGCGCAATACGGTATTGGTTGGGTTGAAACCGATTCCCTTTCCACAGGG 1020
 961 CAATAATTTGGCGGTTATGCCACTAACCAAGCGCACTTGGCTAAGGGAAGGTTGCC
 F I K P R N T V I V I G C G E T P F P G

TGAAAAGGTACTGAGCGGTGCGACTGTTGCGCTGATTTGGGGAAAACAGCGACGAAAGT 1080
 1021 ACTTTTCCATGACTCGCCAGCGTGAACCGACTAACACCCCTTTTGTGCTGCTTTCA
 E K V L S G A T V A L I V G K T A T K V

ACGTTGAAGAAGATGCGGCAGAGTACATCGCCGATATCGCGTGGCTAACGATGTCAGCCT 1140
 1081 TGCACCTTCTTCTACCGCGTCTCATGTAGCGGCTATACCGGACCGATTGCTACAGTCGGA
 R E E D A A E Y I A G Y A L A N D V S L

GCCGGAAGAGAGCTTTTACCGCCCGCAATCAAGCAAAATGTCGTGATGGATTTCGCC 1200
 1141 CGGCCCTTCTCGAAAATGGCGGGCGTTAGTTTCGTTTACAGCACTACCTAAGACGGG
 P E E S F Y R P A I K A K C R D G F C P

CATTGGCGAAACCGTGGCTCTCAGCAATGTCGATAATCTGACCATCTATACCGAGATCAA 1260
 1201 GTAACCGCTTTGGCAGCGAGAGTGTACAGCTATTAGACTGGTAGATATGGCTTAGTT
 I G E T V A L S N V D N L T I Y T E I N

CGGGGCTCTCGCATCTGGAACACCCCGATTTACAACGTAACGCCGACAGCTGCT 1320
 1261 GCCCGCAGGACGGTAGTGACCTTGTGGCGGTAATGTTGCATTGCGGGCTGCGAGA
 G R P A D H W N T A D L Q R N A A Q L L

GAGCGCCCTGAGCGAATTTGCCACGCTGAACCCAGGCGATGCCATTCTGCTCGGCAGCC 1380
 1321 CTCGCGGACTCGCTTAAACGGTGGCACTTGGGTCCGCTAACGAGCAGCGCTGGGG
 S A L S E F A T L N P G D A I L L G T P

ACAGCGCGCGTGAATAACAGCCAGCGCATCGCTGCGTGTCTCGCAGAAGTTTCCC 1440
 1381 TGTCCGCGCGCACCTTTATGTCGGTCCGCTAGCGCACGCACAGAGCGTCTTCAAAGGG
 Q A R V E I Q P G D R V R V L A E G F P

GCCGCTGGAATAATCCGCTAGTGGCAGCAAGTGAAGTGACCCAGCGCAAGAGCTTCCCAAC 1500
 1441 CGGGACCTTTTAGGCCATCACCTGCTGCACCTTCTGCTGCGGCTTCTCGAAGGGTIG
 P L E N P V V D E R E V T T R K S F P T

GCTGCCACACCGCAGCGTACGCTGTTTGCCTCGGCTGAACTACGCCAGCCAGCCAG 1560
 1501 CGACGGTGGCGGTGCGATGCGCAAAACGGGACCGGACTTGTAGCGGGTGGTGGGGT
 L P H P H G T L F A L G L N Y A D H A S

CGAACTGGAATTTAAGCCACCGGAAGAGCGCGTGGTGTCTCGAAGCGCGCAATACCT 1620
 1561 GCTTGACCTTAAATTCGGTGGCTTCTCGGCACCAAGGACTTTCGCGCTTATGGGA
 E L E F K P P E E P L V F L K A P N T L

CACTGGCGATAACCGACCTCCGTCGCGCGCAACAATATTGAATACATGCACTACGAAGC 1680
 1621 GTGACCGCTATGGTCTGGAGGACCGCGGCTTGTATAACTTATGATACGTGATGCTTCC
 T G D N Q T S V R P N N I E Y M H Y E A

GGAGCTGGTGTAGTTATGGCAAGCAGGGCGTAACTGACGGAAGCGGATGCCATGGA 1740
 1681 CCTCGACCCATCAATAACCGTTCGTCGCGCATGTCAGTCCGCTTCCGCTACGGTACCT
 E L V V V I G K Q A R N V S E A D A M D

TTATGTCGCGGCTACACCGTGTGTAACGACTACGCCATTCCGCGACTATCGGAAAATA 1800
 1741 AATACAGCGCCGATGTCGACACATGCTGATGCGGTAAGCGCTGATAGACCTTTTGTAT
 Y V A G Y T V C N D Y A I R D Y L E N Y

CTACCGCCCTAACCTGCGGTAATAAAGCGCGACCGGACTGACCGCGATGCTTCAACCAT 1860
 1801 GATGGCGGATGGACGCCATTTTTCGGCGTGGCTGACTGCGGCTACGAAAGTTGGTA
 Y R P N L R V K S R D G L T P M L S T I

CGTGGCAAGAGGGGATCCCGGACCGCATAATCTGACCTTCCGACCTTCTGCAACGG 1920
 1861 GCAOGGCTTCTCGCTAGGGCTGGCGTATTAGACTGGAAGCGTGAAGCAGTGTGCC
 V P K E A I P D P H N L T L R T F V N G

CGAGTTAGCCAGCAAGGACCCAGCGCATCTGATCTTACGCGTCCCTTCTGATCGC 1980
 1921 GCTCAATGCGGTCGTTCCGTGGTGGCGCTAGACTAGAAGTCGACGGGAAGGACTAGCG
 E L R Q Q G T T A D L I F S V P F L I A

CTACTTAAGCAATTTATGACCTGAATCGGGGACATGATCGCCACCGGCACACCAAAA 2040
 1981 GATGAATTCGTTAAATCTGGACTTAGGCCCGCTGACTAGCGGTGGCGTGTGGTT
 Y L S E F M T L N P G D M I A T G T P K

AGGCTTATCTGACGTGGTGGCGATGAAGTGGTGGTGGAAAGTAGAAGCGTGGGGCG 2100
 2041 TCCGAATAGACTGCACCGGACGCTACTTACCACCACTTCACTTCCGCAACCGCGC
 G L S D V V P G D E V V V E V E G V G R

CCTGGTGAACCAATTTGAGTGGGAAACAGCGAAATGAAAAAGTAAATCATTGGATC 2160
 2101 GGACCACTTGGCTTAACTACTCTCTTGTGCTTTACTTTTTTCAATTAGTAACCTAG
 L V N R I V S E E T A K *
 hpaE → M K K V N H W I

161 AACGGTAAAAATGTTGCAGGTAACGACTACTTCCAGACCACCAATCGGCCAACGGGTGAA 2220
 TTGCCATTTTACAACGTCATGCTGATGAAGGTCTGGTGGTTAGGCGGTGCCCACTT
 N G K N V A G N D Y F Q T T N P A T G E

221 GTGCTGGCGGATGTGGCTCTGGCGGTGAAGCGGAGATCAATCAGGCGGTAGCGGACGG 2280
 CACGACCGCTACACCGGAGACCGCACTTGGCTCTAGTTAGTCCGCATCGCGGTGCG
 V L A D V A S G G E A E I N Q A V A A A

281 AAAGAGGCGTCCCGAAATGGGCAATCTGCCGATGAAAGAGCGTCCGCCGCTGATCGCG 2340
 TTTCTCCGCAAGGGCTTTACCGGTTAGACGGTACTTTCTCGCACGGCGGACTACGCG
 K E A F P K W A N L P M K E R A R L M R

341 CGTCTGGCGATGTGATCGACCAGAACGTGCCAGAGATCGCCGCGATGAAACCGCGGAC 2400
 GCAGACCGCTAGACTAGTGGTCTTGCAAGGCTCTAGCGCGGCTACCTTTGGCGCGT
 R L G D L I D Q N V P E I A A M E T A D

401 ACCGCGCTGCCGATCCATCAGACCAAAAATGTGTGATCCACGCGCTTCCCACAACATTT 2460
 TGGCGGACCGCTAGGTAGTCTGGTTTTTACACAACAGGTTGGCGGAAGGGTGTGAAA
 T G L P I H Q T K N V L I P R A S H N F

461 GAATTTTTCGCGGAAGTCTGCCAGCAGATGAACGGCAAGACCTATCCGGTTGTGCAAG 2520
 CTTAAAAAGCGCTTTCAGACGGTCTCTACTTGGCGTCTGGATAGGCCAACAGCTGTTC
 E F F A E V C Q Q M N G K T Y P V V D K

521 ATGCTCAACTACACGCTGGTGACCGCTGGCGGTTTTGTGCGTGGTATCGCCGTGGAAC 2580
 TAGGAGTTGATGTGGACACCGTCCGCGCACCGCAACACCGCACCATAGCGGCACTTG
 M L N Y T L V Q P V G V C A L V S P W N

581 GTACCGTTTTATGACCGCCACATGGAAGGTCGCGCGGTGTCTGGCGTGGGCAATACCGGG 2640
 CATGGCAAACTACTGGCGGTGTACTTCCAGCGGACACAGACCGCGACCGGTTATGGCG
 V P F M T A T W K V A P C L A L G N T A

641 GTACTGAAAATGTCGGAACCTCCCGCTGACCGCTGACCGCTGGGTGAGCTGGCGCTG 2700
 CATGACTTTTACAGCCTTGAGAGGGCGGACTGGCGAGCTGGCGGACCCACTCGACCGGAC
 V L K M S E L S P L T A D R L G E L A L

701 GAAGCGGTATTCCGCGAGCGTCTGCAAGCTGGTACAGGGTACCGCGCAACCGCAGGG 2760
 CTTCCGCCAATAGGCGGTCCGCGACACTGACAGTGTCCGATGCCCGTGTGGCGTCCC
 E A G I P A G V L N V Q G Y G A T A G

761 GATGCGCTGGTTCGTATCATGACGTACGTCGCCGTGTCTGTCACCGCGGTACGGCCACC 2820
 CTACGCGACCAAGCAGTAGTACTGATGCAACCGCACAGCAAGTGGCGCCATGCGCGTGG
 D A L V R H H D V R A V S F T G G T A T

821 GGGCGCAACATCATGAAAAACCGCGGGCTGAAAAAATACTCCATGGAACGGGCGGTAAA 2880
 CCGCGTGTAGTACTTTTTGGCGCCGACTTTTTTATGAGTACCTTGACCCGCCATTT
 G R N I M K N A G L K K Y S M E L G G K

881 TCGCGGTGCTGATTTTTGAAGATCCGATATTGAACCGCGCTGGACCGCGCTGTT 2940
 AGCGGCCAGACTAAAACTTCTACGGCTATAACTTCCGCGGACCTGCGGCGGACAAG
 S P V L I F E D A D I E R A L D A A L F

941 ACCATCTTCTCGATCAACGGGGAACGCTGCACCGCGGTTCCGCGATCTTTATTCAGCAA 3000
 TGGTAGAAGAGCTAGTTGCCGCTTGGCAGTGGCGGCAAGCGGTGAGAATAAGTCGTT
 T I F S I N G E R C T A G S R I F I Q Q

1001 AGCATCTACCGGAATTCGTTAAACGCTTTCGCCAACCGCGCAACCGCTCTGCGCGTGGC 3060
 TCGTAGATGGGCTTAAAGCAATTTGCGAAACGGTTGCGCGTGGCAGACCGCACCGG
 S I Y P E F V K R F A E R A N R L R V G

1061 GATCCGAACGATCCGAATACCCAGGTTGGCGCGCTTATTAGCCAGCAGCAGCTGGGAAAAA 3120
 CTAGGCTTGTAGGCTTATGGGTCCAACCGCGCAATAATCGTGTCTGTGACCTTTTT
 D P N D P N T Q V G A L I S Q Q H W E K

1121 GTCTCCGGCTATATCGTCTCGSCATTGAAGAAGCGCAACCTGCTGGCGGGCGGCGCC 3180
 CAGAGCCGATATAGGCAGACCGTAACCTTCTCCGCTTGGGAGCAGCCCGCGCGGGC
 V S G Y I R L G I E E G T L L A G G P

1181 GATAAACCGTCCGACCTGCTGCGCACCTGAAAGCGGCAACTTCTGCGCCCAACCGTG 3240
 CTATTTGGCAGGCTGGACCGGCGTGGACTTTCCGCGTGAAGGACGCGGTTGGCAC
 D K P S D L P A H L K G G N F L R P T V

3241 CTGGCAGACGTTGATAACCGTATGCGAGTCCGCCAGGAAGAGATTTTGGGCGGTGCGC 3300
 GACCGCTGCGCACTATTGGCATAACGCTCAGCGGGTCTTCTCTAAAAGCCCGCCAGCGG
 L A D V D N R M R V A Q E E I F G P V A

3301 TGCTGTGCGGTTTAAAGACGAAGCGAAGGCTTACGCTGCGCAAACGAGGTGGAGTAC 3360
 ACGGACGACCGGAAATTTCTGCTTCCGCTTCCGAATGCGACCGTCTGCTGCACCTCATG
 C L L P F K D E A E G L R L A N D V E Y

3361 GGCTCGCGTCTGATCTGGACACAGGATGTCAGCAAAGTGTACGCTGGCGGTGGC 3420
 CCGGAGCGCAGCATGATAGCCTGTCTCAGCTCGTTTCAATCGGACCGCGCACCG
 G L A S Y I W T Q D V S K V L R L A R G

3421 ATTGAAGCTGGCAGTGGTTCGTTCAACACCAGAACGTCGCTGACCTGGCCAGCCATTT 3480
 TAACTTCGACCGTACCACAAGCAGTTGTGGTCTTCCAGCGCAGGACCGGTCGGTAAA
 I E A G M V F V N T Q N V R D L R Q P F

3481 GGGCGGTAAAAGCCTCCGCGCACCGGGGTGAAGGCGGTGAGTACAGCTTGAAGTGTTC 3540
 CCGCGCATTTTCGGAGGCGTGGCGCGCACTTCCGCCACTCATGTCGAAGTTTCAAG
 G G V K A S G T G R E G G E Y S F E V F

3541 GCGGAAATGAAGAAGCTCTGCATTTCCATGGCGCACCATCCAATTCGAAATGGGGAGTC 3600
 CGCTTACTTCTTGCAGAGTAAAGGTACCGCTGGTGGTAAAGGCTTTACCCCTCAG
 A E M K N V C I S M G D H P I P K W G V

3601 TGATATGGGTAAGTTAGCGTTAGCCGCAAGATCAGCAGCAGTACCGCTGATGTATCTCTC 3660
 ACTATACCCATCAATCGCAATCGCGGTTCTAGTGGTGCATGGCAGTACATAGAGAG
 hpaD → M G K L A L A A K I T H V P S M Y L S

3661 TGAACCTGCCAGGAAAAACCGGTTCCGCGGAGGCGCGATCGACGGGCATAAAGAGAT 3720
 ACTTACCGGTCCTTTTTGGTGGCAACCGCGTCCCGCGCTAGCTGCCGHTTTCTCTA
 E L P G K N H G C R Q G A I D G H K E I

3721 CAGCAAGCGTTGCCGGAAATGGCGCTGATACCATTTATCGTTTTGATACCCACTGGCT 3780
 GTCGTTCCGCAACCGCCTTTACCCGCGCTATGGTAATAGCAAAAGCTATGGGTGACCGA
 S K R C R E M G V D T I I V F D T H W L

3781 GGTCAACAGTCTTATCACAATCACTGTGCGAGACCATTTTGAAGCGCTTACACCCAGCAA 3840
 CCAGTTGTCAGCAATAGTGTGACAGCTCTGGTAAAACCTCCGCGATGTTGGTCTGTT
 V N S A Y H I N C A D H F E G V Y T S N

3841 CGAGCTACCGCATTTTATCCGCGATATGACCTACAACCTACGAGGGCAACCGGAGTTGGG 3900
 GCTCGATGGCGTAAAATAGGCGCTACTGATGATGATGCTCCGCTGGGCTCAACCC
 E L P H F I R D M T Y N Y E G N P E L G

3901 GCAGCTTATGCGGATGAAGCCTTAAAGTCCGCGTGGCGGCAAAAGCGCACAACTTCC 3960
 CGTCAATAACCGCTACTTCCGAATTTCCAGCGCACCGCGCTTTTCCGCTGTTGTAAGG
 Q L I A D E A L K L G V R A K A H N I P

3961 CAGCCTGAAACTGGAATACGGCACGCTGTTGCCAATGGCTACATGAATGAAGACAAGCA 4020
 GTGGACTTTGACCTTATCCGCTGGACACCGGTTACGGATGTACTTACTTCTGTTCTG
 S L K L E Y G T L V P M R Y M N E D K H

4021 CTTCAAGTGGTTTCCATTTGAGCTTCTGCACTGTTGACGATTTTGGCGACAGCCGCAA 4080
 GAAGTTTACCAGGTAAGTGAAGACGCTGACAAAGTCTAAAACGGCTTCCGCGTTC
 F K V V S I S A F C T V H D F A D S R K

4081 GCTGGCGGAGCGGATTCGAAAGCGATGAAACAGTACGACCGGACCGTGGCGGCTCCTTGC 4140
 CGACCGCTCCGCTAAGACTTTCGCTAGCTTGTGATGCTGCGGTGGCACCGCGAGGAAAG
 L G E A I L K A I E Q Y D G T V A V L A

4141 CAGCGGTCGTTATCGCACCGCTTTATTGACGATCAGCGAGCGGAAGAAGGGATGAACAG 4200
 GTCGCCAAGCAATAGCGTGGGAAATAACTGCTAGTCTGCTCCGCTTCTTCCCTACTTGT
 S G S L S H R F I D D Q R A E E G M N S

4201 CTACACCGCGAGTTCCGACCGCAGATGGACGAGCTGTGTGAAGCTGTGGCGCGAAGG 4260
 GATGTGGCGCTCAAGCTGGCGGCTACCTGCTCGCACCCACTTCGACACCGCGCTTCC
 Y T R E F D R Q M D E R V C K L W R E G

4261 CCAGTTCAAGAGTTCTGCAATATGCTGGCGGAGTACCGCGACTACTGCTACCGGGAAGG 4320
 GGTCAAGTTTCTCAAGACGTTATACGACCGGCTCATGCGGCTGATGACGATGCCGCTTCC
 Q F K E F C N M L P E Y A D Y C Y G E G

121	CAATATGCACGACACGGTGATGCTGCTGGGGATGCTCGGCTGGGATAAATACGACGGCAA GTTATACGTGCTGTGCCACTACGACACCCCTACGAGCCGACCCCTATTATGCTGCCGTT NMHDTVMLLGLGLGWDKYDGR	4380	5401	GATCCGGAAACCAGCGCCCGCGCAAAGTGTTCGACACCATTTCTGATAAAGCCGCCAAC CTAGGCCCTTTGGGTCGGGGCGGCTTTCACAAGCTGTGGTAAAGACTATGGCGCGGTTG DPEETQRPKRKVFDTISDHAAN	5460
181	GTTGGAGTTTATCACCGAGCTATTC ⁰⁸⁴⁴ CCAAAGCT ⁰⁸⁴⁵ CTCGCACCGGTCAAGTTAACGGTGT CCACCTCAAATAGTGGCTCGATAAGGGTTCGAGACCGTGGCCAGTCCAATTGCGACAAA VEFITELFLPS ⁰⁸⁴⁴ SGTGVNAVF	4440	5461	GCGGGGTGATCCTCGGTGCTGCCATTAAGCCCGATGAGTTGGATCTGGCTTGGATC CGGCCCTACTAGGAGCCACGAGGGTAATTCCGGCTACTCAACCTAGACCGCAACCTAG AGVILGGRPIKPELELDRWI	5520
141	CCCGCTTCCCGGTAAGGACGTTTTATCGCCGCACTTTATCGTTGAATGCAGTGACAACAT GGGCGAAGGGCGCATCTCTGCAAAATACCGCGTGAATAGCAACTTACGCTACTGTTGTA PLPA* hpaF → MPHFIIVECSNI	4500	5521	TCGCCCTGATGTATCGCAATGGCGTATGAAGAAACCGCGCTCCCGCTGGTGTGCTG AGGGCGACTACATAGCGTTACCGCACTAACTCTTTGGCGCAGCGGGACCACACCGAC SALMYRNGVIEETGVAAAGVL	5580
301	CCGCGAAGAACCGACCTCCCGGGTGTTCGCCAAAGTGAATCGAGCGTGGCAGCCAC GGCGTCTCTCGGTGGACGGCCCAACAGCGGTTTCACTTAGGCTGCGACCGTCCGGT READLPLGLFAKVTGLNAV	4560	5581	AATCATCCGGCAACCGGTGGCTGGCTGGCGAACAACCTCGCACCGTATGACGTACAA TTAGTAGGCGCTTTCGCGCACCGGACCGCGCTTGTGTAGCGTGGCATACTGCATGTT NHPAANGVAVWLANKLAPYDVQ	5640
561	GGGTATTTTTCCGCTGGCGGTTATTCGACGCGCGTGCATTTGGTTCGATACTGGCAGAT CCCATAAAAGGGACCGCCCATAGCGTCCGGCGACGTAACCCAGCTATGGACCGTCTA GIFPLAGIRSRVHWVDTWQM	4620	5641	CTGGAAGCGGGCAAATCATTTCCGGCGTTCGTTACCCCGCCGGTTCGGCGCGTAAG GACCTTCGGCCGTTTAGTAAGAGCCCGCAAGTGGCGGGCCAAAGCCCGCGCATTC LEAGQIILGGSFTRPVPAK	5700
621	GGCGACGGGCGAGTATTATCGCTTCGTCATATGACGTTGAAAATCGGCGCAGTGG CCGCTGCGCGTCCGACTAATACGGAAGCAGTATACTGCAACTTTTAGCCGGTCCGAGC ADGQRDYAFVHMTLKIAGAR	4680	5701	GGGACACCTTCCACGTCGATTACGGCAACATGGGCTCCATTAGTCCCGCTTTGTTAA CCGCTTGGAAAGTGCAGTAAATCCGTTGTACCGGAGTAAATCGACGGCGAAACAAAT GDTFHVVDYGNMGSSISCRFV*	5760
681	CAGCTTGGAAAGCGTCCGACGCGGGTGAATGCTGTTTGAAGTATTAAAACGCACTT GTCGGACCTTTCGGCAGTCCGCGCCACTTTACGACAAACTGACTAATTTTGGTGA SLESRQQAQEMLELELIKTHF	4740	5761	GGAGAGACGATGGAAAACAGTTTTAAAGCGGGCGTAAAGCAGCGCCCGCCGATCCG CCTCTTCTGACTACCTTTTGCAAAATTTCCGCGGACTTTCTGTCGGCGGGGCTAGCC hpaI → MENSEFKAAELKAGRPVI	5820
741	CGCGCCCTGATGGAGACCGCTGCTGGCGTGTGCTTTGAGATTGAAGCGTGCATCC GGCGGGGACTACCTCTCGCGGACGACCGCAACAGCAAATCTAATCTTCGACGTAGG AALMESRLLLALSFEIEELHP	4800	5821	ATTATGGCTGGGGCTGAGTAGCAGTACAGCGCAGAGTACTGCCCGGAGCAGGATTCGA TAATACCGACCCCGACTCATGTCGATGTCGGCTCAATGACCGGCTCGTCTAAGCT LWLG LSSSYSAELLAGAGFD	5880
801	GACGCTGAATTTAAACAAAACAGTGCACGCACTGTTTAAAGTAGCGCGATTGCC CTGCGACTTAAATTTGTTTGTTCGACGTCGCGTAACAAATTCATCGCGCTTAAACGG TLNFKQNNVHALFK*	4860	5881	CTGTTATGATCGACGGTACGACCGCGGAAATACGTCGAAACCGTGCACCCAGCT GACCAATAACTAGTCCCACTGTCGCGCGCTTATGACGTTTGGCAGGATGGGTGGA WLLIDGEGHAPNNVQTVLTL	5940
861	GTTGGCGCTGCGCTTACCGGCGCTACAAAACCCAAACCGTACACCGTAGGCGGATAAG REP 1 → CCACCGCAGCGAATGGCCCGGATGTTTGGGGTTTGGCATGTGGCATCCGGCTATTC	4920	5941	ACAGGGGATGCGCCCTACCCCGCAGCGCGGTGGTACGTCGTCGTAACGATCCGGT TGTCGCTAACCGGGATGGGTGCTGGCCACCATGACGAGCAGCCTTGTAGGCCA QAIAPYPSQPVVVRPSWNDPV	6000
921	GCGACCGCCATCCCGCAATGCCACAGGATATCGCTATGTTGATAAACAATACCCACACC ← REP 2 CGCGTCCGCGTAGGCGGTTACGGTGTCTTATAGCGATACAAGCTATTGTTATGGGTGG hpaH → MFDKHTHT	4980	6001	GCAAAACAACAACCTGCTGGACGTCCGACACAAACCTTGTGTTGCCGATGGTACAAAA CGTTTAGTTTGTGACGACCTGACGCGGTGTTTGGAAACGACCCAGCTACCATGTTT QIKQLLDVGTQTL LVP MVQN	6060
981	CTGATCGCCAGCGCTGGATCAGGACGAAAACAGCGCAACAGATCCCGCGCATCTCA GACTAGCGGGTCCGACAGCTAGTCCGCTTTTTTTCGCGCTTGTCTAGGCGGCTAGAGT LIAQRLDQAEKQREQIRAI	5040	6061	CGGTGACGAAAGCGGTGAAGCGGTACGCGCCACCGCTTATCCCGCCGGTATTGCGGG GGGACTGCTTCGGGCACTTCGCCATGCGCGTGGGCAATAGGGGGCGGCCATAAGCGCC ADEAREAVRATRYPPAGIRG	6120
3041	CTGGATTACCGGAGATACCATCGAAGACGCTTACCGGTGACGCGGAAATGGGTTCCG GACCTAATGGCCCTTAGTGGTAGCTTCTGCGAATGCGCCAGTCCGCTTACCCAAAGCG LDYPEIITIEDAYAVQREWVR	5100	6121	TGTGGCAGTCCGCTGCGCCGAGCTCCGCGTGGAAATCGTATTCTCTGATTACCTGCAAAA ACACCGCTCAGCGACCGGGCTCGGAGCGGACCTTAGCATAAGGACTAATGGACGTTT VGSALARASRWNRIPDY LQK	6180
3101	CTGAAAATCGCGAAGTCCGACGCTGAAAGGCCAAAAATCGCCCTGACCTCGAAAGCG GACTTTTAGCGGCTCCAGCGTGCAGCTTTCGCGTGTTTTAGCGGACTGGAGCTTTCCG LKIAEGR T L K G H K I G L T S K A	5160	6181	AGCCAACGATCAAATGTCGCTGCTGTCAGATCGAAACCGGTGAGGCAATGAAGAACT TCGGTTGCTAGTTTACACGACGACACCTTCTAGCTTTCGCGACTCCGTTACTTCTGAA ANDQMCV L V Q I E T R E A M K N L	6240
5161	ATGCAGGCAAGCTCGCAGATCAGCGAACCGGATACGGCGCGCTGCTGGATGACATGTT TACGTCGTTTCGAGCGTCTAGTCCGTTGGCTAATGCGCGCGGACCTACTGTACAAG MQAS SQISEPDY GALLDDMF	5220	6241	ACCGCAGATTCTGGACGTGGAAGCGTTCGACCGGCTGTTTATCGGCCCGCGGATCTGAG TGGCGTAAAGACCTTCCGACGCTTCGCGCACAATAGCCGGGCGCCCTAGACT PQIL DVEGV D G V E I G P A D L S	6300
5221	TTCCACGATGGCAGGATATCCGACCGATCGCTTTATCGTCCGCGCATTGAAGTGGAG AAGTGCTACCGTCTATAGGGTGGCTAGCGAAATAGCACGGCGCTAACTTCACTC FHDGSDIPTDRFIVPRIEVE	5280	6301	CGCGCATATGCGGTTATGCGGTAATCCGACGACCCGGAAGTACAGCGCCGCTTAGGCA GGGCTATACCAATAGCGCCATAGGCGTTCGTTGGGCTTCTATGTCGGCGGTAACCTG ADMGYAGN P Q H P E V Q A A I E Q	6360
5281	CTGGCTTCTGCTGGCAAAACCGCTCGTGGGCCAAACTGCAAGCTGTTTCGACGCTAC GACCGAAAGCAGCAGCGTTTGGCGACGACCGGTTTACGTCGCGACAAGCTGACGATG LAFV LAKPLRGP NCT LFDVY	5340	6361	GGGATCGTGCAGATCCGTAATCGGGCAAGCGCGGGGATCCGATCGCCAAATGAGCA CCGCTAGCAGCTTAGGCACTTAGCCGTTTCGCGGCCCTTAGGACTAGCGGTTACTCGT AIVQIRESGKAPGILIANEQ	6420
5341	AACGCCACGACTACGTGATCCCGCGCTGGAGCTGATCGAAGCTCGCTGCCACAACATC TTGCGTGCCTGATGCACTAGGGCGGACCTCGACTAGCTGCGAGCGAGTGTGTAG NATD Y V I P A L E L I D A R C H N I	5400	6421	ACTGGCAAAACGCTATCTGAACTGGCGCGCTGTTTGTGCGCGTCCGCGTTCGACCCAC TGACCGTTTTCGATAGACCTTAGCCCGCGGACAAACAGCGGACCGCGCAACTGCGT LAKRYLELGALFVAVGVDTT	6480

481 CCTGCTCGCCGCGCCGCTGAAGCGCTGGCAGCAGCGTTTGGTGGCAGGCCACCGCGT 6540
⁰³⁸⁵
GGACGAGCGGCGCGCGGACTTCGGCAGCCGCTGTCGCAAAACCGCGTCCGGTGGCGCA
 L L A R A A E A L A A R F G A Q A T A V

541 GAAGCCCGCGGTGATTAATGCCGGATGGCTGGCTTATCCGGCTACACTCGCACCA 6600
 REP 3 →
 CTTCGGCGCCACATAAATACGGCCTACGGACCGGAATAGGCGGATGTAGCGTGGTT
 K P G V Y *

601 ACCGTAGGCGGATAAGGCGTTTACCGCGCTTCCGGCAAAAGCTGTACCAAAATCGGGA 6660
 ← REP 4
 TGGCATCCGGCCTATTCGCAAAATCGGGGAAGGCGCTTTTCGACATGGTTTAGCGCCT

661 CCGGATAAGGCGTTACGCCGATCCGCCATAAACCTTGTCTACCCCTACAAAAATCCCAT 6720
⁰³⁸⁶ REP 5
GGCCTATCCGCAATCGCGCTAGGCGGTATTTGGAACAGCATGGATGTTTTAGGGTA

721 TAGAGGAAGAAAAATGAGCGACACTCACCTGCCATACCGGAGAGTATCGATCCGGCGAA 6780
 ATCTCCTCTTTTACTCGCTGTGGATGGACGGTATGGCCTCTCATAGCTAGGCGGCTT
 hpaX → M S D T S P A I P E S I D P A N

781 TCAGCATAAAGCGCTGACTGCCGACAAAGCGGGTTATTAAGAAGCTATTTCGGCGCCT 6840
⁰³⁸⁷
 AGICGTATTCGGGACTGACGGCTGTTCGGCAATAATCTTCGATAAAGCGGGGA
 Q H K A L T A G Q Q A V I K K L F R R L

841 GATCGCTTTCTGTTCGTGCTGTTTATCTTCTGCTTCTTGATCGCATCAACATCGGCTT 6900
 CTAGCAAGAACAGCAGCAGCAATAGAAGAGCAAGAACTAGCGTAGTGTAGCCGAA
 I V F L F V L F I F L D R I N I G F

901 TGCCGACTCAGATGGACGCGACCTCGGCTGAGCGCCACCATGTTTGGCCTCGCTAC 6960
⁰³⁸⁸
 ACCGCCGTGAGTGTACCTCGCGTGGAGCCAGACTCGCGGTGGTACAAACCGGAGCGATG
 A G L T M G R D L G L S A T M F G L A T

961 CACCCTGTTCTATGCCGCTTATGTCATCTTCGGCATTCGCCAACAATATGCTGAGTAT 7020
⁰³⁸⁹ REP 6
GTGGGACAAGATACGGCAATACAGTAGAAGCCGTAAGGGTGGTGTAAATACGACTCATA
 T L F Y A A Y V I F G I P S N I M L S I

7021 TGTGCGTGGCGCGCTGGATCGCCACCATCATGCTGCTTGGGATCGCCTCTACTGC 7080
 ACAGCCACCGCGCGACCTAGCGGTGGTAGTACCACGAGACCCGCTAGCGGAGATGAGC
 V G A R R W I A T I M V L W G I A S T A

7081 CACCATGTTGCCACTGGCCCCACAGCTTGTACGTACTGCGTATAGTGTGGCATTAC 7140
 GTGGTACAAACGGTAGCCGGGTGGTGAACATGCATGACGCATATGACCAACCGTAATG
 T M F A T G P T S L Y V L R I L V G I T

7141 CGAAGCGCGCTTCTGCGGTATTCTGCTGTATTTAACCTTCTGGTTCCGGCCTATTT 7200
 GCTTCGGCCGAAGACGACCATAAGACGACATAAATGGAAGACCAAGGCGGATAAA
 E A G F L P G I L L Y L T F W F P A Y F

7201 CCGCGCCGTCGCAACGCTTGTATGTTAGCAATCGCGGTAACGACAGCGTTGGGCTC 7260
 GGCGCGGCGACGGTTGGGAACAATAACATCGTTACGGCCATTTGCTGTGCAACCCGAG
 R A R A N A L F M V A M P V T T A L G S

7261 GATCGTTTCCGGCTACATTTTGTGCTGGATGGCGTATGGCATTAAAGGCTGGCAGTG 7320
⁰³⁸⁹
 CTAGCAAGGCGGATGTAACACAGGACCTTACCGCAGTACCGTAATTTCCGACCGTAC
 I V S G Y I L S L D G V M A L K G W Q W

7321 GCTGTTTTTGTGGAAGGCTTCCCGTGGTATTACTCGCGCTCATGTTGTGTTCTGGCT 7380
 CGACAAAAACGACCTTCCGAAGGCGCCATAATGAGCGCAGTACCACCAAGACCGA
 L F L L E G F P S V L L G V M V W F W L

7381 TGATGACTCACCGACAAGCTAAGTGGCTGACGAAAGAAGACAAAAATGCCTGCAAGA 7440
 ACTACTGAGTGGCCTGTTTCGATTACCGACTGCTTTCTTCTGTTTTTACGACGTTCT
 D S P D K A K W L T K E D K K C L Q E

7441 GATGATGGATAACGATCGTCTGACCGTGGTTCAGCCAGAGGGACCATCAGCCACCACGC 7500
 CTACTACCTATTGCTAGCAGACTGCGACCAAGTCCGCTCCCTCGGTAGTCCGTTGGCG
 M M D N D R L T L V Q P E G A I S H H A

7501 CATGCAACACCGAGTATGTCCGGGAGATCTTCACTCCGGTGGTATGATGATACCCCT 7560
⁰³⁹⁰ REP 7
⁰³⁹¹
 GTACGTGTTGGCTATACACCGCCCTCTAGAAGTGGGCGCCACTACTACATATGGGA
 M Q Q R S M W R E I F T P V V M M Y T L

7561 GGGTATTTCTGCTGACCAACACACTTAGTGGCATCAGCATCTGGACACCGCAGATCCT 7620
 CCGCATAAAGACGGACTGGTTGTGTGAATCAGCGTAGTCTAGACCTGTGGCGTCTAGGA
 A Y F C L T N T L S A I S I W T P Q I L

7621 GCAAAATTTTAAATCAGGGCAGCAGTAAATACCATCGGCCCTGCTGGCGCGTACCGCCA 7680
⁰³⁹²
 CGTTTCAAATTAGTCCCGTGGTATTATAGTGGTAGCGGACGACCGGGCGCATGGGT
 Q S F N Q G S S N I T I G L L A A V P Q

7681 GATTTGTACCATTCTGGGATGGTCTACTGGAGCGTCACTCAGATCGCCGCGAGAAAG 7740
 CTAACATGGTAGAGCCCTACCAGATGACCTCGGCAGTGGTCTAGCGCGGTCTTGC
 I C T I L G M V Y W S R H S D R R Q E R

7741 AAGGCATCACACCGCCCTTCTTATTGTTGCTGCGCGAGGTTGGTTACTGGCTTCGGC 7800
⁰³⁹³
 TTCGCTAGTGTGGCGGAAGGAATAAACAGCGACCGGCTCCAACCAATGACCGAAGCGG
 R H H T A L P Y L F A A A G W L L A S A

7801 AACTGATCACACATGATCCAGATGCTGGGATCATTATGGCTTCGACCGGATCATTGAG 7860
 TTGACTAGTGTGTACTAGTCTACGACCCCTAGTAATACCGAAGCTGGCTAGTAAGTC
 T D H N M I Q M L G I I M A S T G S F S

7861 CGCAATGGGATTTTCTGGACAACCGGATCAGTCCATCAGCCTGCGGCGCAGCGAT 7920
⁰³⁹⁴
 CGTTTACCCTAAAGACCTGTTGCGCTAGTCAAGTAGTGGACGCGCGGCTGCTCGTA
 A M A I F W T T P D Q S I S L R A R A I

7921 CGGTATGGCGTATCAACCGCACGCGCAACATGGTTGAGCGTTAAGTCCGTTTATGAT 7980
⁰³⁹⁵ CAP
GCCATAACGCCACTAGTTCGGGTGACCGTGTAAACCAAGTCCGATTGAGCAATACTA
 G I A V I N A T G N I G S A L S P F M I

7981 CGGCTGGTGAAGATTGACCGCGCAGCTTAAACAGTGGATTGTTGGTTTGTTCGGCGCT 8040
⁰³⁹⁶
 GCGGACCAACTTCTAACTGGCGTGGAAATGTCACCTAACACCAACCAACCGCGCGA
 G W L K D L T G S F N S G L W F V A A L

8041 GCTGGTATGGTGGGGATTATCTGGCGGATCCAATGACGCTCCTCCGCTCGCGGAGC 8100
 CGAOCCTAACCAACCGCCCTAATAGACCGCGTAAAGTTACGTCAGGAGGCGCGCTCG
 L V I G A G I I W A I P M Q S S R P R A

8101 GACCCCGTAAGGAACGAGATGTGTGACCGCTCAGATGGCAATATGATATCAGCAAGA 8160
 CTGGGGCATTCTTGTGTACACACTGCGAGTCAACGGTTATAACTATAGTCGTTTCT
 T P * hpaA → M C D R Q I A N I D I S K E

8161 GTACGATGAAGCGCTGGGCAGGAGATGTGCATTATCAGTCTGCGCCGATGGCGCC 8220
^{PAJ39A}
 CATGCTACTTTCCGACCGCTGCTGCTACAGTAAATAGTCAAGAGCGGGCGTACCGCGG
 Y D E S L G T D D V H Y I P M Q S F A R M A P

8221 TTTTTTGGCCAGTATGCTGCCACATCGCCACGACAGTACTTTCAGATGCAITTCCTCAA 8280
 AAAAAACCGCTCATACGCGGTGAGCGTCTTGTGATGAAAGTCTACGTAAAGAGGT
 F L A S M L P H R H E Q Y F Q M H F L N

8281 TAGCGGACAGATTGAGCTACAACCTGACGATCATCGTACTCGGTGGAAGCGCCCTGTT 8340
 ATCGCTGTCTAACTCGATGTGAACTGCTAGTAGCGATGAGCCACCTTCGCGGGACAA
 S G Q I E L Q L D D H R Y S V E A P L F

8341 TGTCTGACCGCGCGTCACTACCTCATGGCTTATTACGGAGTCTGACCGCGACGGTCA 8400
 ACAGGACTCGCGCGGAGTATGGAGTACGCAATAATGCTCAGACTGCGGCTGCCAGT
 V L T P P S V P H A F I T E S D A D G H

8401 TGTGTTGACGGTACGGGAAGATCTGATCTGGCCCTGCTGGAAGTCTTTTACCGGGCAC 8460
 ACACAAGTCCATGCCCTTCTAGACTAGACCGGGGACGACCTTCAAGAAATAGCCCGTG
 V L T V R E D L I W P L L E V L Y P G T

8461 TGGGAAACCTTCGGCTCGCGGGGATTTCGCTGCTACTGGCAGATAAACCCGACGAAGT 8520
 AGCCCTTGGAAAGCGGACGGCCCTAAACGGACAGTACCGCTTATTGGGCTGCTTGA
 R E T F G L P G I C L S L A D K P D E L

8521 GGCGCGCTGGAACACTATTGGCAACTGATAGAGCGGAAATCGGTAGAACAAGTCCGCTGG 8580
 CCGCGGACCTTGTGATAACCGTGTACTATCTCGCCCTTAGCCATCTGTTGACGGACC
 A A L E H Y W Q L I E R E S V E Q L P G

8581 ACGGGAACACACCTGACGTTACTGGCACAGGAGTATTACCCTACTGCTGCGTAAATGC 8640
^{PAJ39B}
 TGCCTTGTGTTGGGACTGCAATGACCGTTCGCTCATAGTGGGATGACGACGATACG
 R E H T L T L L A Q A V E T L L L R N A

641	AAAAC TCGAGACCATGCCGCCAGCGGAATCGCGGAGAATTAAAAC TGTCCAGCGTTT TTTTGAGCTGCTGGTACGGCGGTGCGCCTTACGCGCTCTAATTTTGACAAGSTGCAAA K L D D H A A S G M R G E L K L F Q R F	8700	9721	ACGCGATTGTTAACCACCAGATCGATCGTCATTTGCGGACCGATAAAGTAAAAGACGTTT TGCGTAAACAATTTGGTGGCTAGCTAGCAGTAAACGGCTGGCTATTTCATTTCTGCAAA A I V N P P I D R H L P T D K V K D V Y	9780
701	TAATATGCTGATTGAAAGCCATTTTCATCAGCAGCTGGACAGTACCGGATTACGCTAAGCA 333A ATTATAGCTAACTTTCCGTAAGTAGTCTGACCTGTCTATGGCCTAATGGGATTGCT N M L I E S H F H Q H W T V P D Y A N E	8760	9781	ACATCAAGCTGGAAAAGAGACTGACCGCGGATTATCGTCAGCGGTGGAAAAGTGGTTG PACEN TG TAGTTCGACCTTTTCTCTGACTGCGGCCCTAATAGCAGTCCGACGCTTTCCACCAAC I K L E K E T D A G I I V S G A K V V A	9840
761	ACTGCATATCACCGAATCACGACTCACGGACATCTGCCGTCGCTTGCCAAACCGTCCGCC TGAGCTATAGTGGCTTAGTGTGAGTGCCTGTAGACGGCAGCGAAACGGTTGGCAGGCGG L H I T E S R L T D I C R R F A N R P P	8820	9841	CCACCAACTCGCGCTGACTCACTACAACATGATTGGCTTCGGCTCGGCACAAAGTAAATGG GGTGGTTGAGCCGCGACTGAGTGTGTGTACTAACCGAAGCGGCGTTCATTACC T N S A L T H Y N M I G F G S A Q V M G	9900
821	AAAACGGTTGATTTTCGACAGGCGACTACGAGAAGCCAGCGGCTGCTGCTTTTCTGA TTTTGCCAACTAAAAGCTGTCGCTCGATGCTCTCGGTTGCGCAGCAGCACAAGAACT K R L I F D R Q L R E A K R L L L F S D	8880	9901	GCGAAAACCGGACTTCGCACTGATGTTGTTGCGCAATGGATGCCGATGGCGTCAAAT CGCTTTTGGGCTGAAGCGTACTACAAGCAACGGCTTTACCTACGGCTACCGCAGTTTA E N P D F A L M F V A P M D A D G V K L	9960
881	TAACGCCGTAACAATATTGCTGGCAACTCGGTTTTAAGGATCCGGCTATTTTGCGCG ATTGCGGCACTTGTATAACGGCAGCTGAGCAGCAAAATCCTAGCCGAATAAAACGGCG N A V N N I A W Q L G F K D P A Y F A R	8940	9961	TAATCCCGCGCTCTTATGAGATGGTCCGGGTGCTACCGGCTCACCGTATGACTACC ATTAGAGGGCGCGGAAATACTCTACCAGCGCCACGATGGCGAGTGGCATACTGATGG I S R A S Y E M V A G A T G S P Y D Y P	10020
941	CTTTTTAATCGCTTAGTCCGTTGCTCCGCCAGTCTTATCGTGCAAAAAGTACCTGT GAAAATAATAGCGAATCAGCCAACGAGCGGTACGAAATAGCACGGTTTTTTCATGACAA F F N R L V G C S P S A Y R A K K V P V	9000	10021	CGCTCCAGCGCTCTGATGAGAACGATGCGATTCTGGTGTGATAAAGTGGTGTGATCC GCGAGAGTCCGCGAAGTACTCTTGTACCGCTAAGACCACTACCTATTGACGACTAGG L S S R F D E N D A I L V M D N V L I P	10080
1001	GACGTGAAATTCCTTAGCCTTAACGGGAAAACGAGCCACCTGCTATTCCCTTTCTTT CTGCACITTAAGGGAATCGGAATTCGCCCTTTGGTCCGTGGTGGAGATAACCCCAAGAA T *	9060	10081	CATGGGAAAACGCTGCTCTACCGGATTTTGTATCGTGCCGCTCGTGGACGATGGAAAG GTACCTTTTTCAGCAGGAGATGGCGTAAACTAGCGCGGCGGACCTGCTACCTTC W E N V L L Y R D E D R C R R W T M E G	10140
1061	CGCCTGGAGCAAGGGGATTTAGTAACTGTAATAAATCGCTAAAAGCGACCTCGATCACAAA GCGGACCTCGTTCCTAAATCATGTCAATTTTTAGCGATTTTCGCTGGAGCTAGTGTTT	9120	10141	GCGGTTTCGCCGATGATATCGCTGCAAGCGCTGTGTGCGCTGGCAGTGAACCTCGACT CGCCAAAGCGGGATACATAGGCGAGCTTCGGACACAGCGGACCGTCACTTTGAGCTGA G F A R M Y P L Q A C V R L A V K L D F	10200
1121	ACCGTACCGACCTGAAAAGTACCAGCCATGTCCTCCAAAAGTCTCTTCCGTAATCGCCC -35 TGGCATGGGCTGGACTTTTCATGCTCGGTAAAGGGTTTTTCAGAGAAGGGCATTAGCGGG	9180	10201	TCATTACGGCACTGCTGAAAAAATCACTCGAATGTACCGGCAACCTGGAGTTCCGCTGGT AGTAATCCGCTGACGACTTTTTTAGTGTGCTTACATGCGCGTGGGACCTCAAGGCACCA I T A L L K K S L E C T G T L E F R G V	10260
1181	CTGCAAGCTTCAATGGAATAACAAAAACAACACACAAAAATAACAATCGTATTCGATTA -10 +1 GACGTTCAAGTTAACTTATGTTTTGTTGTGTGTTTTAATGTTAGCATAAGGCTAAT	9240	10261	TGCAGCCGATCTCGGTGAAGTGGTGGCGTGGCGCAACCTCTTGGCATTGAGTACT ACGTCGGCTAGAGCCACTTACCACCGCACCCTGGTGTGGAAGACCGCTAAGTCACTGA Q F A R M Y P L Q A C V R L A V K L D S	10320
1241	ATACTGTAGAGTGCACATGAAACAGAAAGATTTCCGCGCAGTACCACCGTCCGTTCA HIDROPHO TATGACATCTCCAGCTGACTTTGGTCTTCTAAAGGCGCGGTATGGGTTGACGGCAAGT hpaB → M K P E D F R A S T Q R P F T	9300	10321	CGATGTGTTCTGAAGCGACCGCTGGGTCAACGGGCTTATTTACCGGATCATGCCGAC GCTACACAAGACTTCGCTGCGCACCCAGTTGCCCGCAATAAATGGCTAGTACGGCGTG M C S E A T P W V N G A Y L P D H A A L	10380
1301	CCGGGAAGAGTATCTGAAAAGCTGCAAGGATGGTCCGAGATCTATATCTATGCGGAGC GGCCCTTCTCATAGACTTTTCGACGCTCTACCAGCGCTCTAGATATAGATACCGCTCG G E E Y L K S L Q D G R E I Y I Y G E R	9360	10381	TGCAAACTATCGCTACTGGCACCAATGGCTACGCGAAGATCAAAAACATTATCGAAC ACGTTTGGATAGCGCATGACCGTGGTTACCGGATCGCTTCTAGTTTTTGTAAAGCTTG Q T Y R V L A P M A Y A K I K N I E R	10440
1361	GAGTGAAGACCTCACTACTCATCCGCTTTTCTGTAATGGCGTGGCTGCTGTTGCCAAC CTCAGTTTCTGAGTGTGAGTGGCGTAAAGCATTACGCGGACCGCAGCAACCGGTTG V K D V T T H P A F R N A A A S V A Q L	9420	10441	GCAACGTTACCAGTGGCTGATCTATCTCCCTCCAGTGGCGTGGCTGACCTGAACAATCCG CGTTGCAATGGTCAACCGACTAGATAGAGGGAAGTACGCGGCACTGGACTTGTAGCGG N V T S G L I Y L P S S A R D L N N P Q	10500
1421	TGTACGACGCTTACACAAACCGAGATGCAGGACTCTCTGCTGGAACACCGACACCG ACATGCTGCGGATGTTTGGCCCTACGCTCTGAGAGACAGCCTTGTGGCTGTGGC Y D A L H K P E M Q D S L C W N T D T G	9480	10501	AGATCGACAGTATCTGGCGAAGTATGTCGGCTTCGAAACGATGATGATCAGTCCAGC TCTAGCTGGTATAGACCGCTTCATACACCGCCAAAGCTTGCATACCTAGTGCAGGTCG I D Q Y L A K Y V R G S N G M D H V Q R	10560
1481	GCAGCGCGCTATACCCATAAATCTTCCGCTGGCGAAAAGTGGCGGACCGTGGCGC CGTCCGCGCATATGGGATTTAAGAAGCGCAGCCTTTTCAGCGCTGCTGGACGCGG S G G Y T H K F F R V A K S A D D L R H	9540	10561	GCATCAAGATCTCAAACGATGTTGGGATGCCATTGGCAGCGAGTTGGTGGTCGTCAGG CGTAGTTCTAGGAGTTGACTACACCTACGGTAAACGCTCGCTCAAACACCGAGCTGC I K I L K L M W D A I G S E F G G R H E	10620
1541	ACGAAACGATGCCATCGCTGAGTGGTCAACGCTGAGCTATGGCTGGATGGGCGTACCC PAJ25D TGCTTGGCTACGGTAGCGACTCACAGTGGGACTCGATACCGACTACCGGATGGG E R D A I A E W S R L S Y G W M G R T P	9600	10621	AACTGATGAAATCAACTACTCCGCTAGCCAGGATGAGATTCCGCTGAGTGTCTGCGCC TTGACATACTTAGTGTGATGAGGCCATCGGCTCTACTCTAAGCGGACGTACAGAGCGGG L Y E I N Y S G S Q D E I R L Q C L R Q	10680
1601	CAGACTACAAAGCTGCTTCCGTTGCGCACTGGGGGAACTCCGGGCTTTTACGGTCACT GTCTGATGTTTCGACGAAAGCAACCGGTGACCGCCTTGAGGCCGAAAATGCCAGTCA D Y K A A F G C A L G G T P G F Y G Q F	9660	10681	AGGCACAAAGCTCCGCAATATGGAACAAGATGATGGGATGTTGATCGCTGCTGCTCGG TCCGTTTTCGAGGCGGTTATACCTGTTCTACTACCGCTACCAACTAGCGACGGACAGCC A Q S S G N M D K M M A M V D R C L S E	10740
1661	TCGAGCAGAAACCGCTAACTGGTACACCGCTATTCCAGAACTGGCCCTACTTTAAAC AGCTCGTCTTGGCGGCAATGACCATGTGGGCATAAGTCTTTGACCGGAGATGAAATGG E Q N A R N W Y T R I Q E T G L Y F N H	9720	10741	AATAGCACCAGAACCGCTGGACTGTGCCGACCTGCACAACAACGACGATATCAACATGC TTATGCTGCTTGGCGGACTGACACGGCGTGGAGCTGTTGTTGCTGCTATAGTTGTAGC Y D Q N G W T V P H L H N N D D I N M L	10800

801	TGGATAAGCTGCTGAAATAACGACAGCAGGAGTTAAGATGCAATTAGATGAACAACGCCT ACCTATTGACGACTTTATTGGTTCCTCCAATTCTACGTTAATCTACTTTGTTGGGA D K L L K * hpaC → M Q L D E Q R L	10860	11881	TATCTGGTGGCGTATCGCTACTACAGTTTGTATATCGCCAGAAGGTGATGAAACTCGAC ATAGACCACCGCATAGCGATGATGTCAAACATATAGCGGGTCTTCCACTACTTTGAGCTG Y L V A Y R Y Y S L Y I A Q K V M K L D	11940
861	CGCGTTTCGTGACCAATGCCAGCCTGTGCGCAGCGTAAATATTATCACCACCGAGGG CGCGAAAGCACTGCGTTACCGGTGGACAGCGCTGGCCATTATAATAGTGGTGGCTCCC R F R D A M A S L S A A V N I I T T E G	10920	11941	CGCACGCGGACCAATGCCGATTACGGCAACGGCGTCTTGCTCGGTACGGGATACACCACC GCTGGCGCTGTACGCCCTAATGCCGTTCGCCGAGAACGAGCCAGTGCCTATGTGTGGT D A D N A G L R Q R P S C S V T D T P P	12000
921	CGACGCGGACCAATGCCGATTACGGCAACGGCGTCTTGCTCGGTACGGGATACACCACC GCTGGCGCTGTACGCCCTAATGCCGTTCGCCGAGAACGAGCCAGTGCCTATGTGTGGT D A D N A G L R Q R P S C S V T D T P P	10980	12001	CGGTACGTGTGTTGGTACCACCTTCGCGCTATCGCGGTGCTGGTCCGCTGGTGGT GCAATGCACAACAAACCGTGGTGAAGCGCGATAGCGGCAGGACCAGCGACCAACCA R Y V L F G H H E A A I A G A G P L V G	12060
981	ATCGCTGATGGTGTGCATTAACGCCAACAGTGCATGAACCCGCTTTTTCAGGGCAACGG TAGCGACTACCACAGTAATGCCGTGTGTCAGCTACTTGGCCAAAAGTCCCGTGGCC S L M V C I N A N S A M N P V F Q G N G	11040	12061	CGCGTTCGCGCGCAGATGGGCATCCTACCTGGCAGCTGTGGCTGTGGCGGGTA GGCCAAAGAGCGCGCGTCTACCGTAGGATGGACCGTGGCAGCAGCGACCGCCCAT P V L A A Q M G I L P G T L W L L A G V	12120
041	TAAGTGTGCGTCAACCTCCTCAACCATGAGCAGGAAGTGTGGCAGCCACTTCGCGGG ATTCAACACGAGTTCAGGAGTTGGTACTCGTCTTCACTACCGTGGCGTGAAGCGCC K L C V N V L N H E Q E V M A R H F A G	11100	12121	GTACTGGCGGTGGGTTCCAGGACTTATGGTGTCTTATCTCTCTCGCGTAAACCGA CATGACCGGCCACGCCAAGTCTGAAATACCAGCAAAATAGAGGAGAGCGGATTGCGCT V L A G A V Q D F M V L F I S S R R N G	12180
101	CATGACAGCCATGGGATGGAAGAGCGTTTTAGCCTCTCATGTGGGAAAAGGTCGCGT GTACTGTCCGTACCGTACTTCTCGAAAATCGGAGAGTACGACCGTTTTTCCAGCGCA M T G M A M E E R F S L S C W Q K G P L	11160	12181	CCATCTCTCGGTGAGATGATCAAAGAAGAGATGGGACCCTACCGGGACTATCGCGTG GGTAGAGAGCCACTCTACTAGTTTCTTCTACCTCGGTGATGGCCCTGATAGCGGAC P S L G E M I K E E M G P L P G T I A L	12240
161	GGCGCAGCGCGTGTAAAAGGTTTCGTCGGCAGCTTGAAGGTGAGATCGCGATGTGCA CGCGTGGCCACGATTTTCAAGCGACGGTGCAGAACTTGCACCTAGGCGGTACAGCT A Q P V L K G S L A S L E G E I R D V Q	11220	12241	TTTGGCTGTTCTTAATCATGATCATCATCTCGCGCTCTCGCGCTGATTGTGGTTAAA AAACCGACAAGAATTAGTACTAGTAGTAGGAGCGGACCGGACTAAGACCAATTT F G C F L I M I I I L A V L A L I V V K	12300
1221	GGCAATTGGCACACATCGGTGTATCTGGTGGAGATAAAACATCATCTCAGTGCAGA CCGTTAACCGTGTGTAGACCACATAGACCACCTCTAATTTTGTAGTAGGAGTACCGTCT A I G T H L V Y L V E I K N I I L S A E	11280	12301	GCCCTGGCCGAAAGTCCGTTGGGGTCTTCCACCGTTTGGTCAACCGTACCGATTGCGCTG CGGACCGCGTTTCAGGCACCCACAGAAAGTGGCAAACGAGTGGCATGGTAAACCGGAC A L A E S P W G V F T V C S T V P I A L	12360
1281	AGGTCACGACTTATCTACTTTAAACCGCGTTTCCATCGGGTGTGCTGGAATGGAAGC TCCAGTGCCTGAATAGATGAAATTTGGCGCAAAGTATGGCCACTACGACCTTTACCTTCG G H G L I Y F K R R F H P V M L E M E A	11340	12361	TTTATGGTATCTACATGCGCTTATCCGTCGGGGCGTGGGTGAAGTCTCTGTGCTATT AAATACCCATAGATGTACGCGAAATAGCGAGCCCGCACACCCACTTCAGAGACAGTAA F M G I Y M R F I R P G G V I A H D P Y	12420
1341	TGCGATTAAAGCTCCAGCCCTTACCGCATTTCGCGAAGGGCATTTTTATTGTGATCTC ACGCTAAATTCGAGGTCGGGAAAGTGCCTAAGCAGCTTCCCGTAAAATAACACTAGAG A I *	11400	12421	GGTATCGTCTGCTGGTTCCTCTATCTACTTCGGTGGCGTATTGCTCAGCATCGGTAC CCATAGCAGCAGCAACCGGAGATAGATGAAGCCCGGCACTAAGCAGTGTAGGCATG G I V L L V A S I Y F G G V I A H D P Y	12480
1401	AATCACATATCCACCCGATGAATAAATCTTGCCCAACCGTACCGCCATTCTGACCGCT TTAGTGTATAGTGGGCTACTTATTAAGAACGGGTGGCAGTGGCGTAAGGACTGGCGA	11460	12481	TGGGGTCCGGCAGTACCTTTAAAGACACCACATTACCTTCGCGCTGATTGGCTATGCG AGCCACAGCGGTGACGACGGCCACACCGACTAGGAGCGTGGCGCTGATAGACCGTGG W G P A L T F K D T T I T F A L I G Y A	12540
1461	TACATCCCTAAAATAACCACTCAGTTATTTACCTTACTTTACGCGCGTATCTCTGGCA ATGTAGGGATTTTATTGGTGGTCAATAAATGGAATGAAATCGCGCGCATAGAGACCGT	11520	12541	TTTGTTCGCGACTGCTGCGGTGTGGCTGATCTCGCACCGCGGACTATCTGGCAACC AAACAAGGGGTGACGACGGCCACACCGACTAGGAGCGTGGCGCTGATAGACCGTGG F V S A L L P V W L I L A P R D Y L A T	12600
1521	ACATCACTGCAGGATAGCGGTCAATTTACCTCCTCAAACACACGCAAACTAGAACGGC TGTAGTGACGCTCTATCGCCAGTTAAATGGAGAGTTTGTGTTGCGTTGGATCTTGGCG	11580	12601	TTCCTGAAAATCGCGTATCGTGGCGTGGCGTGGGTATCGTGGTGTGAACCGGAA AAGGACTTTTAGCCGCAATAGCAGCGGACCGCGACCCATAGCACCAGCTTGGGCTT F L K I G V I V G L A L G I V V L N P E	12660
1581	TTGCGCTATGATTAACAATACTCTATGATTTTGGTTTGGCAGGAAGCAGTGG AAGCCGATACTAATTTTAAATGAGATACCTAAAACCAAACCGTCTTCCGCTGCTTAC	11640	12661	CTGAAAATGCTGCGATGACCCAGTACATGACCGTACTGGCCCGCTGTGGAAGGGGCT GACTTTTACGGACGCTACTGGGTATGTAACCTGATGACCGGGGACACCTTTCCGCGA L K M P A M T Q Y I D G T G P L W K G A	12720
1641	AGTGAATCCCGGGAGCTTGCACAGTAAGTACAGGGGTGAAGCAACGCGAGTCCCGCA TCACTTAGGGGCGCTCGAACGTTGTCATTCACTGTCCCACCTGCTTGGCTCGACGGCGT	11700	12721	CTGTTCCCGTTCCTGTTTACCATCGCCTGCTGGTGGGTATCTGGCTTCCACCGCTG GACAAAGGCAAGGACAAGTAGTGTAGCGGACACCGCCATAGACCGAGGTGCGCGAC L F P F L F I T I A C G A V S G F H A L	12780
1701	CCTGTAAGCCAAAAGACGAGTAAAATGCCAGTTTACTATGGATACGAAAAACTA GGACATTCGGTTTTCTGCTGCTCATTTCAGCGTCCAAAATGATACCTATGCTTTTTTGAT OREF12 → M P G F T M D T K K L	11760	12781	ATCTCTCCGGTACGACGCCAAAAGTGGTGGTAAACGAAACCGAGCGCGTTTTCATTGGC TAGAGAAGCCATGCTCGGTTTTGACGACCGGATTGCTTTGGCTGCGCGCAAAGTAAACG I S S G T T P K L L A N E T D A R F I G	12840
1761	TTCAAGCACATACCGTGGTGTCTCGGAATCATCGGTGCATCTGCTCGCGGTAGTT AAGTTCGTGTAGGGACCCACTAAGAGCCTTAGTAGCCAGTAAAGACGGAGCCGATCAA F K H I P W V I L G I I G A F C L A V V	11820	12841	ATCGCCGTTACGACTACCTCAGGAAGCAGCGCTAATACCGGACCAACGTCGACGAGT Y G A M L M E S F V A I M A L V A A S I	12900
1821	CGTAAATGACGCCCCCTCGTGCAGTCCGGGACACCTAGCACGACGGGAGACACAGCCAC A L R R G E H V S A L W I V V A S V S V	11880	12901	ATCGAACGGGTCTTTACTTCGGATGAATACCCCGCTGCTGGCTTGGCATCCATG TAGCTTGGCCGAAATGAAGCGTACTTATGGGGCGGACCGGAAACCGTAGTGGTAC I E P G L Y F A M N T P P A G L G I T M	12960

361	CCTAACCTGCATGAAATGGGTGGCGAGAACCGCCGATCATCATGGCCAGCTGAAAGAC GGATTGGACGTACTTTACCCACCGCTCTTGGCGGGCTAGTAGTACCGCGTGCAGCTTTCTG P N L H E M G G E N A P I I M A Q L K D	13020	14041	AACCATCCCACAAAGCCGTACATGAGCTATGAAGAATTTCTTCGCGAACCCGAGAACGCA TTGGTAGGGGTGTTCCGGCATGTACTCGATACTTCTTAAGAAGGCGCTTGGCGTCTTGGCT ORF14 → M K N S S A N A R T H	14100
321	GTTACCGCACACGGCGAGACCGTCACTCTCTGGGGCTTCGTGATTTCGCCAGAGCAGATC CAATGGCGTGTGGCGCTCTGGCAGTGGAGGACCCGGAAGCACTAAAGCGGTCTCGTCTAG V T A H A Q T V T S W G F V I S P E Q I	13080	14101	CGCTACGGGGCGATGGTAAAGGGGTATGGGCTGTGTTAATGGAGAGAGCATGAACCC GOGATGCCGCGCTACCATTTCCGCCATACGGGACAACAATACCTCTCTCGTACTTGGG I A T A A M V K A V C A V N G E S M N P	14160
081	CTGCAACCCGCAAGACATTTGGTGAAGCTTCTGTCTCTGAACCGTGCAGGTGGTGGCGCA GAGCTTTGGCGCTTTCTGTAACTACTCGGAAGACAGGACTTGGCAGCTCCACCACGGCT L Q T A K D I G E P S V L N R A G G A P	13140	14161	GATTGCAGTTACCTACTCACGGTTTTTTAGGCGCGGGCAAAACACCTGCTGGCCCA CTAACGTCAATGGGATGAGTGGCAAAAAATCCGCGCCGCTTTTGGTGGGACGACGGGT I A V T L L T G F L G A G K T T L R H	14220
141	ACGCTGGCGTAGGTATCGCTCACGTGTTCCACAAGTCTGCCAGTGGCTGACATGGGC TGCACCCGCTCATAGGGAGTGCACAAGTGTTCACGACCGTACCGACTGTACCCG T L A V G I A H V F H K V L P V A D M G	13200	14221	TATTCTTAATGAACAACCGCTACAAGATTGCCGTGATTGAAAACGAATTCCGGGAAGT ATAAGAATTACTTGTGTGGCGATGTTCTAACGGCACTAATTTTGTCTAAGCGCTTCA I L N E Q H G Y K I A V I E N E F G E V	14280
201	TTCTGGTATCACTTCGGTATCTGTTCGAAGCCGTTCATCCTGACCGCTGGATGGC AAGACCATAGTGAAGCCATAAGACAAGCTTCGGGACAAGTAGGACTGGCGCACTACCG F W Y H F G I L F E A L F I L T A L D A	13260	14281	CTCCGTGATGATCAATTGATTGGTATCGCGCAACGAGATCAAAAACGCTGACCAACGG GAGGCACTACTAGTTAACTAACCTAGCCGCTGGCTGAGTTTGGGACTGGTGGC S V D D Q L I G D R A T Q I K T L T N G	14340
261	GGTACTCGTTCGCCGCTTTATGCTGCAAGACCTGCTGGTAACTTCATCCGCTTCCTG CCATGAGCAAGACGGGCAAAATACGACGTTCTGGACGACCCATTGAAGTAGGGCAAGGAC G T R S A R F M L Q D L L G N F I P F L	13320	14341	CTGCATCTCGTGTTCGCCGCTCAACGAGCTGGAGGACGGCTACTCGACCTGCTGGATAA GACGTAGACGACAAGCGGAGGTGCTCGACCTCCTGCGCGATGAGCTGGACGACCTATT C I C C S R S N E L E D A L L D L L D N	14400
321	AAAAAACCGACTCTCTGGTTCGCCGCTATCATCGGTACTGGCGGCTGTGGGTCTGTGG TTTTTTGGCTGAGAGCAACCGCCATAGTAGCCATGACGCGCGACACACCCAGACACC K K T D S L V A G I I G T A G C V G L W	13380	14401	TCTCGACAAGGCAATATTCAGTTCGATCGTCTGGTCAATGAATGCACCGGATGGCCGA AGAGCTGTTCCCGTTATAAGTCAAGCTAGCAGACCACTAAGTACGGTGGCGTACCGCT L D K G N I Q F D R L V I E C T G M A D	14460
381	GGTACCTGCTGATCAGGGCGTGGTGCATCCGCTGGGCGGTTAAGACCTGTGGCCCG CCGATGGACGACATAGTCCCGACAGCTAGCGACCCGCGCAATTCCTGGACACGGCGG G Y L L Y Q G V V D P L G A L R A C G R	13440	14461	TCTGGCCGATATTTCACAACTTTTTCTCCCATGAGATTTTATGCCAGCGCTACCTGCT AGGACCGGGTAATAAGTTTGGAAAAGAGGGTACTTAAATACGGTCCGATGGACGA P G P I I Q T F F S H E I L C Q R Y L L	14520
441	TGTTCCGTTATTTCCAACAGATGCTGGCAGCGTAGCGTGGTACTGGCACCGTTGTGC ACAAGCCATAAAGTTGGTCTACGACCGTCCGCACTCCGACCATGACCCGTTGGCAACAGC C S V F P T R C W Q P *	13500	14521	GGACGGCGTATTGCGCTGGTGGATCGGTCATGCGGATGAGCAGATGAACCGATTAC CCTGCCGACTAACGGCACCTACGCCATGTACGGCTACTCGTCTACTGGCTAAGTG D G V I A L V D A V H A D E Q M N R F T	14580
501	TGATTAAGATGAAGCGCACCAATACATCTGGGTAACGTTGTTCCGGCTGTATGGCTGC ACTAATTTACTTTCGGTGGGTTATGATAGCCATGACACAAGGCGGACATACCGAAG I K M K R T Q Y I W V T V V P A V W L L	13560	14581	CATCGCCAGTCCAGGTTGGTACGCGCCAGCCTATTCTGCTGACTAAAACCGACGTCGC GTAGCGGGTCCAGGTCACCGATGCGGGTGGCATAAGACGACTGATTTGGCTGCAGCG I A Q S Q V A T P D R I L L T K T D V A	14640
561	TTATCTGCACCACCTGGGCGTGGGCTGAAACTGTTACGACCAACCCGAGATGGAAG AATAGACGTGGTGGACCCGACCCAGACTTTGACAAGTCTGGTGGGCGTCTACCTTC I C T T W A L G L K L F S T N P Q M E G	13620	14641	AGGGAAGCAGAAAACTCGGTCACCGCTGGCGCGCATCAACGCCGTCACCGGCTA TCCGCTTCTTTTGGACGAGTTCGGGACCGCGGTAGTTGGGGCAGTGGCCAGAT G E A E K L R Q R L A R I N A R A P V Y	14700
621	GCTTCTTCTAGATGGCAAGCAGTACAAGAGAAGATTGCTAACGGTACTGACCTGACGG CGAAGAAGATGACCGTTTCGTGATGTTTCTCTTCAACGATTGCCATGACTGGACTGCC F E Y M A K Q Y K E K I A N G T D L T A	13680	14701	CACCGTCAACCGCGGACATGACCTGGCCGCTGCTGTTCAACCAACCGTTTTATGCT GTGGCAGTGGTGGCGCTGTAGCTGGACCCGACGACAAGTGTGGTGGCAAAAATACGA T V T H G D I D L G L L F N T N G F M L	14760
681	CGCACGAGATGGCAACATGAACACATCGTTGTGAACAACACTACCAACCGGGCCCTGA CGGTGCTCAACGGTTGACTTGGTGTAGCAACTTGTGTGATGGTTGGCCCGGACT H E I A N M N H I V V N N Y T N A G L S	13740	14761	CCAGAAAACGTCGTGAGCAACAAACCGGTTTCCACTTTATCGCGGATAAACAACGA GGTCTTTTGGCAGTGTGGTTTGGGCAAGGTGAAATAGCGCTATTGTTTGGT Q E N V V S T K P R F H F I A D K Q N D	14820
741	GTATTCTGTTCTGATTGGTGTACAGCATCATCTTCTACGGTTTCAAACCTGGCTTG CATAAGACAAGGACTAACACCACATGCTGATGAGAAGTGCACAAGTTTGGACCGAAG I L F L I V Y S I I F Y G F K T W L A	13800	14821	TATTCGTGATTTGGTGAACACTGGATTACCCGGTAGATATCAGCGAAGTTTCCCGGT ATAAAGCAGCTAACACCACCTTACCTAATGGGCCATCTATAGTCGCTTCAAAGGGCCCA I S S I V V E L D Y P V D I S E V S R V	14880
801	CGGTGGTAAACAGCGACAAACGTAAGTACTGACAAGAACAACCGTACGTTCCAATCCCGGAG GCCACGCTATGTCGCTGTTTCATGACTGTTTCTTTGGGATGCAAGGTTAGGCCTTC V R N S D K R T D K E T P Y V P I P E G	13860	14881	GATGGAAAACCTGCTGCTGGAAATCGGCGGATAAAGTCTGCGTTACAAGGGGATGCTGTG CTACCTTTGGACGACGACTTAGCCGCTATTTCACGACGCAATGTTCCCTACGACAC M E N L L L E S A D K L L R Y K G M L W	14940
861	GCGGCGTGAAGATCTCTTCGACCCACTAACCGTATTTAGCCCGCTTCGCGGGGCTTTG CGCCGCACTCTAGAGAAGCGTGGTATTGGCATAAATCGGGGCAAGCCGCCCGCAAC G V K I S S H H *	13920	14941	GGTTGACGGCAACCTAACCGCTGCTGTTCCAGGGCGTCCAGCGCTCTACACCGCCGA CCAATGCGCGTGGATTGGCGGACGACAAGTCCGCGAGTCCGCGAGATGTCCGGCT V D G E P N R L L F G G V Q L Y S A D	15000
921	TTCTTATCAGAGTGAACATGTTTGGTAACTTAGGACAGGCAAAAAATATCTGCCCGAG AAGAATAGTCTCACTTGATACAAACCATGAATCCTGTCGCTTTTTTTATAGAGCGGGTC	13980	15001	CTGGGACAGCCATGGGGGATGAAACACCGCATAGCAGATGGTGTATTATCGGATTCA GACCGTCCGGTACCCGCTACTTTGGGCGTATCGTGTCTACCACAAATAGGCTAAGT W D R P W G D E T P H S T M V F I R I Q	15060
981	GCGGGAAGATGTTGATTGGCATTCCAGACTATGACAACACTCGTACGACATATGAAGACC CGCGCTTCTACAACCTAACCGTAAGGTCGATACTGTTGATGACGCTCGTATACTTCTGG	14040			

RESULTADOS

61	GTTGCCGGAGGATGAGATTGGGGCGCGTTTGGGGGACTCAAAAAGTACCAGCCCTCCA CAACGGCCTCCTACTCTAAGCCCGCCGCAAACGCCCTGAGTTTTTCATTTGGTCGGGAGGT L P E D E I R A A F A G L K K *	15120	16141	CCAGTTAGCAATATATCCGACTGGGGTTATACACCTTAGGGTTCATTTCAAAGGCCAA GGTCAATCGTTATATAGGGCTGACCCCAATATGTGGAAATCCCAAGTAAAGTTTTCCGTT W N A I Y G S Q P N Y V K P N M E F P L	16200
121	ACATCACAATACCTGAGACTTTATCACCCTCATATTCCTCCAGAATATGAGGGCTTTCCG TGTAGTGTATGAGACTCTGAAATAGTGGGAGTATAAGGAGGCTTATACTCCCGAAAGCC	15180	16201	TAGCCCTTCCAGTCCCATTTTCCCGTACCAGGAACGGGTAATCGGGATCATGGCCTGA ATCGGAAAGGTGAGGTAAAAGGGCATGGTCCTTCCCATAGCCCTAGTACCAGACT L G K W D W K G T G P V P L R P D H G S	16260
181	GCTTAATTAFTCTGAACGTGCAACACTTCCTGCGACTCCTTATGCGCCTCCACATCCTG CGAATTAATAGAGACTTGCACGTTGTGAAGGACGCTGAGGAATACGGGAGGTGAGGAC * R Q V H L V E Q S E K H A E V D Q	15240	16261	TTGACGATCTGGATAAGCACCAGTATGAACATAACCAATATTGCCGTTTACATCAGCATA AACTGTAGACCTATTCGTGGTCATACCTTGTATTGGTTATAACGGCAAATGTAGTCGTAT Q R D P Y A G T H V Y G I N G N V D A Y	16320
241	CTTCGTTAACCAGAGCGACTTACGGCCAAAATTTTCGTACATTTTCAGTGATCTTCATA GAAGCAATGGTCTCGTGAATGCGGTTTAAAAGCATGAAAAGTCAGACTAGAAGTAT K T L W L S K R G F N E Y M K L Q D E Y	15300	16321	GTACCAGTTGATGGTCAGTGTGTTTTCGCTGCCTGCTGTGCCACTCCTGCCAATTTTT CATGGTCACTACCAGTACGAAAGGCGAGGACACAGGTGAGGACGGTAAAAA Y W N I T L A Q K A A Q Q T W E Q W N K	16380
301	GTGCTTATCAACTGTTCCATCGGAGCAATAAACCCACTCTGACCGGGTGCAGCACATC CACGAATAGTTGACAAGGTAGCCCTCGTTATTGGGTGAGACTGGCCACGCTGGTGTAG H K D V T G D P A I F G S Q G P A V V D	15360	16381	GGCCTTCATCTGATGAGTCCAGGCCAGAAAGACCCCTCTTACCATCCCATGCGCG COGGAAGTAGACTACTCAGGTCCGGTCGTTTCTCGGGTGGAGAAATGGTAGGGTACGGC A K M Q H T W A L L S A V E K G D W A R	16440
361	CCAGGCAAGCAGCAGGATCGTTCGTGTTGGTGAAGAAACAATCATATCGTTTTCTGT GGTCCGTTTCGTCTCTGCTAGCGAACGACCACTCTTTTGTAGTATAGCAAAGACA W A L V P R D S T T P S F V I M D N E T	15420	16441	GGATTAGCGTAAGCCGTTTGTGCTGCTGGTCAGTTTGGAGAAATGTTGCCATGCACCGT CCTAAATCGCATTCCGCAACACAGCAGACCAGTCAAACTCTTACAACGGTACGTGGCA S K A Y A T Q T T Q D T Q L I N G H V T	16500
421	TCCAGGTTTIGTACTCCGCTGATGACCGGTTTCTCCGCTGCGGCTCGGGTACACC AGGTGCCAAAATAGAGCGGACTACTGCGCAAGAGGCGAGCCGCGGACGCCATGG G R N Q Y E A Q H R T E E A A A Q P V G	15480	16501	ACGCCAGACAGTAAAGTCTCTGCTGACCATTTTTTCAACGTAATGGTTTCTCAGCGCT TGCGGTCTGTCAITTCAGAGACGGACTGGTAAAAAGTGGCATTACCAAGGAGTGGCGA R W V T F T E A Q G N K V T I T E E R S	16560
481	AAAGAAATTAATTTGCCCGGAACGTTAAGGCCATTCAGAGTGTTCACGTTACTCACATT TTTCTTAATAAAGCGGCTTGAATTCGGTAAAGTCCCAAAAGGTCAATGAGTGTAA F E N N A R F T L A M A P T K W N S V N	15540	16561	TAACATTTTCAACCTTACCATTATGCAAGTAGTACCGTGTTCCTCTGCCGACAGCG ATTGTAAGTGGTGAATGGTAATACGTTTCAATCATCGGACCAAGAGACGGCTGTGGC L M K V W K G N H L Y Y G P K E A S L R	16620
541	ATTGCCATAGCGTTTGGAAAGAGTCTCCAGGATCTCTCCAGCGAGCAACACAACTC TAACGGTATCGCAAACTTTCTCAGAGGTCATAGAAAGTTCGGTTCGTTGTGGAG N G Y R K S L T E W T D E L A A L V E	15600	16621	TTCAGCAAAAATATCGACATCATCGCGAACTTCCGCTGATCCCGGAAATCACACC AAGTCGTTTTTATAGCTGTAGTAGCGGCTTGGAGCGCAACTAGGGTCTCTTAGTGTGG E A F I D V D D G P Y A F P T N G S G S I V G	16680
601	CTGCTGTGGTTTCCAGCAACAGATCAACCGCTGTGGATTGGTGATTTGTCTCCCTG GACGACCAAAAGGTCGTTTGTCTAGTTGGCGGACCCCTAACCACTAAACAGAGGGAC Q Q P K G A F L D V A Q P I P S A K D G Q	15660	16681	ATTATGACCAAAAACAGCCAGGATAGGCAATGGTGTATTGCCAGTGACATCATAACC TAATAGCTGTTTTGGTCCGGTCTATCCGTTTACCACATAACGGTCACTGTAGTATTGG N H G F V L G P Y A F P T N G S G S I V G	16740
661	CACCGCTCATACAAAATTTTGTCCAACTTATATTGAGCAACAGTTGGGCCGTC GTGGCGGAGTATGTTTAAAAACGAGTGTGAATATAAGTGGCTTGGTCAACCCGCGAG V A E Y L I K A G V S I N L S G T P G D	15720	16741	AGCACCGTGCAGCAATACCATAAGTATACGAGCGCATACCAGCAAACTGGGGACC TCGTGGCAGCTGTGGTTATGGTATTTCATATCGCTCCGGTATGGTCGGTTTACGCGTGG A G H L G I G Y T Y A P A Y W G F Q P G	16800
721	CTGGSTTGTTCGTAGCCACTGGCGCTGTACCATTATCAAAATGGCATAGTACGGCAGC GACCAACAAGCATCGGTGACCGGACATGGTGAATAGTTTACCGTATCCATGCGCTG Q T T E Y G S A S Y W K D F P M P V A A	15780	16801	ATTTACCATGATTCGCTTCCGATCCTGGGCTTTCGTTTTGCCGATAACCCACATATTGCT TAAATGGTACTAACGAAAGCGTAGGACCCGAAACGAAACGGCTATTGGGTGTATAACGA N V M I A K A D Q A K S K G I V W M N S	16860
781	CACTACGGTACGCTTCAACATACTGTCAGCCAAACGTTCCAGATGGCAGAGCCTGGCTG GTGATGCCATGCAAGTGTATGACCACTGGTTTGAAGTCTACCGTCTGGACCGAC V V T R K L M S T L W V N L I A S G P Q	15840	16861	GGTCGTTGGATACCCCGCCAGCATTGGCACCACCCTGTGCAAAATGTGCAACAATAGT CCAGCAACCTATGGGGCGGTCTGGTAACCGTGGTGGGACAGTTTAAACAGTGTATTCA T T P Y G A L G N A G G Q A E Q A V I T	16920
841	CTGCCAGTTTTACCATCATATTAAGCAAATGATGCCATCCCAAGTGTAAATGTTTC GACGGTCCAAAATGGTAGTAATTCGTTTAACTACGGTAGGGTGCACAATTACAAAG Q W T K G D D N L L N I G D W R T L T E	15900	16921	TTCCCGTCTCTCCCTGCTGTTAACGCCAGCAGTGGCCATCCGCCCTTTTGTGGTGTG AAGGGCAAGAAGGGACGACAATTCGGTTCGTCAGCGGTAGGGGGGAAACGACCAGC E R N K G A T L A L L A G D A G K A P R	16980
901	TACCAACTGACGACGGGATCGCTCTGTCTCAAACAGATGTCGCTGCTGACAGAGTAGG ATGGTTGACTGCTGCGCTAGCAGACAGTGTGGTCTACAGGACGAACGCTCTCATCC V L Q R R P D S Q T L G S T A A Q L T P	15960	16981	GTCAAGCATTGGTGCAGGTAATCGTAGCGTGGCAACAGAGTGTGTTGGCGAGTTTTG CAGTTCGTAACCACTCCATTTAGCATCCAGCGTTGCTTCGACGACAACCGCTCAAAC D L M P A P L D Y R P L L A A T Q S N Q	17040
961	TAAAAAAGCCTCAGGTTAAGATCCTGACGACTGTTTGGGGAATAACATCCCATGCCTG ATTTTTTTCGGAGTCCAATTCAGGACTGCTGACCAAACCGCTTATTGTAGGTTACGGAC L F L R L N L D Q R S T Q R I V D W A Q	16020	17041	CTGATTAATTAAGTGGGTAGTTACTCTCTGACGCAATAGTGGTGGCGCTGATGG GACTAATTTAATTCACCCATCAATGAGAGAATCGCGTATACCAACCCGGACTACC Q N F K L P Y N S E Q V A I T T P A S P	17100
1021	ATCAGCAGTTAAGCGTGGCTTTGCTCAAGCAGTCCGTCGATCTCCGTAACCGGATCTGC TAGTCGTCATTCGACCCGAAACAGGTTTCGTCAGCCAGTACAGGCAATGCGCTAGACG D A T L R P K Q E L L R D I E T V R D A	16080	17101	GTTTACCAGCCATTTCAACTGATTAATACCAGCATGCCTGTGATACACCATATTTATC CAAATGGTCGGTAAAGTGTACTAATTTATGGCGGTACGGAACACTATGGGTATAAATAG N V L W K L Q N E V A M G Q S V G Y K D	17160
1081	ACCACCCCAAAAAGGCAACAGATCTGAAGCGGGATAATCTTTTGGGGAGAATGTGT TGGTGGGGTGTTCGTTGCTAGACTTCGCCCTATTAGAAAAACCCCTCTTAACAA G G W L F A F L D S A P Y D K Q P S N N	16140	17161	TTTTAAGCGTTAGCAGTCCAGATATCAATTTCCGCTAGTGTATCAGAGAAGCGGTT AAAATTTCCGCAATCGTCAGGCTAATAGTTAAAGCGATCAGATAGTCTCTCGCCAA K L A T L L A L N D I E S T S D S F R N	17220

7221	TGCCATGGTGCCCAAAATATCATCGCGACATCAAACGGTTC	17280	17521	CATCTGAAAAGGGCATCTTGCTACTACATAGCCATAGCCATAAAATAGGTGCCATGT	17580
	ACGGTACCACGGGTGTTTATAGTAGCGCTGTAGTTTGCCAAGGGTCGCGAATCCTCATT			GTAGACTTTTTCCGCTAGAACACGATGATGTATCGGTATCGGTATTTATCCACGGTACA	
	A M T G V F I M A V D F P E W R K P T F			M Q F L R D Q A V V Y G Y G Y F L H W T	
7281	GCCAAATGTATTAACGTGTTTGGTAAGAGCGTCTCTGGATTGGTATTTACCTTATCAAT	17340	17581	ATCATTGGCATAAAATATGGCGCATGCCGTATTTCATCGGAACAATCTTTATCTCACTTGA	17640
	CGGTTTACATAATTGACAAAACCATCTCGCAGAGACCTAACCATAAATGGAATAGTTA			TAGTAACCGTATTTATACCGCTACGGCATAAGTAGCGCTTGTAGAAATAGAGTGAAC	
	G F T N F Q K P L L T E P N T N V K D I			D N A Y I H P M G Y E D R V I K I E S S	
7341	CCAGGCATTCATCCATCAGCGTAGCCTTGACAGAAATGGACATATCCTCTGGGAAAGGSC	17400	17641	CGACTGCTCAGCCAGTGCAGGTAAGCTCCAATAATACATCAGGGAAGCAGTAACACAGTT	17700
	GGTCCGTAAGTAAGGTAGTCGCATCGGAACGTCCTTACCTGTATAGGAGACCCCTTCCCG			GCTGACGAGTGGTCACGTCCATTTCGAGGTTATTATGTAGTCCCTTCGTCATTGTGTCAA	
	W A N M G D A Y G Q L I S M D E P S L A			S Q E A L A P L S W Y Y M L S A T V C N	
7401	AGCAATTTGCGCCGGATAGCATCGGCCAGTAGTTACGACGGATATCTTTATCAAATTT	17460	17701	CACGATCATAAGTATTTCTATTTTTCATTGTATCCTCTGGCAGGTGTATAATTAGCGAATT	17760
	TCGTTAAACGCGGGCTTCTAGGCGGTTCATCAAGTCTGCCTATAGAAATAGTTTAAA			GTGCTAGTATGCTAAAGATAAAAAGTAACATAGGAGACCGTCCACATATTAAATCGCTTAA	
	A I Q A R I A D P W Y N R R I D K D F K			V I M R N R N K M ← <i>pac</i>	
7461	CACAAAATCTTTGCCAAGCACTTCGGGACAGTCCCTTGAGTACTGCGAOGTGCCATTTC	17520	17761	GATACTAGCAACGAAGCTT	17779
	GTGTTTGTAGAAACGGTTCGTGAAGCGCTGTGAGGAACTCATGACGCTGCACGGTAAAG			CTATGATCGTTGCTTCGAA	
	V F D K G L V E A V T G Q T S R R A M E				

Figura 23. Secuencia de nucleótidos y aminoácidos de los genes y enzimas que componen la ruta catabólica del 4-HPA. En la figura está indicada la secuencia de nucleótidos de los genes de la ruta del 4-HPA así como la secuencia de aminoácidos de sus productos: *hpaR* (represor del operon *meta*); *hpaG* (5-oxo-pent-3-en-1,2,5-tricarboxilato (OPET)/2-hidroxi-hept-2,4-dien-1,7-dioato (HHDD) descarboxilasa/isomerasa); *hpaE*, (5-carboximetil-2-hidroxi-mucónico semialdehído (CHMS) deshidrogenasa); *hpaD* (homoprotocatecuato (HPC) dioxigenasa); *hpaF* (5-carboximetil-2-hidroxi-mucónico (CHM) isomerasa); *hpaH* (2-oxo-hept-3-en-1,7-dioato (OHED) hidratasa); *hpaI* (2,4-dihidroxi-hept-2-en-1,7-dioato (HHED) aldolasa); *hpaX* (codifica para una proteína perteneciente a la superfamilia de transportadores transmembranales); *hpaA* (regulador del operón *hpaBC*); *hpaB* (componente de la 4-HPA-hidroxilasa); *hpaC* (proteína cooperadora de la 4-HPA-hidroxilasa). También está indicada la secuencia de nucleótidos de los genes situados en las regiones flanqueantes al *cluster*: *tsr* (codifica para la proteína quimiorreceptora de serina); ORFs 12, 13 y 14 (codifican para proteínas de función desconocida); *pac* (codifica para la penicilina G acilasa). En la figura aparece subrayada la secuencia correspondiente a los *primers* utilizados en la secuenciación: 3232A, 32RA, 333A, HIDROPRO, O38B-S, OB1A-I, OB2A-H, OB2B3-9, OB2K3 y 5, OB41-7, OB61-5, OD61-3, ORIA-F, PACEN, PAJ19A-B, PAJ19AC, PAJ25R, PAJ26D, PAJ32B, PAJ32T, PAJ33B, PAJ36R, PAJ37A-C, PAJ38A. También está subrayada la secuencia correspondiente a los posibles sitios de unión de la CAP y los posibles terminadores de la transcripción. Las secuencias REP se indican con una línea discontinua. Los sitios de la iniciación de la transcripción están señalados (+1), así como las secuencias promotoras -10 y -35. La secuencia presentada aparece en la base de datos GeneBank/EMBL con el número de acceso Z37980.

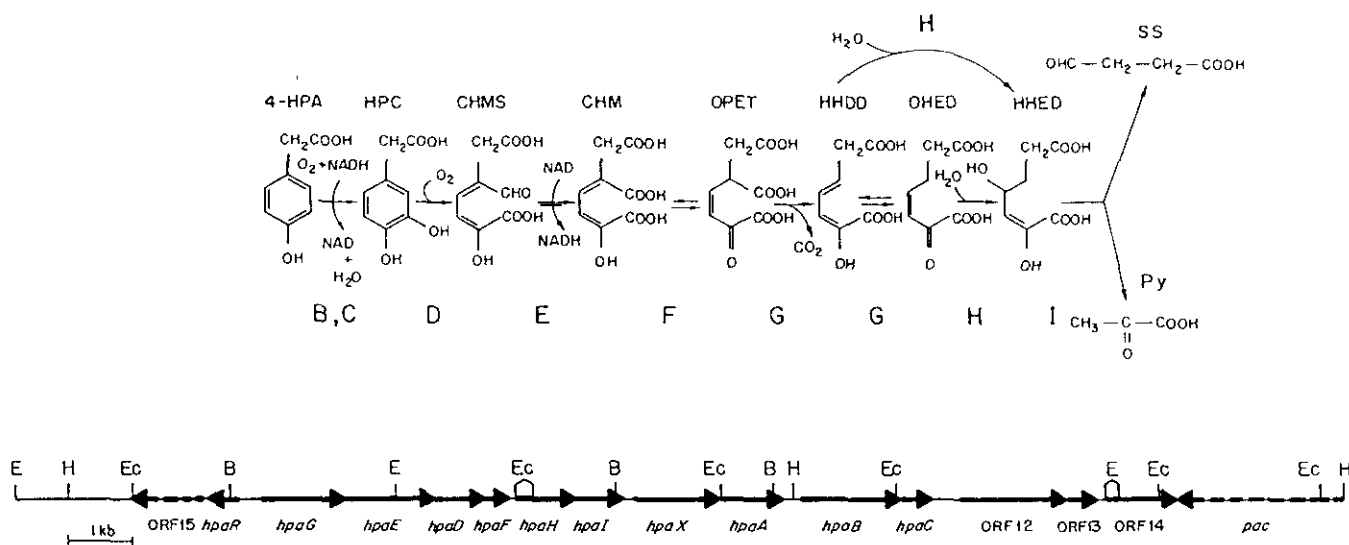


Figura 24. Pasos enzimáticos propuestos de la ruta catabólica del 4-HPA. Los pasos enzimáticos fueron propuestos por Roper y cols., (1993). Las flechas indican la dirección de transcripción de los genes. Las flechas en línea discontinua indican los genes que han sido parcialmente secuenciados en este trabajo. Los sitios de restricción son como siguen: B, *Bam*HI; E, *Eco*RI; Ec, *Eco*RV; H, *Hind*III; Los metabolitos intermedios de la ruta del 4-HPA están indicados en la parte superior de la figura. Abreviaturas: HPC, homoprotocatecuato; CHMS, 5-carboximetil-2-hidroxi-mucónico semialdehído; CHM, 5-carboximetil-2-hidroxi-muconato; OPET, 5-oxo-pent-3-en-1,2,5-tricarboxilato; HHDD, 2-hidroxi-hept-2,4-dien-1,7-dioato; OHED, 2-oxo-hept-3-en-1,7-dioato; HHED, 2,4-dihidroxi-hept-2-en-1,7-dioato; Py, piruvato; SS Succinato semialdehído. B, C, D, E, F, G, H e I representan las enzimas codificadas por los respectivos genes.

HpaD	MGKLALAAKITHVPSMYLSELPGKNHGCRQGAI DGHKEI SKRCREMVDTIIV	53
HpcB	MGKLALAAKITHVPSMYLSELPGKNHGCRQGAI DGHKEI SKRCREMVDTIIV	53
HpaD	FDTHWL VNSAYHINCADHFEGVYTSNELPHFIRDMTYNYEGNPELGQLIADEALKLVRA	113
HpcB	FDTHWL VNSAYHINCADHFEGVYTSNELPHFIRDMTYNYEGNPELGQLIADEALKLVRA	113
HpaD	KAHNIPSLKLEYGTLVPMRYMNEDKHFKVVSISAFCTVHDFADSRKLGEAILKAIEQYDG	173
	: : :	
HpcB	KAHNIPSLKLEYGSVPMRYMNEDKRFKVVSISAFCTVHDFADSRKLGERIVKAIEQYDG	173
HpaD	TVAVLASGSLSHRFIDDQRAEEGMNSYTRFDQRQMDERVVKLWREGQFKEFCNMLPEYAD	233
HpcB	TVAVLASGSLSHRFIDDQRAEEGMNSYTRFDQRQMDERVVKLWREGQFKEFCNMLPEYAD	233
HpaD	YCYGEGNMHDTVMLLGMLGWDKYDGKVEFITELEFPSSGTGQVNAVFPPLPA	283
	: : : : : :	
HpcB	YCYGEGNMHDTVMLLGMLGWDKYDGKVSLSPSYSQASWHRSG	276

Figura 27. Alineamiento de la secuencia de aminoácidos de las dioxigenasas HpaD y HpcB. HpcB es la HPC-dioxigenasa de *E. coli* C (Roper y Cooper, 1990a). | y : indican, respectivamente, aminoácidos idénticos y conservados.

HpaI	MENSFKAALKAGRPQIGLWLGLSSSYSSELLAGAGFDWLLIDGHEAPNNVQTV	53
HpaI	LTQLQAIAPYPSQPVVRPSWNDPVQIKQLLDVGTQTLVPMVQNADEAREAVRATRYPPA	113
Rnpb1	DIGFYNFLIPFVETKEEAELAVASTRYPPE	30
HpaI	GIRGVGSALARASRWNRI PDYLQKANDQMCVLVQIETREAMKNLPQILDVEGVDGVFIGP	173
Rnpb1	GIRGVSVS-HRANMFGTVADYFAQSNKNITILVQIESQQGVDNVDIAAATEGVDGIFVGP	89
HpaI	ADLSADMGYAGNPQHPEVQAAIEQAI VQIRESGKAPGILIANEQ LAKRYLELGFVAVG	233
Rnpb1	SDLAAALGHLGNASHPDVQKAIQHI FNRSAH GKPSGILAPVEADARRYLEWGATFVAVG	149
HpaI	VDTLLARAAEALAARFGAQA TAVKPGVY	262
Rnpb1	SDLGVERSATQKLADTFKK	168

Figura 29. Alineamiento de la secuencia de aminoácidos de la aldolasa HpaI con una proteína de función desconocida de *E. coli*. La proteína Rnpb1 truncada está codificada por un gen de función desconocida localizado cerca del gen *rnpB* (Komine e Inokuchi, 1991). | y : indican, respectivamente, aminoácidos idénticos y conservados.

HpaR	MINCFYINDFEFFRGKIMHDSLTIALLQAR	30
HpcR	MINCFYINDFEFFRGKIMHDSLTIALLQAR	30
HpaR	EAAMSYFRPIVKRHNLTQQWRIVRILAESPMDFHDLAYRACILRPSLTGILTRMERDG	90
HpcR	EAAMSYFRPIVKRHNLTQQWRIVRILAESPMDFHDLAYRACILRPSLTGILTRMERDG	90
HpaR	LVLRLKPIVDQKLYISLTKEGQALYNRAQTQIEEAYRQIEAQFTAQEKMQQLTHLLEEFI	150
HpcR	LVLRLKPIVDQKLYISLTKEGQALYNRAQTQIEEAYRQIEAQFTAQEKMQQLTHLLEEFI	150
HpaR	ALGNSRQEDIPGDNE	165
HpcR	ALGNSRQEDIPGDNE	165

Figura 30. Alineamiento de las secuencias de aminoácidos de las proteínas HpaR y HpcR. HpcR es la proteína represora de la transcripción del operón *meta* de la ruta catabólica del HPC de *E. coli* C (Roper y cols., 1993). | indica aminoácidos idénticos.


```

HpaA      MCDRQIANIDISKEYDESLGTDVHYQSFARMAPFLASMLPHRHEQYFQM 50
          ||:::  ::  :  ::  ::::  ::::  |  :::::  :|:|:|  :::
MelR      MCSSEDEKQTRSPLSLYSEYQRMEIEF----RAPHIMPT--SHWHGQ-VEV 49

HpaA      HFLNSGQIELQLDDHRYSEAPLFVLTPPSPVPHAFITESDADGHV---LT 97
          ::  :|:|:|  :::::  :::::  ::|  :::|:|:|:|:|:|:|:|  |:
MelR      NVPFDGDVVEYLINNEKVNINQGHITLFWACTPHQLTDTGTCQSMAlFNLP 99

HpaA      VREDLIWPLLEVLYPGTRETFFGLPGICLSLADKPDELAALAHYWQLIERE 147
          ::  |  |||  :  |:  ::  |  |:  ::  |||  :::|:|:|  :  |  :
MelR      MHLFLSWPLDKDLI--NHVTHGM--VIKSLA--TQQLSPFEVRRWQQELN 143

HpaA      S----VEQLPGREHTLTL--LAQAVFTLLLRNAKLDDHAASGMRGELKLF 191
          |  :  :|:|  |  :|  |  ::  ::::  :|  |  :  :|  :|  :  ::
MelR      SPNEQIRQLAIDEIGLMLKRFSLSGWEPIILVNKTSRTHKNSVSRHAQFYV 193

HpaA      QRFNMLIESHFHQHWTPDYANELHITESRLTDICRRFANRPPKRLIFDR 241
          ::  :|:|:|:|  :|:|  |  |:|:|:|:|:|  :|  :|:  :  :  |:|  ::
MelR      SQMLGFIAENYDQALTINDVAEHVKLNANYAMGIFQRVMQLTMKQYITAM 243

HpaA      QLREAKRLLLFSDNAVNNIAWQLGFKDPAYFARFFNRLVGCSPSAYRAKK 288
          :::::  ||  :|:|:|  :|:|  ||:|:|  |  :  |:|:|  ||  :||
MelR      RINHVRRLLSDTDKSILDIALTAGFRSSRFYSTFGKYVGMSPQQYRKLS 290
          **  **  **  *  *  **  *

HpaA      VPVT      I-DIA---GF-S--YF---F----G-TPS--R
          :
          :
MelR      QQ
  
```

Figura 31. Alineamiento de las secuencias de aminoácidos de las proteínas HpaA y MelR de *E. coli*. MelR es un miembro de la familia de reguladores XylS/AracC. La secuencia consenso del motivo conservado en la región C-terminal de esta familia está indicada en la parte inferior de la figura. La posible región implicada en la unión al DNA está subrayada (Gallegos y cols., 1993). | y : indican, respectivamente, aminoácidos idénticos y conservados. Los asteriscos indican los aminoácidos idénticos a la secuencia consenso.

HpaX	MSDTSPAIPESIDPANQHKALTAGQQAVIKKLFRRLLIVFLFVLFIFSFLDRINIG	55
	: : : :: : :::: : :	
Pht1	MTTAAATDAVPHLLQRSHERIEKVYRKVTLRLMTFIFVAWVNLNYLDRVNIS	51
HpaX	FAGLTMGRDLGLSATMFGGLATTLFYAAYVIFGIPSNIMLSIVGARRWIATIMVLWGIAS	115
	: : : : : : : : : : : : : : : :	
Pht1	FAQVYLKHDLGMSDADLRLRRKLVFHLRHRIGNTQYAYLQKIGARLTI TRIMVLWGLISA	111
HpaX	ATMFATGPTSLSYLRLIVGITEAGFLPGIILLYLTFWFPAYFRARANALFMVAMPVTTALG	175
	: : : : : : : : : : : : : : :	
Pht1	SMAEMTTPTEFYIARALLGAAEAGFWPGIILYLYWYPGARRARITSRFLLAIAAAGIIG	171
HpaX	SIVSGYILS-LDGMALKGWQWFLLEGFPSVLLGVMVFWLDDSPDKAKWLTKEKDKCL	234
	: : : : : : : : : : : : : : : : : : : : : :	
Pht1	GPLSGWILTHFVDVMGMKNWQWFMFILEGLPAAVMGVMAYFYLVDPKPEQAKWLDDEEKSI	231
HpaX	QEMMDNDRLTLVQPEGAI SHHAMQQRSMWREIFTPVVMYTLAYFCLTNTLSAISIWTPQ	294
	: : : : : : : : : : : : : : : : : : : : : :	
Pht1	LDALAADRAG-KKPVTDKRHAVLAALKDPR-VYVLAAGWATVP-LCGT----ILNYWTPT	284
HpaX	ILQSFNQSSNITIGLLAAVPQICTILGMVYWSRHSRRRQERRHHTALPYLFAAGW-LL	353
	: : : : : : : : : : : : : : : : : : : : : : :	
Pht1	IIRN-TGIQDVLHVGLLSTVPYIVGAIAMILIARSSDIRLERRKHFFFSIAFGALGACLL	343
HpaX	ASATDHNMIQMLGIIMASTGSFSAMAI FWTTDPQISLARAIGIAVINATGNIGSALS	413
	: : : : : : : : : : : : : : : : : : : : : :	
Pht1	PHVVDSAIISITCLAMIAVSYFGAAAI IWSIPPAYLNDESAAGGISAISSLGQIGAFCAP	403
HpaX	FMIGWLKDLTGSFNSGLWFVAALLVIGA-GIIWAI PMQSSRPRATP	458
	: : : : : : : : : : : : : : : : : :	
Pht1	IGLGWINTVTGSLAIGLTIIGALVLAGGMVAVLIAVPANALSEKPLTDE	451

Figura 32. Alineamiento de la secuencia de aminoácidos de la proteína HpaX con la proteína Pht1. Pht1 es una proteína de la ruta de degradación del fitato de *P. putida* perteneciente a la superfamilia de transportadores transmembranales (Nomura y cols., 1992). | y : indican, respectivamente, aminoácidos idénticos y conservados.

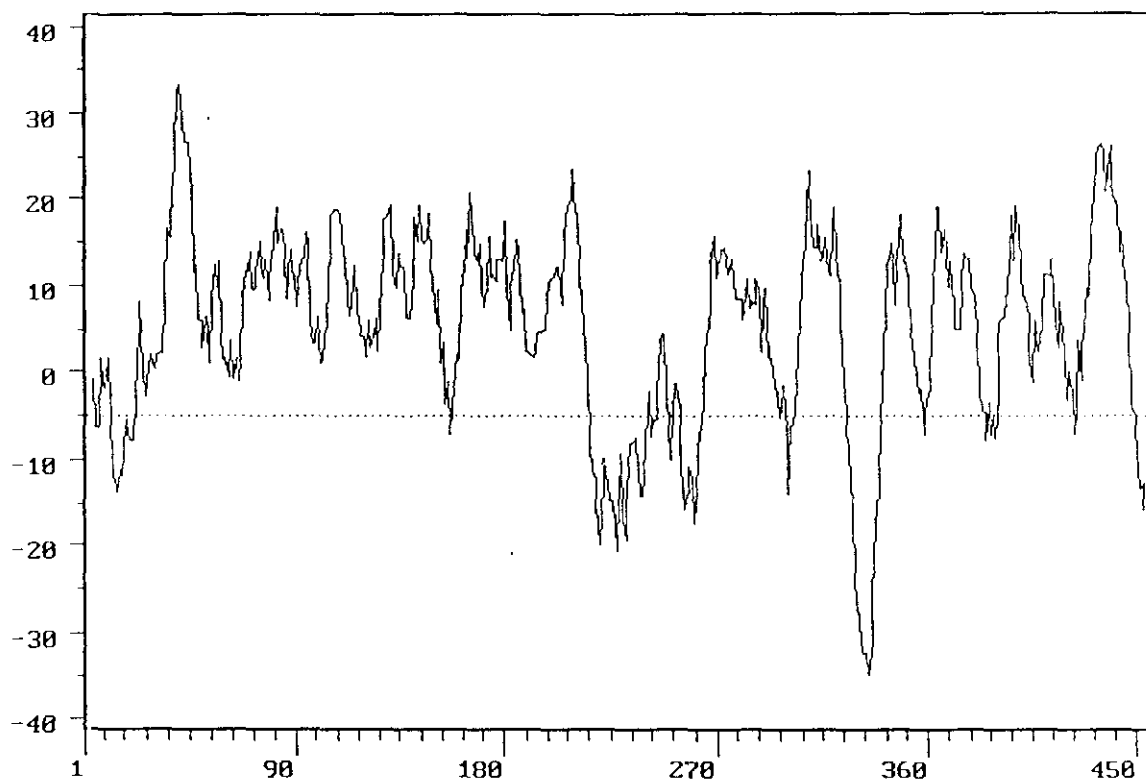


Figura 33. Perfil hidrofóbico de la proteína HpaX. Se llevó a cabo por el método de Kyte y Doolittle (1982).

3.3. Análisis de las regiones flanqueantes a la ruta del 4-HPA y su localización en el cromosoma de *E. coli*

El análisis de las regiones flanqueantes mostró la presencia de cinco ORFs (figuras 23 y 24). La ORF15 estaba incompleta y era similar a la proteína quimiorreceptora de serina codificada por el gen *tsr* de *E. coli* K12 (Ames y Parkinson, 1994). Esta ORF también fue localizada a 119 pb del codon de terminación del gen *hpcR* de *E. coli* C (Roper y cols., 1993). La ORF12 era similar a la proteína CstA de *E. coli* MC4100, la cual se expresa durante la fase estacionaria de crecimiento y en estados carenciales de fuente de carbono (Schultz y Matín, 1991), y a los primeros 567 aminoácidos de la ORF721 de *E. coli* MG1655 (GenBank/EMBL ac. U14003) (figura 34). La ORF13 era similar a una proteína de función desconocida (Cst2) que forma parte del mismo operón que la proteína CstA (Schultz y Matín, 1991) y al C-terminal de la ORF721 de *E. coli* MG1655 (GenBank/EMBL ac. U14003) (figura 34). Por tanto, la ORF721 de la estirpe MG1655 sería equivalente al producto de la fusión de las ORFs 12 y 13 de *E. coli* W. Esta fusión podría estar provocada por una mutación en la ORF721, que originara un cambio de fase mediante la inserción de un nucleótido en la posición equivalente al nucleótido 13.421 de la figura 23. En este sentido, Schultz y Matín (1991) demostraron, mediante la técnica de maxicélulas, la producción de las dos proteínas del operón de la CstA equivalentes a las ORFs 12 y 13 de *E. coli* W.

La ORF14 era muy similar al producto del gen *yjiA* de función desconocida, localizado en la región 5' del gen *mrr* en diferentes estirpes de *E. coli* K12 (Waite-Rees y cols., 1991; GenBank/EMBL ac. U14003), y a la proteína P47K de *P. chlororaphis* B23 implicada en el metabolismo del nitrilo (Nishiyama y cols., 1991) (figura 35). Las ORFs 12, 13 y 14 se transcriben en la misma dirección que los genes que componen los operones *hpaBC* y *meta*. Por el contrario, el gen *pac*, localizado a 82 pb del codon de terminación de la ORF14, se transcribe en el sentido opuesto (figura 22). El gen *pac* no fue secuenciado en su totalidad ya que su secuencia había sido determinada anteriormente por otros autores (Oh y cols., 1987).

El hecho de que el gen *cstA* haya sido localizado en el minuto 14 del mapa genético del cromosoma de *E. coli* W3110, mientras que los genes *tsr* y *mrr* se mapean en los minutos 99 y 98,5 respectivamente (Bachmann, 1990), indicaría que la adquisición de los genes de la ruta del 4-HPA se había producido simultáneamente a un reordenamiento de estos genes en el cromosoma de *E. coli* W, o que el cromosoma de esta estirpe es muy diferente al de *E. coli* K12 W3110. Sin embargo, en la estirpe MG1655 de *E. coli* K12, los genes *tsr*, ORFf721, *yjiA* y *mrr* están situados en una posición adyacente en el cromosoma entre los minutos 92,8 y 0,1 (GenBank/EMBL ac. U14003) (figura 36). De acuerdo con este dato, se podría sugerir que el *cluster* de genes de la ruta del 4-HPA se habría insertado en el cromosoma entre el gen *tsr* y la ORFf721, en las posiciones correspondientes a los nucleótidos 73 y 11.450 de la figura 23 (figura 36). Estos nucleótidos están localizados en regiones no codificantes, a 57 y 101 pb de los codones de terminación de los genes *hpaR* y *hpaC*, respectivamente. Además la comparación de las secuencias de las estirpes MG1655 y W reveló la existencia de una inserción entre los nucleótidos 11.602-11.727 de la figura 23, en la región inmediatamente anterior al codon de iniciación de la ORFf721 (figura 36).

Por otra parte, en el apartado 2.7 se comentó que se había localizado la hidroxilasa de *E. coli* C en uno de los extremos del fragmento *HindIII* del plásmido pHCl. La secuenciación del otro extremo reveló la presencia de una ORF similar a la ORF14. La secuencia de nucleótidos de esta región era prácticamente idéntica a la de *E. coli* ATCC 11105 hasta el codon de terminación de la ORF14. A partir de este punto, se pierde completamente la similitud entre las secuencias de las estirpes MG1655, W (figura 36) y C (datos no mostrados). Asimismo, se midió la actividad PGA en las estirpes B y C de *E. coli* utilizando las mismas condiciones en el ensayo que para la PGA de *E. coli* ATCC 11105 y se analizó, mediante Southern blot, la presencia de genes homólogos al gen *pac* en estas estirpes (datos no mostrados). Estos experimentos indicaron que el gen *pac* no está presente en las estirpes B y C. Estos resultados sugerirían que el gen *pac* podría haberse insertado en el cromosoma de *E. coli* W en la región adyacente al codon de terminación del gen *yjiA*.

CstA	MNKSgKYLWVTVLSVMGAFALGYIALNRGEQINALWIVVASVCYIYLIARFYG	53	CstA	AAQVSS---WGFSITPDTLNQIASEVGEQSII--SRAGGAPTL--AVGMAYILHGALG-	463
ORF12	MPGFTMDTKKLFKHI PWVILGIIGAFCLAVVALRRGEHVSALWIVVASVSVYLVAIRYYS	60	ORF12	NAPI IMAQLKDVTAH~AQT VTSWGFVISPEQILQTAKD IGEPSVLNRAGGAPTLAVGIAH	479
ORFf721	MPGFTMDTKKI PKHI PWVILGIIGAFCLAVVALRRGEHISALWIVVASVSVYLVAIRYYS	60	ORFf721	NAPI IMAQLKDVTAHAATAATVSSWGFVISPEQILQTAKD IGEPSVLNRAGGAPTLAVGIAH	480
CstA	LYIAKNVLAVDPTRMTPAVRHNDGLDYVPTDKKVLFGHHFAAIAGAGPLVGPVLAQMGI	113	CstA	-----GMDVAFWYHFAILFEALFILTAVDAGTRAAREFMLQDLLGVVSPGLKRTDSLPA	517
ORF12	LYIAQKVMKLDPTTRATPAVINNDGLNYVPTNRYVLFQHHFAAIAGAGPLVGPVLAQMGI	120	ORF12	VFHKVL PVADMGEFWYHFGILFEALFILTALDAGTRSARFMLQDLLGNFIPFLKKTDSLVA	539
ORFf721	LYIAQKVMKLDPTTRATPAVINNDGLNYVPTNRYVLFQHHFAAIAGAGPLVGPVLAQMGI	120	ORFf721	VFHKVL PMADMGEFWYHFGILFEALFILTALDAGTRSARFMLQDLLGNFIPFLKKTDSLVA	540
CstA	LPGMIWLLAGVVLAGGVDQDFMVL FVSTRDRGRSLGELVKEEMGPTAGVIALVACFMIMVI	173	CstA	NLLATALCVLAWGYFLHQGVVDPLGALTLGCRCLVLPTRCWQWR	562
ORF12	LPGLTWLLAGVVLAGAVQDFMVLFISSRRNGRSLGEMIKKEEMGPLPGTIALFGCFLIMI I	180	ORF12	GIIGTAGCVLWGYLLYQGVVDPLGALRACGRCSVFPTRCWQP	582
ORFf721	LPGLTWLLAGVVLAGAVQDFMVLFISSRRNGASLGEMIKKEEMGPPGTIALFGCFLIMI I	180	ORFf721	GIIGTAGCVLWGYLLYQGVVDPLGGVKS LWPFLGINSQMLAAVALVLGTVVLIKMKRTQ	600
CstA	ILAVLAMI VVKALTHSPWGTYTVAFTIPLALFMGIYLYRYPGRIGEVSVIGLVFLIFAI	233	Cst2	MLCAVVLFKMKRQR	14
ORF12	ILAVLALIVVKALAESPWGVFTVCSTVPIALFMGIYMRIFIRPGRVGEVSVIGIVLLVASI	240	ORF13	MLAAVALVLGTVVLIKMKRTQ	21
ORFf721	ILAVLALIVVKALAESPWGVFTVCSTVPIALFMGIYMRIFIRPGRVGEVSVIGIVLLVASI	240	Cst2	YAWVALVPTAWLLICTLTAGWQKAFSPDAKV~GFLAIANKFDAMIDSGNIPSYTESQLA	73
CstA	ISGGWVAESPWAPYFDFTGVQLTWMLVGYGFVAAVL PVWLLAPRDYLSFLKIGTIVG	293	ORF13	YI WVTVVPAVWLLICTT WALGLKLFSTNPQMEGFFYMAKQYKEKIANGTDLTAHEIANMN	81
ORF12	YFGGVIAHDPYWGPAITFKDTTITFALIGYAFVSALLPVWLLAPRDYLATFLKIGTIVG	300	ORFf721	YI WVTVVPAVWLLICTT WALGLKLFSTNPQMEGFFYMASQYKEKIANGTDLTAQQIANMN	660
ORFf721	YFGGVIAHDPYWGPAITFKDTTITFALIGYAFVSALLPVWLLAPRDYLATFLKIGTIVG	300	Cst2	QLVFNRLDAGLTIFFM VVVV LALPSIKTALAALKDPKPTAKETPYEFPENVEEIVAQ	133
CstA	LAVGILIMRPLTMPALTKEVDGTGPVWTGNLFPFLFITIACGAVSGFHALISSGTTPKM	353	ORF13	HIVVNNYTNAGLSILFLIVYSIIFYGKTLWAVRNSDKRTDKETPYVPIPEGGVKISSH	141
ORF12	LALGIVVLPNPELKPAMTQYIDGTGPLWKGALFPFLFITIACGAVSGFHALISSGTTPKL	360	ORFf721	HIVVNNYTNAGLSILFLIVYSIIFYGKTLWAVRNSDKRTDKETPYVPIPEGGVKISSH	720
ORFf721	LALGIVVLPNPELKPAMTQYIDGTGPLWKGALFPFLFITIACGAVSGFHALISSGTTPKL	360	Cst2	AKGAH	136
CstA	LANEQQACFIGYGGMLMESFVAIMALVSACIIDPGVYFAMNSPMAVL--APAGTADVVAS	411	ORF13	H	142
ORF12	LANETDARFIGYGAMLMESFVAIMALVAASIIIEPGLYFAMNTPPAGLGITMPNLHEMGGE	420	ORFf721	H	721
ORFf721	LANETDARFIGYGAMLMESFVAIMALVAASIIIEPGLYFAMNTPPAGLGITMPNLHEMGGE	420			

Figura 34. Comparación de las secuencias de aminoácidos de las ORFs 12 y 13 con proteínas que se inducen en la fase estacionaria de crecimiento de *E. coli*. CstA y Cst2 forman un mismo operón en *E. coli* W3110 y la ORFf712 pertenece a la estirpe MG1655 de *E. coli*. | y : indican, respectivamente, aminoácidos idénticos y conservados.


```

MG1655 --tsr--CAT--53n--TTTATGAATATCCC--140n--AAACTGACC--124n--TCGGCCAACAT
      =   |||   =   ||| ||| ||| ||| |   ||| |||   =   ||| ||| | |||
W      -ORF15-CAT--53n--TTTATGAAAAGGTG-11369n-TTCCTGACC--124n--TCGGCCTATGAT
      ←   In.           |           |
                        | CLUSTER 4-HPA |

MG1655 TAATCAATAC---Δ---ATG---ORFf721---yjiA---TAAGGGGTT--39n--TCA--mrr
      ||| |||           |||   =           =   ||| |           | |
W      TAACAAATTA--124n--ATG-ORF12-ORF13--ORF14--TAACCAGCC--71n--TTA--pac
                        In. →           →           →           Ter.           Ter. ←

```

Figura 36. Comparación de las secuencias flanqueantes al *cluster* de genes de la ruta del 4-HPA con la región correspondiente del cromosoma de la estirpe MG1655 de *E. coli* K12. Las flechas indican la dirección de transcripción de los genes; | y : indican, respectivamente, aminoácidos idénticos y conservados; Δ, significa delección de secuencia; = indica regiones muy conservadas; In., indica codon de iniciación; Ter., indica codon terminador.

3.4. Las regiones intergénicas

El análisis de las regiones intergénicas reveló la presencia de cinco secuencias palindrómicas similares a la secuencia consenso de las REP (secuencia extragénica, palindrómica y repetida (figuras 37 y 38A). Este tipo de palíndromes se ha encontrado a lo largo del cromosoma de algunas enterobacterias pero su función es aún un tema controvertido, aunque se ha demostrado que pueden estar implicados en la expresión génica e influir en la estabilidad del mRNA (Stern y cols., 1984). En la figura 37 se puede observar que dos REP están localizadas entre los genes *hpaF* y *hpaH*, la región intercistronica más larga del operón *hpaGEDFHI*, mientras que las otras tres están localizadas en la región 3' del gen *hpaI*, el último gen del operón *meta*. Las REP 1 y REP 2 están invertidas y, como se muestra en la figura 38B, forman una estructura de bucle con una energía libre de -65.9 kcal/mol. La posibilidad de que este bucle se comporte como un terminador de la transcripción dividiendo al operón *meta* en dos policistrones aún no ha sido comprobada. En el apartado 2.4 se comentó que habían sido localizados los posibles terminadores de la transcripción de los operones *hpaBC* y *hpaXA*, en cambio, entre los genes *hpaI* y *hpaX*, los únicos palíndromes localizados son las REPs 3, 4 y 5 por lo que no se descarta que estas secuencias sean las responsables de la terminación de la transcripción del operón *meta*. Por otra parte, se ha detectado la presencia de estructuras semejantes a las REP en la ruta *orto* de degradación de catecol (*cat*) de *P. putida*, aunque como en el caso de las REP de las enterobacterias, su función no ha sido determinada (Houghton y cols., 1995).

El sitio de iniciación de la transcripción del operón *meta* se localizó comparando la secuencia de nucleótidos de esta región con la equivalente del operón *hpcECBDGH* de *E. coli* C (Roper y cols., 1993) (figura 37). Esta comparación también permitió localizar las secuencias promotoras -35 y -10 y un sitio de unión a la CAP (proteína receptora de cAMP).

```

R K L M ← ORF15
ACGTTTAAACATGGTTTCTCTTTCTGTGGACCGATGCAAGATTTTCGGCCTGGAAAGCAAAAACCTTTATGAAAAGGTGAACCCGCCATCCATGGAAAA 100

* E N D G P - 138aa - L S D H M ← hpaR
ACCAGGCGCTGGTTATCAGATACTAAAAAGTTATTCATTATCGCCCGG--414n--TAGTGAGTCGTGCATTATCTTTCCCCTGAAAAATTCAAATCA 605
TTAATATAGAAACAGTTAATCATCACGTTACCGGGACTGCGCTGTCCCTGCAATTCTGATTTTATCAAATGAGCATGTTTTATTAACTCACTGATAT 705

CAP -35 -10
TTAATATATTCTTATCTTTAAATGTAATAGTTTGGTTAATTAGATCACATTTACATCACTTACTTTTGCCAAACTACGCAATCGGGAATAAAAGTAATCA 805

+1 hpaE → M K K V N H
hpaG → M K G T I - 419aa - E E T A K *
TTAACATATTAATGATTAAGATGATCGAATCCGAGGAGTGGTACCATGAAAGGCACATC--819n--GAGGAAACAGCGAAATGAAAAAGTAAATCA 2153

hpaD → M G K L A - 273aa - F P L P A *
W I N G - 473aa - P K W G V * hpaF → M P H
TTGGATCAACGGT--1419n--CCGAAATGGGGAGTCTGATATGGGTAAGTTAGCG--819n--TTCCCCTTCCCGGTAAGGACGTTTTATGCCGCA 4473

F I V E - 114aa - H A L F K * REP 1 →
CTTTATCGTTGAA--342n--CACGCATTGTTAAGTAGCGCGCAGATTGCCCGGTGGCGCTGCGCTTACCGGGCTACAAAACCCCAAACCGTACACC 4906

← REP 2 hpaH → M F D K - 259aa - C R F V * hpaI → M
GTAGGCCGGATAAGGCGCAGCCGCATCCGGCAATGCCACAGGATATCGTATGTTTCGATAAA--777n--TGCCGCTTTGTTTAAAGGAGAGAACGATGG 5774
E N S F K - 251aa - K P G V Y * REP 3 →
AAAACAGTTTTAAA--753n--AAGCCCGCGTGTATTAATCGCGGATGCGCTGCGCTTATCCGGCCTACACTCGCACCAAACCGTAGGCCGGATAAGG 6618

← REP 4 ← REP 5
CGTTTACGCGCTTCCGGCAAAAAGCTGTACCAAATCGCGGACCGGATAAGGCGTTACGCGCATCCGCCATAAAACCTTGTGCTACCCCTACAAAAATCCC 6718

hpaX → M S D T S - 448aa - P R A T P * hpaA → M C D R Q - 285aa - K V P
ATTAGAGGAAGAAAAATGAGCGACACCTCA--1344n--CCGCGAGCGACCCCGTAAGGAACGACGATGTGTGACCGTCAG--855n--AAAGTACCTG 8999
V T *
TGACGTGAAATTCCTTAGCCTTAACGGGAAACCAGGCACCACCTGCTATTCCCCTTTCTTCGCCTGGAGCAAGGGGATTTAGTAACGTTAAAAATCGCT 9099
AAAAGCGACCTCGATCACAAAACCGTACCCGACCTGAAAAGTACCAGCCATTGTCCCAAAGTCTCTTCCCGTAATCGCCCTGCAAGCTTCAATTGAAT 9199

hpaB → M K P E D - 510aa - D K L L K *
AACAAAAACAACACACAAAATTAACAATCGTATTCGATTAATACTGTAGAGGTCGACATGAAACCAGAAGAT--1530n--GATAAGCTGCTGAAATAA 10820

hpaC → M Q L D E - 160aa - M E A A I *
CGCAGCAGGAGTTAAGATGCAATTAGATGAA--480n--ATGGAAGCTGCGATTTAAGCTCCAGCCCTTCACGCATTCGTGCAAGGGGCATTTTTAT 11391
TGTGATCTCAATCACATATCCACCCGATGAATAATTCTTGCCCAACCGTCACCGCATTCCTGACCGCTTACATCCCTAAAATAACCACTCAGTTATTTA 11491
CCTTACTTTACGCGCGCTATCTCTGGCAACATCACTGCAGGATAGCGGTCAATTTACCTCCTCAAACACAACGCAAACCTAGAACGGCTTCGGCCATG 11591
ATTAACAAATTACTCTATGGATTTTGGTTTGGCAGGAAGGCAGCAAGTGAGTGAATCCCCGGGAGCTTGCAACAGTAAGTACAGGGGTGAACGAACGCA 11691

ORF12 → M P G F T - 568aa - V F P T R C W Q P *
ORF13 → M L A A V A L V L
GCTGCCGACCTGTAAGCCAAAAGACGACGAGTAAAATGCCAGGTTTACT--1704n--GTATTTCCAACCAGATGCTGGCAGCCGTAGCGCTGGTACT 13486

G T V V L - 123aa - I S S H H *
GGGCACCGTTGTGCTG--369n --ATCTTTCGCACCACTAACCCTATTAGCCCCGCTTCGGCGGGGCTTGTCTTATCAGAGTGAACCTATGTTTGG 13946
TAACCTTAGGACAGGCAAAAAAATATCTCGCCAGCGCGGAAGATGTTGATTGGCATCCAGACTATGACAACTACGTCGAGCATATGAAGACCAACCAT 14046

ORF14 → M K N S S - 336aa - A G L K K *
CCCGACAAGCCGTACATGAGCTATGAAGAATTCTTCC--1008n--GCGGACTCAAAAAGTAACCAGCCCTCCAACATCACAATACCTGAGACTTTATC 15145

* R Q V K L - 836aa - N R N K M ← pac
ACCCTCATATTCTCCAGAATATGAGGGCTTTCGGGCTTAATTATCTCTGAACGTGCAA--2508n--ATTCTATTTTTTCATTGTATCCTCTGGCAGGT 17744
GTATAATTAGCGAATTGATACTAGCAACGAAGCTT 17779

```

Figura 37. Regiones intergénicas del cluster de genes de la ruta del 4-HPA. Sólo se muestran los extremos 5' y 3' de las regiones codificantes. Las flechas indican la dirección de transcripción de los genes. Las secuencias REP están subrayadas con doble línea. Los posibles terminadores de la transcripción y el presunto sitio de unión a la proteína CAP del operón *meta* están subrayados. También está indicado el sitio de iniciación de la transcripción (+1) del operón *meta* y las secuencias promotoras -10 y -35. Los asteriscos señalan el codon de terminación de las proteínas.

A

```

REP 1      GCCCGGTGGCGC-----TGCGCTTACCGGGCCTAC
REP 2      GCCGGATG-CG--GCT----GCGCCTTATCCGGCCTAC
REP 3      GCCGGATG-CGC-----TGCGCTTATCCGGCCTAC
REP 4      GCCGGAAG-CGGCGTAA---ACGCCTTATCCGGCCTAC
REP 5      GCGGATG-CGGCGTA----ACGCCTTATCCGGTCCGC

CONSENSO  GCCGGATG-CGCCGC-----GCGCCCTTATCCGGCCTAC
           T       A  T       A  T       A
    
```

B

```

-----REP1----->
  C - G - - - G   AAAACCCCA
GCC GG TG CG CTGCGC TTA CCGG CCTAC      A
CGG CC AC GC GACGCG AAT GGCC GGATG      A
  - T - C   G  A   -   CCACATGCC
-----REP2----->
    
```

Figura 38. Alineamiento de las secuencias REP. Panel A, alineamiento de las REPs localizadas en las regiones intercistrónicas del *cluster* de genes de la ruta del 4-HPA con la secuencia consenso (Stern y cols., 1984). Panel B, estructura de bucle formada por las REP 1 y 2.

4. REGULACIÓN DEL OPERÓN *hpaBC*

4.1. Localización del promotor P_{BC}

Para estudiar la regulación de la transcripción del operón *hpaBC* era necesario disponer de un sistema que detectara clara y específicamente la producción de la 4-HPA-hidroxilasa tanto en medio sólido como en medio líquido. Inicialmente se pensó en utilizar el fenotipo negro provocado por la producción de esta enzima, pero éste sólo era evidente en los clones hiperproductores. Por ello, para determinar la presencia de un posible promotor en la región de 254 pb comprendida entre el gen *hpaA* y el operón *hpaBC*, se fusionó dicha región al gen trazador *lacZ*. Este sistema tiene la ventaja de que la expresión del gen fusionado se puede detectar *in vivo*, con el sustrato cromogénico X-Gal, o *in vitro* determinando la actividad β -galactosidasa. En la figura 39 se detalla la construcción de esta fusión en el plásmido pROH1. Este plásmido se utilizó para transformar células competentes de la cepa *E. coli* MC1116 seleccionando los transformantes a 30 °C en placas de medio LB suplementado con kanamicina y X-Gal. A las 12 horas de incubación, las colonias presentaban un fenotipo azul a diferencia de la cepa control MC1116 (pRS551) cuyas colonias eran blancas. En la figura 40A se muestra la expresión de la β -galactosidasa en la cepa MC1116 (pROH1) durante la fase exponencial de crecimiento, comparada con la cepa control. Una vez comprobado que en esta región había un promotor, al que se denominó P_{BC} , se determinó la actividad β -galactosidasa de las células cultivadas en un medio con 4-HPA. Estos resultados demostraron que, en la cepa MC1116 (pROH1), la expresión de la fusión P_{BC} -*lacZ* no se inducía en presencia de 4-HPA (figura 40A). Para comprobar si este promotor estaba regulado por una proteína existente en la estirpe salvaje, se transformaron las células competentes de *E. coli* ATCC 11105 con el plásmido pROH1. Al cultivar la cepa *E. coli* ATCC 11105 (pROH1) en medio M9 con glicerol en presencia o ausencia de 4-HPA se comprobó que los niveles de β -galactosidasa aumentaban 10 veces en presencia de 4-HPA, lo que demostraba que la expresión de la fusión P_{BC} -*lacZ* estaba regulada positivamente en esta estirpe (figura 40B).

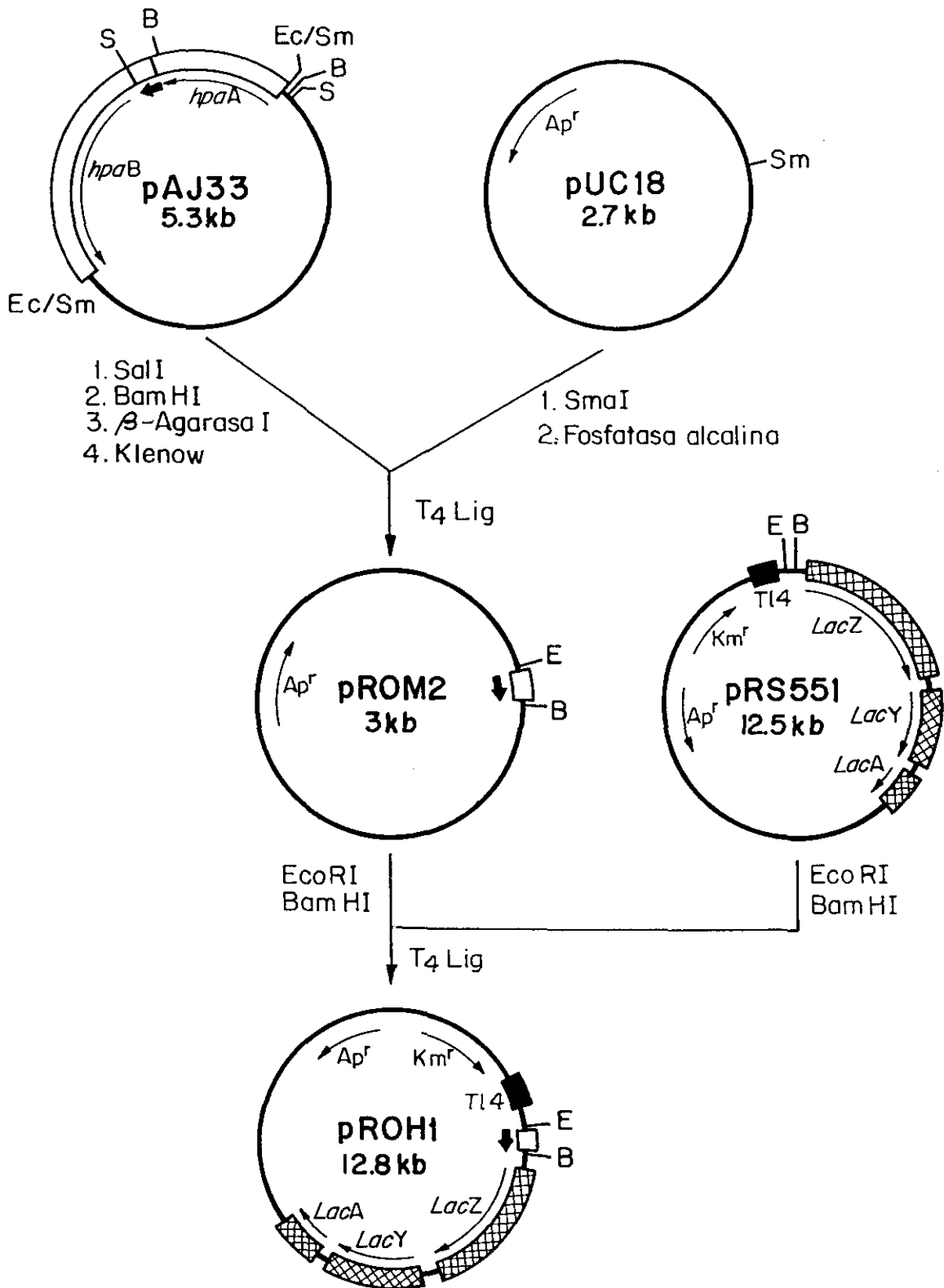


Figura 39. Construcción del plásmido pROH1. Las flechas finas indican la dirección de la transcripción de los genes y las gruesas el promotor P_{BC} . Abreviaturas: Ap^r , resistencia a ampicilina, Km^r , resistencia a kanamicina; Tl_4 , terminadores de la transcripción; B, *Bam*HI; E, *Eco*RI; Ec, *Eco*RV; S, *Sal*I; Sm, *Sma*I.

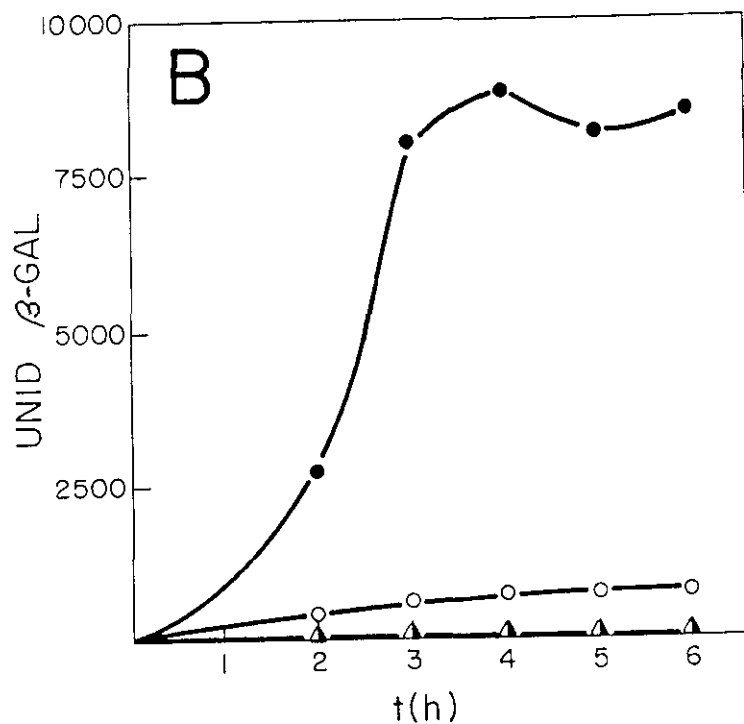
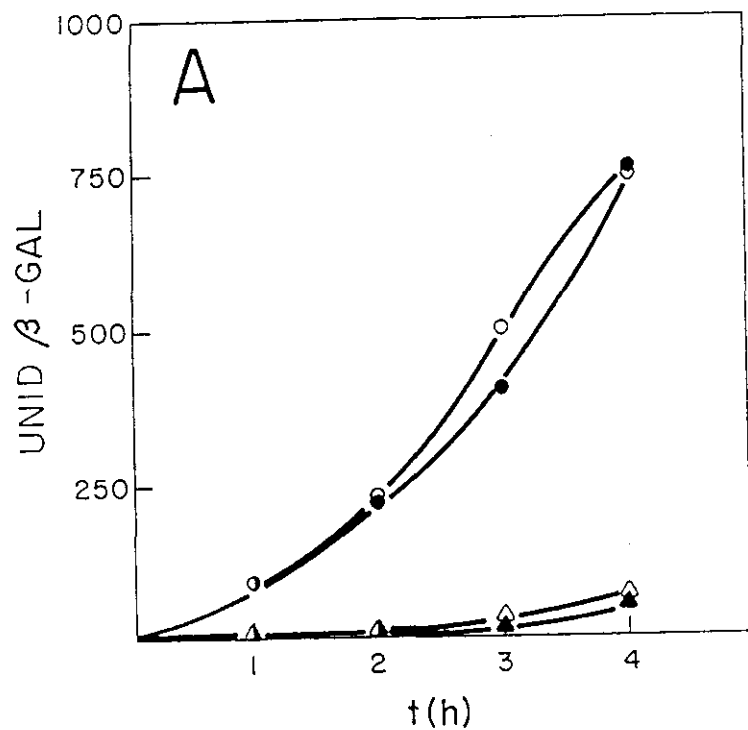


Figura 40. Expresión de la fusión P_{BC} - $lacZ$ en las estirpes K12 y W de *E. coli*. El panel A muestra la producción de β -galactosidasa por la cepa MC1116 transformada con el plásmido (pROH1) (O, ●) ó pRS551 (Δ , \blacktriangle). Las células se incubaron en medio LB a 30 °C en presencia (●, \blacktriangle) ó ausencia (O, Δ) de 4-HPA 1 mM. El panel B muestra la producción de β -galactosidasa por la cepa *E. coli* ATCC 11105 transformada con el plásmido (pROH1) (O, ●) ó pRS551 (Δ , \blacktriangle). Las células se cultivaron a 30 °C en medio M9 suplementado con glicerol 20 mM en presencia (●, \blacktriangle) ó ausencia (O, Δ) de 4-HPA 1 mM.

4.2. Determinación de los inductores del operón *hpaBC*.

La purificación de la proteína HpaB permitió obtener un suero anti-HpaB (apartado 19. de Materiales y Métodos) y por tanto, detectar la producción de esta enzima mediante la técnica de Western blot. Esta técnica se empleó para determinar qué compuestos inducían la producción de la 4-HPA-hidroxilasa en la estirpe salvaje. Para ello, se cultivó *E. coli* ATCC 11105 a 30 °C en medio M9 conteniendo glicerol y uno de los siguientes compuestos: 4-HPA, 3-HPA, PA, L-tirosina, L-fenilalanina, L-metionina y L-alanina. Se tomaron alícuotas de los cultivos durante la fase exponencial de crecimiento y se analizaron por Western blot (figura 41). De esta manera se demostró que de los compuestos ensayados sólo el 4-HPA, 3-HPA y PA inducen la producción de la hidroxilasa.

Los inductores de la producción de 4-HPA-hidroxilasa también fueron analizados mediante el estudio de la producción de β -galactosidasa en la cepa *E. coli* ATCC 11105 (pROH1), incubada en medio M9 suplementado con glicerol y diferentes compuestos aromáticos (tabla 12). Los resultados de estos experimentos permiten concluir, en primer lugar, que los inductores de la hidroxilasa son el 4-HPA, el 3-HPA y el PA, y, en segundo lugar, que algunos sustratos de la enzima como el HPC, el 2,5-dihidroxifenilacetato, el 3-cloro-4-HPA, el 4-cloro-fenilacetato, el *p*-cresol, el fenol, el 3-cloro-fenol y el 4-cloro-fenol, no se comportan como inductores.

4.3. Localización de la región reguladora del operón *hpaBC*

Los resultados derivados de la secuenciación de la ruta metabólica del 4-HPA revelaron que esta ruta está formada por dos operones catabólicos, los genes reguladores *hpaR* y *hpaA*, y el gen *hpaX* que parece cotranscribirse con el gen *hpaA* (figura 22). Como ya se ha mencionado, el gen *hpaR* es homólogo al gen *hpcR* de *E. coli* C, que actúa como represor del operón *meta* en esta estirpe. Según los trabajos de Roper y cols. (1993), la transcripción del operón *meta* se induce en presencia de HPC y/o 4-HPA. Para comprobar si la transcripción del operón *hpaBC* estaba regulada también por la proteína HpaR se construyó el plásmido pHCR2 ligando el fragmento *EcoRI* de 5,9 kb del

plásmido pHCR1 al vector pACYC184 digerido con la misma enzima (datos no mostrados). La mezcla de ligación se utilizó para transformar las células competentes de MC1116 (pROH1) con el fin de expresar en la misma célula la fusión P_{BC} -*lacZ* y el gen *hpaR*. Los niveles de β -galactosidasa en la cepa MC1116 (pROH1, pHCR2) eran similares a los de la cepa MC1116 (pROH1, pACYC184) (datos no mostrados) lo que indicaba que, aparentemente, la proteína HpaR no estaba implicada en la regulación del operón de la hidroxilasa. Para comprobar la influencia ejercida por las proteínas HpaA y HpaX en la regulación de la hidroxilasa, se construyó el plásmido pRE1 según se detalla en la figura 42 y en el esquema de la figura 43. Con este plásmido se transformó la cepa MC4100 (pROH1) que presentaba el mismo fenotipo que la cepa MC1116 (pROH1) pero crecía más rápidamente en medio mínimo, facilitando los estudios de producción de β -galactosidasa en la fase exponencial de crecimiento. La cepa MC4100 (pROH1, pRE1) al ser incubada en medio M63 con glucosa presentaba un nivel basal de β -galactosidasa de 40.000 unidades que aumentaba al cultivar las células en presencia de 4-HPA (figura 44). Para confirmar que el incremento del nivel basal de β -galactosidasa observado en la cepa MC4100 (pROH1, pRE1) respecto a la cepa MC4100 (pRS551, pRE1) no se debía a un aumento en el número de copias del plásmido pROH1, se midió la actividad β -lactamasa de dichas cepas. En ambos casos, la actividad β -lactamasa a las 4 horas de incubación era de $1-4 \times 10^{-3}$ μ moles de cefaloridina hidrolizados/minuto, lo que indicaba que el aumento de la actividad β -galactosidasa en la cepa MC4100 (pROH1, pRE1) no era debido a un aumento del número de copias del plásmido pROH1 sino a un incremento en la expresión de la fusión P_{BC} -*lacZ*, aparentemente provocado por las proteínas HpaA y/o HpaX.

4.4. Implicación de las proteínas HpaA y HpaX en la regulación del operón *hpaBC*

Para determinar si la transcripción del promotor *hpaBC* estaba regulada por la proteína HpaA, la proteína HpaX, o ambas, se construyeron los plásmidos pRA1 y pRX1 (figuras 42 y 43). A continuación se transformó la cepa MC4100 (pROH1) con cada uno de ellos, seleccionando los recombinantes en medio LB con kanamicina y cloranfenicol. La actividad β -galactosidasa que presentaban las cepas MC4100 (pROH1, pRA1) y

MC4100 (pROH1, pRX1) demuestra que la proteína HpaA regula positivamente la expresión de la fusión P_{BC} -*lacZ* en presencia de 4-HPA (tabla 13).

El plásmido pRA2 (figura 43) se construyó ligando el fragmento *SalI* de 1,4 kb de pRA1, al vector pRS551 digerido con la enzima de restricción *Bam*HI, tratando previamente ambos fragmentos con el fragmento Klenow de la DNA polimerasa. La mezcla de ligación se utilizó para transformar la cepa MC4100 seleccionando, en placas de LB con kanamicina, X-Gal y 4-HPA, los clones que presentaban el inserto orientado en la dirección que permitía expresar el gen *lacZ* mediante el promotor P_{BC} . Las colonias positivas presentaban un fenotipo azul muy intenso. En la cepa MC4100 (pRA2), el gen *hpaA* y la fusión P_{BC} -*lacZ* se expresan en *cis* lo que permitió estudiar la producción de la 4-HPA-hidroxilasa en un sistema semejante, aunque en multicopia, al existente en el cromosoma de la estirpe salvaje. En la figura 45A se muestra la producción de β -galactosidasa en esta cepa cultivada en presencia de 4-HPA, 3-HPA, PA y 2-HPA, comparándola con la de *E. coli* ATCC 11105 (pROH1) (figura 45B) cultivada en las mismas condiciones. La respuesta de ambas cepas ante la presencia de los inductores es similar, aunque la actividad β -galactosidasa en la cepa MC4100 (pRA2) es aproximadamente 6 veces mayor debido, probablemente, a que en la estirpe salvaje sólo hay una copia del gen *hpaA*.

Por otra parte, la diferencia observada entre los niveles basales de β -galactosidasa de las cepas MC4100 (pROH1, pRA1) y MC4100 (pROH1, pRE1) (tabla 13) podría justificarse si la proteína HpaX activara a la proteína HpaA. Sin embargo, la producción de β -galactosidasa en las cepas MC4100 (pRA2, pACYC184) y MC4100 (pRA2, pRX1) era muy similar (datos no mostrados).

4.5. Represión por catabolito de la 4-HPA-hidroxilasa

Los trabajos de Roper y cols. (1993) demostraron que el operón *meta* de la ruta de degradación del 4-HPA en la estirpe C de *E. coli* estaba sujeto a represión catabólica. Por tanto era lógico pensar que, puesto que la hidroxilación del 4-HPA es un paso previo a las reacciones mediadas por las enzimas codificadas en el operón *meta*, la expresión de

el operón *hpaBC* también dependiera de la fuente de carbono utilizada en el medio de cultivo. Para comprobar esta hipótesis, se cultivaron las células de *E. coli* ATCC 11105 (pROH1) en medio M9, en presencia y ausencia de 4-HPA, utilizando glucosa ó glicerol como fuente de carbono (figura 46A). La producción de β -galactosidasa en esta cepa era 10 veces menor cuando la glucosa estaba presente en el medio de cultivo sugiriendo que, efectivamente, el operón de la hidroxilasa también estaba sujeto a represión catabólica. En la figura 46B se muestra la producción de β -galactosidasa en las cepas MC4100 (pRA2) y SBS688 (pRA2) utilizando las condiciones de cultivo descritas anteriormente. En este caso, la presencia de glucosa en el medio sólo afectaba a la cepa MC4100 (pRA2), sugiriendo fuertemente que la represión catabólica estaba mediada por la proteína CAP (*crp*) ya que la cepa SBS688 era un mutante Δcrp de MC4100.

4.6. Determinación de los sitios de iniciación de la transcripción de los promotores

P_A y P_{BC}

Ya se ha comentado que los genes *hpaA* y *hpaX* estaban prácticamente juntos en el cromosoma, lo que parecía indicar que ambos genes formaban parte de la misma unidad de transcripción (figura 23); sin embargo, los resultados obtenidos hasta ese momento en las células que contenían el plásmido pRA1 ó el plásmido pRA2, demostraban la existencia de un promotor (P_A) situado en la región de 203 pb anterior al codon de iniciación del gen *hpaA*. El sitio de iniciación de la transcripción del promotor P_A se determinó mediante la técnica de “primer extension” (figura 47A) para lo cual, el oligonucleótido O38A (figura 23) se hibridó con el RNA total de la cepa MC4100 (pROH1, pRE1). Este RNA se utilizó posteriormente para determinar el sitio de iniciación de la transcripción del promotor P_{BC} (figura 47B) utilizando en este caso el oligonucleótido OLACZ50, de 20 nucleótidos, situado en la cadena no codificante del vector pRS551, a 50 pb de la diana *Bam*HI del mismo (apartado 22 de Materiales y Métodos). Los resultados obtenidos en ambos casos se confirmaron repitiendo el experimento según el método descrito en el apartado 22.2 y en el caso del promotor P_{BC} , utilizando el RNA total de la cepa MC4100 (pRA2) cultivada en presencia de 4-HPA. En la figura 47C se muestra la secuencia de nucleótidos correspondiente al promotor P_{BC} , así como dos posibles sitios de iniciación de la transcripción en la región 5' del gen

hpaA, estando uno de ellos (σ_1) precedido por un posible sitio de unión a la proteína CAP. Para comprobar la funcionalidad de ambos sitios de iniciación de la transcripción sería necesario realizar una serie de experimentos genéticos que hasta el momento no han sido desarrollados.

4.7. El promotor P_x

Se ha demostrado que la hiperproducción de algunos reguladores pertenecientes a la familia XylS/AraC provoca la inducción de la transcripción del gen al que regulan independientemente de la presencia del efector (Marqués y Ramos, 1993). Por tanto, la diferencia apreciada entre los niveles basales de β -galactosidasa de las cepas MC4100 (pROH1, pRA1) y MC4100 (pROH1, pRE1) (tabla 13) podrían indicar que la proteína HpaA estaba hiperproducida en esta última. La determinación del sitio de iniciación de la transcripción del promotor P_A demostró que la expresión del gen *hpaA* estaba controlada por este promotor tanto en la cepa MC4100 (pROH1, pRE1) como en la cepa MC4100 (pRA2) (apartado 4.6). Estos resultados sugerían que la supuesta hiperproducción de HpaA en las células que contienen el plásmido pRE1, podría deberse a la formación de otro transcrito a partir de un promotor más fuerte que P_A . Para comprobarlo se construyó el plásmido pRX2 (figura 43) ligando un fragmento *Eco47III* de 288 pb del plásmido pAJ38 (figura 22) al vector pRS551 digerido con la enzima *BamHI* y tratado con Klenow. Este fragmento (nucleótidos 6.502-6790 de la figura 23) abarca toda la región intergénica situada entre el operón *meta* y el gen *hpaX*, por lo que se consideró que debía contener el posible promotor de este último (P_x). La cepa MC4100 (pRX2) producía β -galactosidasa constitutivamente hasta un nivel de 100.000 unidades, igual que las cepas MC4100 (pRX2, pRA1) y MC4100 (pRX2, pHCR2) (datos no mostrados).

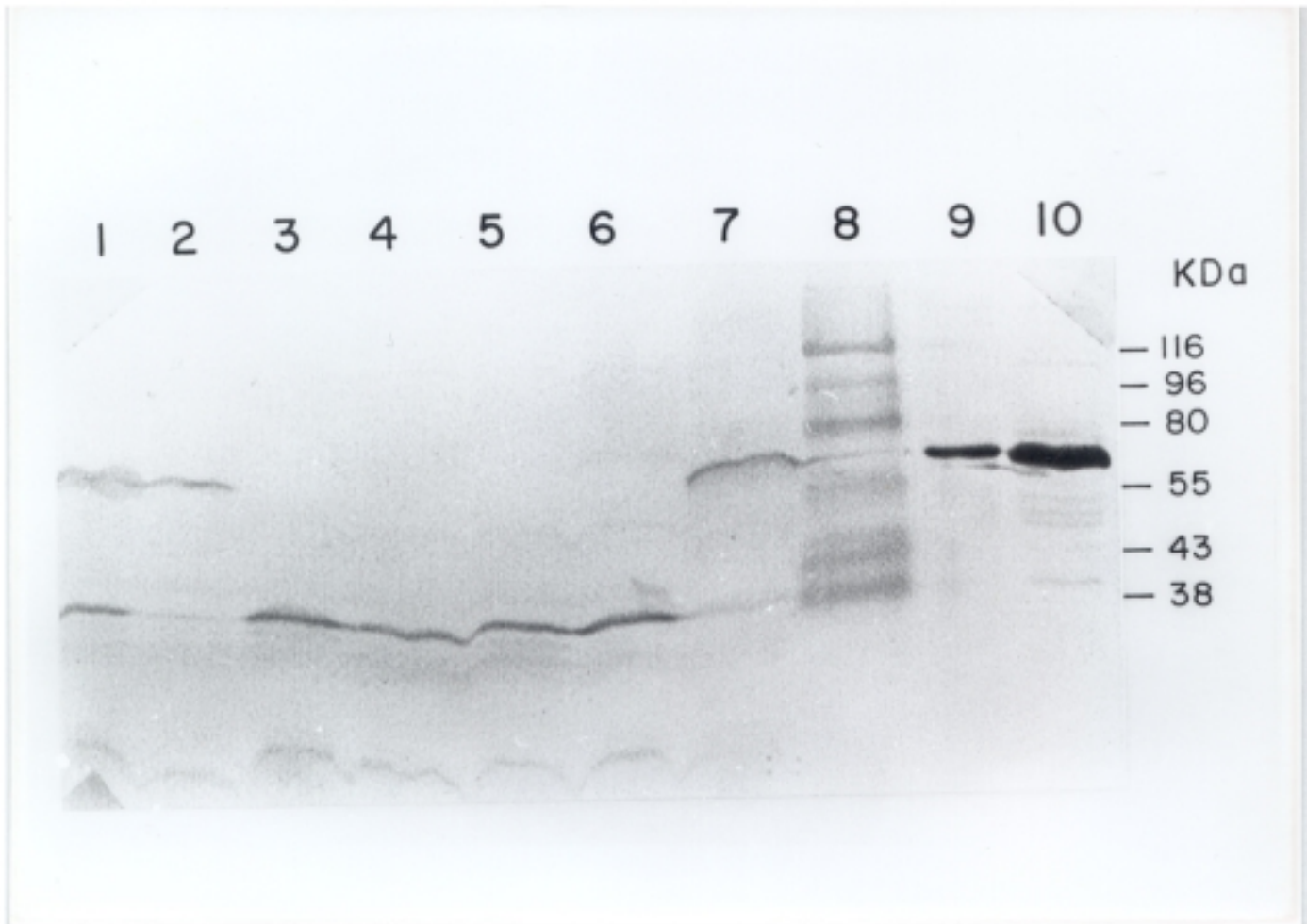


Figura 41. Inductores de la producción de 4-HPA-hidroxilasa. El análisis por Western blot fue llevado a cabo con anticuerpos anti-HpaB y células de *E. coli* ATCC 11105 cultivadas en medio M9 suplementado con glicerol 20 mM, en presencia de 4-HPA (línea 1), PA (línea 2), L-Phe (línea 3), L-Ala (línea 4), L-Tyr (línea 5), L-Met (línea 6) y 3-HPA (línea 7) a una concentración final de 5 mM. Las líneas 9 y 10 corresponden a la proteína HpaB purificada y al extracto crudo de la cepa DH1 (pAJ221). La posición de los marcadores de peso molecular se muestra en (kDa) en el margen derecho.

Tabla 12. Inductores del operón *hpaBC*

SUSTRATO	ACTIVIDAD β -GAL ^a
4-Hidroxifenilacetato	+++
3-Hidroxifenilacetato	++
2-Hidroxifenilacetato	N.D. ^b
fenilacetato	+
3-Cloro-4-hidroxifenilacetato	N.D.
3-Clorofenilacetato	N.D.
Homoprotocatecuato	N.D.
2,5-Dihidroxifenilacetato	N.D.
<i>p</i> -Cresol	N.D.
<i>m</i> -Cresol	N.D.
<i>o</i> -Cresol	N.D.
Fenol	N.D.
4-Clorofenol	N.D.
3-Clorofenol	N.D.
2-Clorofenol	N.D.

^a Actividad β -gal: para ensayar la actividad β -galactosidasa, *E. coli* ATCC 11105 (pROH1) se cultivó durante 4 horas a 30 °C en medio M9 suplementado con glicerol 20 mM y los compuestos aromáticos mostrados en la tabla a una concentración final de 1 mM.

^bN.D.: no detectado

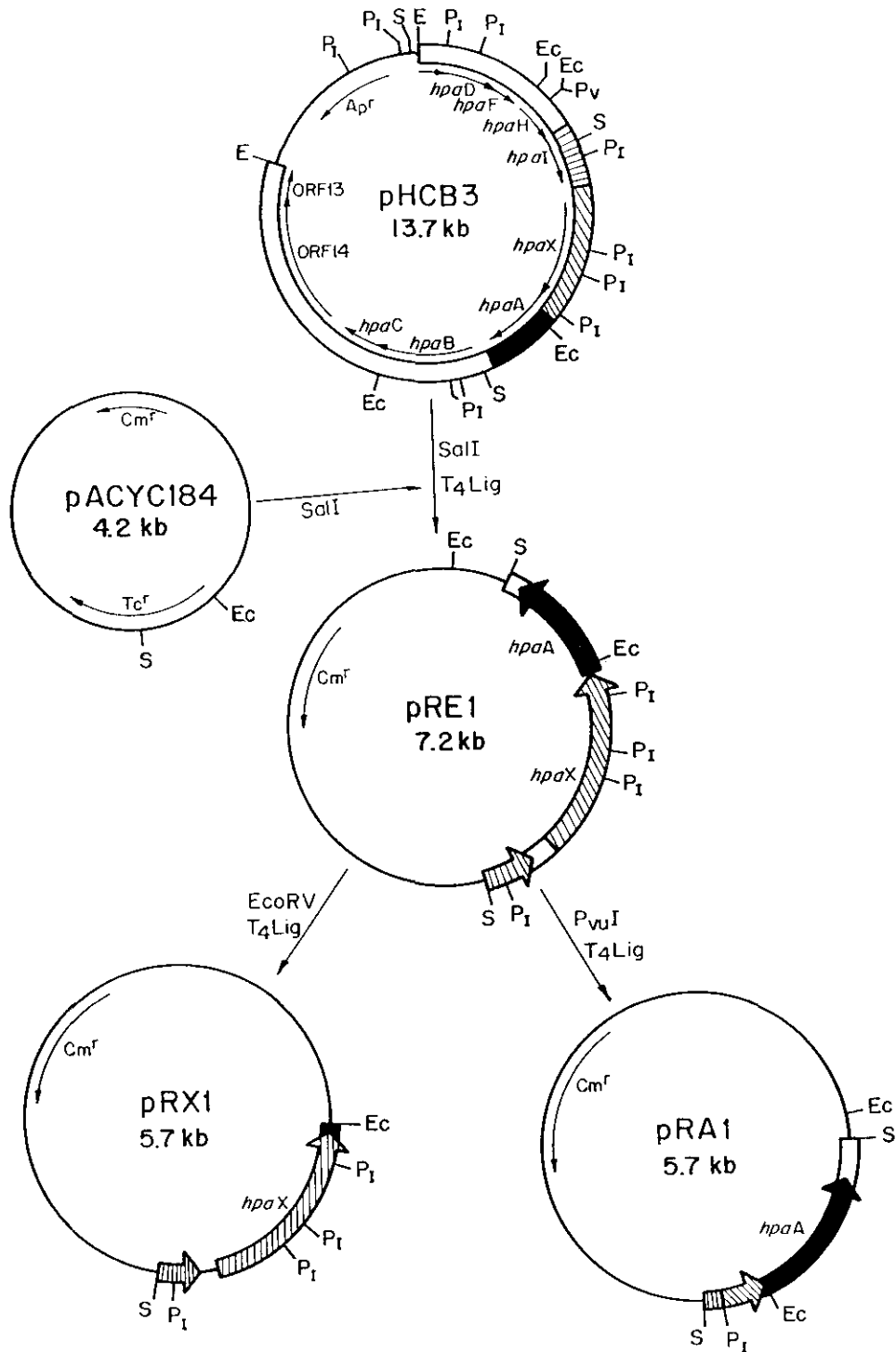


Figura 42. Construcción de los plásmidos pRE1, pRX1 y pRA1. ■, gen *hpaA*; ▨ gen *hpaX*; ▩, gen *hpaI*; □, resto de los genes contenidos en el plásmido pHCB3. Las flechas indican la dirección de transcripción de los genes. Abreviaturas: Ap^r, resistencia a ampicilina; Cm^r, resistencia a cloranfenicol; Tc^r, resistencia a tetraciclina; B, *Bam*HI; E, *Eco*RI; Ec, *Eco*RV; P_I, *Pvu*I; S, *Sal*I.

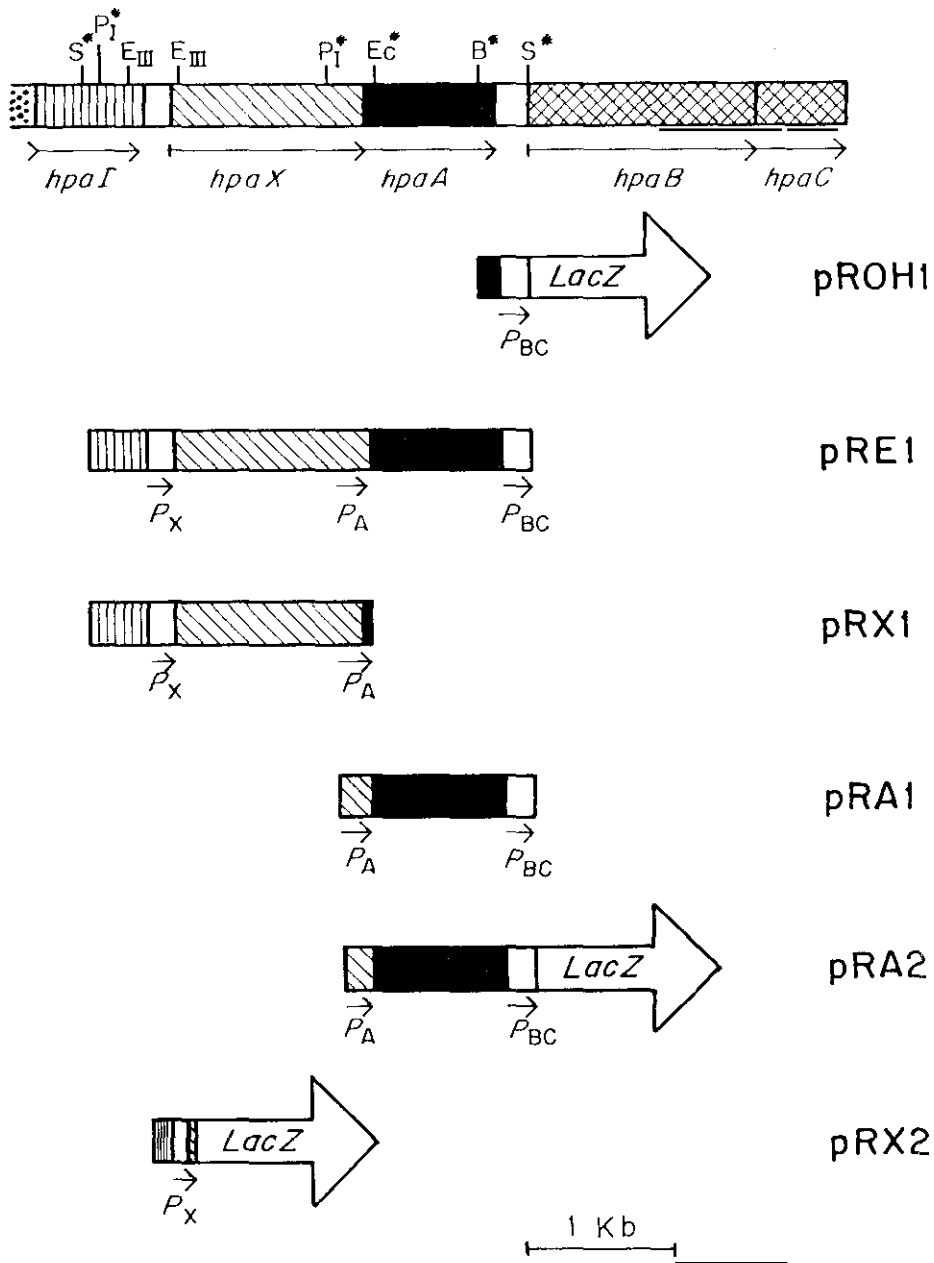


Figura 43. Fragmentos clonados en los plásmidos pROH1, pRE1, pRX1, pRA1, pRA2 y pRX2. ■, gen *hpaA*; ▨, gen *hpaX*; ▤, gen *hpaI*; ▩, operón *hpaBC*; ▧, resto de los genes del operón *meta*, □, operón *lac*; Las flechas indican la dirección de transcripción de los genes. En la figura se indican la localización de los promotores P_A , P_{BC} y P_X , y los sitios de restricción utilizados para la construcción de los plásmidos. Los asteriscos indican los sitios de restricción que no son únicos en el fragmento mostrado. Abreviaturas: B, *Bam*HI; Ec, *Eco*RV; E_{III} , *Eco*47III; P_I , *Pvu*I; S, *Sal*I.

Tabla 13. Estudio de la expresión de la fusión P_{BC} -*lacZ* en presencia de HpaA y HpaX.

Cepa	Unidades β -gal x 10^{-3} ^a	
	+ 4-HPA	- 4-HPA
MC4100 (pROH1, pACYC184)	0,70	0,62
MC4100 (pROH1, pRE1)	63,5	38,7
MC4100 (pROH1, pRA1)	41,4	0,5
MC4100 (pROH1, pRX1)	0,95	0,94

^a Actividad β -gal: para ensayar la actividad β -galactosidasa las células se cultivaron durante 4 horas a 30 °C en medio M63 suplementado con glicerol 20 mM en presencia o ausencia de 4-HPA 1 mM.

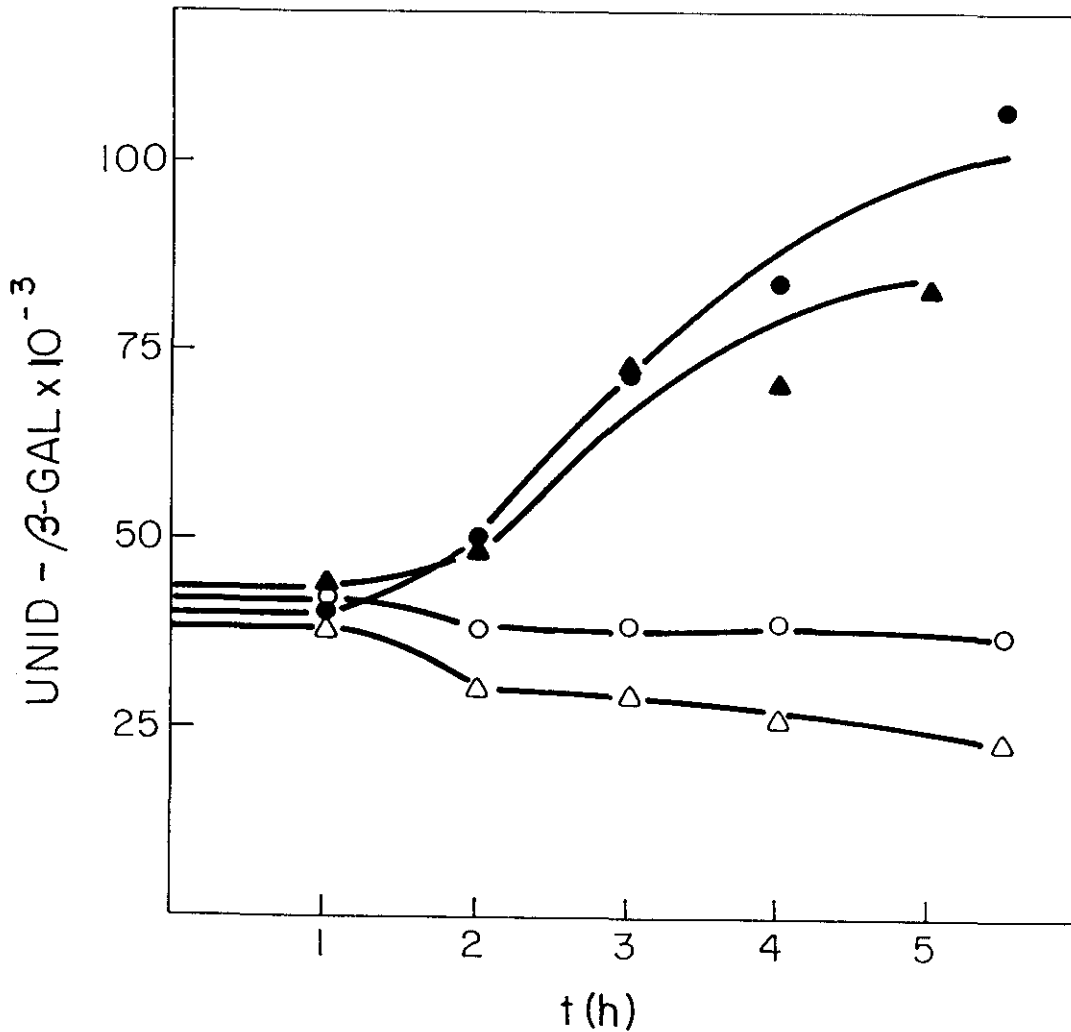


Figura 44. Producción de β -galactosidasa por la cepa *E. coli* MC4100 (pRE1, pROH1). Las células se cultivaron a 30 °C en medio M63 suplementado con glucosa 10 mM (Δ , \blacktriangle) ó glicerol 20 mM (O, \bullet), en presencia (\bullet , \blacktriangle) ó ausencia (O, Δ) de 4-HPA 1 mM.

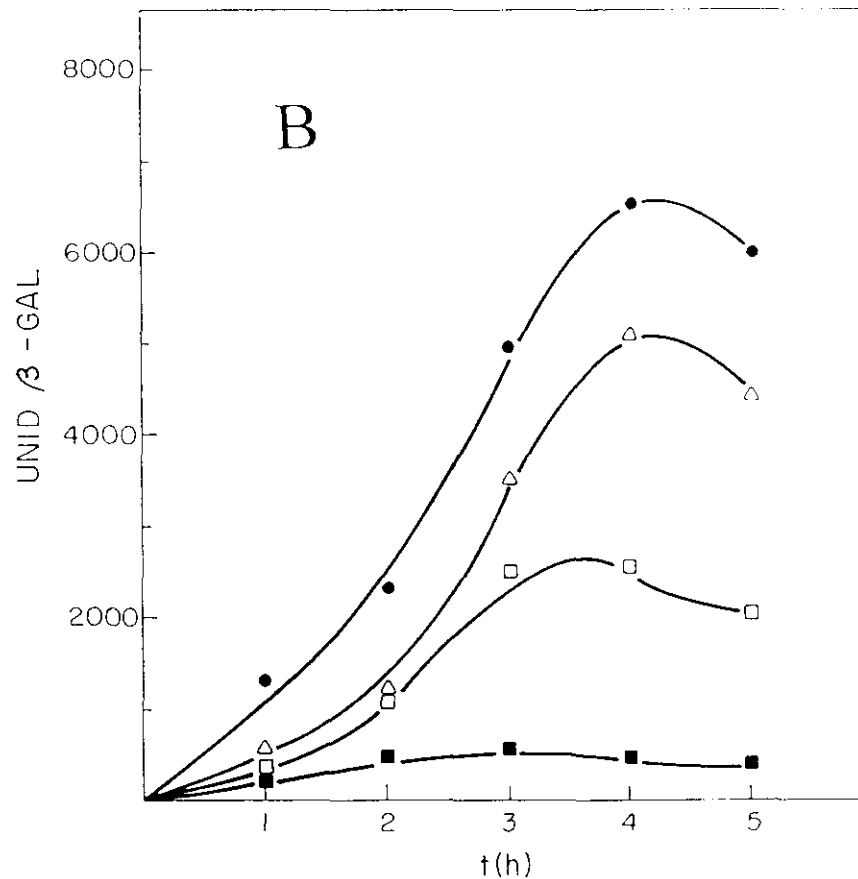
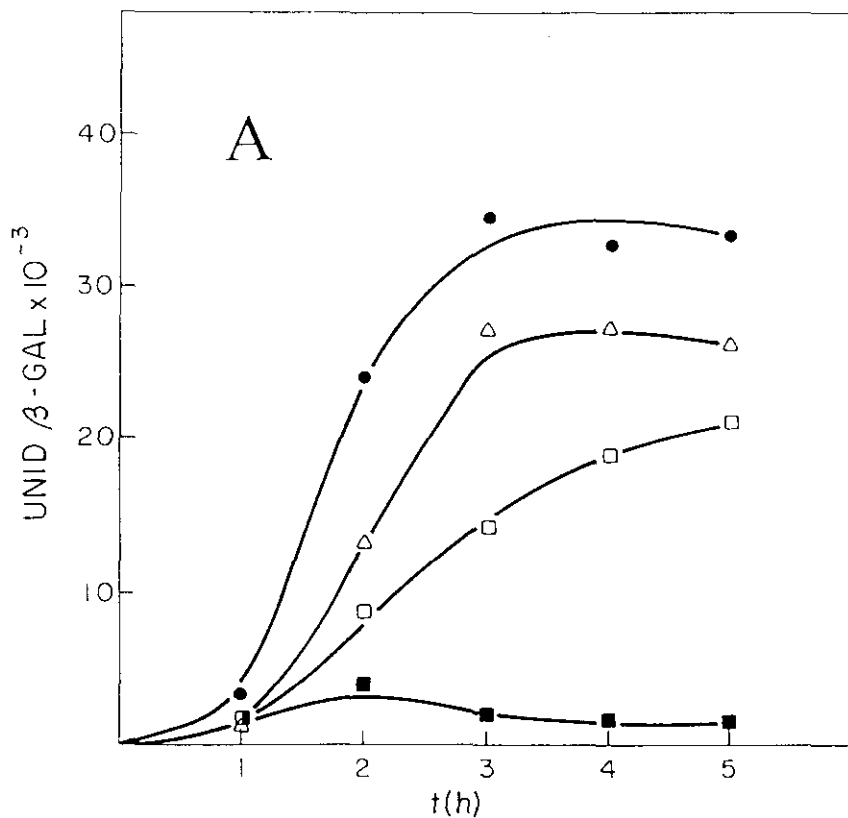


Figura 45. Inducción de la producción de β -galactosidasa en las cepas de *E. coli* MC4100 (pRA2) y ATCC 11105 (pROH1). Las cepas *E. coli* MC4100 (pRA2) (Panel A) y *E. coli* ATCC 11105 (pROH1) (Panel B) fueron cultivadas a 30 °C en medio M63 suplementado con glicerol 20 mM y 4-HPA (●), 3-HPA (Δ), PA (□) y 2-HPA (■) a una concentración final de 1 mM.

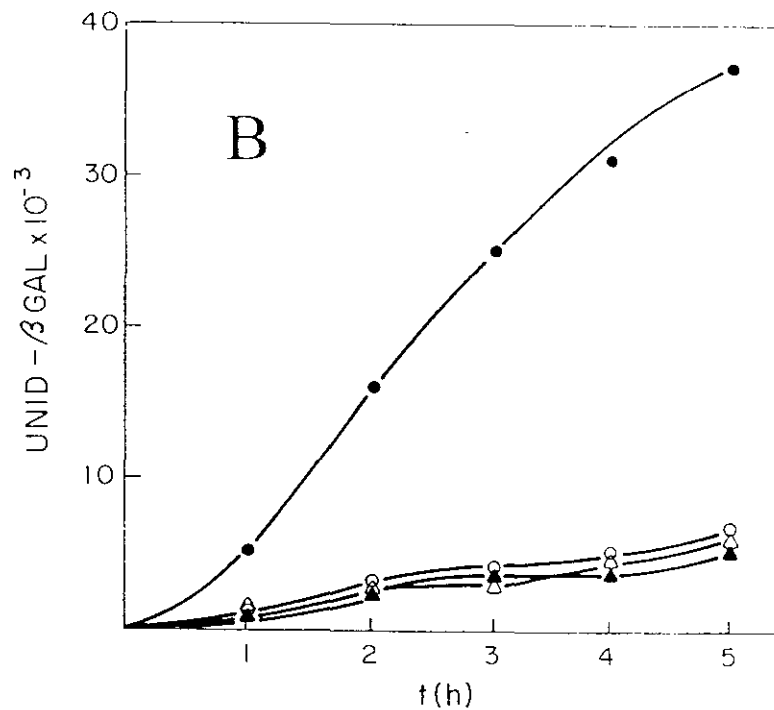
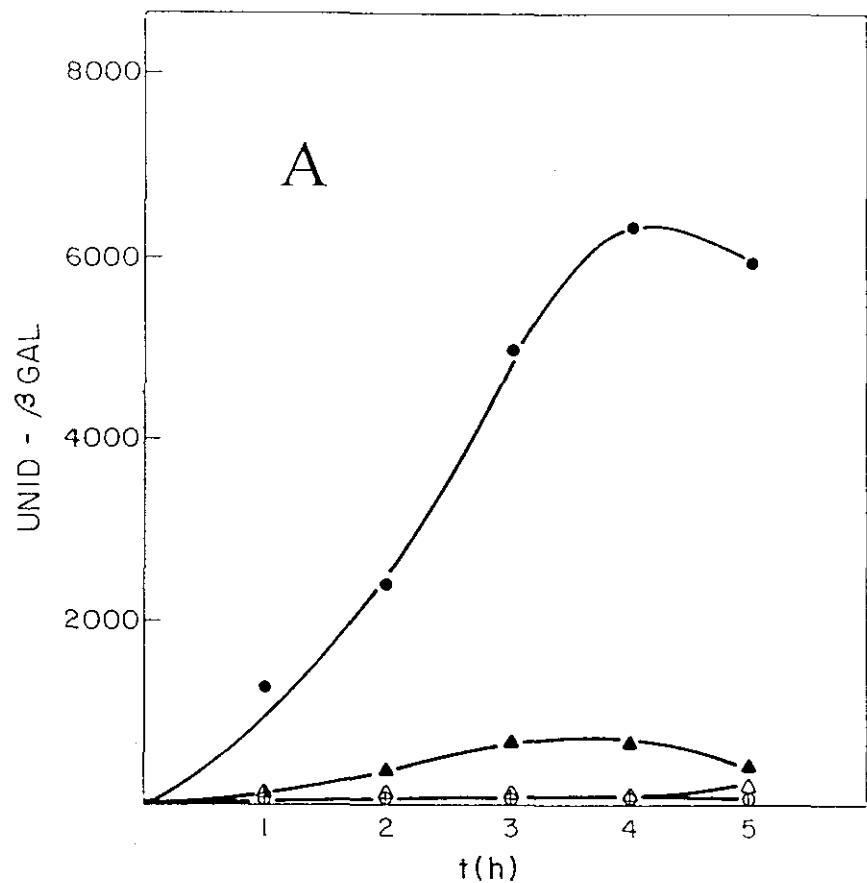
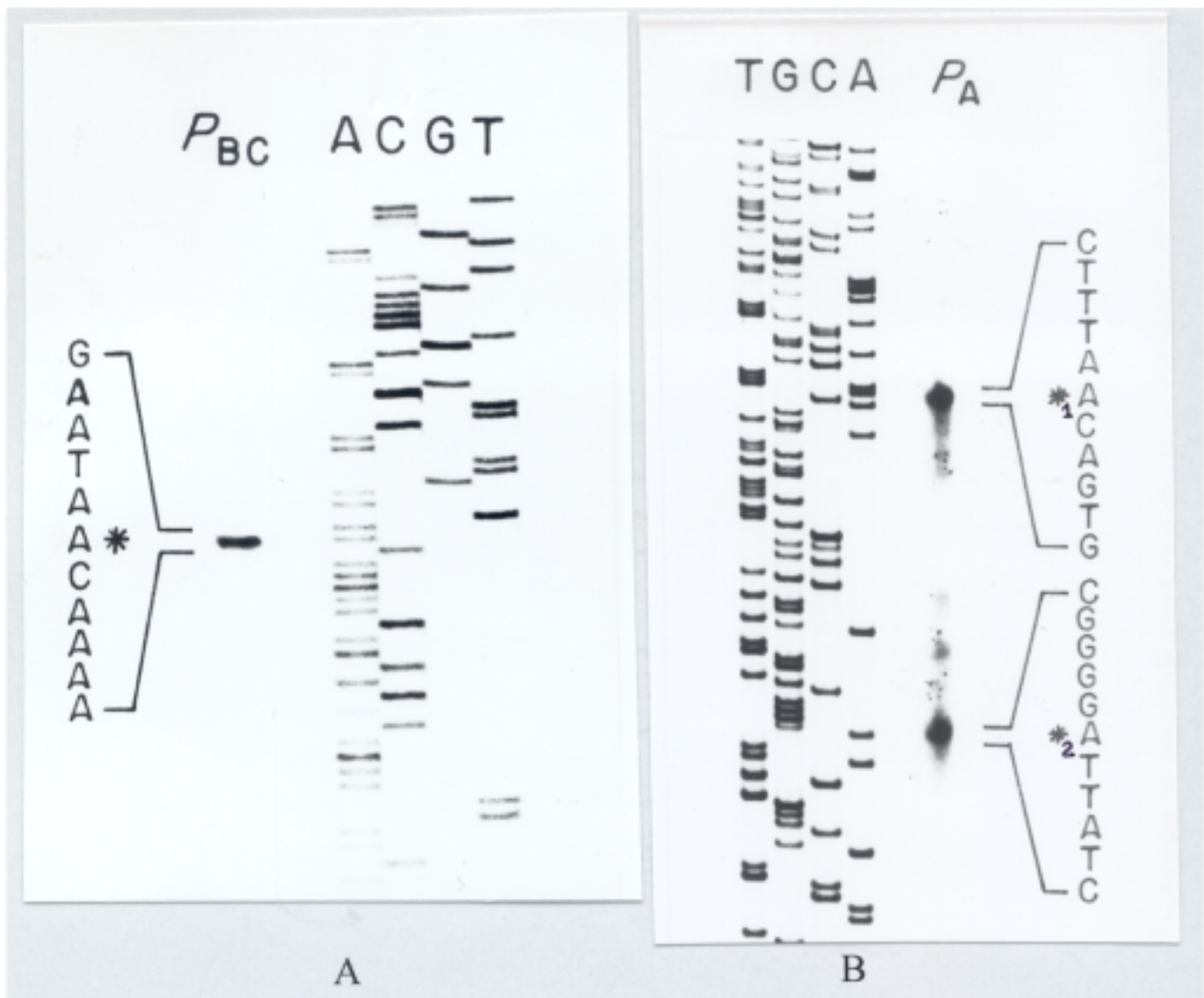


Figura 46. Represión por catabolito de la 4-HPA-hidroxilasa. Panel A: *E. coli* ATCC 11105 (pROH1) cultivada a 30 °C en medio M9 suplementado con glucosa 10 mM (Δ , \blacktriangle) ó glicerol 20 mM (O, \bullet), en presencia (\bullet , \blacktriangle) ó ausencia (O, Δ) de 4-HPA 1mM. Panel B: *E. coli* MC4100 (pRA2) (O, \bullet), y *E. coli* SBS688 (pRA2) (Δ , \blacktriangle) cultivadas a 30 °C en medio M63 suplementado con 4-HPA 1 mM y glucosa 10 mM (O, Δ) ó glicerol 20 mM (\bullet , \blacktriangle).



hpaX → M S D -- 380aa -- T P D Q S I S L R A R A I G I A V I N A T G N I
 GGAAGAAAAATGAGCGAC---1140n---ACACCGGATCAGTCCATCAGCCTGCGGGCACGAGGATCGGTATTGCGGTGATCAACGOCCTGGCAACAT 1229

P_A → CAP
 G S A L S P F M I G W L K D L T G S F N S G L W F V A A L L V I G
 TGGTTCAGCGTTAAGTCCGTTATGATCGGCTGGTTGAAAGATTTGACCGGCGAGCTTTAACAGTGGATTGTGGTTTGTTCGCCGCTGCTGGTGGATTGGT 1329

-35 -10 +1(+₁) -35 -10
 A G I I W A I P M Q S S R P R A T P * *hpaA* → M C D R Q - 285aa - K V P
 GCGGGGATTATCTGGGGGATTCCAATGCAGTCCCTCCCGGAGCGACCCCGTAGGAGCGACGATGTGTGACCGTCAAG--855n---AAAGTACCTG 2275
 +1(+₂)

V T *
 TGCAGTGAATTCCTTAGCCTTAACGGGAAACCGCCACCACTGCTATTCCCTTTCTTCGCTGGAGCAAGGGGATTAGTAACGTTAAAAATCGCT 2375

P_{BC} →
 AAAAGCGACCTCGATCACAACCGGTACCCGACCTGAAAGTACCAGOCATTGTCCCAAAAGTCTCTCCCGTAATCGCCCTGCAAGCTTCAATGAAAT 2475
 -35 -10

hpaB → M K P E D - 510aa - D K L L K *
 AACMAAACACACACAAATTAACAATCGTATTCCGATTAATACTGTAGAGGTGACATGAAACGAGAAGAT--1530n---GATAGCTGCTGAATAA 4096
 +1

C

Figura 47. Localización de los promotores P_{BC} y P_A . Panel A: sitio de iniciación de la transcripción del operón *hpaBC*. Panel B: sitios de iniciación de la transcripción del gen *hpaA*. Panel C: secuencia de nucleótidos de las regiones correspondientes a los promotores P_{BC} y P_A ; los sitios de iniciación de la transcripción (+1) y las regiones promotoras -10 y -35, el posible sitio de unión de la proteína CAP, los RBS de *hpaA* y *hpaB*, y el posible terminador de la transcripción del gen *hpaA* se muestran subrayados. *₁ y *₂ indican los dos posibles sitios de iniciación en la región promotora P_A .

5. EXPRESIÓN DE LOS GENES DE LA RUTA DEL 4-HPA EN ORGANISMOS HETERÓLOGOS

El plásmido pAJ40 se construyó con el fin de comprobar que la ruta completa del 4-HPA estaba formada por los genes *hpa* y que no era necesario ningún gen adicional para la mineralización de este compuesto. Para construir el plásmido pAJ40 se ligó el fragmento *EcoRI* de 11 kb del plásmido pHCB3 al plásmido PHCR3 digerido con la enzima de restricción *EcoRI* (tabla 10, figura 22). La mezcla de ligación se utilizó para transformar las células competentes de *E. coli* ET8000. La selección de los clones recombinantes se llevó a cabo a 30 °C, en medio M9 sin glucosa suplementado con ampicilina y 4-HPA 5 mM. Sólo los clones que pudieran utilizar el 4-HPA como única fuente de carbono serían capaces de crecer en este medio. Después de dos días de incubación en estas condiciones, se obtuvieron dos colonias que fueron reaisladas en el mismo medio selectivo en presencia de ácido nalidíxico, ya que la cepa ET8000 es resistente a este compuesto. De esta forma se aisló la cepa ET8000 (pAJ40) a partir de la cual se obtuvo el plásmido pAJ40, que se utilizó posteriormente para transformar células competentes de *E. coli* DH1 y así comprobar que la capacidad de utilizar el 4-HPA y, por tanto, de expresar todos los genes implicados en su ruta de degradación, no era dependiente de cepa. La cepa *E. coli* DH1 (pAJ40) también era capaz de mineralizar este compuesto tanto en medio sólido como en medio líquido (tabla 14). Mediante estos experimentos se demostró que los genes necesarios para la degradación del 4-HPA estaban contenidos en el fragmento clonado en el plásmido pAJ40.

5.1. Construcción de módulos transponibles con los genes de la ruta del 4-HPA

Los minitransposones derivados de Tn5 se han diseñado en un plásmido (pGP704) cuya replicación (*oriRK6*) depende específicamente de la proteína π (*pir*) (Herrero y cols., 1990), por lo que sólo puede mantenerse de forma estable en cepas que produzcan esta proteína, y que además contiene el origen de transferencia conjugativa (*oriTRP4*) del plásmido promiscuo RP4. Por tanto, los vectores-transposones mini-Tn5, son estables en bacterias lisógenas λ_{pir} y pueden ser movilizados por conjugación a una bacteria receptora mediante las funciones de transferencia del plásmido conjugativo RP4.

En el caso de que la cepa receptora no produzca la proteína π el plásmido no podrá ser mantenido y, con una cierta frecuencia, habrá transposición del mini-Tn5 o sus derivados a un replicón diferente, normalmente el cromosoma de la célula receptora. El gen que codifica para la transposasa (*tnp**) no está incluido dentro de la unidad móvil, por lo que, si la bacteria receptora no es capaz de mantener el plásmido transferido sólo se integrará una copia del segmento móvil en su cromosoma. Los transconjugantes así obtenidos son muy estables desde el punto de vista genético y fáciles de seleccionar en base al marcador de resistencia incluido en el segmento móvil (de Lorenzo y Timmis, 1994).

Uno de los objetivos de esta tesis era la expresión de la ruta catabólica del 4-HPA en otros microorganismos diferentes a *E. coli* W para conferirles la capacidad de metabolizar este compuesto y así ampliar su capacidad biodegradativa. Con este fin fueron construidos los plásmidos pAJ224 y pAJ402 como se muestra en la figura 48. El primero es un derivado de pCNB5 que tiene clonado, dentro del segmento móvil, el operón *hpaBC* orientado de forma que su expresión pueda ser controlada por el promotor *P_{trc}*, así como el represor *lacI^q* y el marcador de resistencia a kanamicina. El plásmido pAJ402 es un derivado de pUTmini-Tn5-Km2 y fue construido para poder insertar la ruta completa del 4-HPA en el cromosoma de otros microorganismos diferentes a *E. coli* W, seleccionando los transconjugantes en base a su resistencia a kanamicina.

5.2. Integración de la ruta del 4-HPA en el cromosoma de *E. coli* K12

Las células competentes de *E. coli* S17-1 λ *pir* se transformaron con el plásmido pAJ402 que le confería resistencia a ampicilina y a kanamicina. Esta estirpe contiene en su cromosoma los genes que codifican para las funciones de transferencia del plásmido RP4 y para la proteína π (Herrero y cols., 1990), por lo que podía ser empleada como cepa donadora en los experimentos de transposición. Como organismo receptor se eligió la cepa ET8000 de *E. coli* K12. La conjugación se llevó a cabo según el método descrito por de Lorenzo y Timmis (1994) utilizando, para la selección de los transconjugantes, medio M9 sin glucosa suplementado con kanamicina y 4-HPA como única fuente de

carbono. La cepa donadora era auxotrofa para L-prolina, por lo que no podía crecer en el medio de selección. Los transconjugantes así obtenidos se volvieron a seleccionar en base a su sensibilidad a la ampicilina, descartándose así los procesos de cointegración, y a su capacidad de mineralizar el 4-HPA en presencia de kanamicina y ácido nalidíxico, un marcador específico de la cepa ET8000. Como resultado de estos experimentos se obtuvo la cepa ET4025 capaz de utilizar 3- y 4-HPA como única fuente de carbono pero no 2-HPA (figura 49, tabla 14). Al incubar durante toda la noche a 30 °C la cepa ET4025 en medio mínimo suplementado con glicerol y fenol, en presencia y ausencia de 4-HPA, se comprobó que, en esta cepa, la transformación *in vivo* de fenol en catecol dependía de la presencia del inductor 4-HPA (tabla 15). Estos resultados demuestran que el plásmido pAJ402 contenía no sólo las enzimas catabólicas de la ruta del 4-HPA sino también los genes reguladores.

5.3. Producción de la 4-HPA-hidroxilasa en *P. putida* KT2442

Como se ha demostrado en los apartados 2.1., 2.2. y 2.3., la 4-HPA-hidroxilasa es una enzima que actúa sobre una gran variedad de compuestos aromáticos monohidroxilados, pero sólo el 3- y 4-HPA son metabolizados completamente por la estirpe salvaje. La producción de esta enzima en otros microorganismos capaces de mineralizar metabolitos producidos por la 4-HPA hidroxilasa, permitiría al organismo receptor utilizar nuevos sustratos aromáticos como fuente de carbono, lo que supondría una expansión vertical de su capacidad biodegradativa.

El organismo receptor elegido para comprobar esta hipótesis fue *P. putida* KT2442 ya que era capaz de metabolizar el benzoato mediante la ruta *orto* del catecol (ver apartado 3.2.2.2. de la Introducción) y, por tanto, debería ser capaz de mineralizar el catecol procedente de la transformación de fenol mediada por la 4-HPA-hidroxilasa. Como organismo donador se utilizó la cepa *E. coli* S17-1 λ pir transformada con el plásmido pAJ224 (figura 48), llevándose a cabo la conjugación según el método descrito por de Lorenzo y Timmis (1994). Para seleccionar los transconjugantes se utilizó medio LB suplementado con kanamicina y rifampicina dado que *P. putida* KT2442 es resistente a rifampicina a diferencia de la cepa donadora. La nueva cepa de *P. putida*, a la

que se denominó KTH2, presentaba un fenotipo negro similar al de los clones de *E. coli* que expresan el operón de *hpaBC* (figura 50), lo que demostraba que la enzima se producía constitutivamente en esta cepa.

A continuación se investigó la capacidad de *P. putida* KTH2 para mineralizar el fenol. Para ello, se incubaron las células a 30 °C con agitación durante toda la noche en medio M9 suplementado con glicerol 20 mM, IPTG 3 mM y fenol 1 mM. Este cultivo se utilizó para inocular, el medio compuesto por M9, IPTG 3 mM y fenol 3 mM como fuente de carbono, hasta una DO_{600} de 0,1. Después de nueve horas de incubación, el cultivo alcanzó la fase estacionaria de crecimiento con una DO_{600} de 0,55. Una nueva adición de fenol 3 mM a los cultivos en fase estacionaria provocó que la cepa KTH2 iniciara un nuevo proceso de crecimiento. Sin embargo, si el fenol se añadía inicialmente a una concentración superior a 3 mM no se detectaba crecimiento, probablemente debido a la gran toxicidad de este compuesto. Como control, se repitió paralelamente el experimento con la cepa *P. putida* KT2442 cuyo cultivo presentaba a las nueve horas una densidad óptica similar a la inicial (figura 51).

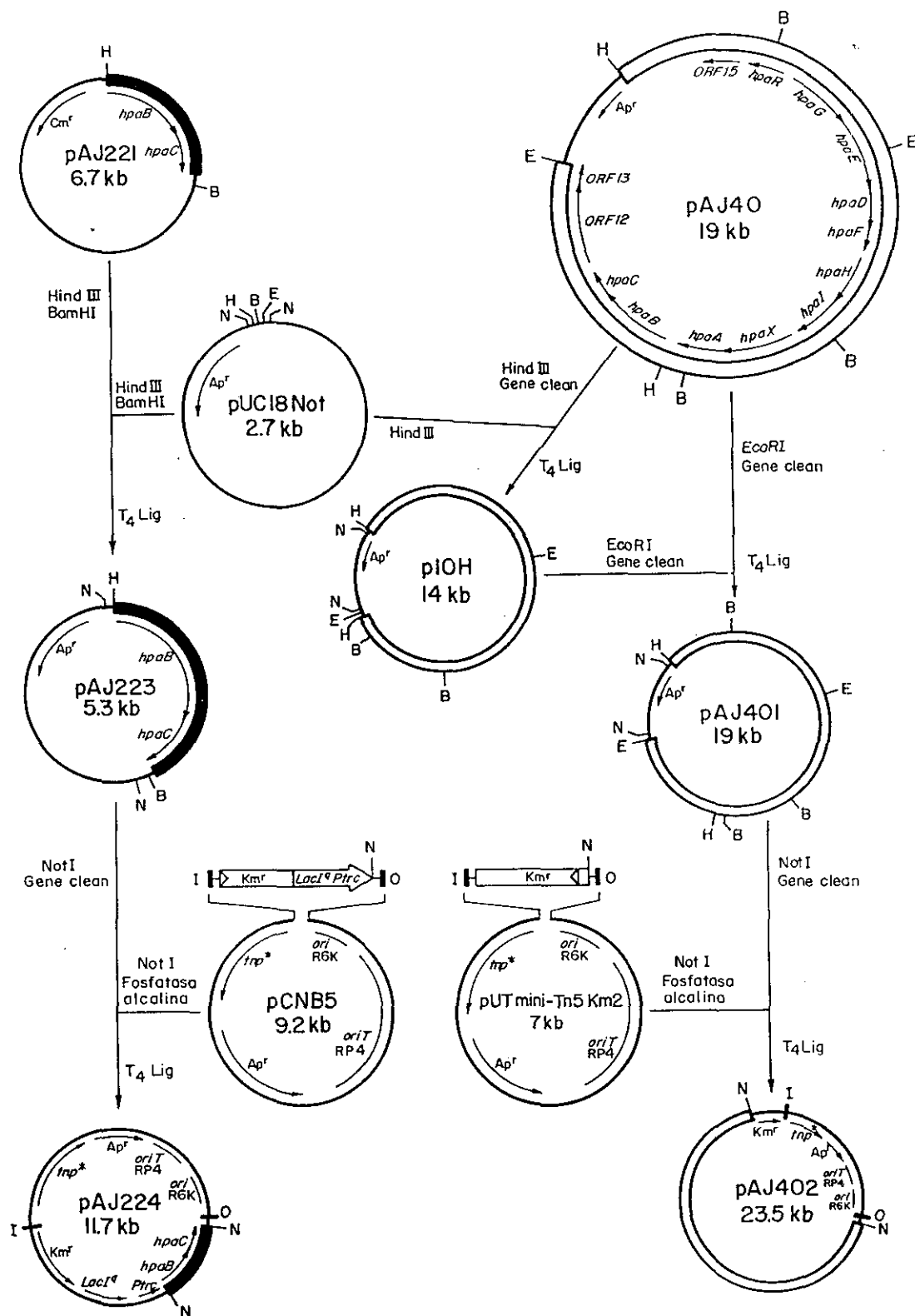


Figura 48. Construcción de los plásmidos pAJ402 y pAJ224. ■, operón *hpaBC*; □ fragmento de DNA que contiene la ruta completa para la degradación de 4-HPA. Abreviaturas: *Ap^r*, resistencia a ampicilina; *Km^r*, resistencia a kanamicina; *Cm^r*, resistencia a cloranfenicol; B, *Bam*HI; E, *Eco*RI; H, *Hind*III; N, *Not*I; *P_{trc}*, promotor *trc*; *trp**, transposasa Tn5 desprovista de la diana *Not*I; I y O, secuencias terminales (19 pb) del minitransposón.

Tabla 14. Capacidad de mineralizar el 4-HPA

Cepa de <i>E. coli</i>	DO₆₀₀^a
ET8000	0,3
ATCC 11105	0,8
ET8000 (pAJ40)	0,85
DH1 (pAJ40)	1,0
ET4025	0,95

^a D.O_{600 nm}: los cultivos se inocularon a una DO₆₀₀ = 0,3 y se incubaron durante toda la noche en medio M9 suplementado con 4-HPA 5 mM.

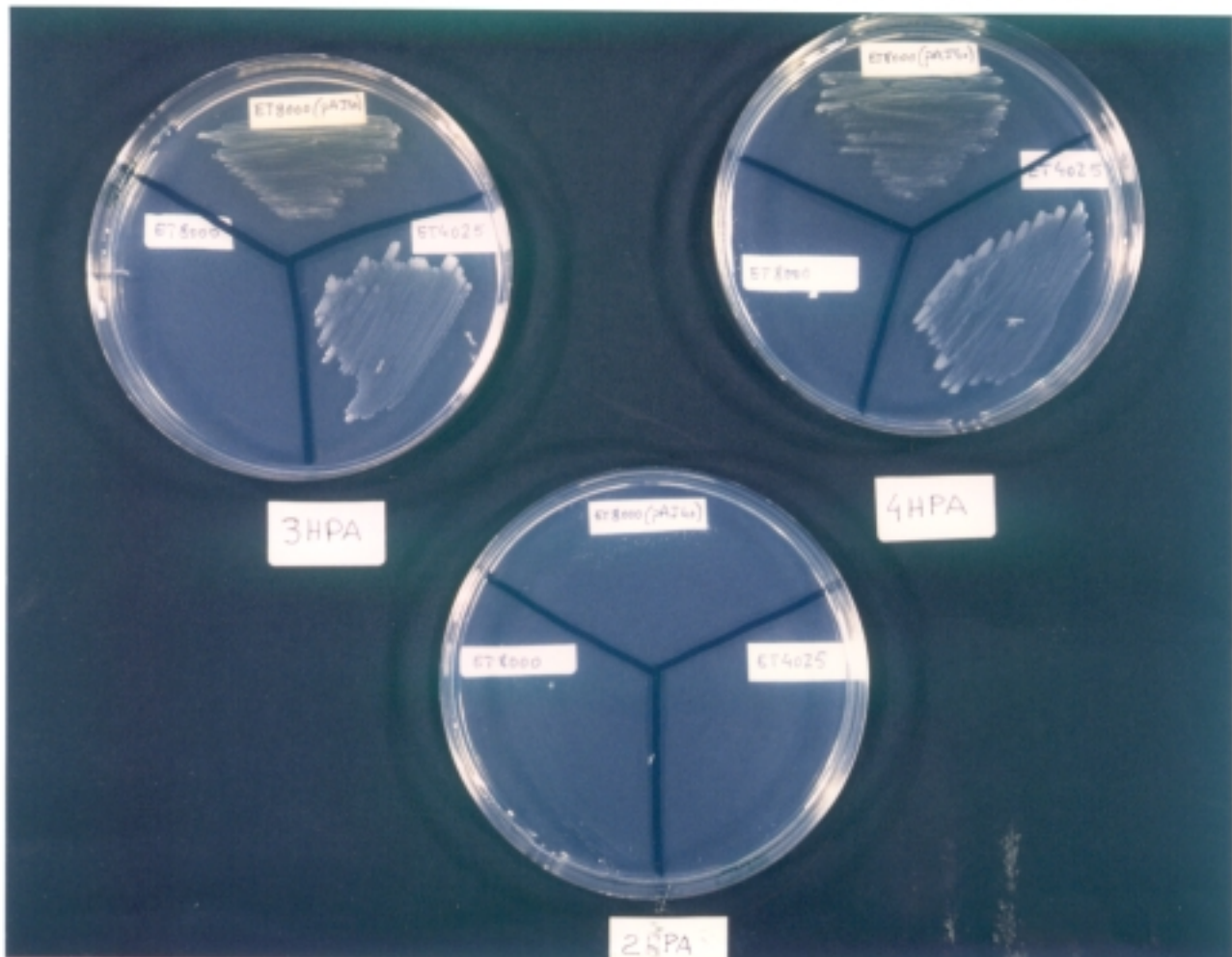


Figura 49. Mineralización de 3- y 4-HPA por las estirpes de *E. coli* K12 que producen las enzimas de la ruta del 4-HPA. Las estirpes ET8000, ET8000 (pAJ40) y ET4025 fueron cultivadas en medio M9 suplementado con glicerol 20 mM y 4-HPA, 3-HPA y 2-HPA a una concentración final de 5 mM.

Tabla 15. Inducción de la 4-HPA-hidroxilasa en diferentes estirpes de *E. coli*

Cepas	Catecol (nmol/ml) ^a	
	-4-HPA	+4-HPA
K12 ET8000	<1	<1
K12 ET4025	<1	329
W ATCC 11105	<1	695

^a Catecol (nmol/ml): La inducción de la actividad hidroxilasa fue determinada mediante la medida *in vivo* de la transformación de fenol en catecol. Las células fueron cultivadas durante toda la noche a 30 °C en medio M9 suplementado con glicerol 20 mM en presencia o ausencia de 4-HPA 1 mM.

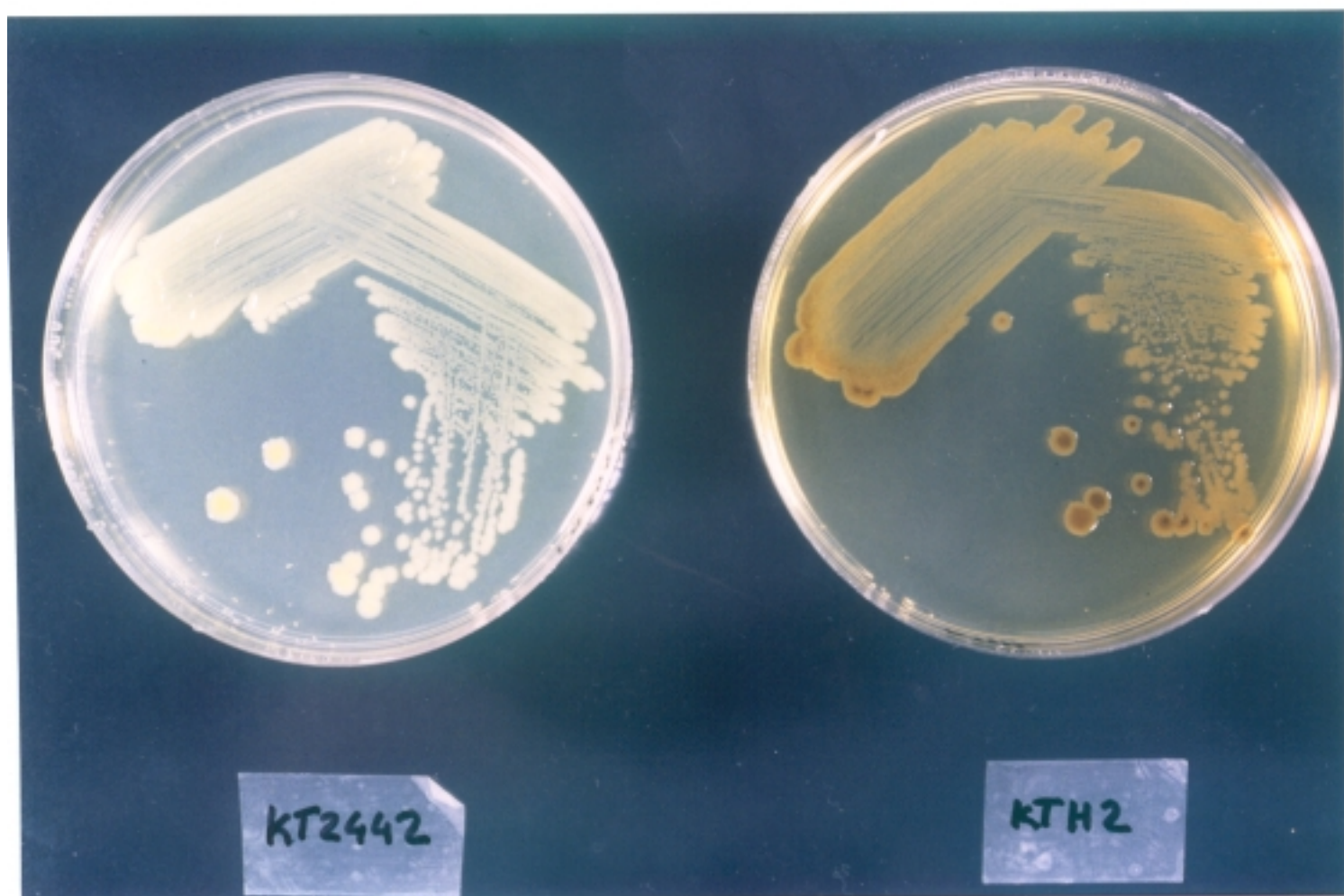


Figura 50. Fenotipo de la cepa *P. putida* KTH2. Las estirpes KT2442 y KTH2 de *P. putida* fueron cultivadas en medio LB a 30 °C.

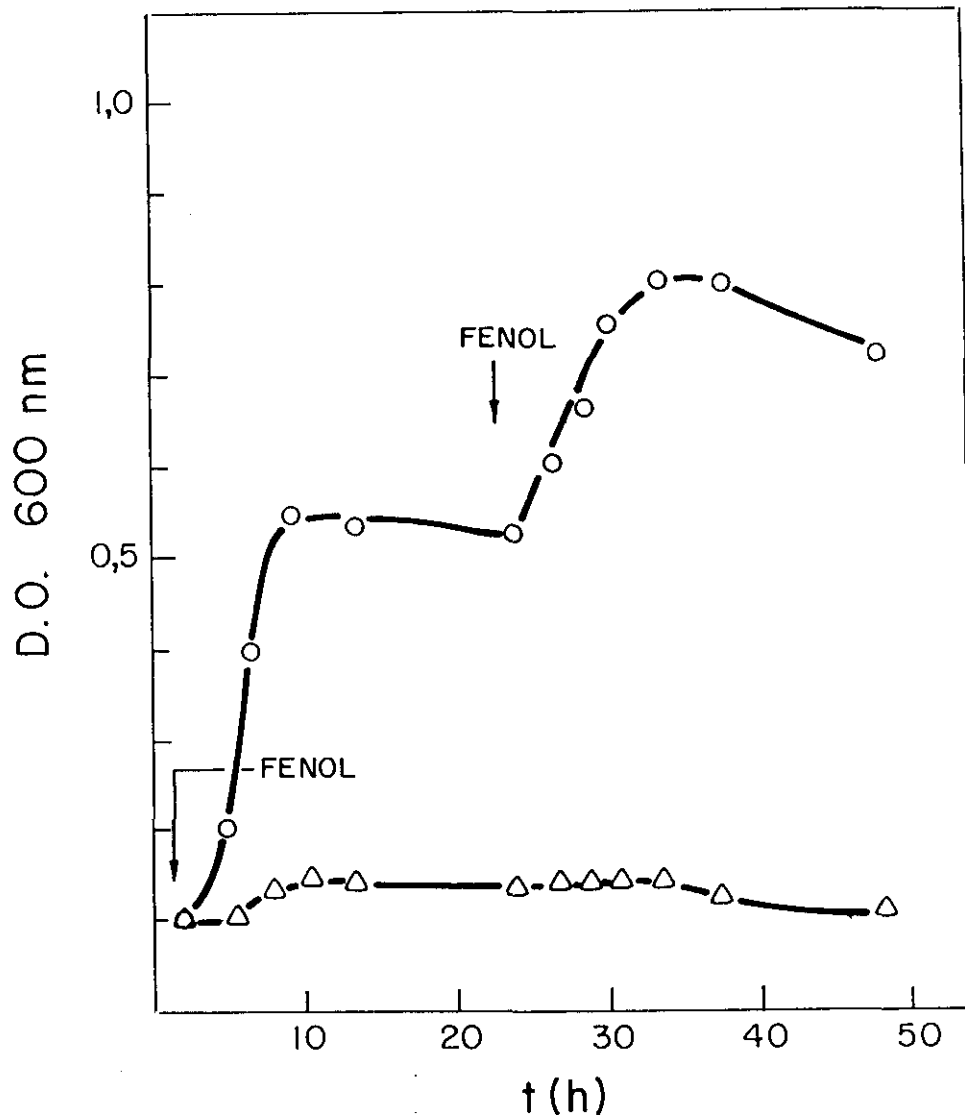


Figura 51. Curva de crecimiento de *P. putida* KTH2 utilizando fenol como fuente de carbono. Las cepas *P. putida* KT2442 (Δ) y *P. putida* KTH2 (O) fueron cultivadas en medio M9 suplementado con IPTG 3 mM y fenol, que se añade a una concentración de 3 mM al iniciarse el cultivo y después de 24 horas de incubación.

IV. DISCUSIÓN

1. CARACTERIZACIÓN DE LA 4-HPA-HIDROXILASA DE *E. coli* W

La posible implicación del gen *pac* en el catabolismo de compuestos aromáticos fue postulada por Valle y cols., (1991) basándose en la capacidad de la PGA para hidrolizar distintos ésteres y amidas del PA y algunos derivados de éste que, posteriormente, pueden ser utilizados como sustratos por los microorganismos productores. En este sentido, se sabe que la cepa de *E. coli* ATCC 11105, productora de PGA, es capaz de utilizar como única fuente de carbono algunos productos de la reacción catalizada por la PGA como el PA y el 4-HPA (Cooper y Skinner, 1980). Sin embargo, hasta el momento no ha sido presentada ninguna evidencia que confirme esta suposición, por lo que el papel fisiológico de estas enzimas todavía no ha sido determinado. Esta hipótesis y la posibilidad de que el clon HB101 (pAEC19), que contiene el gen *pac* de *E. coli*, pudiera producir derivados catecólicos a partir de compuestos aromáticos monohidroxilados, sirvieron de base para el esquema de trabajo en esta Tesis, consistente en la caracterización del gen responsable de la transformación de compuestos aromáticos por este clon así como el estudio de su implicación en el metabolismo de los derivados del PA.

La identificación de los productos de la transformación de fenol, L-tirosina, 4-HPA y 3-HPA en los clones que contienen el plásmido pAJ19 mediante distintos procedimientos, ha permitido confirmar que la enzima codificada por el gen *hpaB* es una hidroxilasa de compuestos aromáticos dependiente de FAD y NADH. Esta enzima ha sido clasificada como una 4-HPA-hidroxilasa ya que presenta la máxima actividad sobre este compuesto, aunque su rango de especificidad de sustrato es muy amplio, siendo capaz de hidroxilar otros compuestos aromáticos como el 3-HPA, el fenol, la L-tirosina y otros derivados de ésta (tabla 9). El hecho de que esta enzima fuera capaz de producir HPC a partir de 3- ó 4-HPA sugirió que podría ser similar a la hidroxilasa dependiente de NADH detectada por Cooper y Skinner (1980) en el extracto crudo de *E. coli* C. También se ha descrito en otros microorganismos la presencia de otras 4-HPA-hidroxilasas que utilizan FAD y NADH ó NADPH como cosustratos (Adachi y cols., 1964; Hareland y cols., 1975; van den Tweel y cols., 1988; Martín y cols., 1991; Arunachalam y cols., 1992). La especificidad de sustrato de cada una de estas enzimas es

diferente; por ejemplo, las 4-HPA-hidroxilasas de *P. acidovorans* y *P. putida* no actúan sobre el 3-HPA (Hareland y cols., 1975; Arunachalam y cols., 1992) mientras que en el caso de *K. pneumoniae* se han descrito dos hidroxilasas diferentes específicas para el 3- y 4-HPA, respectivamente (Martín y cols., 1991). A diferencia de las anteriores, la 4-HPA-hidroxilasa de *Flavobacterium* sp. (van den Tweel y cols., 1988) produce 2,5-DHPA (ácido 2,5-dihidroxifenilacético) en lugar de HPC.

Por otra parte, existen otras monooxigenasas más inespecíficas como la fenol-hidroxilasa (*tbuD*) de *P. picketti* PKO1 que, además del fenol, utiliza como sustratos el *o*-, *m*- y *p*-cresol, el catecol, el clorofenol, el bromofenol y el resorcinol (Kukor y Olsen, 1990). El guaiacol y el 3,4-dimetilfenol se comportan como falsos sustratos de esta enzima, ya que provocan la oxidación del cosustrato (NADPH) independientemente de la incorporación del grupo hidroxilo en el anillo aromático (Kukor y Olsen, 1990). Este desacoplamiento entre las reacciones de oxidación e hidroxilación también ha sido observado en otras oxidasas de función mixta (Beadle y Smith, 1982), por lo que no hay que descartar la posibilidad de que el 3-Cl-4-HPA, el 4-Cl-PA, el *p*-cresol, el 3- y 4-clorofenol, la L-dopa y los derivados dihidroxilados del fenol y del PA (tabla 9), cuyos productos de reacción no han sido determinados en este trabajo, puedan comportarse como falsos sustratos de la 4-HPA-hidroxilasa de *E. coli* W.

Inicialmente se pensó que la 4-HPA-hidroxilasa estaba compuesta únicamente por la proteína HpaB, ya que la cepa DH1 (pAJ27) presentaba actividad 4-HPA-hidroxilasa. Además, la mayoría de las monooxigenasas dependientes de FAD y NADH (ó NADPH) son proteínas monocomponentes con una masa molecular semejante a la calculada para la proteína HpaB (figura 12) (Kukor y Olsen, 1990; Neujahr y Gaal, 1973; Beadle y Smith, 1982; Harayama y cols., 1992). Sin embargo, se han obtenido suficientes pruebas genéticas y bioquímicas que descartan esta suposición. Por una parte, los resultados derivados de la expresión de los genes *hpaB* y *hpaC* en *cis*, en *trans* y en células diferentes (tabla 11), demuestran que la presencia de la proteína HpaC en la mezcla de reacción origina un incremento en la actividad 4-HPA-hidroxilasa de la proteína HpaB. Además, el análisis de la secuencia de nucleótidos de estos genes sugiere que ambos forman parte de un mismo operón (*hpaBC*) (figura 23). En consonancia con estos

resultados, se puede afirmar que la 4-HPA-hidroxilasa es un sistema enzimático formado por dos componentes; la proteína HpaB de 58.781 Da y la proteína HpaC de 18.679 Da.

Las 4-HPA-hidroxilasas de *P. ovalis* y *P. putida* son los únicos sistemas enzimáticos de este tipo descritos hasta el momento, aunque todavía no han sido estudiados a nivel molecular (Adachi y cols., 1964; Arunachalam y cols., 1992). Como se ha comentado en el apartado 3.2.1.1 de la Introducción, la 4-HPA-hidroxilasa de *P. putida* está formada por una flavoproteína dimérica con dos subunidades idénticas de 30,7 kDa y una proteína cooperadora de 38,5 kDa. En este sistema enzimático, el acoplamiento entre las reacciones de oxidación e hidroxilación es llevado a cabo por la proteína cooperadora ya que, en ausencia de ésta, la flavoproteína oxida el NADH en presencia del sustrato y de otros compuestos que se comportan como falsos sustratos, no dando lugar en ningún caso al producto dihidroxilado. La aparente ausencia de centros redox en la proteína cooperadora descarta la posibilidad de que actúe como oxidasa terminal de este sistema, por lo que hasta el momento se desconoce la función de esta proteína. Los dos componentes de las 4-HPA-hidroxilasas de *P. putida* y *P. ovalis* no parecen formar complejos muy estables ya que, en ambos casos, pueden ser separados por una simple precipitación fraccionada con sulfato amónico (Adachi y cols., 1964; Arunachalam y cols., 1992).

En el caso de la 4-HPA-hidroxilasa de *E. coli*, el componente HpaB es capaz de producir catecol a partir de fenol independientemente de la presencia de la proteína HpaC, aunque este proceso es más eficiente cuando ambas proteínas actúan conjuntamente. Además, la oxidación del NADH dependiente de sustrato aumenta en presencia del componente HpaC hasta unos niveles similares a la formación de producto dihidroxilado (tabla 11), lo que demuestra que la oxidación del cosustrato no es un proceso llevado a cabo únicamente por la proteína HpaB. Por otra parte, se ha comprobado durante el proceso de purificación del componente HpaB que el Cibacrón blue actúa como ligando de adsorción de esta proteína, sugiriendo su capacidad para unirse a coenzimas como el FAD y/o el NADH (figura 18B). En cambio, la proteína HpaC eluye durante el proceso de lavado de la columna lo que indica que, como en el caso de las hidroxilasas de *P. putida* y *P. ovalis*, el complejo formado entre ambos

componentes no es muy estable, y además, que esta proteína no debe poseer ningún sitio de unión a NADH o FAD.

Todos estos resultados sugieren que la proteína HpaB es el componente flavoproteico del sistema y que la proteína HpaC actúa como proteína cooperadora del mismo aunque, aparentemente, la formación del complejo FAD-enzima requiere la presencia de ambas (figura 17). No obstante, la comparación de la secuencia de aminoácidos del componente HpaB con la de otras flavoproteínas secuenciadas no reveló ninguna similitud significativa, incluso en las regiones implicadas en la unión a FAD y NADH (Rossman y cols., 1974; Wierenga y cols., 1986; Eggink y cols., 1990; Di Marco y cols., 1993). En este sentido, Howell y cols. (1972) demostraron que algunas flavoproteínas hidroxilasas, que requieren donadores de electrones externos como el NADH o NADPH, presentan un fenómeno común de control según el cual, la reducción de la flavina por NAD(P)H aumenta considerablemente tras la formación del complejo enzima-sustrato. Este mecanismo parece estar limitado a aquellas hidroxilasas que utilizan compuestos aromáticos como sustratos (Ballou, 1982), sugiriendo una correlación entre la presencia de un sitio atípico de unión de FAD/NAD(P)H y el incremento en la afinidad de la unión de NAD(P)H provocada por la presencia del sustrato. Por otra parte, Arunachalam y cols. (1992), han propuesto que la naturaleza bicomponente de la 4-HPA-hidroxilasa de *P. putida* supone un mecanismo de defensa secundario que evita el consumo improductivo de NADH. Por tanto, y de acuerdo con estas observaciones, es coherente suponer que esta nueva familia de hidroxilasas bicomponentes puedan tener un sitio atípico de unión a FAD/NADH que hasta el momento no ha sido caracterizado.

El mecanismo por el cual la proteína HpaC favorece la hidroxilación del sustrato mediada por el componente HpaB aún no ha sido esclarecido. Esta proteína se parece a la ORF6 del *cluster* de genes de la síntesis de la actinorrodina de *S. coelicolor* (figura 15) (Fernández -Moreno y cols., 1992) y cuya función no se ha determinado exactamente. Se ha propuesto que esta proteína podría actuar como una dimerasa uniendo dos moléculas del último precursor de la actinorrodina para producir la estructura final del antibiótico, aunque no se descarta que pudiera actuar como una proteína reguladora.

2. COMPARACIÓN DE LA RUTA DEL 4-HPA DE *E. coli* W CON OTRAS RUTAS CATABÓLICAS

Generalmente, para relacionar dos rutas metabólicas desde el punto de vista evolutivo se tienen en cuenta dos criterios; por una parte, la similitud en la estructura primaria de las enzimas equivalentes y por otra, la conservación en la organización física de los genes (Williams y Sayers, 1994). Se puede afirmar que dos enzimas derivan de un mismo gen ancestral cuando cumplen los siguientes requisitos: i) catalizan la misma reacción, ii) las subunidades de éstas tienen una masa molecular similar, iii) sus secuencias de aminoácidos pueden ser alineadas sin introducir huecos y iv) el porcentaje de identidad entre ambas proteínas es superior al 30 % (Harayama y Timmis, 1992). Teniendo en cuenta estos criterios, en este apartado se va a estudiar la relación entre la ruta metabólica del 4-HPA de *E. coli* W y otras rutas microbianas de degradación de compuestos aromáticos.

2.1. Comparación de los genes *hpa* con los genes estructurales de otras rutas metabólicas de derivados aromáticos

Los primeros trabajos en el estudio de la capacidad de *E. coli* para metabolizar compuestos aromáticos fueron realizados por Cooper y Skinner (1980). Estos autores establecieron que la estirpe C de *E. coli* mineraliza el 3- y 4-HPA por una ruta inducible de tipo *meta*, donde el HPC actúa como intermediario central dihidroxilado. A pesar de que el operón *meta* de degradación del HPC (*hpcECBDGH*) de esta estirpe ha sido clonado y secuenciado (Roper y cols., 1993), el operón de la 4-HPA-hidroxilasa, así como los genes localizados entre ambos operones (*hpaA* y *hpaX*) (figura 24), no habían sido clonados hasta el desarrollo de esta Tesis. Dado que *E. coli* W también es capaz de metabolizar el 3- y 4-HPA, existía la posibilidad de que las rutas catabólicas de estos compuestos en ambas estirpes pudieran ser homólogas. Los resultados decisivos que permitieron la verificación de esta hipótesis fueron: i) la confirmación de la existencia de secuencias homólogas a *hpaB* en *E. coli* C (figura 20), ii) la clonación del operón *hpaBC* de *E. coli* C en el plásmido pHc1 y la secuenciación parcial del mismo, iii) la demostración de la presencia de genes homólogos a la HPC-dioxigenasa de *K.*

pneumoniae tanto en la estirpe C como en la estirpe W (figura 21) y iv) la clonación y secuenciación del operón *meta* de degradación del HPC de *E. coli* W (figuras 22-24). Los resultados obtenidos mediante análisis por Southern blot (figuras 20 y 21) también sugieren que otros microorganismos capaces de utilizar el 4-HPA como única fuente de carbono como *E. coli* B, *K. pneumoniae* y *K. citrophila* podrían tener una ruta catabólica similar a la de las estirpes C y W de *E. coli*.

La ruta catabólica del 4-HPA de *E. coli* W está compuesta por once genes; ocho genes estructurales organizados en dos operones catabólicos, el operón *hpaGEDFHI* u operón *meta* y el operón *hpaBC*, dos genes reguladores *hpaR* y *hpaA*, y el gen *hpaX* de función desconocida, que está relacionado con la superfamilia de transportadores transmembranales MFS (figura 24). Ya se ha mencionado que las enzimas del operón *meta* son homólogas a las de la estirpe C de *E. coli*, sin embargo, la comparación de la secuencia de aminoácidos de las enzimas que componen tanto el operón *meta* como el operón *hpaBC*, ha revelado que sólo los genes *hpaE*, *hpaG* y *hpaH* codifican para proteínas que se parecen a las enzimas equivalentes en otras rutas catabólicas de compuestos aromáticos (figuras 25 y 26).

La deshidrogenasa HpaE pertenece a la superfamilia de las aldehído deshidrogenasas (ADH) igual que las deshidrogenasas XylG y XylC de la ruta TOL (figura 4) y la proteína DmpC de la ruta de degradación de fenol de *Pseudomonas* CF600 (Dmp) (figura 26) (Horn y cols., 1991; Nordlund y Shingler, 1990). La comparación de HpaE con otros miembros de las ADH indica que en esta enzima están conservados todos los motivos característicos de la superfamilia. Por ejemplo, la Cys-278 de HpaE parece ser un residuo esencial implicado en la unión del NAD⁺, además, la región anterior a esta cisteína, incluyendo la Phe-271, está muy conservada en estas proteínas. Otra secuencia característica, LGGK, está situada a continuación del Glu-244 de HpaE. Este aminoácido parece estar implicado en el relé de carga o bien se trata del aminoácido que actúa como nucleófilo. Por otra parte, en el motivo localizado entre la Gly-224 y la Gly-229, la secuencia FTGGTATG se corresponde con la secuencia consenso (F/Y)TG(S)(T)XXG presente en esta familia, que también parece estar implicada en la unión del NAD⁺. Asimismo, el motivo GTGREG localizado en la

posición 454 de HpaE parece ser el responsable del reconocimiento del cosustrato. Por último, el Glu-470 de HpaE podría participar en el relé de carga aunque este residuo no está conservado en todos los miembros de las ADH (Horn y cols., 1991).

La proteína HpaE es equivalente a las deshidrogenasas XylG y DmpC que son las primeras enzimas de la rama del 4-oxalocrotonato (rama deshidrogenasa/descarboxilasa) de la ruta TOL y de la ruta Dmp (figuras 2 y 4). Es interesante resaltar que, a diferencia de estas rutas que pueden degradar los derivados catecólicos por la rama del 4-oxalocrotonato o por la rama hidrolítica (Powlowski y Shingler, 1994), la ruta del 4-HPA de *E. coli* es la única descrita hasta el momento que únicamente consta de la rama deshidrogenasa/descarboxilasa.

Las descarboxilasas e hidratasas de las rutas de tipo *meta* se caracterizan por poseer un origen común a pesar de tener funciones diferentes (Harayama y cols., 1989; Shingler y cols., 1992). Estas enzimas están codificadas por genes diferentes cuyos productos presentan un porcentaje de identidad superior al 30%, por lo que se supone que provienen de la duplicación génica de un gen ancestral. Los genes *xylJ* y *dmpE* codifican para las 2-oxopent-4-enoato-hidratasas de las rutas TOL y Dmp. En cambio, la actividad descarboxilasa sólo es evidente en los clones que expresan conjuntamente los genes *xylI* y *xylJ* de la ruta TOL, y *dmpE* y *dmpH* de la ruta Dmp, no detectándose en los clones que producen únicamente las descarboxilasas XylI ó DmpH. Además, todos los intentos de purificación de la 4-oxalocrotonato-d Descarboxilasa activa han dado como resultado una copurificación de las enzimas XylI/XylJ ó DmpE/DmpH (Harayama y cols., 1989; Shingler y cols., 1992).

La descarboxilasa HpaG y la hidratasa HpaH se parecen a esta familia de proteínas pero, a diferencia de las descarboxilasas anteriores, la proteína HpaG parece haber sufrido dos procesos evolutivos; una duplicación génica y una fusión posterior de ambos genes, ya que sus extremos N- y C- terminal son muy similares (figura 25). Además, los experimentos realizados por Roper y Cooper (1993) con la proteína homóloga (HpcE) de *E. coli* C indican que esta proteína es la única responsable de la actividad descarboxilasa. Por otra parte, estos autores atribuyen a la proteína HpcE una

doble función OPET-descarboxilasa/HHDD-isomerasa. En este sentido, Harayama y cols., (1989) demostraron que, en la ruta TOL, la reacción equivalente a la isomerización llevada a cabo por esta enzima entre los compuestos HHDD (forma enol) y OHED (forma ceto) (figura 24), es un paso enzimático dispensable, teniendo en cuenta que, debido a su gran inestabilidad, la forma enol se transforma espontáneamente en la forma ceto. Para llevar a cabo estos experimentos, estos autores sintetizaron químicamente el 2-hidroxipent-2,4-dienoato (forma enol) y demostraron que este compuesto es el sustrato real de la hidratasa XylJ. Sin embargo, Roper y Cooper, (1993) demostraron la actividad isomerasa de la enzima HpcE utilizando el HHDD obtenido mediante la descarboxilación enzimática del OPET en un clon que contenía el gen *hpcE* en el que no se detectó actividad HHDD-isomerasa (Jenkins y Cooper, 1988). La consideración de los procedimientos de obtención de la forma enólica utilizados en ambos casos, hace necesario confirmar los resultados obtenidos para HpcE mediante un método similar al utilizado por Harayama y cols. (1989).

El resto de los genes que componen el operón *meta*, incluso su gen regulador *hpaR*, no se parecen a los genes equivalentes de otras rutas. Hasta el momento se conoce la secuencia completa de los genes estructurales de las rutas TOL, Dmp, TOD (ruta de degradación del tolueno de *P. putida* F1) y algunas rutas de degradación de policlorobifenilos de distintas estirpes del género *Pseudomonas* (Horn y cols., 1991; Powlowski y Shingler, 1994; Wang y cols., 1995; Furukawa, 1994) (figura 2). La comparación de las enzimas que componen estas rutas ha demostrado que todos los genes equivalentes tienen un origen común y que, aparentemente, mutaciones producidas a lo largo de la evolución han causado cambios en la especificidad enzimática provocando diferencias en la capacidad biodegradativa de los microorganismos (Williams y Sayers, 1994; Harayama y Rekik, 1993). El origen de los genes *hpaD*, *hpaF* y *hpaI* parece no estar relacionado con el de las correspondientes dioxigenasas, isomerasas y aldolasas de las rutas citadas anteriormente, lo que implica un claro ejemplo de evolución convergente.

La presencia de permeasas en las rutas de degradación de compuestos aromáticos no ha sido estudiada hasta hace relativamente poco tiempo. Ultimamente se han

secuenciado algunos genes integrados en operones de rutas catabólicas o situados inmediatamente al lado de estos, que codifican para proteínas que presentan una similitud muy alta con proteínas transportadoras. Por ejemplo, el gen *todX* forma parte del operón *tod* (figura 2) y su producto se parece a la proteína FadL de *E. coli* implicada en el transporte a través de la membrana celular de ácidos grasos de cadena larga. La función de este gen no está todavía muy clara, pero se ha propuesto que puede actuar como transportador del tolueno o como proteína receptora de tolueno, responsable de la activación de las proteínas reguladoras (Wang y cols., 1995). Otros ejemplos de este tipo de proteínas son la proteína Pht1 de la ruta de degradación del ftalato de *P. putida* y la proteína PcaK de la ruta *orto* de degradación de 4-hidroxibenzoato de *P. putida* PRS2000 (Nomura y cols., 1992; Harwood y cols., 1994). Ambas pertenecen a la superfamilia de transportadores transmembranales (MFS) (Marger y Saier, 1993) aunque todavía no se conoce exactamente su función. En el caso de la proteína Pht1 se ha sugerido que podría actuar como transportador de ftalato o como regulador positivo de los genes *pht* (Nomura y cols., 1992). En cambio, la enzima PcaK podría actuar, además de como transportador de 4-hidroxibenzoato, como proteína quimiorreceptora de sustrato, y por tanto, estar implicada en la respuesta quimiotáctica que *P. putida* experimenta en presencia de este compuesto (Harwood y cols., 1994).

La comparación de la secuencia de aminoácidos de la proteína HpaX y el análisis de su perfil hidrofóbico (figuras 32 y 33) sugieren que esta enzima es un miembro de la MFS como los casos anteriormente expuestos. La mayoría de los miembros de esta familia contienen el motivo (R/K)XXX(R/K) entre las α -hélices 2-3 y 8-9 que da lugar a una región altamente hidrofílica (Marger y Saier, 1993). Entre las hélices 8-9 de HpaX se han localizado dos motivos consecutivos similares al consenso entre la Arg-318 y la Arg-327 (RHSDRRQERR), en cambio, en la región situada entre las hélices 2 y 3, el carácter hidrofílico es aportado por el motivo VGARR situado entre la Val-96 y la Arg-100 (figura 32).

Hasta el momento no se ha clonado ningún gen que codifique para una 4-HPA-permeasa, aunque se ha detectado la presencia de estas proteínas en algunos microorganismos que mineralizan el 4-HPA como *K. pneumoniae* y *E. coli* C (Allende y

cols., 1992; Cooper y Skinner, 1980). Los resultados obtenidos del análisis de la secuencia de HpaX sugieren que esta enzima podría actuar como un transportador de 4-HPA aunque, hasta que se disponga de las pruebas bioquímicas que lo demuestren, no se descarta que pueda estar implicada en procesos de regulación y/o quimiotaxis como se ha propuesto para las proteínas TodX, PcaK y Pht1.

2.2. Los genes reguladores de la ruta del 4-HPA

Generalmente, las rutas bacterianas de derivados aromáticos están reguladas positivamente por proteínas pertenecientes a las familias LysR, NtrC ó XylS/AraC (Inouye y cols., 1988; Schell, 1993; Gallegos y cols., 1993). Sin embargo, los trabajos de Roper y cols. (1993) demuestran que la expresión de los genes del operón *hpcECBDGH* de *E. coli* C está regulada negativamente por la proteína HpcR que, como ya se ha mencionado, no se parece a ninguna proteína secuenciada hasta el momento. La represión mediada por esta proteína depende de la presencia de HPC y/o 4-HPA que actúan como inductores. Además, estos autores determinaron que la expresión del operón *meta* está sujeta a represión catabólica, lo que justifica la presencia de un posible sitio de unión a la proteína CAP localizado en la región promotora del operón (figura 37). Aunque aún no se han realizado los experimentos bioquímicos y genéticos necesarios para extrapolar estos resultados al operón *hpaGEDFHI* y su regulador *hpaR*, dada la homología que presentan las enzimas, que en el caso de las proteínas HpaR y HpcR supone un 100% de identidad (figura 30), se asume que el mecanismo de regulación es similar en ambas estirpes.

La proteína HpaA es un miembro de familia de reguladores XylS/AraC (Gallegos y cols., 1993) que regula positivamente la expresión del operón *hpaBC* utilizando como efectores 3-HPA, 4-HPA y PA. Los miembros de esta familia se caracterizan por estar estructurados en dos dominios: el C-terminal que presenta un motivo muy conservado implicado en la unión a DNA (figura 31) (Gallegos y cols., 1993; Michán y cols., 1995) y el N-terminal, menos conservado entre los miembros de esta familia, que podría estar implicado en la unión enzima-efector aunque todavía no existen las pruebas suficientes para confirmarlo.

La comparación de la secuencia de aminoácidos de la proteína HpaA reveló que MelR es el miembro de esta familia que más se le parece (figura 31). Esta proteína regula positivamente la transcripción del operón *melAB* implicado en el catabolismo de la melobiosa en *E. coli* (Webster y cols., 1989). Las diferencias más notables en la transcripción de los genes *hpaA* y *melR* son, por una parte, que *melR* se transcribe en sentido contrario al operón al que regula, como la mayoría de los miembros de esta familia (Webster y cols., 1988), sin embargo, el gen *hpaA* se transcribe en el mismo sentido que el operón *hpaBC*. Además, los experimentos realizados con la cepa MC4100 (pROH1, pRE1) y MC4100 (pRA2) (figuras 44 y 45) indican que la transcripción del gen *hpaA* puede estar mediada por los promotores P_A y/o P_X (figura 43). Estos resultados sugieren que este gen podría formar parte del operón *hpaXA* aunque esta hipótesis no podrá confirmarse hasta que se analice el número de unidades de transcripción generadas en esta región. Por otra parte, la expresión del operón *hpaBC* está sujeta a represión por catabolito (figura 46). Dado que se ha localizado un posible sitio de unión a la CAP en el promotor P_A (figura 47C), y que este efecto no se observa en las cepas MC4100 (pROH1) y MC4100 (pRX2) (datos no mostrados), se podría asumir que la represión catabólica observada es debida a la intervención de la proteína CAP en la expresión del gen *hpaA* mediada por el promotor P_A . En este sentido, se sabe que la expresión de *melR* también está regulada por la CAP (Webster y cols., 1988).

Las proteínas HpaA y XylS son los únicos miembros de esta familia implicados en la regulación de rutas catabólicas de compuestos aromáticos (Gallegos y cols., 1993). Como se mencionó en el apartado 3.2.4. de la Introducción, XylS regula positivamente la expresión del operón *meta* de la ruta TOL y su producción está controlada por el regulador XylR y el factor σ^{54} de la RNA polimerasa (Marqués y Ramos, 1993). Se ha propuesto que la forma activa de la proteína es un dímero cuya formación se favorece en presencia de los efectores y/o cuando XylS está hiperproducida (Marqués y Ramos, 1993). Los ensayos realizados con la cepa MC4100 (pRX2) demuestran que P_X es un promotor muy fuerte, por lo que los elevados niveles de β -galactosidasa producidos por la cepa MC4100 (pROH1, pRE1) en ausencia del efector (figura 44), podrían deberse a

una hiperproducción de HpaA, la cual, como en el caso de XylS (Marques y Ramos, 1993), provocaría la dimerización de la proteína y por tanto activaría el promotor P_{BC} .

Desde el punto de vista cualitativo, la activación del promotor P_{BC} cuando el gen *hpaA* se expresa mediante el promotor P_A , es similar a la de la cepa *E. coli* ATCC 11105 (pROH1) (figuras 44 y 45). Todos estos resultados sugieren que la expresión mediada por el promotor P_X podría estar controlada por algún represor en la estirpe salvaje el cual no está presente en la cepa MC4100. En este sentido, los ensayos realizados con la cepa MC4100 (pRX2, pHCR2), donde se expresan conjuntamente la fusión P_X -*lacZ* y el gen *hpaR* (datos no mostrados), indican que el regulador HpaR no parece estar implicado en la expresión mediada por P_X . No obstante, hay que tener en cuenta que, a pesar de que los experimentos desarrollados en este trabajo aportan información muy valiosa sobre el mecanismo de regulación del promotor P_{BC} , los sistemas multicopia no son los más apropiados para analizar promotores desde el punto de vista cuantitativo, puesto que el nivel de expresión del gen *lacZ* depende del número de copias del plásmido. En cambio, un sistema en el que, tanto el gen regulador como la fusión P_{BC} -*lacZ*, hubieran sido insertados en el cromosoma, reproduciría más exactamente la situación existente en la cepa salvaje (Kessler y cols., 1992). Por ello, hasta que no se disponga de un sistema monocopia en el que el gen *hpaA* se exprese, por una parte a partir de P_A y, por otra, a partir de P_X , no podremos confirmar cuál es el promotor implicado en la expresión de *hpaA*.

Los únicos sustratos de la 4-HPA-hidroxilasa que se comportan como efectores de la proteína HpaA son el 3-HPA y el 4-HPA. En cambio, el PA induce la producción de la 4-HPA-hidroxilasa pero no es transformado por esta enzima. Dado que *E. coli* W es capaz de utilizar el PA como única fuente de carbono, la inducción de la 4-HPA-hidroxilasa por este compuesto en la estirpe salvaje (figura 41, tabla 12) podría sugerir la existencia de una PA-monooxigenasa mediante la cual el PA sería transformado en 3- ó 4-HPA, que serían los verdaderos responsables de la inducción. Esta enzima permitiría conectar las rutas del PA y del 4-HPA. Sin embargo, el hecho de que el PA también provocara la activación de la proteína HpaA en la cepa MC4100 (pRA2) (figura 45) excluye esa posibilidad ya que, como se ha demostrado en esta Tesis, *E. coli* K12 no

contiene la ruta del 4-HPA, y por otra parte, la transformación de PA en esta estirpe no origina la acumulación de 4-HPA (Cooper y cols., 1985). Por tanto, es más coherente sugerir que, en este caso, la proteína HpaA se comporta de una forma inespecífica en el reconocimiento del efector. Además, este efecto también se ha observado en la proteína XylS, ya que algunos efectores son compuestos metabolizables por la ruta TOL como el benzoato y sus derivados 3-metil-, 3-etil-, 4-metil y 3,4-metil-benzoato. Sin embargo, el 2-metil-benzoato y ciertos derivados halogenados del benzoato son capaces de activar la proteína pero no son mineralizables por esta ruta (Ramos y cols., 1986).

2.3. La organización física de los genes

Como se mencionó al comienzo del apartado 2., uno de los criterios utilizados para relacionar las rutas catabólicas desde el punto de vista evolutivo es la conservación en la organización física de los genes. Esta premisa permite confirmar el origen común de los genes que componen la ruta *meta* de algunas rutas catabólicas. Por ejemplo, la secuencia de reacciones que conllevan a la mineralización del catecol de las rutas TOL, Dmp y de la ruta del naftaleno de *P. putida* PpG7 (NAH) (figura 2) es la misma, ya que las tres contienen las enzimas que codifican para la ramas hidrolasa y deshidrogenasa/descarboxilasa, y el porcentaje de similitud entre las enzimas equivalentes es muy elevado (Harayama y Rekik, 1993). Además, el orden de transcripción de los genes *xyITEGFJQKIH* (ruta *meta*) de la ruta TOL, es idéntico al de los genes equivalentes de la ruta Dmp y prácticamente igual al de la ruta NAH. Sin embargo, esta similitud disminuye entre las enzimas que codifican para las rutas periféricas, cuyos genes están localizados en la región inmediatamente anterior al primer gen de la ruta *meta* (*xylT*, *dmpQ* y *nahT*) formando parte del mismo operón (figura 2) (Williams y Sayers, 1994). Estos genes suelen estar relacionados con los equivalentes de otras rutas, así por ejemplo, en el caso de la ruta TOL, los genes que codifican para la toluato dioxigenasa (*xylXYZ*), son muy similares a los genes que codifican para la benzoato dioxigenasa de *A. calcoaceticus* (*benABC*) (figura 2) (van der Meer y cols., 1992). En consecuencia, algunos autores han sugerido que la evolución de las rutas metabólicas se produce mediante el ensamblaje de módulos funcionales, en el que están implicados procesos de

duplicación génica, recombinación y transposición (van der Meer y cols., 1992; Williams y Sayers, 1994; Wyndham y cols., 1994).

Los resultados obtenidos en esta Tesis sugieren que, teniendo en cuenta solamente la similitud entre proteínas equivalentes, los únicos genes de la ruta del 4-HPA que podrían tener un origen ancestral común a otros genes equivalentes de las rutas de tipo *meta*, son los genes *hpaG*, *hpaH*, *hpaE* y *hpaA*. En cambio, el orden de transcripción de los genes de esta ruta no permite establecer ninguna relación evolutiva con ninguna ruta catabólica que no sea la de *E. coli* C (figuras 2 y 24).

Por otra parte, se ha demostrado que el *cluster* de genes de la ruta del 4-HPA ha sido adquirido por la estirpe W como un *cassette* catabólico (figuras 36 y 37). Este proceso de adaptación microbiana permite relacionar esta ruta con otros sistemas, ya que la adquisición de rutas metabólicas completas mediante procesos de transferencia genética ha sido ampliamente demostrada en otros microorganismos (Reineke y Knackmuss, 1979; van der Meer y cols., 1992). De hecho, se han descrito algunos casos en los que todos los genes responsables de la mineralización de un compuesto están incluidos en un transposón, como por ejemplo la ruta TOL y la ruta NAH (Wyndham y cols., 1994). Hasta el momento, no se ha encontrado ninguna evidencia que permita especular sobre el mecanismo implicado en la adquisición de la ruta del 4-HPA por la estirpe W, aunque la ausencia de un gen que codifique para una típica transposasa, así como la ausencia de secuencias invertidas en los extremos del *cluster*, sugieren que no se ha adquirido mediante un proceso típico de transposición.

Por último, se ha descrito la tendencia de algunos genes a insertarse en el cromosoma en una posición muy cercana a la de otros genes que codifican para enzimas relacionadas fisiológicamente con las proteínas producidas por los primeros. Aunque todavía no se conocen los factores que determinan estos procesos, algunos autores han sugerido que no se trata de un hecho aleatorio (Houghton y cols., 1995). Por tanto, no debería descartarse la posibilidad de que el proceso de inserción del gen *pac* en la región del cromosoma próxima a la del *cluster* de la ruta del 4-HPA, haya sido controlado por algún proceso de adaptación microbiana, dirigido a ampliar la capacidad biodegradativa

de *E. coli* W. Por otra parte, sería interesante investigar si existe alguna relación funcional entre genes de la ruta del 4-HPA y las ORFs 12 y 13 ya que, aunque la función del producto del gen *cstA* no ha sido determinada con exactitud, hay que tener en cuenta que este gen se ha localizado en *E. coli* MC4100 a 600 pb del operón *entCEBA-P15*, que es responsable de la síntesis del compuesto aromático enterobactina, y que la expresión de algunos genes de este tipo se induce en presencia de ciertos compuestos aromáticos (Schultz y Matín, 1991; Blom y cols., 1992).

3. PERSPECTIVAS DE FUTURO EN LA UTILIZACIÓN DE LOS GENES DE LA RUTA DEL 4-HPA EN LA DESCONTAMINACIÓN MEDIOAMBIENTAL

Aunque el 3- y 4-HPA no son compuestos xenobióticos, la acumulación masiva de éstos puede acarrear problemas de toxicidad además de provocar otros efectos nocivos, desde el punto de vista medioambiental, derivados de su mal olor (Spoelstra, 1978). Ya se ha comentado que, puesto que *E. coli* W es capaz de mineralizar estos compuestos, este microorganismo podría ser utilizado en los procesos de depuración biológica de residuos con alto contenido en 4-HPA, como el alpechín (apartado 5.2 de la Introducción). En este sentido, la caracterización de la ruta del 4-HPA realizada en este trabajo, ha permitido determinar cuáles son los genes implicados en esta ruta, y así disponer de la información necesaria para desarrollar un sistema que permita transferir a otros microorganismos la capacidad de metabolizar el 3- y 4-HPA (figuras 48-50). La viabilidad del uso de mini-transposones derivados de Tn5 en procesos de transferencia genética se ha comprobado en diferentes especies de los géneros *Pseudomonas*, *Escherichia*, *Klebsiella*, *Salmonella*, *Proteus*, *Vibrio*, *Bordetella*, *Actinobacillus*, *Rhizobium*, *Rhodobacter*, *Agrobacterium* y *Alcaligenes* (de Lorenzo y Timmis, 1994). Por tanto, la construcción de un módulo transponible con los genes de la ruta del 4-HPA en derivados mini-Tn5, hace posible que los microorganismos antes mencionados puedan ser también considerados como candidatos en el diseño de biorreactores para la depuración biológica del alpechín.

En esta Tesis también se ha demostrado que la producción de la 4-HPA hidroxilasa de *E. coli* W en *P. putida*, permite a este microorganismo utilizar el fenol como única fuente de carbono (figura 51), constituyendo un claro ejemplo de la aplicación de la ingeniería genética en la expansión de la capacidad metabólica microbiana. La baja especificidad de sustrato que presenta esta enzima deja abierta la posibilidad de utilizar el vector-transposón pAJ224 para insertar el operón *hpaBC* en el genoma de otros microorganismos capaces de mineralizar algún producto de la reacción mediada por la 4-HPA-hidroxilasa, y así ampliar el espectro de sustancias metabolizables por dichos microorganismos.

Por otra parte, el catecol se utiliza en la industria farmacéutica y alimentaria como compuesto base para la síntesis de algunos medicamentos y aromatizantes. Actualmente, la obtención del catecol se lleva a cabo mediante un proceso de oxidación química del fenol que da lugar a una mezcla de productos aromáticos hidroxilados. Esta mezcla se somete posteriormente a una serie de procesos de purificación que elevan considerablemente el coste total del proceso. En este sentido, sería interesante estudiar la posibilidad de obtener catecol enzimáticamente mediante el uso de la 4-HPA-hidroxilasa de *E. coli* e incluso mediante procesos de fermentación bacteriana.

Por último, la activación efector-dependiente de la proteína HpaA permite considerar a esta proteína como un candidato para el desarrollo de biosensores (van der Lelie y cols., 1994) que permitan la detección de la presencia de PA, 3- y 4-HPA en un entorno determinado, como por ejemplo en alimentos, puesto que, como ya se ha comentado, la concentración de PA y alguno de sus derivados es un parámetro a tener en cuenta en la clasificación de ciertos alimentos (Stegg y Montag, 1988).

V. CONCLUSIONES

De los resultados presentados en esta memoria se derivan las siguientes conclusiones:

1. La ruta catabólica del 4-HPA de *E. coli* W está compuesta por once genes; ocho genes estructurales organizados en los operones *hpaGEDFHI* y *hpaBC*, los genes reguladores *hpaR* y *hpaA*, y el gen *hpaX* que podría actuar como una 4-HPA-permeasa.
2. El operón *hpaBC* codifica para la hidroxilasa HpaB y la proteína cooperadora HpaC, siendo ambos componentes necesarios para que se manifieste la máxima actividad hidroxilasa
3. La 4-HPA-hidroxilasa es un sistema enzimático dependiente de FAD y NADH que presenta un rango muy amplio de especificidad de sustrato ya que además del 3- y 4-HPA puede hidroxilar otros compuestos aromáticos.
4. La 4-HPA-hidroxilasa no presenta similitud con ninguna de las monooxigenasas secuenciadas hasta la fecha, por lo que puede considerarse como el primer miembro de una nueva familia.
5. La expresión del operón *hpaBC* está regulada positivamente por la proteína HpaA, perteneciente a la familia de reguladores XylS/AraC. Entre los distintos compuestos ensayados en este trabajo, sólo el 4-HPA, el 3-HPA y el PA se comportan como efectores de la proteína HpaA. La expresión de la proteína HpaA está sujeta a represión catabólica.
6. El operón *hpaGEDFHI* está situado en una posición adyacente al operón *hpaBC* y es similar al operón *meta* de la ruta catabólica del HPC de *E. coli* C.
7. El gen *hpaR* es similar al gen *hpcR* de *E. coli* C que codifica para el represor del operón *meta* en esta estirpe. Los productos de ambos genes no se parecen a ninguna proteína reguladora secuenciada hasta el momento.

8. La descarboxilasa HpaG, la deshidrogenasa HpaE y la hidratasa HpaH presentan una similitud significativa con las enzimas equivalentes de otras rutas catabólicas de compuestos aromáticos. Por el contrario, la dioxigenasa HpaD, la isomerasa HpaF y la aldolasa HpaI sólo muestran similitud con las enzimas equivalentes de la ruta del HPC de *E. coli* C.

9. El análisis de las regiones flanqueantes de los genes de la ruta catabólica del 4-HPA de *E. coli* W ha permitido concluir que este *cluster* se ha insertado en el cromosoma de esta estirpe como un módulo funcional. La comparación de esta región con la secuencia de nucleótidos localizada entre los minutos 92,8 y 0,1 del cromosoma de la estirpe MG1655 de *E. coli* K12, ha permitido determinar que la ruta del 4-HPA se ha insertado exactamente a 61 pb del codon de iniciación del gen *tsr* de *E. coli*, que codifica para la proteína quimiorreceptora de serina.

10. Todos los genes necesarios para que *E. coli* K12 mineralice el 4-HPA están contenidos en este *cluster*, ya que su expresión en este microorganismo le permite utilizar 4-HPA como fuente de carbono.

11. La expresión del operón *hpaBC* en *P. putida* KT2442 capacita a este microorganismo para mineralizar fenol, lo que supone una expansión vertical de su capacidad degradativa.

VI. BIBLIOGRAFÍA

- Adachi, K., Takeda, Y., Senoh, S., and Kita, H. 1964. *Biochim. Biophys. Acta* 93, 483-493.
- Adams, S.S. 1992. *J. Clin. Pharmacol.* 34, 317-323.
- Afring, P.R., Chalker, B.E. y Taylor, B.F. 1981. *Appl. Environ. Microbiol.* 41, 1117-1183.
- Aldrich, T.L. Frantz, J.F., Gill, J., Killbane, J.J. y Chakrabarty A.M. 1987. *Gene* 52, 185-195.
- Altenschmidt, U., Oswald, B., y Fuchs, G. 1991. *J. Bacteriol.* 173, 5494-5501.
- Allende, J.L., Gibello, A., Martín, M. y Garrido-Pertierra, A. 1992. *Arch. Biochem. Biophys.* 292, 583-588.
- Ames, P. y Parkinson, J. 1994. *J. Bacteriol.* 176, 6340-6348.
- Anson, J.G. y Mackinnon, G. 1984. *Appl. Environ. Microbiol.* 48, 868-869.
- Arunachalam, U. y Massey, V. 1994. *J. Biol. Chem.* 269, 11795-11801.
- Arunachalam, U., Massey, V. y Miller, S. 1994. *J. Biol. Chem.* 269, 150-155.
- Arunachalam, U., Massey, V. y Vaidyanathan, C.S. 1992. *J. Biol. Chem.* 267, 25848-25855.
- Assinder S.J. y Williams, P.A. 1988. *J. Gen. Microbiol.* 134, 2769-2778.
- Babbitt, P.C., Kenyon, G.L., Martin, B.M., Charest, H. y Sylvestre, M. 1992. *Biochemistry* 31, 5594-5604.
- Bachmann, B.J. 1990. *Microbiol. Rev.* 54, 130-197.
- Bak, F., y Widdel, F. 1986a. *Arch. Microbiol.* 146, 177-180.
- Bak, F., y Widdel, F. 1986b. *Arch. Microbiol.* 146, 170-176.
- Balba, M.T. y Evans, W.C. 1980a. *Biochem. Soc. Trans.* 8, 625-627.
- Balba, M.T. y Evans, W.C. 1980b. *Biochem. Soc. Trans.* 8, 452-454.
- Balba, M.T. y Evans, W.C. 1980c. *Biochem. Soc. Trans.* 8, 632-635.
- Balíce, V. y Cera, O. 1984. *Grasas y Aceites* 35, 178-180.
- Ballou, D.P. 1982. En *Flavins and Flavoproteins* ed. Massey, V. y Williams, C.H. Elsevier, Amsterdam. pp. 301-310
- Barbero, J.L., Buesa, J.M., Buitrago, G.G., Méndez, E., Pérez-Aranda, A. y García, J.L. 1986. *Gene* 49, 69-80.
- Batie, C.J., Lahaie, E. y Ballou, D.P. 1987. *J. Biol. Chem.* 262, 1510-1518.
- Beadle, C.A. y Smith, A.R.W. 1982. *Eur. J. Biochem.* 123, 323-332.

- Berkel., W.V., Westpha, A., Eschrich, K., Eppink, M. y Dekok, A. 1992. *Eur. J. Biochem.* 210, 411-419.
- Bernhardt, F.H., Pachowsky, H. y Staudinger, H. 1975. *Eur. J. Biochem.* 57, 241-256.
- Berry, D.F., Madsen, E.L. y Bollag, J. 1987. *Appl. Environ. Microbiol.* 53, 180-182.
- Bestetti, G., Galli, E., Ruzzi, M., Baldacci, G., Zennaro, E. y Frontali, L. 1984. *Plasmid* 12, 181-188.
- Biegert, T. Altenschmidt, U. Eckerskorn, C. y Fuchs, G. 1993. *Eur. J. Biochem.* 213, 555-561.
- Blake, C.K. y Hegeman, G.D. 1987. *J. Bacteriol.* 169, 4878-4783.
- Blakley, E.R. 1972. *Can. J. Microbiol.* 18, 1247-1255.
- Blakley, E.R., Kurz, W., Halvorson, H. y Simpson, F.J. 1967. *Can. J. Microbiol.* 13, 147-157.
- Blom, A., Harder, W. y Matin, A. 1992. *Appl. Environ. Microbiol.* 58, 331-334
- Bolívar, F. 1978. *Gene* 4, 121-136.
- Boruff, C.S. y Buswell, A.M. 1934. *J. Am. Chem. Soc.* 56, 886-888.
- Bossert, I. y Young, L.Y. 1986. *Appl. Environ. Microbiol.* 52, 1117-1122.
- Bottomley, P.J. 1993. *Curr. Opin. Biotechnol.* 4, 318-322.
- Boyd, S.A. y Shelton, D.R. 1984. *Appl. Environ. Microbiol.* 47, 272-277.
- Boyd, S.A., Shelton, D.R., Berry, D. y Tiedje, J.M. 1983. *Appl. Environ. Microbiol.* 46, 50-54.
- Bradford, M.M. 1976. *Anal. Biochem.* 72, 248-256.
- Brenner, D.J. 1984. En *Bergey's Manual of Systematic Bacteriology*. ed. Holt, J. G. Baltimor, London. vol. 1, pp. 408-516.
- Bugg, T.D.H. 1993. *Biochim. Biophys. Acta.* 1202, 258-264.
- Burlingame, R. y Chapman, P. J. 1983. *J. Bacteriol.* 155, 113-121.
- Burligame, R., Wyman, L. y Chapman, P. 1986. *J. Bacteriol.* 168, 55-64.
- Caron, J., Coffield, M. y Scott, J. 1989. *Proc. Natl. Acad. Sci. USA* 86, 963-967.
- Cieslak, P.R., Barrett, T.J., Griffin, P.M., Gensheimer, K.F., Beckett, G., Buffington, J. y Smith, M.G. 1993. *Lancet* 342, 8867.
- Clark, F.M. y Fina, L.R. 1952. *Arch. Biochem. Biophys.* 36, 26-32.
- Coco, W., Rothmel, R., Henikoff, S. y Chakrabarty, A.M. 1993. *J. Bacteriol.* 175, 417-427.

- Cole, M. 1964. *Nature* 203, 519-520.
- Cooper, R.A., Jones, D.C.N. y Parrott, S. 1985. *J. Gen. Microbiol.* 131, 2753-2757.
- Cooper, R.A., Knowles, P.F., Brown, D.E., McGuirl, M.A. y Dooley, D.M. 1992. *Biochem. J.* 288, 337-340.
- Cooper, R.A. y Skinner, M.A. 1980. *J. Bacteriol.* 143, 302-306.
- Cornelis, G., Sluiters, C., de Rouvroit, C.L. y Michielis, T. 1989. *J. Bacteriol.* 171, 254-262.
- Crawford, R.L. 1975. *Appl. Microbiol.* 30, 439-444.
- Crawford, R.L. Hutton, S.W. y Chapman, P.J. 1975. *J. Bacteriol.* 121, 794-799.
- Crooks, G.P. y Copley, S.D. 1993. *J. Am. Chem. Soc.* 115, 6422-6423.
- Chakrabarty, A.M. 1972. *J. Bacteriol.* 112, 815-823.
- Chandry, P.S., Davidson, B.E. y Hillier, A.J. 1994. *Microbiol.* 140, 2251-2261.
- Chatterjee, B.K. y Chakrabarty, A.M., 1983. *J. Bacteriol.* 153, 532-534.
- Chen, C.C. y Wu, C.M. 1984. *J. Food Sci.* 49, 1208-1209.
- Ching, L., Poh, C.L., Shingler, V. 1995. *J. Bacteriol.* 177, 1485-1490.
- Dangel, W., Brackmann, R., Lack, A., Mohamed, M., Kock, J., Oswald, B., Seyfried, B., Tschsch, A. y Fuchs, G. 1991. *Arch. Microbiol.* 155, 256-262.
- Daumy, G.O., Danley, D. y McColl, A.S. 1985. *J. Bacteriol.* 163, 1279-1281.
- de Andrés, C., y García, J.R. 1982. *Bol. Inform. Medio Ambiente* 22, 28-35.
- de Lorenzo, V. y Timmis, K. 1994. *Methods Enzymol.* 235, 386-405.
- de Lorenzo, V., Eltis, L., Kessler, B., y Timmis, K.N. 1993. *Gene* 123, 17-24.
- de Lorenzo, V., Herrero, M., Jacobzik, U., y Timmis, K.N. 1990. *J. Bacteriol.* 172, 6568-6572.
- de Lorenzo, V., Herrero, M., Metzke, M. y Timmis, K.N. 1991. *EMBO J.* 10, 1159-1169.
- de Lorenzo, V. y Rojo, F. 1994. En *Avances en ingeniería genética*. ed. CSIC. coord. Vicente, M. Madrid, España. pp. 333-358.
- Devi, N.A., Kutty, R.K., Vasantharajan, V.N. y Subba R.V. 1975. *J. Bacteriol.* 122, 866-873.
- Di Marco, A.A., Averhoff, B. A., Kim, E.E., y Ornston, N.L. 1993. *Gene* 125, 25-33
- Díaz, E., Munthali, M., de Lorenzo, V. y Timmis, K.N. 1994. *Mol. Microbiol.* 13, 855-861.

- Don, R.H. y Pemberton, J.M. 1981. *J. Bacteriol.* 145, 681-686.
- Dorn, E., Hellwig, M., Reineke, W. y Knackmuss, H.-J. 1974. *Arch. Microbiol.* 99, 61-70.
- Doten, R.C., Ngai, K.L., Mitchell, D.J. y Ornston, L.N. 1987. *J. Bacteriol.* 169, 3168-3174.
- Downing, R.G. y Broda, P. 1979. *Mol. Gen. Genet.* 168, 203-204.
- Duggleby, C.J., Bayley, S.A., Worsley, M.J., Williams, P.A. y Broda, P. 1977. *J. Bacteriol.* 130, 1274-1280.
- Duggleby, H.J., Tolley, S.P., Hill, C.P., Dodson, E.J., Dodson, G. y Moody, P.C.E. 1995. *Nature*, 373, 264-268.
- Dutton, P.L. y Evans, W.C. 1967. *Biochem. J.* 104, 30-31.
- Eggink, G., Enegel, H., Vriend, G., Terpstra, P. y Wilholt, B. 1990. *J. Mol. Biol.* 212, 135-142.
- Eiglmeier, K., Boos, W. y Cole, S. T. 1987. *Mol. Microbiol.* 1, 251-258.
- Elder, D.J.E. y Kelly, D.J. 1994. *FEMS Microbiol. Rev.* 13, 441-468.
- Elder, D.J.E., Morgan, P. y Kelly, D.J. 1992. *Arch. Microbiol.* 157, 148-154.
- Ensley, B.D. y Gibson, D.T. 1983. *J. Bacteriol.* 155, 505-511.
- Evans, W.C. y Fuchs, G. 1988. *Annu. Rev. Microbiol.* 42, 289-317.
- Fawcett, T., Garrido-Pertierra, A., y Cooper, R. A. (1989) *FEMS Microbiol. Lett.* 57, 307-312.
- Fernández-Moreno, M.A., Martínez, E., Boto, L., Hopwood, D.A. y Malpartida, F. 1992. *J. Biol. Chem.* 267, 19278-19290.
- Ferrer, E., y Cooper, R.A. 1988. *FEMS Microbiol. Lett.* 52, 155-160.
- Fiestas, J.A., Navarro, R., León, R., García A.J. y Maestrojuán, G. 1982. *Grasas y Aceites* 33, 265-270.
- Frantz, B. y Chakrabarty, M. A. 1986. En *The biology of Pseudomonas*. ed. J.R. Sokatch. Academic Press, Inc, New York. pp. 295-323.
- Furukawa, K. 1994. *Biodegradation* 5, 289-300.
- Gallegos, M.T., Michán, C. y Ramos J.L. 1993. *Nucl. Acids Res.* 4, 807-810.
- García, J.L. y Buesa, J.M. 1986. *J. Biotechnol.* 3, 187-195.
- Garrido-Pertierra, A., y Cooper, R. A. 1981. *Eur. J. Biochem.* 117, 581-584.
- Ghosal, D. y You, I. S. 1988. *Mol. Gen. Genet.* 211, 113-120.

- Gianotti, S. y Stefano, R. 1991. *Enotecnico* 27, 61-64.
- Gibello, A., Ferrer, E., Martín, M. y Garrido-Pertierra, A. 1994. *Biochem. J.* 301, 145-150.
- Gibson, D.T., Yeh, W.K., Liu, T.N. y Subramanian, V. 1982. En *Oxygenases and Oxygen Metabolism*, ed. Nozaki, M., Yamamoto, S. Ishimura, Y., Coon, M.J., Ernster, R.W., Estabrook, London. pp. 51-62.
- Gibson, K.J. y Gibson, J. 1992. *Appl. Environ. Microbiol.* 58, 696-698.
- Grant, D.J.W. 1967. *J. Gen. Microbiol.* 46, 213-222.
- Grbic-Galic, D. y Vogel, T.M. 1987. *Appl. Environ. Microbiol.* 53, 254-260.
- Harayama, S., Kok, M., y Neidle, L. E. 1992. *Annu. Rev. Microbiol.* 46, 565-601.
- Harayama, S. y Rekik, M. 1989. *J. Biol. Chem.* 264, 15328-15333.
- Harayama, S. y Rekik, M. 1993. *Mol. Gen. Genet.* 239, 81-89.
- Harayama, S., Rekik, M., Ngal, K.-L. y Ornston, N. 1989. *J. Bacteriol.* 171, 6251-6258.
- Harayama, S. y Timmis, K.N. 1992. En *Aerobic biodegradation of aromatics hydrocarbons by bacteria*. ed. Sigel H. y Sigel A., Marcel Dekker Inc., New York. vol. 28, pp 99-156.
- Hareland, W.A., Crawford, R.L., Chapman, P.J., y Dagley, S. 1975. *J. Bacteriol.* 121, 272-285.
- Harnett, C. Neidle, E.L., Ngai, K.L. y Ornston, L.N. 1990. *J. Bacteriol.* 172, 956-966.
- Harpel, M.R. y Lipscomb, J.D. 1990a. *J. Biol. Chem.* 265, 6301-6311.
- Harpel, M.R. y Lipscomb, J.D. 1990b. *J. Biol. Chem.* 265, 22187-22196.
- Harwood, C.S., Nichols, N.N., Kim, M.K., Ditty, J.L., y Parales, R.E. 1994. *J. Bacteriol.* 176, 6479-6488.
- Healy, J.B. y Young, L.Y. 1978. *Appl. Environ. Microbiol.* 35-216-218.
- Healy, J.B., Young, L.Y. y Reinhard, M. 1980. *Appl. Environ. Microbiol.* 39, 436-444.
- Heimhuber, B. y Herrmann, K. 1990. *Deutsche Lebensmittel Rundschau* 86, 205-209.
- Henikoff, S., Haughn, G.W., Calvo, H.M. y Wallace, J.C. 1988. *Proc. Natl. Acad. Sci. USA* 85, 6602-6606.
- Herrero, M., de Lorenzo, V. y Timmis, K. N. 1990. *J. Bacteriol.* 172, 6557-6567.
- Higson, F. 1992. *Adv. App. Microbiol.* 37, 2-19.
- Hofer, B., Backhaus, S., y Timmis, K. N. 1994. *Gene* 144, 9-16.

- Holtel, A., Abril, M.A., Marqués, S. y Timmis K.N. 1990. *Mol. Microbiol.* 4, 1551-1556.
- Hopper, D. J. y Chapman, P.J. 1970. *Biochem. J.* 122, 19-28.
- Hopper, D. J. y Kemp, P.D. 1980. *J. Bacteriol.* 142, 21-26.
- Horn, J.M., Harayama, S. y Timmis, K.N. 1991. *Mol. Microbiol.* 5, 2459-2474.
- Houghton, J.E., Brown, T.N., Appel, A.J. Hughes, E.J. y Ornston, L.N. 1995. *J. Bacteriol.* 177, 401-412.
- Howell, L.G., Spector, T. y Massey, V. 1972. *J. Biol. Chem.* 247, 4340-4350.
- Imai, R., Nagata, Y., Fukada, M., Tagaki, M. y Yano, K. 1991. *J. Bacteriol.* 173, 6811-6819.
- Inouye, S., Nakazagua, A. y Nakazagua, T. 1987. *Proc. Natl. Acad. Sci. USA* 84, 5182-5186.
- Inouye, S, Nakazawa, A. y Nakazawa, T. 1988. *Gene* 66, 301-306.
- Irie, S., Doi, S. Yorifuji, T., Takagi, M. y Yano, K. 1987. *J. Bacteriol.* 169, 5174-5179.
- Jain, R.K., Bayly, R.C. y Skurray, R.A. 1984. *J. Gen. Microbiol.* 30, 3019-3028.
- Jamaluddin, M.P. 1977. *J. Bacteriol.* 129, 690-697.
- Janer, L. 1980. *Grasas y Aceites* 31, 273-279.
- Janssen, D.B., Pries, F. y van der Ploeg, J.R. 1994. *Annu. Rev. Microbiol.* 48, 163-191.
- Jenkins, J., y Cooper, R.A. 1988. *J. Bacteriol.* 170, 5317-5324.
- Kabisch, M. y Fortnagel, P. 1990. *Nucl. Acids. Res.* 18, 3405-3406.
- Kaiser, J.P. y Haselmann, K.W. 1982. *Experientia* 38, 176.
- Kaphammer, B. y Olsen, R.H. 1990. *J. Bacteriol.* 172, 5856-5862.
- Kepner, R.E., Webb, A.D. y Maggiova, L. 1968. *Am. J. Enology Vitic.* 19, 116-120.
- Kessler, B., de Lorenzo, V. y Timmis, K. 1992. *Mol. Gen. Genet.* 233, 293-301.
- Kikuchi, Y., Yasukochi, Y., Nagata, Y., Fukuda, M., y Takagi, M. 1994. *J. Bacteriol.* 176, 4269-4276.
- Knudsen, S. M. y Karlström, O.H. 1991. *Appl. Environ. Microbiol.* 57, 85-92.
- Koch, J., Eisenreich, W., Bacher, A. y Fuchs, G. 1993. *Eur. J. Biochem.* 211, 649-661.
- Koch, J. y Fuchs, G. 1992. *Eur. J. Biochem.* 205, 195-202.
- Köhler, T., Harayama, S., Ramos, J.L. y Timmis, K.N. 1989. *J. Bacteriol.* 171, 4326-4333.
- Komine, Y., y Inokuchi, H. 1991. *J. Bacteriol.* 173, 1813-1816

- Krumholz, L.R. y Bryant, M.P. 1986. *Arch. Microbiol.* 144, 8-14.
- Kubota, T., Oki, Y., Uehara, H. y Haramaki, Y. 1989. *J. Agric. Chem. Soc. Jap.* 63, 49-50.
- Kuhad, R.C. y Singh, A. 1993. *Crit. Rev. Biotechnology* 13, 151-172.
- Kukor, J.J. y Olsen, R.H. 1990. *J. Bacteriol.* 172, 4624-4630.
- Kukor, J.J., Olsen, R.H. y Ballou, D.P. 1988. *J. Bacteriol.* 170, 4458-4465.
- Kushner, S. R. 1978. En *Genetic Engineering*. ed. Boyer, H.B. y Nicosia, S. Elsevier, Amsterdam. pp. 17.
- Kustu, S., North, A.K. y Weiss, S.D. 1991. *Trends Biochem. Sci.* 16, 397-402.
- Kyte, J. y Doolittle, R.F. 1982. *J. Mol. Biol.* 157, 105-132.
- Ladd, J.N. 1962. *Nature* 194, 1099-1100.
- Laemmli, U.K. 1970. *Nature* 227, 680-685.
- Lee, C.W. y Richard, J. 1984. *J. Dairy Res.* 51, 461-469.
- Lei, S.P., Lin, H.C., Heffernan, L. y Wilcox, G. 1985. *J. Bacteriol.* 164, 717-722.
- Leisinger, T. y Brunner, W. 1986. En *Biotechnology* ed. Schonborn, W. Editorial VCH, Berlin, vol. 8, pp. 475-513.
- Li, P. Wang, X.Q., Wang, H.Z. y Wu, Y.N. 1993. *Biomed. Environ. Sci.* 6, 389-398.
- Lin, E. C. C. 1987. En *Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology*. ed. Neidhart, FC. Am. Soc. Microbiol., Washington D.C. pp. 244-284.
- Locher, H.H., Leisinger, T. y Cook, A.M. 1991. *Biochem. J.* 274, 833-842.
- MacNeil, D. 1981. *J. Bacteriol.* 146, 260-268.
- Manja K.S., Kaul, R.K. y Rao, K.M. 1992. *Indian J. Exp. Biol.* 30, 823-825.
- Manson, J.R. y Cammack, R. 1992. *Annu. Rev. Microbiol.* 46, 277-305.
- Marger, M.D. y Saier, M.H., Jr. 1993. *Trends Biochem.* 18, 13-20.
- Margolin, A.L., Svedas, V.K. y Berezin, I.V. 1980. *Biochem. Biophys. Acta.* 616, 283-289.
- Markus, A., Krekel, D. y Lingens, F. 1986. *J. Biol. Chem.* 261, 12883-12888.
- Marqués, S. y Ramos, J.L. 1993. *Mol. Microbiol.* 9, 923-929.
- Martín, M., Gibello, A., Fernández, J., Ferrer, E. y Garrido-Pertierra, A. 1991. *J. Gen. Microbiol.* 132, 621-628.

- Martínez, J., Pérez, J., Moreno, E. y Ramos-Cormenzana, A. 1986. *Grasas y Aceites* 37, 215-223.
- Martínez, L. Ramos, A., García M.P. y Garrido, S.E. 1992. *Grasas y Aceites* 43, 75-81.
- Merino, E. Balbas, P. Recillas, F., Becerril, B., Valle, F. y Bolivar, F. 1992. *Mol. Microbiol.* 6, 2175-2182.
- Mermod, N., Ramos, J.L., Bairoch, A. y Timmis, K.N. 1987. *Mol. Gen. Genet.* 207, 349-354.
- Michán, C., Busby, W. y Hyde, E. 1995. *Nucl. Acids Res.* 23, 1518-1523.
- Miller, J.H. 1972. *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Miyada, C.G., Horwith, A.H., Cass, L.G., Timko, J. y Wilcox, G. 1990. *Nucl. Acids Res.* 8, 5267-5274.
- Mohamed, M. y Fuchs, G. 1993. *Arch. Microbiol.* 159, 554-562.
- Mohamed, M., Seyfried, B., Tschsch, A. y Fuchs, G. 1993. *Arch. Microbiol.* 159, 563-573.
- Molin, S. Klem, P., Poulsen, K.L., Biehl, H., Gerdes, K. y Anderson, P. 1987. *Bio/Technology*, 5, 1315-1318.
- Monticello, J. D. y Finnerty, W.R. 1985. *Annu. Rev. Microbiol.* 39, 371-389.
- Mountfort, D.O. y Bryant, M.P. 1982. *Arch. Microbiol.* 133, 249-256.
- Nakai, C., Kagamiyama, H., Nozaki, M., Nakazawa, H. e Inouye, S. 1983. *J. Biol. Chem.* 258, 2923-2928.
- Nakazawa, T. 1978. *J. Bacteriol.* 133, 527-535.
- Nakazawa, T. y Yokota, T. 1973. *J. Bacteriol.* 115, 262-267.
- Nebert, D.W. y González, F.J. 1987. *Annu. Rev. Biochem.* 56, 945-993.
- Neidle, E.L., Hartnett, C., Bonitz, S. y Ornston, L.N. 1988. *J. Bacteriol.* 170, 4874-4880.
- Neidle, E.L., Hartnett, C. y Ornston, L.N. 1989. *J. Bacteriol.* 171, 5410-5421.
- Neidle, E.L., Hartnett, C., Ornston, L.N., Bairoch, A., Rekik, M., y Harayama, S. 1991. *J. Bacteriol.* 173, 5385-5395.
- Neilson, A.H., Allard, A.S., Lindgren, C. y Remberger, M. 1987. *Appl. Environ. Microbiol.* 53, 2511-2519.
- Neujahr, H.Y. y Gaal, A. 1973. *Eur. J. Biochem.* 35, 386-400.

- Ninfa, A.J., Ninfa, E.G., Lupas, A.N., Stock, A., Magasanik, B. y Stock, J. 1988. *Proc. Natl. Acad. Sci. USA* 85, 5492-5496.
- Nishiyama, M., Horinouchi, S., Kobayashi, M., Nagasawa, T., Yamada, H. y Beppu, T. 1991. *J. Bacteriol.* 173, 2465-2472.
- Nomura, Y., Nakagawa, M., Ogawa, N., Harashima, S. y Oshima, Y. 1992. *J. Ferment. Bioeng.* 74, 333-344.
- Nordlund, I., Powlowski, J. y Shingler, V. 1990. *J. Bacteriol.* 172, 6826-6833.
- Nordlund, I. y Shingler, V. 1990. *Biochim. Biophys. Acta.* 1049, 277-230.
- Nurk, A., Kasak, L. y Kivisaar, M. 1991. *Gene* 102, 13-18.
- Oh, S.-J., Kim, Y.-C., Park, Y.-W., Min, S.-Y., Kim, I.-S., y Kang, H.-S. 1987. *Gene* 56, 87-97.
- Orser, C.S., Dutton, J., Lange, C., Jablonski, P., Xun, L. y Hargis M. 1993a. *J. Bacteriol.* 175, 2640-2644.
- Orser, C.S., Dutton, J., Lange, C. Xun, L., Zhart, T.C. y Schneider, B.J. 1993b. *J. Bacteriol.* 175, 411-416.
- Parales, R.E., y Harwood, C.S. 1993. *J. Bacteriol.* 175, 5829-5838.
- Parker, L.L. y Hall, B.G. 1990. *Genetics* 124, 455-471.
- Parrott, S., Jones, S. y Cooper, R.A. 1987. *J. Gen. Microbiol.* 133, 347-351.
- Perkins, E.J., Gordon, M.P., Caceres, O. y Lurquin, P.F. 1990. *J. Bacteriol.* 172, 2351-2359.
- Powlowski, J. y Shingler, V. 1994. *Biodegradation* 5, 219-236.
- Proctor, M.H. y Scher, S. 1960. *Biochem. J.* 76, 33.
- Que, L., Windom, J. Crawford, R.L. 1981. *J. Biol. Chem.* 256, 10941-10944.
- Rajasekharan, S., Rajasekharan, R. y Vaidyanathan, C.S. 1990. *Arch. Biochem. Biophys.* 278, 21-25.
- Raju, S. G., Kamath, A. V., y Vaidyanathan, C. S. 1988. *Biochem. Biophys. Res. Commun.* 154, 537-543.
- Ramos, J. L., Díaz, E., Dowling, D., de Lorenzo, V., Molin, S., O'Gara, F., Ramos, C., y Timmis, K.N. 1994. *Bio/Technology* 12, 1349-1356.
- Ramos, J. L., González-Carrero, M. y Timmis, K. N. 1988. *FEBS Lett.* 226, 241-246.
- Ramos, J.L., Rojo, F., Zhou, L. y Timmis, K.N. 1990. *Nucl. Acids Res.* 18, 2149-2152.

- Ramos, J.L., Stolz, A., Reineke, W. y Timmis K.N. 1986. *Proc. Natl. Acad. Sci. USA.* 83, 8467-8470.
- Ramos-González, M.I., Duque, E., y Ramos, J.L. 1991. *Appl. Environ. Microbiol.* 57, 3020-3027.
- Reineke, W. y Knackmuss, H.J. 1984. *Appl. Environ. Microbiol.* 47, 395-402.
- Reineke, W. y Knackmuss, H.J. 1979. *Nature* 277, 385-386.
- Roa, A., García, J.L., Salto, F. y Cortés, E. 1994. *Biochem. J.* 303, 869-876.
- Rojo, F., Piepper, D.H., Engesser, K.H., Knackmuss, H. J. y Timmis, K. N. 1987. *Science.* 238, 1395-1398.
- Roper, D.I., y Cooper, R.A. 1990a. *FEBS Lett.* 275, 53-57.
- Roper, D.I., y Cooper, R.A. 1990b. *FEBS Lett.* 266, 63-66.
- Roper, D.I., y Cooper, R.A. 1993. *Eur. J. Biochem.* 217, 575-580.
- Roper, D.I., Fawcett, T., y Cooper, R.A. 1993. *Mol. Gen. Genet.* 237, 241-250.
- Roper, D.I., Stringfellow, J.M. y Cooper, R.A. 1995. *Gene* 156, 47-51.
- Rose, R.E. 1988. *Nucl. Acids. Res.* 16, 355.
- Ross, G. W. y O'Callaghan. 1975. *Methods Enzymol.* 43, 69-85.
- Rossmann, M.G., Moras, D., y Olsen, K.W. 1974. *Nature* 250, 194-199.
- Rothmel, R.K., Aldrich, T.L., Houghton, J.E., Coco, W.M., Ornston, L.N. y Chakrabarty, A.M. 1990. *J. Bacteriol.* 172, 922-931.
- Rudolphi, A., Tschsch, A. y Fuchs, G. 1991. *Arch. Microbiol.* 155, 238-248.
- Saavedra, J.M. 1989. En: *Catecholamines II.* ed. Trendelenburg, U. y Weiner, N. pp 182-201.
- Sambrook, J., Frisch, E.F., y Maniatis, T. 1989. *Molecular cloning: A laboratory manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Samid, D., Ram, Z., Hudgins, W.R., Shack, S., Liu, L., Walbridge, S., Oldfield, E. H. y Myers, C.E. 1994. *Cancer Res.* 54, 891-895.
- Sanger, T., Nicklen, S. y Coulson, A.R. 1977. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
- Sauber, K., Frohner, C., Rosenberg, G., Eberspacher, J. y Lingens, F. 1977. *Eur. J. Biochem.* 74, 89-97.

- Schaechter, M. y Neidhart, F.C. 1987. En *Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology*. ed. Neidhart, FC. Am. Soc. Microbiol., Washington D.C. pp. 1-2.
- Schell, M.A., 1985. *Gene* 36, 301-309.
- Schell, M.A. 1993. *Annu. Rev. Microbiol.* 47, 597-626.
- Schell, M.A. y Sukordhaman, M. 1989. *J. Bacteriol.* 171, 1952-1959.
- Schink, B. y Pfennig, N. 1982. *Arch. Microbiol.* 133, 195-201.
- Schneider, B., Müller, R., Frank, R. y Lingens, F. 1991. *J. Bacteriol.* 173, 1530-1535.
- Schultz, J.E. y Matin, A. 1991. *J. Mol. Biol.* 218, 129-140.
- Shingler, V., Bartilson, M. y Moore, T. 1993 *J. Bacteriol.* 175, 1596-1604.
- Shingler, V., Powlowski, J. y Marklund, U. 1992. *J. Bacteriol.* 174, 711-724.
- Silhavy, T.J., Berman, M.L. y Enquist, L.W. 1984. *Experiments in Gene Fusions*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Simons, R. W., Houman, F. y Kleckner, N. 1987. *Gene* 53, 85-96.
- Skinner, M.A., y Cooper, R.A. 1982. *Arch. Microbiol.* 132, 270-275.
- Sleat, R. y Robinson, J.P. 1983. *J. Microbiol.* 129, 141-152.
- Smith, M.R. 1990. *Biodegradation* 1, 191-206.
- Southern, E. N. 1975. *J. Mol. Biol.* 98, 507-517.
- Spain, J.C., y van Veld. P.A. 1983. *Appl. Environ. Microbiol.* 45, 428-435.
- Sparnins, V.L. y Chapman, P.J. 1976. *J. Bacteriol.* 127, 362-366.
- Sparnins, V.L., Chapman, P.J., y Dagley, S. 1974. *J. Bacteriol.* 120, 159-167.
- Spoelstra, S.F. 1978. *Appl. Environ. Microbiol.* 36, 631-638.
- Stanier, R.Y., Adelberg, E.A. y Ingraham, J.L. 1986. En *Microbiologia*. Editorial Reverté S.A. pp. 686.
- Steeg, E. y Montag, A. 1988. *Zeitsch. Fuer Lebensmittel Unters. Forsch.* 187, 115-120.
- Stellwagen, E. y Baker, B. 1976. *Nature* 261, 719-720.
- Stern, M.J., Ferro-Luzzi, G., Smith, N.J., Robinson, E.C. y Higgins, C.F. 1984. *Cell* 37, 1015-1026.
- Stoner, C.M. y Schleif, R.P. 1982. *J. Mol. Biol.* 154, 649-652.
- Streiwieser, A. y Heathcock, C.H. 1979. En *Química Orgánica*. pp 569-592. Ed. Interamericana.

- Suárez, M., Gibello, A., Allende, J.L., Martín, M., Ferrer, E. y Garrido-Pertierra, A. 1991. *Appl. Microbiol. Biotechnol.* 34, 677-682.
- Suflita, J.M. y Gibson, S.A. 1985. Proc. 2nd Int. Conf. Ground Water Qual. Res., ed N.N. Durham, A.E. Redelfs. Stillwater, Okla: Univ. Cent. Water Res. Okla. State Univ. pp.30-32
- Suflita, J.M., Horowitz, A., Schelton, D.R. y Tiedje, J.M. 1982. *Science* 218,1115-1117.
- Suzuki, M., Hayakawa, T., Shaw, J.P., Rekik, M. y Harayama, S. 1991. *J. Bacteriol.* 173, 1690-1695.
- Swank, R.T. y Munkres, K.D. 1971. *Anal. Biochem.* 39, 462-477.
- Swindol, C.M., Aelion, C.M. y Pfaender, F.K. 1988. *Appl. Environ. Microbiol.* 54, 212-217.
- Szewczuk, A. Siewinski, M. y Slowinska, R. 1979. *Anal. Biochem.* 103, 166-169.
- Taira, K., Hayase, N., Arimura, N., Yamashita, S., Miyazaki, T. y Furukawa, K. 1988. *Biochemistry* 27, 3990-3996.
- Tarvin, D. y Buswell, A.M. 1934. *J. Am. Chem. Soc.* 56, 1751-1755.
- Taylor, B.F., Campbell, W.L. y Chinoy, I. 1970. *J. Bacteriol.* 102, 430-437.
- Thacker, R. y Gunsalus, I.C. 1979. *J. Bacteriol.* 137, 697-699.
- Thibault, A., Cooper, R. M., Figg, W. D., Venzon, D. J., Sartor, A. O., Tompkins, A. C., Weinberger, M. S., Headlee, D. J., McCall, N. A., Samid, D. y Myers, C. E. 1994. *Cancer Res.* 54, 1690-1694.
- Timmis, K., Steffan, R.J. y Untermann, R. 1994. *Annu. Rev. Microbiol.* 48, 525-57.
- Tobin, J.F. y Schleif, R.F. 1990. *J. Mol. Biol.* 211, 75-89.
- Tobin, J.F. y Schleif, R.F. 1987. *J. Mol. Biol.* 196, 789-799.
- Tomasek, P.H. y Crawford, R.L. 1986. *J. Bacteriol.* 167, 818-827.
- Tsai, D.G., y Jones, G.A. 1975. *Can. J. Microbiol.* 21, 794-801.
- Tsai, Y.L. Olson, B.H. 1992. *Appl. Environ. Microbiol.* 58, 754-757.
- Tschech, A. y Fuchs, G. 1987. *Arch. Microbiol.* 148, 213-217.
- Tschech, A. y Schink, B. 1985. *Arch. Microbiol.* 21, 794-802.
- Tsuda, M., Minegishi, K. e Iino, T. 1989. *J. Bacteriol.* 171, 1386-1393.
- Valle, F., Balbás, P., Merino, E. y Bolívar, F. 1991. *Trends Biochem.* 16, 36-40.
- van den Tweel, W.J.J., Smits, J.P., y de Bont, J.A.M. 1988. *Arch. Microbiol.* 149, 207-213.

- van der Lelie, D., Corbisier, P., Baeyens, W., Wuertz, S., Diels, L. y Mergeay M. 1994. *Res. Microbiol.* 145, 67-74.
- van der Meer, J.R., de Vos, W.M., Harayama, S. y Zehnder, A.J.B. 1992. *Microbiol. Rev.* 56, 677-694.
- van der Meer, J.R., Eggen, R.I.L., Zehnder, A.J.B., y de Vos, W.M. 1991b. *J. Bacteriol.* 173, 2425-2434.
- van der Meer, J.R., Frijters, A.C.J., Leveau, J.H.J, Eggen, R.I.L., Zehnder, A.J.B., y de Vos, W.M. 1991a. *J. Bacteriol.* 173, 3700-3708.
- Vandame, E. J. 1985. *Econom. Microbiol.* 5, 467-522.
- Vitovski, S. 1993. *FEMS Microbiol.* 108, 1-6.
- Waite-Rees, P. A., Keating, C. J., Moran, L. S., Slatko, B. E., Hornstra, L. J. y Benner, J. S. 1991. *J. Bacteriol.* 173, 5207-5219.
- Walker, N. y Lippert, K. D. 1965. *Biochem. J.* 95, 5c-6c.
- Wang, L.H., Hamzah, R.Y., Yu, Y., y Tu, S.C. 1987. *Biochemistry.* 26, 1099-1104.
- Wang, Y. Rawlings, M. Gibson, D. T., Labbé, D, Bergeron, H., Brousseau, R. y Lau, P.C.K. 1995. *Mol. Gen. Genet.* 246, 570-579.
- Webster, C., Gardner, L. y Busby, S. 1989. *Gene* 83, 207-213.
- Webster, C., Gaston, K., y Busby, S. 1988. *Gene* 68, 297-305.
- Whittich, R., Rast, H.G. y Knackmuss, H.J. 1988. *Appl. Environ. Microbiol.* 54, 1842-1847.
- Whittle, P. J., Lunt, D.O. y Evans, W.C. 1976. *Biochem. Soc. Trans.* 4, 490-491.
- Wierenga, R.K., Terpstra, P. y Hol, W.G. J. 1986. *J. Mol. Biol.* 187, 101-107.
- Williams, P.A. y Murray, K. 1974. *J. Bacteriol.* 120, 416-423.
- Williams, P.A. y Sayers, J.R. 1994. *Biodegradation* 5, 195-217.
- Williams, R. J. y Evans, W.C. 1975. *Biochem. J.* 143, 1-10.
- Wirth, R., Friesenegger, A. y Stefan, F. 1989. *Mol. Gen. Genet.* 216, 175-177.
- Wyndham, R., Cashore, A. E., Nakatsu, C. H. y Peel, M. C. 1994. *Biodegradation* 5, 323-342.
- Yamaguchi, M. y Fujisawa, H. 1982. *J. Biol. Chem.* 257, 12497-12502.
- Yanisch-Perron, C., Vieira, J. y Messing, J. 1985. *Gene* 33, 103-119.
- Yavas, I. y Rapp, A. 1992. *Fluessiges Oest* 59, 472-476.
- Yen, K.M., Sullivan, M. y Gunsalus, I.C. 1983. *Plasmid* 9, 105-111.

- Yen, K.M., Karl, M.R., Blatt, L.M., Simon, M.J. y Winter, R.B. 1991. *J. Bacteriol.* 173, 5315-5327.
- You., I.S., Ghosal, D. y Gunsalus, I.C. 1991. *Biochemistry* 30, 1635-1641.
- Zeyer, J. y Kearney, P. C. 1984. *J. Agric. Food Chem.* 31, 304-308.
- Zeyer, J. y Kocher, H.P. 1988. *J. Bacteriol.* 170, 1789-1794.
- Ziegler, K. Braun, K., Böckler, A. y Fuchs, G. 1987. *Arch. Microbiol.* 149, 62-69.
- Zylstra, G.J. y Gibson, D.T. 1989. *J. Biol. Chem.* 264, 14940-14946.